

**University of Kelaniya**  
**Sri Lanka**

**PROCEEDINGS**

IEEE International Research Conference on  
Smart Computing and Systems Engineering  
(SCSE 2019)

28<sup>th</sup> March 2019

*Department of Industrial Management,  
Faculty of Science, University of Kelaniya, Sri Lanka*

© University of Kelaniya, Sri Lanka

Proceedings of the IEEE International Research Conference on Smart  
Computing and Systems Engineering  
SCSE 2019

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of the publisher.

ISSN 2613-8662

***Published by***

Department of Industrial Management  
Faculty of Science, University of Kelaniya  
Kelaniya, Sri Lanka

# Contents

	<i>Page</i>
Editorial Board	v
Programme Committee	vii
Organizing Committee	ix
Keynote Speech	
<i>Professor Athula Ginige</i>	xi
<i>Professor Chin-Lin Hu</i>	xiii
List of Papers	xv
Smart Computing	1
Systems Engineering	184



## **EDITORIAL BOARD**

### **IEEE International Research Conference on Smart Computing and Systems Engineering - 2019**

**Chief Editor**                      Dr. Keerthi Wijayasiriwardhane

**Committee**                        Dr. Suren Peter

Dr. Shantha Jayalal

Dr. Chathura Rajapakse

Dr. Dilani Wickramarachchi



## PROGRAMME COMMITTEE

### IEEE International Research Conference on Smart Computing and Systems Engineering - 2019

<b>Prof. Athula Ginige</b>	Western Sydney University, Australia
<b>Prof. Chin-Lin Hu</b>	National Central University, Taiwan
<b>Dr. Mats Isaksson</b>	Swinburne University of Technology, Australia
<b>Associate Prof. Masakazu Takahashi</b>	Yamaguchi University Management of Technology, Japan
<b>Prof. Setsuya Kurahashi</b>	University of Tsukuba, Japan
<b>Dr. Muhammed Badruddin Khan</b>	Al Imam Mohammad ibn Saud Islamic University(IMSUI), Riyadh, KSA
<b>Dr. Irvan Mhd</b>	Tokyo Institute of Technology, Japan
<b>Prof. Takao Terano</b>	Tokyo Institute of Technology, Japan
<b>Dr. Raj Prasanna</b>	Massey University, New Zealand
<b>Dr. Prem Samaranayake</b>	Western Sydney University, Australia
<b>Dr. Malathi Sajeewani</b>	University of Melbourne, Australia
<b>Dr. Antonio Hyder</b>	Hackers and Founders Research, United States
<b>Dr. Ajantha Athukorala</b>	University of Colombo School of Computing, Sri Lanka
<b>Dr. Dhammika Elkaduwe</b>	University of Peradeniya, Sri Lanka
<b>Prof. Janaka Wijayanayake</b>	University of Kelaniya, Sri Lanka
<b>Dr. Lalith Goonatilake</b>	Former Director - Trade Capacity Building, UNIDO, Austria
<b>Dr. Julian Nanayakkara</b>	University of Kelaniya, Sri Lanka
<b>Dr. Annista Wijayanayake</b>	University of Kelaniya, Sri Lanka
<b>Dr. Keerthi Wijayasiriwardhane</b>	University of Kelaniya, Sri Lanka
<b>Dr. Jeewanie Jayasinghe</b>	University of Ruhuna, Sri Lanka
<b>Dr. Suren Peter</b>	University of Kelaniya, Sri Lanka

<b>Dr. Shantha Jayalal</b>	University of Kelaniya, Sri Lanka
<b>Dr. Thashika Rupasinghe</b>	University of Kelaniya, Sri Lanka
<b>Dr. Nirosinie Fernando</b>	Deakin University, Australia
<b>Dr. Ruwan Wickramarachchi</b>	University of Kelaniya, Sri Lanka
<b>Dr. Chathura Rajapakshe</b>	University of Kelaniya, Sri Lanka
<b>Dr. Dilani Wickramarachchi</b>	University of Kelaniya, Sri Lanka
<b>Dr. Sachintha Pitigala</b>	University of Kelaniya, Sri Lanka



## **ORGANIZING COMMITTEE**

### **IEEE International Research Conference on Smart Computing and Systems Engineering - 2019**

<b>Conference Chair</b>	Dr. Shantha Jayalal
<b>Co-chairs</b>	Dr. (Ms.) Dilani Wickramarachchi Dr. Chathura Rajapakse
<b>Conference Secretary</b>	Ms. Kasuni Vidanagamachchi
<b>Members</b>	Dr. Ruwan Wickramarachchi Mr. Buddhika Jayawardena Mr. Nirmala Liyanaarachchi Ms. Samantha Weerabahu Ms. Samya Dimithrie



## *Keynote Speech*

# **Smart Computing and Systems Thinking**

**Professor Athula Ginige**

*Professor of Information Technology  
Western Sydney University (WSU), Australia  
A.Ginige@westernsydney.edu.au*

Smart Computing aims to combine advances in Information and Communication Technologies to create smart systems to make human life better, thus providing a new approach to address many complex and challenging problems faced today by humanity.

Initially people used these advances in ICT to optimize various routine activities commencing with scientific calculations, payrolls and evolving into robots in car assembly lines. Growth of the Internet and resulting connectivity made people to explore ways to optimize processes we use in businesses, as well as daily life. Applications such as Enterprise Resource Planning, Supply chain management, Customer relationship management started to emerge, coupled with new e-commerce models. With rapid growth of mobile phone usage, we started to evolve into a digitally connected society. This has given us opportunity to optimize how we fulfil human needs in a sustainable way at a System level.

Today, humanity is facing many challenges. To facilitate a coordinated approach to address this humanitarian challenge, the United Nations have formulated 17 Sustainable Development Goals (SDGs). The first three being “No Poverty”, “Zero Hunger” and “Good Health and Well-Being”.

We need Systems thinking to develop Smart Computing solutions to face these challenges. We cannot isolate the components and develop these separately. We need to treat these as set of interacting components and develop system level solutions. A very important aspect is to include the user, the human being also as part of the system. Over time, the users’ ability and thinking will evolve. Hence, the system also needs to evolve with such changes happening in the system.

To achieve first three SDGs, effectively and efficiently, the agriculture sector becomes essential, especially for developing countries. Due to multi-level co-ordination failures in agriculture domain in these countries, farmers often fail to get the best possible harvest and about 40% of what is produced is wasted. Our research has shown that the root cause for this co-ordination failure is that all stakeholders in the agriculture domain not being able to get required information at the time of making informed decisions. To address this issue, we initially developed a mobile based Smart Computing solution to enhance the flow of information and empower farmers to make informed decisions. This solution was then extended to all stakeholders in the agriculture domain making it a Digital Knowledge Ecosystem. This system is now evolving to include effective co-ordination of 4 key processes in the agriculture domain; providing finance, input aggregation, market linkages in addition to timely delivery of scientific knowledge, thus becoming a Digital Agribusiness Ecosystem.

From the insights gained from this research, we have developed a Smart Computing Framework based on Systems thinking for achieving Sustainable Development Goals.



## *Keynote Speech*

# **Smart Home with Soft-Intelligence & IoT Technology**

**Professor Chih-Lin Hu**

*Associate Professor*

*Department of Communication Engineering, National Central University, Taiwan (R.O.C.)  
clhu@ieee.org*

Smart home computing focuses on “home” and combines related peripheral industries to jointly, promote smart home services. These include home care, security monitoring, smart appliances and home entertainment, and other integrated services needed to build a safe and secure, energy-saving and sustainable, health care and comfortable, home service. Though ten years have passed, even now, there is no clear definition of “intelligence (or smart).” Philips, a consumer electronics company, believes that intelligence should include sensing, connected, learning, adaptation, and ease to interact. Smart applications and services are still in the development stage; they are not all capable of these five characteristics.

In this presentation titled “Smart Home with Soft-Intelligence & IoT Technology,” many scenarios that differ from the traditional usage which requires the tedious manual operations, thanks to the advance of software intelligence and IoT Technologies. From my personal point of view, I will propose four technical directions to build a "smart home," which I personally call "Enabling Technologies and Effects": home networking, context awareness, recognition, and user interface. For ease of understanding, the speaker will present a set of practical designs, developments, and technologies for enabling smart home services with various networked devices and sensor devices in home computing contexts as well as the resource and services on the cloud. It is believed that the prospect of smart home services could provide better user experience, and inspires new and future applications and services in ubiquitous home network environments.



## LIST OF PAPERS

Smart Computing	1 - 183
Systems Engineering	184 - 253





## Theme 1: Smart Computing

SC-01	<b>An infectious disease medical policy simulation and gaming</b> Setsuya Kurahashi	1
SC-02	<b>Multimodal user interaction framework for e-commerce</b> Supun Hewawalpita and Indika Perera	9
SC-03	<b>Prediction of user intentions using Web history</b> K. T. C. S. Samarasinghe and S. Jayalal	17
SC-04	<b>Occupancy monitoring system for workplace washrooms</b> V. M. P. Godakandage, K. R. M. Kothalawala, E. J. A. Chathumali and A. W. Madhubhashana	24
SC-05	<b>Candidate recruitment based on automatic answer evaluation using WordNet</b> J. Kasundi and G. U. Ganegoda	29
SC-06	<b>Implicit intention and activity recognition of a human using neural networks for a service robot eye</b> M. W. C. N. Moladande and B. G. D. A. Madhusanka	38
SC-07	<b>Face and upper-body emotion recognition using service robot's eyes in a domestic environment</b> T. M. W. Vithanawasam and B. G. D. A. Madhusanka	44
SC-08	<b>MRI based Glioma segmentation using Deep Learning algorithms</b> H. N. T. K. Kaldera , S. R. Gunasekara and M. B. Dissanayake	51
SC-09	<b>Data mining approach for identifying suitable sport for beginners</b> P. T. Amarasena, B. T. G. S. Kumara and S. Jointion	57
SC-10	<b>Distribution cost optimization using Big Data Analytics, Machine Learning and Computer Simulation for FMCG Sector</b> A. M. C. Adikari and T. P. Amalan	63
SC-11	<b>A scalable bioinformatics analysis platform based on Microservices Architecture</b> S. Rajapaksa, A. Wickramarachchi, V. Mallawaarachchi, W. Rasanjana, I. Perera and D. Meedeniya	70
SC-12	<b>Novel machine learning ensemble approach for landslide prediction</b> C. N. Madawala , B. T. G. S. Kumara and L. Indrathilaka	78
SC-13	<b>Communication framework for vehicular ad-hoc networks using Blockchain: A case study of Metro Manila Electric Shuttle automation project</b> A. S. Kulathunge and H. R. O. E. Dayarathna	85

SC-14	<b>A novel computational method to capture FPGA technology trends from Patent information</b> Masashi Shibata and Masakazu Takahashi	91
SC-15	<b>Use of LIME for human interpretability in Sinhala document classification</b> P. K. S. Kumari and P. S. Haddela	97
SC-16	<b>A Blockchain-based decentralized system to ensure the transparency of organic food supply chain</b> B. M. A .L. Basnayake and C. Rajapakse	103
SC-17	<b>Part of speech tagging for Twitter conversations using Conditional Random Fields model</b> A. R. F. S. Fanoon and G. A. I. Uwanthika	108
SC-18	<b>Language identification at word level in Sinhala-English code-mixed social media text</b> Kasthuri Shanmugalingam and Sagara Sumathipala	113
SC-19	<b>VoiceNote: An intelligent tool for monetary transactions with integrated voice support</b> Buddhima Hashan, Yoshani Abeyrathna, Mathisha Kaluaratchi and Samantha Thelijjagoda	119
SC-20	<b>Passion Fruit disease detection using dmage processing</b> S. B. D. H. Dharmasiri and S. Jayalal	126
SC-21	<b>Sinhala handwritten postal address recognition for postal sorting</b> M. F. A. Ifhaam and S. Jayalal	134
SC-22	<b>Devise best practices for mobile application development: Addressing the challenges in designing user interfaces</b> W. P. G. M. Weerasinghe and D. N. Wickramaarachchi	142
SC-23	<b>Software usability improvements for Generation Z oriented software application</b> W. P. N. H. Pathirana and D. N. Wickramaarachchi	151
SC-24	<b>Green Cloud Computing: A review on adoption of Green-Computing attributes and vendor specific implementations</b> J. M. T. I. Jayalath, E. J. A. P. C. Chathumali, K. R .M. Kothalawala and N. Kuruwitaarachchi	158
SC-25	<b>Linguistics analytics in data warehouses using fuzzy techniques</b> P. P. G. D. Asanka and A. S. Perera	165
SC-26	<b>An application of transfer learning techniques in identifying herbal plants in Sri Lanka</b> Y. R. Azeez and C. Rajapakse	172
SC-27	<b>Impact of emerging motorbike swarms on transportation requirements in Sri Lanka: An agent based simulation study</b> E. R. Kulathilake and C. Rajapakse	179

## Theme 2: Systems Engineering

SE-01	<b>GPS guided auto sensing system for motor vehicles</b> D. C. R Ranasinghe, W. K. I. L. Wanniarachchi and U. A. D. N. Anuradha	184
SE-02	<b>Industry 4.0 maturity assessment of the Banking Sector of Sri Lanka</b> O. K. K. Bandara, V. K. Tharaka and A. P. R. Wickramarachchi	190
SE-03	<b>Integrating smart transportation system for a proposed Smart City - A mapping study</b> H. K. S. K. Hettikankanama and S. Vasanthapriyan	196
SE-04	<b>Modelling &amp; simulation of the hydration process of cement for rapid concrete constructions</b> S. M. D. T. Sandaruwan and A. Pallegedera	204
SE-05	<b>A simulation model to analyse Sri Lankan Megacity logistics behaviour: Megapolis logistics of Sri Lanka</b> U. G. M. D. Dias , V. K. Tharaka and J. N. Nanayakkara	212
SE-06	<b>Factors influencing Enterprise Information Systems adoption of Small and Medium Enterprises (SMEs): A case study on Sri Lankan manufacturing sector</b> T. R. Kaluarachchi, W. M. S. K. Weerabahu and L. D. J. F. Nanayakkara	220
SE-07	<b>Process aligned business rule framework for immunization services in Sri Lanka</b> Gilmini Dantanarayana , Prasad Jayaweera and Anusha Walisadeera	228
SE-08	<b>Automated, low-cost pedestrian crossing carriage for efficient traffic control and pedestrian safety</b> B. A. O. S. Yasaswin and T. M. K. K. Jinasena	234
SE-09	<b>Analysis of Factors Influencing the Virtual Learning Environment in a Sri Lankan higher studies institution</b> Rajaratnam Charanya and Manoharan Kesavan	240
SE-10	<b>Predictive Maintenance 4.0 as next evolution step in industrial maintenance development</b> P. Poor, J. Basl and D. Zenisek	245



# **Smart Computing**



## An infectious disease medical policy simulation and gaming

Setsuya Kurahashi  
Graduate School of Business Sciences  
University of Tsukuba  
Tokyo, Japan  
kurahashi.setsuya.gf@u.tsukuba.ac.jp

### Abstract

*This paper analyses a new type of infectious disease by an agent-based simulation and gaming model based on Ebola fever and dengue fever. The mathematical model such as SIR (Susceptible, Infected, Recovered) has been used to model these infectious diseases. Besides, a simulation and gaming model enables to represent the decision-making of each citizen on the computer, and also reveals the pandemic by the contact process among people in the model. The study challenges to design an infectious disease model in which some health policies are introduced including vaccine stocks, antiviral medicine stocks, medical staff and so on. Aside from the policies, a gaming simulation of a new type of infectious disease, which has not yet an effective vaccine, is also implemented in the model. We created a medical policy decision game dealing with infections using a serious game approach. As results of experiments, it has been found that preventive vaccine, antiviral medicine stocks and the number of medical staffs are crucial factors to prevent the spread. Besides, a modern city is vulnerable to dengue fever due to commuting by train. It has also been found that self-control and restraint on immigration are essential, and decision-making for vaccine reserve amount and medical support to the partner country where the infection has spread.*

**Keywords:** *Infectious disease, health policy, Ebola haemorrhagic fever, dengue fever*

### 1. Introduction

Since history, human beings are fearful of infectious diseases, fighting death. Today, people still have the illusion that they have been released from fear of pathogenic microorganisms. Due to the development of science and technology, pathogenic bacteria have been elucidated, and effective vaccines and therapies have been developed for infectious diseases that threatened human society for a long time, such as cholera and plague. Regarding plague, a plague bacterium was discovered in 1894, and in 1943 an antibiotic

streptomycin effective for plague and tuberculosis was discovered. To-day, cases of plague have been confirmed, mainly in Africa, but if proper treatment is performed with an antibacterial agent, it will cure.

On the other hand, infectious diseases still have been serious risk factors in the world. Smallpox has been recorded in human history since more than B.C 1100. Many other infectious diseases have imposed the risk to human such as malaria, cholera, tuberculosis, typhus, AIDS, influenza and so on. Although people have tried to prevent and hopefully eradicate them, unknown infectious diseases including SARS, Ebola fever and dengue fever have emerged in our societies.

Infectious disease models, which are SIR (Susceptible, Infected, Recovered) model and SEIR (S, Exposed, I, R) model, have been studied for years, and widely used to analyse diseases using a mathematical model. Although these models are useful to make an infectious spreading estimation, the SIR model has difficulty to find which measures are effective, because it has one parameter  $\beta$  as an infection rate to represent infectiveness. Therefore, it is difficult to evaluate the effect of the temporary closing of stores and school due to the influenza epidemic. The modeler has difficulty in interpreting the  $\beta$ . The agent-based or the individual-based approach, however, has an ability to solve these problems [1-11]. It enables to reveal the spread of an infection because it simulates the contact process among people and also represents the behaviour of each person.

In this study, we developed a model to simulate Ebola fever and dengue fever based on infectious disease studies using agent-based modelling [12, 13]. The purpose of the study is how to prevent an epidemic of infectious diseases not only using mechanisms of the outbreak but also decision making of health policy. Most importantly, we need to decide a policy where people are on the move frequently worldwide in our modern society. Besides, we should minimize the economic and human loss caused by the epidemic.

## 2. Cases of infectious disease

### A. *Ebola haemorrhagic fever*

There is no sign of converging the epidemic of Ebola haemorrhagic fever in West Africa. Over 8000 people died as of January in 2015. Ebola haemorrhagic fever virus has the strong infectivity and toxicity.

A source of Ebola infection is allegedly by eating a bat or a monkey, but it is unknown whether eating these animals is a source of the disease. The current epidemic, which began in Guinea in Dec. 2013, 23 people have died. The authorities of Guinea, Liberia and Sierra Leone have each launched a state committee of emergency and have taken measures to cope with the situation. The prohibition of entry over the boundary of Guinea is included in these measures.

After that, the epidemic spread to Nigeria and Senegal. In September 2014, the president of MSF (Medecins Sans Frontiers) declared that "Six months into the worst Ebola epidemic in history, the world is losing the battle to contain it." Although infection of Ebola haemorrhagic fever is strong, it is not an airborne infection. It is thought that infection is caused by contact with patients' bodily fluids such as vomit, blood, flesh, saliva, mucus, excreta, sweat, tear, breast milk, semen and so forth.

There is a risk that a cough and a sneeze include the virus, so the infection risk is high within 1 meter in length of the cough or sneeze. The incubation period is normally seven days, and then the person gets infected after showing the symptoms. The symptoms in the early stage are similar to influenza. They are fever, a headache, muscular pain, vomiting, diarrhoea, and a stomachache. The fatality rate is very high; 50 to 90 per cent. There is no effective medical treatment medicine confirmed officially and several medicines are currently being tested. According to a guideline of WHO, the serum of a recovered patient is one of the most effective treatments.

### B. *Dengue fever*

Dengue is an infectious disease caused by being bitten by a mosquito having dengue virus (*Aedes*). There are more than 100 countries where mosquitoes that mediate dengue virus live, mainly in tropical and subtropical regions, and it is said that there are about 100 million people worldwide annually. The first domestic infection in nearly 70 years was reported in Japan in 2014, and attention is needed even in Japan.

- a) Cause and infection route: The pathogen are dengue viruses. Dengue virus has four serotypes (type 1, 2, 3, 4). Even if the virus of the same type is re-infected, it

is mild by immunity, but if it infects different types, immunity may be excessive and severe. An infection will be established by biting a mosquito carrying a dengue virus. In Japan, the *Aedes aegypti* has not been confirmed, and domestic infections are caused by *Aedes albopictus*. It will not infect people directly from humans.

- b) Symptoms: Typically, after a latency period of 2 to 15 days (mostly 3 to 7 days) after being bitten by mosquitoes, high fever (38 to 40 degrees Celsius), headache, orbital pain, joint pain, muscle pain, etc. Symptoms will recover in about a week. It rarely becomes severe and may cause dengue hemorrhagic fever that exhibits bleeding symptoms or shock symptoms.

The related disease is Zika fever. It is an illness caused by Zika virus via the bite of mosquitoes. It can also be potentially spread by sex according to recent reports [17, 18]. Most cases have no symptoms and present are usually mild including fever, red eyes, joint pain and a rash [16], but it is believed that the Zika fever may cause microcephaly which severely affects babies by small head circumference.

## 3. The health policy simulation model of infectious disease

We developed an agent-based simulation model of infectious disease based on Epstein's smallpox model. The model enables to simulate Ebola fever and dengue fever.

Each round consists of interaction with the entire agent population. The call order is randomised, and agents are processed or activated, serially. When an agent is activated, she identifies her immediate neighbours for interaction on each round. Each interaction results in contact, and then, in turn, the contact results in a transmission of the infection from the contacted agent to the active agent with probability.

## 4. Related work

### A. *Smallpox and Bioterrorism simulation*

Epstein [6, 7] made a smallpox model based on 49 epidemics in Europe from 1950 to 1971. In the model, 100 families from two towns were surveyed. This model was designed as an agent-based model, and simulation of infectious disease was conducted using the model. The results of experiments showed that 1) in a base model in which any infectious disease measures were not taken, the epidemic spread within 82 days and 30 per cent of people died, 2) it was difficult to trace all contacts to



patients in an underground railway or an airport, though a trace vaccination measure was effective, 3) although a mass vaccination measure was effective, the number of vaccinations would be huge so it was not be realistic, 4) epidemic quenching was also effective, and reactive household trace vaccination along with pre-emptive vaccination of hospital workers showed a dramatic effect.

### B. Individual-based model for infectious diseases

Ohkusa [8] evaluated smallpox measures using an individual-based model of infectious diseases. The model sup-posed a town including 10,000 habitants and a public health centre. In the model, one person was infected with the smallpox virus at a shopping mall. They compared a trace vaccination measure and a mass vaccination measure. As a result of the simulation, it was found that the effect of trace vaccination dropped if the early stage of infection was high and the number of medical staffs is small, while the effect of mass besides was stable.

### C. Summary of related work

From these studies, the effectiveness of an agent-based model has been revealed, yet these are not sufficient models to consider a relationship between vaccination and antiviral medicine stocks, and the number of support medical staff and medicine from other countries. In addition, authorities need to decide on the blockade, restrictions on outings including cars and railways while considering the economic loss of the policy.

## 5. A health policy simulation model of infectious disease

We designed a health policy simulation model of infectious disease based on Epstein's smallpox model. The model includes Ebola haemorrhagic fever and dengue fever.

### A. The basic model of Ebola hemorrhagic fever

The family includes two parents and two children; thus, the population is each 400 from each town. All parents go to work in their town during the day except 10 per cent of adults who go to another town. All children attend school. There is a public hospital serving the two towns in which every 5 people from each town work. We assume all individuals to be susceptible which means no background of immunity. One hundred families live in two towns.

Each round consists of interaction through the entire agent population. The call order is randomized where each random and agents are processed or activated, serially. On each round, when an agent is activated, she identifies her immediate neighbours for interaction. Each interaction results in con-tact. In turn, that contact results in a transmission of the infection from the contacted agent to the active agent with probability.

The probability of contact at interaction is 0.3 at a work- place and a school, while 1.0 at home and a hospital. The probability of infection at contact is 0.3 at a workplace and a school, while 1.0 at home and a hospital. In the event the active agent contracts the disease, she turns blue to green, and her own internal clock of disease progression begins.

In the event the active agent contracts the disease, she turns a stage from no-infection to latent infection, and her internal clock of disease progression begins. After seven days, she will turn yellow and begins infecting others. However, her disease is not specified in this stage. After three days, she begins to have vomiting and diarrhoea, and the disease is specified as Ebola. Unless the infected individual is dosed with antiviral medicine within three days of exposure, the medicine is ineffective. This is an imaginary medicine to play the policy game. At the end of day 12, individuals are assumed to hospitalise. After four more days, during which they have a cumulative 90 per cent probability of mortality, surviving individuals recover and return to circulation permanently immune to further infection. Dead individuals are coloured black and placed in the morgue. Immune individuals are coloured white.

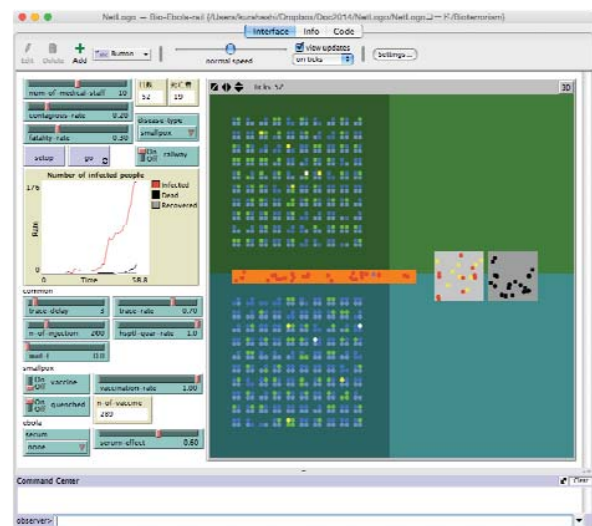


Figure 1. Interface view of a health policy simulation model of smallpox and Ebola fever.

## B. The basic model of dengue fever

About 80% of cases have no symptoms which are called latent infection, but the latent patients can transmit dengue virus to other mosquitos. The incubation period of dengue virus disease is not clear, which is likely to be 3 to 9 days. After the incubation period, the symptoms including fever, skin rashes, conjunctivitis, muscle and joint pain, malaise, and headache occur and last for 6 days. Dengue virus disease is relatively mild and requires no specific treatment, so any strategies are not selected in the model [14, 15].

In this model, a mosquito is designed as an agent as well as an inhabitant agent. Dengue virus is transmitted to people through the bite of an infected mosquito from the Ae-des. This is the same mosquito that transmits Zika. Therefore, the model of dengue fever is based on the life and habits of mosquitoes that transmit Zika. Mosquitos bite and thus spread infection at any time of day. Humans are the primary host of the virus, but it also circulates in mosquitoes. An infection can be acquired via a single bite. A mosquito that takes a blood meal from a person infected with dengue fever, becomes itself infected with the virus in the cells lining its gut. About 10 days later, the virus can be spread to other humans. An adult mosquito can live for 30 days with the virus. This model is designed on the assumption that habitats of mosquitos and climate are around south-east Japan in the Asian monsoon region in summer.

Mosquitoes live around each town, office, and school in dengue fever model. The areas they live in overlap with human regions. Therefore, the dengue virus can be transmitted between mosquitoes and human. Additionally, mosquitoes also live around a rail station in another dengue fever model where people use a train to commute their offices every day.

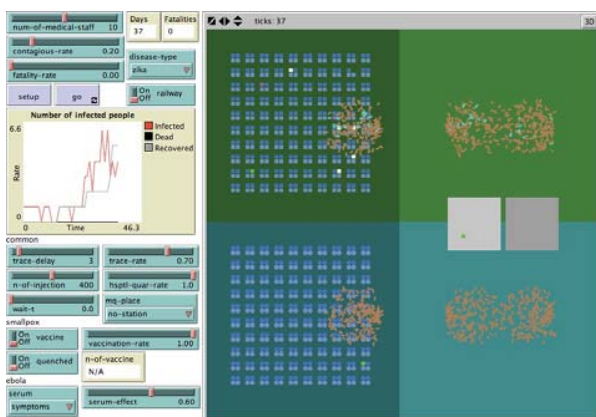


Figure 2. Interface view of a health policy simulation model of dengue fever.

We are concerned about the possibility that mosquitos are infected with dengue virus from people in both towns. Therefore, we conducted an experiment on whether the train station can be the new source of infection or not. A mosquito bite human once per four days. An infection rate from a mosquito to human is set at 0.5 and from human to a mosquito is set at 1.0.

## C. Vaccination strategies for Ebola hemorrhagic fever

The vaccination strategies we can select in the model are mass antiviral medicine and trace serum. Each of them has advantages and disadvantages.

- 1) Mass vaccination: As preemptive vaccination, the mass vaccination strategy adopts an indiscriminate approach. First, all of the medical staff is vaccinated to prevent infection. When the first infected person is recognised, certain per cent of individuals in both towns will be vaccinated immediately. The vaccination rate and the upper limit number of vaccinations per day are set on the model for the strategy.
- 2) Trace serum or antiviral medicine dosing: All of the medical staff is given serum or antiviral medicine as TAP (Target antivirus prophylaxis). Given a confirmed Ebola hemorrhagic fever case, medical staff traces every contact of the infected person and provides the medicine to that group. In addition to the mass vaccination strategy, the trace rate and the delay days of contact tracing are set according to the model.

## 6. Experimental results

### A. base model of Ebola hemorrhagic fever

The process of infection in an Ebola hemorrhagic model is plotted in Figure 3.

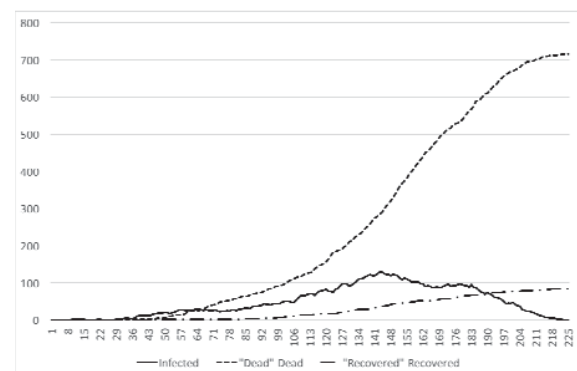
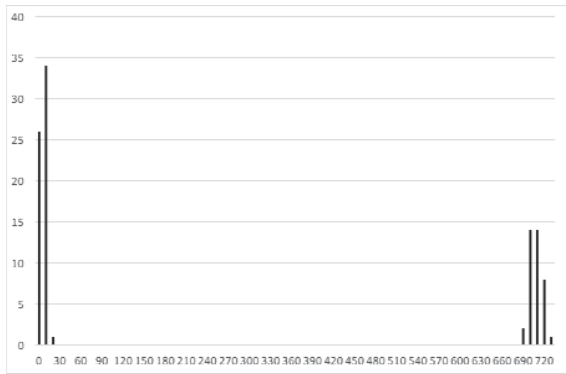


Figure 3. The experimental result of the Ebola base model: non-intervention.

The model employs non-intervention to the disease. A solid line, a dotted line and a line with marker indicate the number of infected, dead and people who have recovered respectively. When a player adopts non-intervention, it takes approximately 231 days until convergence of the outbreak and more than 716 people have died.



**Figure 4. The frequency chart of the Ebola base model: non-intervention.**

Figure 4 shows the results of 100 experiments. The number of fatalities under 20 people was 59 times out of 100 experiments, but the number of fatalities over 700 people in 800 in-habitants recorded 39 times. The result indicates that the infection phenomenon is not based on the normal distribution. It means we need to carefully consider all possibilities and risk beyond the scope of the assumption regarding infection spread.

#### B. Mass vaccination model

The process of infection is plotted with the mass vaccination strategy in which individuals are vaccinated randomly after three days when given a confirmed case. The policy succeeds to prevent the outbreak because the number of vaccinations per day is 600 and three-fourths of inhabitants are vaccinated per day. On the other hand, the low ability of vaccination ends in failure because the number of vaccinations per day is 400 and a half of inhabitants are vaccinated. The ability of vaccination per day bifurcates the results.

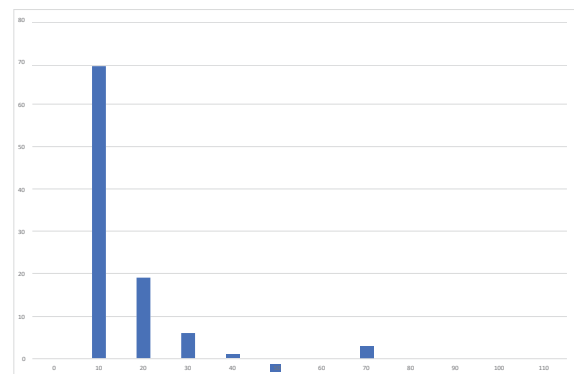
#### C. Trace serum strategy

It was found that the ability of more than 50 serums per day was able to control an epidemic in most cases. In the case of using public transportation to commute, however, it makes a substantial difference. Four hundred serums per day could not prevent the outbreak, while 600 serums per day succeeded to prevent it. Trace vaccination strategy is one of the most effective policies in a town where people commute by car, but a large number of

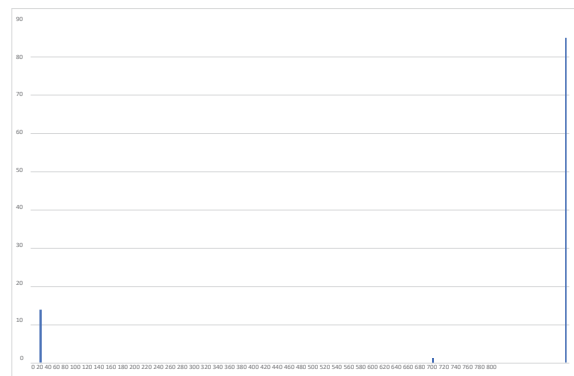
serums per day, at least a half of people, is required if most of the people use public transportation systems like a railway and a bus.

#### D. Isolated Town model of dengue fever with/without a rail station

The process of infection in dengue model of an isolated town is plotted in Figure 5. It was found that most cases are under 30 fatalities, but a few cases are over 70 fatalities. On the other hand, Figure 6 shows that cases are polarised to different results in a modern city with a railway for a commute. The case under 10 infected people is 15 times, and the case in which all people are infected by mosquitoes is 85 times.



**Figure 5. The experimental result of dengue fever model without a railway.**



**Figure 6. The experimental result of dengue fever model with a railway.**

## 7. Infectious disease medical policy game

In the experiment so far, we have seen that it is difficult to prevent the spread of infection in areas with many rail-way users. A contemporary society faces the risk of the infection spreading across borders like Ebola hemorrhagic fever and new influenza. In this case,

international cooperation of medical policy team is important. The World Health Organisation (WHO) is promoting international co-operation for measures against infectious diseases, and frequently taken up at the G7 Summit and other conferences. Therefore, in order to expand the model so far, we developed a cooperative game for countermeasures against infectious diseases and conducted an experiment to promote cooperation between two countries or regions.

In this game, it is supposed that a new type of infectious disease occurs in which the disease is similar to smallpox and Ebola hemorrhagic fever. Vaccine and antiviral medicine for the disease are already developed and provided to the market in this model. Players as authorities of two countries decide the amount of both medicines' stocks according to their restricted budget. They also need to decide the number of medical staffs, blockade and restrictions on outings. The players should consider giving support medicine and staff to countries to prevent or control its epidemic for his/her own country while taking account of economic cost and loss. Travel restrictions have a huge economic impact, while it is very effective in stopping an outbreak. Supporting to another country means decreasing its own preparations. Thus, this game has a complicated structure of trade-offs among cost, effect, cooperation and defense.

#### A. Medical policy model

The medical policy leaders of both countries will make decisions on the following policies. 1) Number of vaccines ordered for stockpiling 2) Number of antiviral drug orders for stockpiling 3) Employment of medical staff who can respond to Ebola hemorrhagic fever Discussions on the collaboration of infectious disease countermeasures with neighbouring countries are underway, and policies that support each other can be implemented. 4) Number of vaccine support to neighbouring countries 5) Number of antiviral drugs supported in neighbouring countries 6) Number of supports for medical staffs who can cope with infectious diseases to neighbouring countries. However, since this assistance will reduce the stock and medical staff of their country, the duty of support is exempted if danger is approaching. On the other hand, as an emergency response, it is possible to instruct the residents to go out and to order restrictions on entry and departure with neighbours. 7) The rate of a voluntary ban on leaving home, 8) The rate of restrictions on immigration, 9) Mass vaccination for all residents, 10) Trace vaccination for infected people, 11) A dose of an antiviral drug to a contact with an infected person, 12) A dose of an antiviral drug to a contact getting fever with an infected person

Players can get a report of infection status including her/his own country and a neighbouring country every

ten days. Although they decide to avoid pandemic in their own country as they cooperate when they get each report, they have to cope with a difficult situation within a pre-determined budget. They are admitted to using additional budget by the government, but the smaller they use the budget, the more advantage they get to win the game. For the sake of their budget management, the following parameters are pre-defined in a game; budget, the price of the vaccine, the price of the antiviral drug, personal expenses of medical stuff, economical cost of a voluntary ban on leaving home, and economical cost of restrictions on immigration. Figure 7 shows Interface view of a health policy game.

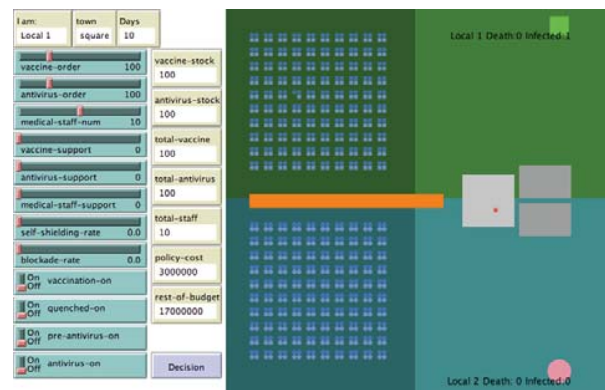


Figure 7. Interface view of a health policy simulation and gaming.

#### B. First experimental gaming result

Two teams of two people (C country, S country) conducted two games. After performing the briefing, the player got used to the operation of the game with a model without railroad use.

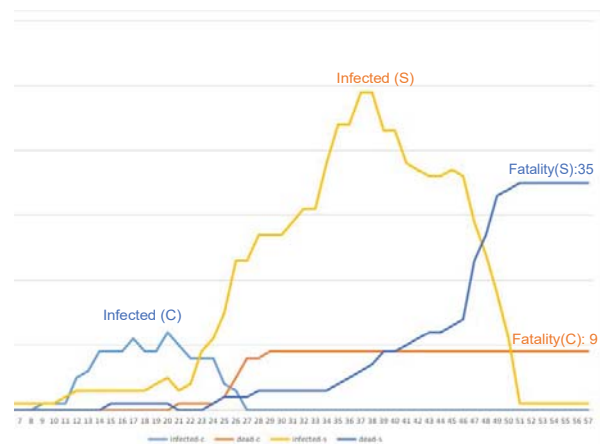
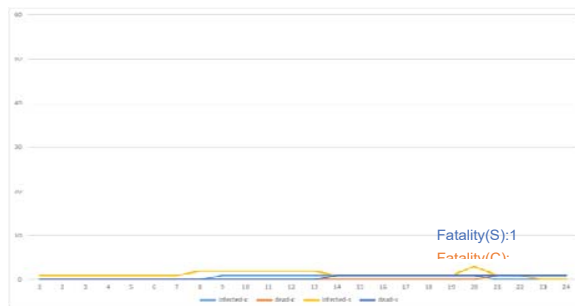


Figure 8. First experimental gaming result.

The first game was conducted with a model with railway use. The purpose of the game was to minimise the number of deaths due to Ebola hemorrhagic fever infection in their country and to minimise policy costs as much as possible. The transition of the first infection in the experiment is shown in Figure 8. The experiment result shows that the death toll in country C was 9 and the death toll in S country was 35. S country was a little delayed to order enough vaccine and supported too much to C country. C country decided to order a plenty of vaccine, antivirus drug corresponding to the epidemic.

### C. Second experimental gaming result

As a result of the second experiment, the spread of infection was suppressed, and the number of deaths decreased sharply by one person each. The transition of the first infection in the experiment is shown in Figure 9.



**Fig. 9. Second experimental gaming result.**

From experience that allowed the spread of infection at the first time, players in the two countries were expected to deepen collaborative relationships, but in fact, the actions taken by both players showed that the level of vaccine and staff support is about the same level as the first or less. Instead, the players of both countries had decided to refrain from going out of their country and early strengthening the blockade of the border. S country ordered vaccine twice as much as the 1st game when the epidemic was found. C country decided to send stuff to S country and a partial voluntary ban on leaving home and blockade.

## 8. Discussion

The results of these experiments are that 1) trace vaccination strategy is one of the most effective policies to Ebola hemorrhagic fever, but 2) both Ebola hemorrhagic fever and Dengue fever can be affected seriously by the case of train commuters, even if the trace vaccination strategy is adopted for them. This result indicates that a blockade of railways is a more effective alternative strategy in modern cities where most people commute by public transportation system such as a train,

subway, bus and so on. Next, we conducted an experiment using a medical policy game considering inter-national collaboration. In this experiment, it is assumed that Ebola hemorrhagic fever infection occurred in two borders, and it is aimed to cooperate and prevent the spread of the virus. As a result of the experiment, the total of 44 deaths were issued in the first experiment, but the second was able to suppress to 2 people. After the first game finished, since the meeting to support each other was voluntarily held, the effect of collaborative learning by the game was seen.

## 9. Conclusion

In this study, we developed the infection process of Ebola hemorrhagic fever and Dengue fever as a deductive model based on ABM. Symptoms after infection and each parameter were defined based on actual data, and residents and regional models were defined as two simple regional models as possible. Experimental results showed that trace serum administration is most effective in the case of Ebola hemorrhagic fever, and complete suppression of infection is difficult when railway use is assumed for commuting as public transportation.

In addition, this study also conducts the model of Dengue fever to understand the mechanism of spreading Dengue virus via mosquitoes. As a result of experiments, it is crucial for preventing a Dengue epidemic to get rid of mosquitoes from a train station where people use it to commute to their office.

Besides, as international cooperation by medical authorities is indispensable today, we have constructed a medical policy game that encourages bilateral coordination and conducted experiments. As a result, experiencing the game prompted a certain degree of cooperative action, plus the prevention of the spread of infections to other countries by forced policies such as self-control and border blockade, these were chosen by the player.

This model can be applied to various institutional evaluations as a deductive model. It also gives ideas that contribute to game designers' cooperative behaviour and emergence, educational effectiveness, and medical policy design.

## 10. Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP17H02035 and JP17K19994.

## 11. References

- [1] Burke, Donald S., et al., "Individual-based Computational Modeling of Smallpox Epidemic Control Strategies", *Academic Emergency Medicine*, vol. 13, no. 11, 2006, pp. 1142–1149.

- [2] Longini Jr, Ira M., et al., “Containing a large bioterrorist smallpox attack: a computer simulation approach”, *International Journal of Infectious Diseases*, vol. 11, no. 2, 2007, pp.98–108.
- [3] Gilbert, Nigel. Agent-based models. No. 153. Sage, 2008.
- [4] Easley, D., and Kleinberg, J., Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press, 2010.
- [5] Okabe, N., “Risk and Benefit of Immunization: Infectious Disease Prevention with Immunization”, *Iryo & Shakai*, vol. 21, no. 1, 2011, pp. 33–40.
- [6] Epstein, Joshua M. et al., Toward a containment strategy for smallpox bioterror: an individual-based computational approach, 2002.
- [7] Epstein, Joshua M., Generative social science: Studies in agent-based computational modeling. Princeton University Press, 2006.
- [8] Ohkusa, Y., “An Evaluation of Counter Measures for Smallpox Outbreak using an individual based model and Taking into Consideration the Limitation of Human Resources of Public Health Workers”, *Iryo & Shakai*, vol. 16, no. 3, 2007, pp. 295–284.
- [9] Hunder, E., Namee, B. M., Kelleher, J. D., “A Taxonomy for Agent-Based Models in Human Infectious Disease Epidemiology”, *Journal of Artificial Societies and Social Simulation* 20 (3) 2, 2017.
- [10] Yoneyama, T., Das, S., Krishnamoorthy, M., “A Hybrid Model for Disease Spread and an Application to the SARS Pandemic”, *Journal of Artificial Societies and Social Simulation* 15 (1) 5, 2012.
- [11] Bruch, E., Atwell, J., “Agent-based Models in Empirical Social Research”, *Sociol Methods Res.* 44(2), 2015, pp. 186-221.
- [12] Kuhlman, C. J., “Hybrid Agent-based Modeling of Zika in the United States”, in *Proceedings of the 2017 Winter Simulation Conference*, 2017.
- [13] Wiratsudakul, A., Suparit, P., Modchang, C., “Dynamics of Zika virus outbreaks: an overview of mathematical modeling approaches”, *PeerJ* 6, e4526, 2018
- [14] ECDC, *Aedes aegypti* - Factsheet for experts, Health Topics: Vectors: Mosquitos. European Centre for Disease Protection and Control. Retrieved June 2018.
- [15] CDC, Zika Virus, Centers for Disease Control and Prevention, U.S. Department of Health & Human Services, Retrieved June 2018
- [16] WHO, Zika virus, World Health Organization, Media centre, Fact sheets, January 2016.
- [17] Chen, Lin H. and Hamer, Davidson H., “Zika Virus: Rapid Spread in the Western Hemisphere”, *Annals of Internal Medicine*, 2 Feb. 2016.
- [18] Hennessey, M., Fischer, M., and Staples, J. E., “Zika Virus Spreads to New Areas - Region of the Americas”, May 2015 - January 2016, Centers for Disease Control and Prevention, vol. 65, no. 3, 2016, pp. 55-58.

## Multimodal user interaction framework for e-commerce

Supun Hewawalpita  
Department of Computer Science and  
Engineering  
University of Moratuwa, Sri Lanka  
supungs.10@cse.mrt.ac.lk

Indika Perera  
Department of Computer Science and  
Engineering  
University of Moratuwa, Sri Lanka  
indika@cse.mrt.ac.lk

### Abstract

*E-commerce has grown to be a major user of e-services and online purchases through e-commerce are largely preferred over the traditional brick and mortar purchasing. Yet it is challenging for the consumers to fully experience the products or services with limited senses, lack of tangibility and sense of presence. Therefore, the objective of this study is to assess how multimodal interactions can be used in e-commerce, to improve the consumer experience. This research describes the design and implementation of multimodal interactions enabled prototype framework for e-commerce developed to address that issue. The framework supports three interaction modes: speech inputs, gesture inputs and interactive 3D product presentations. A case study and a usability study conducted on this framework showed satisfactory results and also revealed some interesting social and cultural barriers to multimodal interactions that involve language and psychological factors.*

**Keywords:** *Multimodal interactions, consumer experience, e-commerce, usability study*

### 1. Introduction

Electronic commerce has become an essential part of the day to day lives with the rapid advances in telecommunication technologies and the Internet. Unlike traditional brick and mortar retail operations, electronic shopping is not constrained by the distance or time, providing more control to the consumer. On the other hand, e-commerce services are less tangible than the same services delivered physically in person. As a result, the consumer cannot use all the senses to experience the product or service as in traditional shopping and rather has to rely mostly on sight. It is challenging for the consumers as well as for the designers of electronic commerce systems. In the end, the quality of e-commerce services depends on how the users perceive and experience the service.

To overcome this issue, e-commerce services are now looking to adopt multimodal interfaces to enable free and natural interaction using the natural human capabilities of communication [1]. Having multiple modes of interaction in a system supposed to provide more affordance and greater accessibility to the users and brings more bandwidth to the communication.

Developing or updating an e-commerce platform with such interactions and context awareness features can be a complex and tedious task. A specifically designed framework with reusable components and guidelines can simplify the task for some extent, but there is hardly any such framework. This paper discusses an experimental multimodal interaction framework that was designed to address this issue.

### 2. Related work

One main problem of e-commerce is, the users have to experience the provided services or products through a limited number of senses and are not able to feel, touch or manipulate through the web interface. It is an indirect experience provided through symbolic representations and mediated interactions using textual and graphical descriptions. In contrast, a direct experience should provide unmediated interaction between the consumer and product or service allowing using his full sensory capacity [2]. Indirect experiences are in most cases less effective than direct experiences on the consumer [3]. Multimodality can be used to address this primary issue by expanding the modes of communications and consumer experience.

There is a number of interaction modes that are researched and experimented in literature. Some of the interaction modes have found their way into the mainstream consumer applications while some remains limited to the research contexts.

Speech is probably the most natural communication mode for humans and has been researched for a while. Early studies have been conducted on developing prototype e-commerce systems combining HTML and VoiceXML technologies [4, 5]. They allowed the users to interact with the website in the traditional way while interacting with the voice service over the phone. Now

speech recognition has become inbuilt in most of the mobile and desktop operating systems and with the availability of APIs [6], developing multimodal voice interactive services is more probable than before.

3D visualization is one technique that can be used when there are physical objects or products involved. This is achieved by using several photos from different angles or a 3D model of the product/object. Allowing viewing all parts of the object interactively have an effect on fast reasoning and understanding of the object by the consumer. It has been found that in e-commerce services, 3D interactive product presentation generates a more positive attitude toward products [7]. Particularly the 360° views using the actual photos are found to be more appealing to the consumers than 3D models [8]. Consumers interacting with such products are more likely to experience an increased sense of presence [9, 10]. Some other studies have investigated the effects of interactive image technology (IIT) and suggest that higher level of IIT such as virtual 3D models are suitable for online apparel retailers to improve consumer perception and shopping experience [11, 12].

Another related area of 3D based merchandise is the augmentation of products with 3D product models through augmented reality, particularly in mobile devices. There have been some efforts on fusing e-commerce with in-store shopping using augmented reality visualizations [26]. Going a further step, several types of research [27, 28] suggest completely virtual 3D e-commerce environment. Users are represented as 3D avatars that can walk in the environment and interact with products, sellers and other users. But the practicability of using it for a real business scenario is very limited. However, virtual reality visualization might be suitable for some specific cases such as luxury brands online trading [29] and largescale products such as real estate and apartments [30].

Gesture-based interactions have been an area of research for a long time and are a key modality in the realisation of a natural interaction paradigm [13]. Recent innovations such as Microsoft Kinect [14] and Leap Motion [15] have made accurate gesture recognitions possible and readily available to use in any application.

There have been other researches on using virtual agents [16, 17] and human-like avatars [18, 19] as multimodal interaction modes. They are particularly useful for building trust towards the system and as interaction support for older users.

### 3. Methodology

In this research, the importance of multimodality in e-commerce and its impact on the overall quality of services were analyzed first. The next section highlights some experiments conducted for this purpose. Then

suitable contextual factors and how they affect the mode of interactions and service outcomes were identified and defined based on the finding of those experiments.

Then an interaction framework was modelled and implemented in a prototype level. This framework consists of a set of tools, components, and guidelines that can be reused to implement multimodal interactive e-services. Then an evaluation scheme was modelled and the performance of the prototype was evaluated under a selected use case.

## 4. Preliminary studies

Before designing the discussed multimodal interaction framework, two experiments were conducted to identify the effects of multimodal interactions on users in the context of e-services.

### 4.1. Experiment 1

From the literature survey, 3D visualization was identified to have interesting potential in e-commerce as an interaction mode. But there was a lack of research about its effects on consumers' preference. Thus an experiment [20] was conducted to study these effects. Users were asked to interact with a prototype e-commerce system where products were displayed in 3D visualizations and normal 2D images. Various feedbacks including the general attitude toward the product, the likelihood of purchase were collected from the users after the experiment and analyzed using the analysis of variance technique.

The results of this experiment showed a significant ( $p < 0.05$ ) difference of attitudes toward the product presentation depending on the format products presented on a computer. It was found that participants have a favourable attitude towards products when they are displayed in an interactive 3D view than the typical 2D image format. Interestingly perceived value of the products also significantly ( $p < 0.05$ ) increased for products displayed in the 3D view.

However, it was also found that with multiple degrees of freedom, those positive effects reduce and sometimes have negative effects. It was assumed that controlling the 3D visualization with mouse or touch inputs is difficult for some users and thus results in negative effects. The longer learning curve of this modality had been a challenge for novice users without much prior 3D experience and had negatively affected their user experience.



**Table 1. Summary of perceived product values in the preliminary study.**

	Treatments		
	2D Images	3D View (1DOF)	3D View (2DOF)
N	220	173	122
$\Sigma X$	476454	577570	246165
Mean	2165.7	3338.5549	2017.7459
$\Sigma X^2$	1684031928	3241195300	833687291
Std. Dev.	1725.6797	2762.8622	1668.8415

## 4.2. Experiment 2

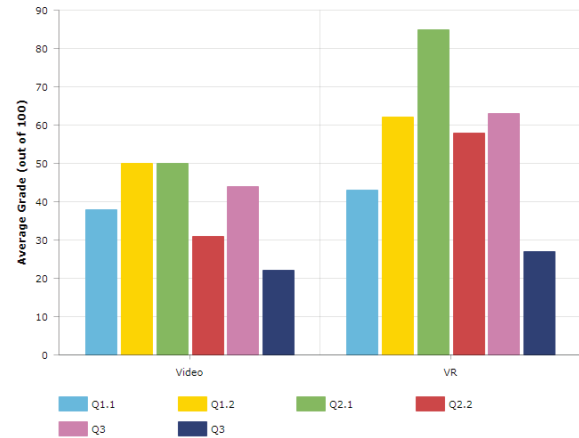
Looking for alternative interaction approaches another experiment was conducted to study the effects of virtual reality (VR) and gesture interactions on learner experience in a massive open online course (MOOC) environment. For the experiment, a MOOC was deployed with two courses with the same content. In one course the content was delivered as interactive virtual reality experiences while the content in the other course was delivered as traditional video content.



**Figure 1. A student engaged in an interactive VR activity.**

The virtual reality experiences were delivered via Oculus Rift [21] VR headset paired with Leap Motion hand tracking device [15] for gesture interactions. Four student groups followed the two courses for a number of weeks and their performances were measured through weekly quizzes. When marks of the students in each

activity were analyzed, some interesting results were found. The test groups which had course content delivered in virtual reality, showed significantly ( $p < 0.05$ ) better performance in total marks for the quizzes (Figure 2). Moreover, analysis of students' feedback suggested that the students who initially followed a VR based MOOC are more likely to try another MOOC course.



**Figure 2. Average grade obtained for quizzes by the 4 student groups.**

## 5. Framework

From the analysis of the two preliminary experiments and literature studies, the following three interaction modes were selected to be supported by the framework, considering their practicality and effectiveness.

- Speech interactions
- Interactive 3D visualizations
- Gesture interactions

### 5.1. Overview

The architecture of the framework was carefully designed to cater to the basic requirements of an e-commerce system as well as the selected multimodal interaction capabilities. At a higher level, the architecture follows the typical Model-View-Controller (MVC) pattern.

Figure 3 shows the basic architecture design for the framework. Most of the e-commerce capabilities such as product catalogue, searching, checking out are in the models and controllers whereas multimodal interactions are mostly handled in the views. The interaction handlers interact with e-commerce services to retrieve information required to cater to some user interactions such as voice-based searching.

## 5.2. Implementation

The multimodal interaction framework was developed on top of the CodeIgniter [22] framework. It was specifically chosen for its MVC based design, small footprint and simplicity. E-commerce functionalities of the framework were implemented mainly on the server side adhering to the traditional MVC pattern while the multimodal interaction components were mostly implemented on the client side.

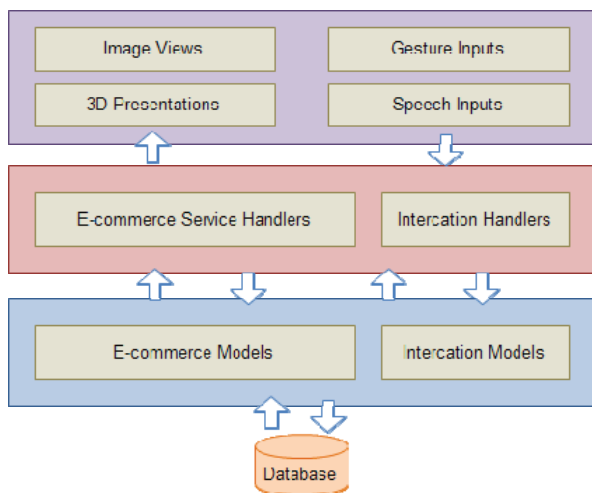


Figure 3. The architecture of the framework.

The 3D product presentations were made up of a number of actual photos taken around the product in even angles and the resulting visualization could be smoothly rotated using the mouse giving the perception of a 3D object. Previous research [8] have found out that the consumers have a more favourable attitude towards these pseudo-3D views made up of actual images, compared to 3D models, due to the realism and the interaction possibilities. The framework supported 12, 24 or 36 frames per 3D presentation taken around a single or two axes. With more frames, the smoother the presentation to interact.

The framework supported interacting with the 3D visualization with mouse or hand gestures. For the hand gesture mode, a video stream from a web camera was used and its preview was displayed on the top corner of the page as a reference for the users. The gesture used for the interaction was moving the hand fist sideways and up-down in front of the camera. The 3D product presentation rotates according to movements of the hand. When recognized, hand fist was highlighted in the camera preview. Accessing and processing the webcam stream was implemented using JavaScript available

through open source libraries. As of yet only a few of the major web browsers allow accessing the webcam with JavaScript due to security and other concerns.

The framework also supported Leap Motion based gesture interactions to control the 3D visualization. Leap Motion is a hardware device capable of accurately tracking hand and finger using infrared light. With their JavaScript API hand and finger positions, movements and speed data were monitored and using a simple algorithm rotation of the 3D presentation was calculated. It allowed controlling the 3D scene more smoothly and accurately by moving an open hand sideways and up-down but needed the device to be plugged into the computer.

The search function in the framework was speech-enabled and users could press a button and speak what they want to search. Web Speech API [6] provided by the Chrome web browser was used to implement this and language mode was set to English (India) en-IN which was the closest for the English speaking style of the users. The native language of the users is not supported by the Web Speech API at the moment.

Although the Web Speech API specification has been published in 2012 by W3C Community, still only a few mainstream web browsers such as Chrome and Firefox support this specification, making this feature of the framework only available in those two browsers.

While speaking intermediate partial results of speech recognition were immediately displayed on the search input box so users get the feedback whether their words were recognized. Upon successful recognition, the framework automatically searched for products and displayed the results.

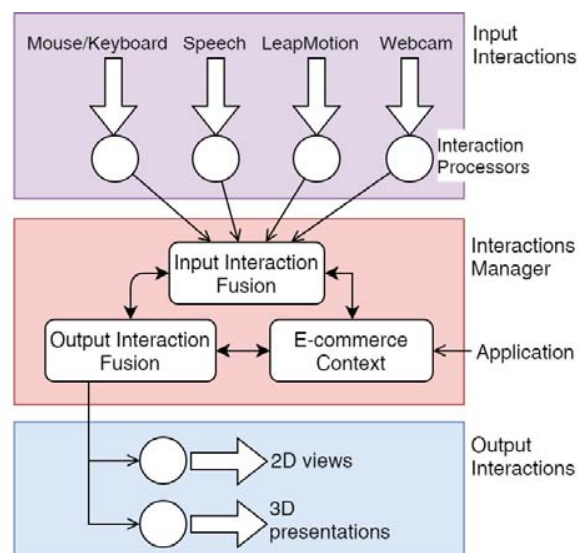


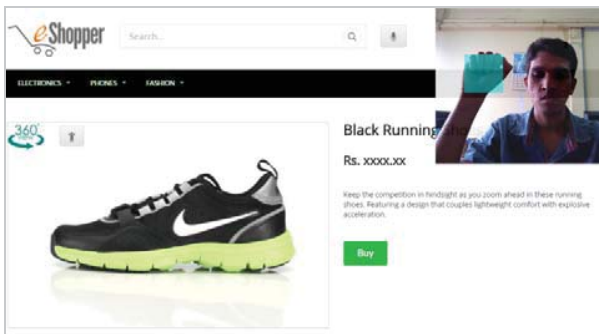
Figure 4. Interaction architecture for the framework.

Figure 4 shows the interaction architecture for the framework. Each input interaction was processed by the associated libraries or frameworks and the interaction manager integrates with the e-commerce context including product details and information to provide the output interactions.

## 6. Case study

To study the development aspects of the framework a case study was conducted by using the framework to develop an actual e-commerce system. The requirements for the proposed system can be summarized as below.

- Should support the general e-commerce functionalities
- Should provide a new experience for consumers
- Features to increase user engagement and purchase intentions



**Figure 5. Interacting with the 3D product presentation with gestures.**

The requirements mostly aligned with the capabilities of the developed multimodal interactions framework. The two preliminary studies had already suggested that multimodal interactions indeed provide a new user experience for users. From the result of the first experiment, it was evident that 3D product presentations can increase the product engagement, perceived value of products as well as the purchase intentions.

The proposed e-commerce website was implemented extending the framework. Its MVC based modular architecture allowed to rapidly develop the system supporting three forms of interaction modalities: speech, gesture and 3D visualization interactions. Due to the usability issues of 3D visualizations with two degrees of freedom, identified in the first experiment, only one degree of freedom was enabled in the system. Both webcam and Leap Motion based gesture interactions were enabled as an alternative to mouse interactions with the 3D visualizations.

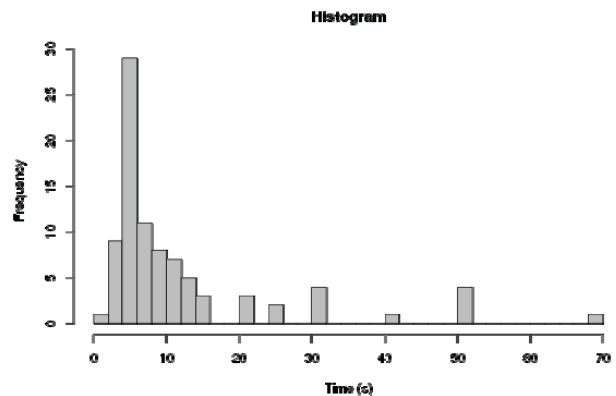
## 6.1. User study

To study the usability aspects of the developed prototype e-commerce system, a user study was conducted. 15 test users (7 males and 8 females) aged between 21 and 28 were recruited for the evaluation. All of them were non-native English speakers and had previous experiences in using e-commerce websites. The subjects were given a set of tasks to perform with the system after a brief introduction to the system. In the first task, users were asked to use speech to search for products from a given set of keywords in English and they could attempt up to 10 times. The number of successful and failed attempts was recorded with the time taken for each attempt.

In the second task, users were asked to interact with the 3D product presentation using hand gestures as well as a mouse or touchpad. Evaluation sessions were video recorded and monitored to find the difficulties faced by the users. Users were interviewed after the experiment and asked a predefined set of questions to identify their opinions and difficulties they faced during the evaluations.

## 6.2. Results

All the users could successfully search using speech inputs and the mean success rate of a person per 10 attempts was 6.2 ( $\sigma = 2.14$ ). The time taken for a single successful search varied from 4 seconds to 70 seconds with a mean value of 11.54 seconds ( $\sigma = 12.99$ ). Fig. 6 shows the histogram for the distribution of time taken for successful search using speech.



**Figure 6. Distribution of time taken for a successful search using speech.**

In the second task, all the users were successful in controlling the 3D product view, mouse or laptop touchpads. But some difficulties could be identified when using the touchpad. 62.5% of the users ( $f = 10$ ) reported that they were unable to use webcam based hand gestures

to control the 3D view up to a satisfactory level. Their hands were not recognized by the system properly and thus wasn't able to control the 3D view. On the contrary, all the users were successful in using Leap Motion based gesture interactions up to a satisfactory level.

**Table 2. Summary of qualitative feedbacks gathered from the interviews.**

Difficulties Faced	Reasons for not Using in Public
Speech	
Doesn't recognize my words	I might attract attention
Miss recognize my words	Might distract others
Too slow	
Gestures	
Doesn't recognize my hand	I might look ridiculous
Miss recognize my face as a hand	I might attract attention
Difficult to control the view	I might distract others
View rotates too fast	
My hand hurts	
The camera view is confusing	
Browser blocks webcam access	

In the post-experiment feedbacks (Table 2), only a minority of users (f=4) expressed that they would use speech or gesture interactions when other people are around whereas 50.0% users (f = 8) said they would not and others (f = 4) were unsure. Majority of the users had reported hand fatigue as a difficulty they faced while gesture interacting (Table 2), especially for webcam-based gestures.

## 7. Discussion

The case study and its user study revealed some interesting observation and challenges for the framework and multimodal interactions in general.

Although all the users were able to use speech interactions successfully, some of them had some difficulties. A negative correlation could be detected between the success rate and the time took per successful attempt. It is understandable as the less successful users had to try multiple times to get a correct search thus consuming more time. It was noted that users with clear and sharp voices were more successful compared to the others. English was not the native language of any of the users and their accents were mixed and did not belong to

any supported English modes. Some users were uncomfortable speaking to the computer while been monitored.

No specific difficulty could be identified nor reported by users while interacting with the 3D visualization using a mouse. When the preliminary study showed some negative results in user experience when the visualization allowed two degrees of freedom in control, it was assumed that users find it difficult to rotate in two directions with the mouse. But we could not identify such difficulty in the user study. However, users who interacted using laptop touchpad, instead of the mouse, did face some issues. They could not rotate the product a full cycle in one touch motion and had to perform several touch-motion swipes to view the products' full 360 degrees. But with the mouse, the product could be rotated entire 360 degrees in a single motion.

Webcam-based gesture interactions found out to be the most difficult for users and the majority of them could not even get their hands recognized by the system. Even though other users' hands were recognized for some extent, their success in controlling the 3D view with the hand was minimal. Backgrounds behind the users were noted to be largely affecting the performance of gesture recognition and when there was a monotonous and clear background behind, the gesture recognition was more accurate whereas with a cluttered background gesture recognition was extremely poor. Whereas powerful computer vision libraries such as OpenCV [23] are available for desktop applications no such advance libraries for web applications are available. Also, the fact that a web application running on a browser has limited capabilities and resources makes it difficult to achieve a satisfactory level of visual gesture recognition.

Majority of the users reported hand fatigue after trying the gesture interactions for a small period of time. With the typical positioning of the webcam, users had to raise their fist to face level and keep unsupported for the duration of gesture interactions. Thus the fatigue in hands is understandable. This effect is well known as Gorilla Arm Syndrome [24] and is a common phenomenon with vertical touch displays and hand gesture interactions. Holding arm in mid-air, for gestures cause arm and shoulder fatigue in a short period and with prolonged durations, it results in physical discomfort and decreased performance [25]. There are recently developed matrices to quantitatively characterize this effect, such as Consumed Endurance [26].

When asked about using speech or gesture interactions while other people are around, the majority of the users have stated that they would not. Only 25% of the users have stated that they would use the multimodal interactions, while the others were unsure. They expressed that they will not be comfortable using these interactions in such environment and they might look

awkward. Even during the user study, some users were not comfortable speaking and gesturing to the computer while being monitored. This reveals an interesting social barrier for multimodal interactions. Shyness and social awkwardness are found all over the world but differ depending on the cultures and geographic regions [27]. Behaviour in public can be viewed as acceptable, awkward or offensive depending on these factors. As evident in this study, users from some cultural backgrounds might find it uncomfortable to use some interaction modes in fear of unwanted attention and being negatively judged by others, in a social environment.

## 8. Conclusion

The goal of this research was to design and implement multimodal interactions enabled prototype framework for e-commerce, developed to improve the consumer experience. The framework supported three interaction modes: speech inputs, gestures inputs and interactive 3D product presentations. Resource limitations and lack of support for technology standards such as Web Speech API in web browsers had imposed major constraints on the capabilities of the framework. Still, the user study conducted on a prototype e-commerce website built with the framework showed satisfactory results.

Webcam-based visual gesture recognition in the framework found out to be less effective and improving its accuracy is difficult due to the technical and resource limitations. Leap Motion has shown to be an effective alternative for that with good accuracy and a lot of interaction possibilities but requires an external hardware device to be plugged in.

The study revealed some interesting social and cultural barriers to multimodal interactions that involve language and physiological factors. Some non-native English speakers shown to be having difficulties interacting with speech and enabling speech interactions for e-commerce systems with a diverse and multicultural user base can be extremely complicated. User acceptance is also found out to be affected by cultural shyness and social anxiety. As a result, users from some cultural backgrounds can be uncomfortable and hesitate from using some interaction modes while other people are around, in fear of being negatively judged.

From the findings of this research, we suggest that further research focus should be on overcoming these practical barriers and bringing multimodal interactions to mass usage in electronic commerce platforms.

Going further with the development of the framework, we expect to integrate virtual reality as another interaction mode. Although proper equipment is expensive, it is still effective even with basic forms of virtual reality tools such as smartphones and VR glasses. The VR modality will be specifically useful for

visualizing and experiencing larger products such as automobiles and apartments. Accessibility is another feature in consideration, to aid differently abled people to experience e-commerce and online shopping.

## 9. References

- [1] Stivers, T. and Sidnell, J., "Introduction: multimodal interaction," *Semiotica*, vol. 2005, no. 156, pp. 1–20, 2005.
- [2] Gibson, J. J., "The senses considered as perceptual systems," 1966.
- [3] Millar, M. G. and Millar, K. U., "The effects of direct and indirect experience on affective and cognitive responses and the attitude-behavior relation," *J. Exp. Soc. Psychol.*, vol. 32, no. 6, pp. 561–579, 1996.
- [4] Trabelsi, Z., Cha, S. H., Desai, D. and Tappert, C., "A voice and ink XML multimodal architecture for mobile e-commerce systems," in *Proceedings of the 2nd international workshop on Mobile commerce*, 2002, pp. 100–104.
- [5] Rouillard, J., "A multimodal E-commerce application coupling HTML and VoiceXML," in *Proceedings of the Eleventh International World Wide Web Conference, Waikiki Beach, Honolulu, Hawaii, USA*, 2002.
- [6] "Web Speech API Specification." [Online]. Available: <https://dvcs.w3.org/hg/speech-api/raw-file/tip/speechapi.html>. [Accessed: 24-Jun-2017].
- [7] Suh, K. S. and Lee, Y. E., "The effects of virtual reality on consumer learning: an empirical investigation," *Mis Q.*, pp. 673–697, 2005.
- [8] Moritz, F., "Potentials of 3D-Web-Applications in E-Commerce - Study about the Impact of 3D-Product-Presentations," in *2010 IEEE/ACIS 9th International Conference on Computer and Information Science*, 2010, pp. 307–314.
- [9] Li, H., Daugherty, T. and Biocca, F., "Impact of 3-D advertising on product knowledge, brand attitude, and purchase intention: The mediating role of presence," *J. Advert.*, vol. 31, no. 3, pp. 43–57, 2002.
- [10] Hassenzahl, M., Beu, A. and Burmester, M., "Engineering joy," *IEEE Software*, vol. 18, no. 1, pp. 70–76, 2001.
- [11] Fiore, A. M., Kim, J. and Lee, H. H., "Effect of image interactivity technology on consumer responses toward the online retailer," *J. Interact. Mark.*, vol. 19, no. 3, pp. 38–53, 2005.
- [12] Kim, J., Fiore, A. M. and Lee, H. H., "Influences of online store perception, shopping enjoyment, and shopping involvement on consumer patronage behavior towards an online retailer," *J. Retail. Consum. Serv.*, vol. 14, no. 2, pp. 95–107, 2007.
- [13] Kavanagh, S., "Facilitating natural user interfaces through freehand gesture recognition," in *Proceedings of the CHI*, 2012, vol. 12.
- [14] "Kinect Sensor."

- [15] "Leap Motion." [Online]. Available: <https://www.leapmotion.com>.
- [16] Lucente, M. "Conversational interfaces for e-commerce applications," *Commun. ACM*, vol. 43, no. 9, pp. 59–61, 2000.
- [17] Liew, T. W. and Tan, S. M., "Exploring the effects of specialist versus generalist embodied virtual agents in a multi-product category online store," *Telemat. Inform.*, Oct. 2017.
- [18] Rigas, D. and Gazepidis, N., "A further investigation of facial expressions and body gestures as metaphors in e-commerce," in *Proceedings of the 7th Conference on 7th WSEAS International Conference on Applied Informatics and Communications*, 2007, pp. 148–153.
- [19] Liew, T. W., Tan, S. M. and Ismail, H. "Exploring the effects of a non-interactive talking avatar on social presence, credibility, trust, and patronage intention in an e-commerce website," *Hum.-Centric Comput. Inf. Sci.*, vol. 7, no. 1, p. 42, Dec. 2017.
- [20] Hewawalpita, S. G. S. and Perera, I., "Effect of 3D Product Presentation on Consumer Preference in E-Commerce," presented at the Moratuwa Engineering Research Conference (MERCon), University of Moratuwa, Sri Lanka, 2017.
- [21] "Oculus Rift." [Online]. Available: <https://www.oculus.com/>.
- [22] "CodeIgniter." [Online]. Available: <https://codeigniter.com>.
- [23] "OpenCV library." [Online]. Available: <http://opencv.org/>. [Accessed: 24-Jun-2017].
- [24] Carmody, T., "Why 'gorilla arm syndrome' rules out multitouch notebook displays," *Wired Oct*, vol. 10, 2010.
- [25] Jude, A., Poor, G. M. and Guinness, D., "Personal Space: User Defined Gesture Space for GUI Interaction," in *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, New York, NY, USA, 2014, pp. 1615–1620.
- [26] Hincapié-Ramos, J. D., Guo, X. Moghadasian, P. and Irani, P. "Consumed endurance: a metric to quantify arm fatigue of mid-air interactions," in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, 2014, pp. 1063–1072.
- [27] Rubin, K. H. and Coplan, R. J., *The Development of Shyness and Social Withdrawal*. Guilford Press, 2010.

## Prediction of user intentions using Web history

K.T.C.S. Samarasinghe  
 Department of Statistics & Computer  
 Science  
 Faculty of Science  
 University of Kelaniya  
 Sri Lanka  
 csenaraths@live.com

S. Jayalal  
 Department of Industrial Management  
 Faculty of Science  
 University of Kelaniya  
 Sri Lanka

### Abstract

*In the present internet has become much more necessary thing to humans and we use it as a way of sharing information and way of communication. If the networks can identify the user's intentions, it will be affecting to increase productivity and personalization. Predicting user intention(s) is interesting and useful for many applications such as threat identification, imposing restrictions and caching web details. The aim of this research is to develop a method to predict user intention using supervised machine learning methods with user's past historical behaviours. Experiments in this study used access log on a local server and focused on creating single user prediction and multiuser generalize prediction models. Experimental models were created based on several multi-classifier algorithms, such as Support Vector Machine (SVM), Multilayer Perceptron (MLP) and K-Nearest Neighbor (KNN). KNN based models outperform other used algorithms. Also results in this study show that there is some sort of behavioural patterns for peoples to use the internet according to the time and the groups they interact.*

**Key words:** Access Log, Behavioral Patterns, Historical Behaviors, Multi Classifier, Proxy Server, User Intentions

### 1. Introduction

At present, the internet is an essential part of the day to day life of people all around the globe and people used it as a way of sharing information and way of communication. Within the past few years, the internet has become an integral part of everyone's life because of the capabilities of mobile communication devices. Then, data which produce on the internet about users become much more personalized. Analyzing these data is generally known as Web mining. The basic purpose of web mining is discovering the knowledge from data on the web. Web mining can be categorizing into three main research fields based on which type of data to use for web mining as content mining, web structure mining and web usage

mining. Web content mining used to extract useful information from unstructured, semi-structured or structured contents of web documents and web structure mining is the process of analyzing the connection structure of the web sites by utilizing the link structure of the web documents. Web Usage Mining concentrates on the techniques that could predict the navigational pattern of the users while the users interact with the web-based on users browsing behaviours. Within the Internet, information about the user's behaviours is usually contained in three places. They are browsing history of a web browser, access log of a proxy server or local server and request log of a web server. Also, when increasing users who use the internet and amount of usage cause to greater impact on network utilization. Then, to keep speediness and latency on a productive level as users expect features of these networks should be enhanced. These enhancements may be software or hardware upgrades or techniques upgrades. Currently, there are many available techniques for local networks performance enhancements like bandwidth management, compression, Protocol acceleration etc. With bandwidth management network administrators have ability to change bandwidth allocation for each user /client or application. There are several techniques for bandwidth management. They are traffic shapers, QOS (Quality of service), WAN optimization, web caching. WAN optimization and web caching are based on caching and compression techniques. Traffic shapers have ability to limit the bandwidth of users or application. While QOS is to provide priority to networks, including dedicated bandwidth, controlled jitter, low latency and improved loss characteristics. But, in both cases traffic shapers and QOS network administrator decide to which application, the user should get priorities and what are the limitations. Usually, this option is the best practice. But, if the network contains numerous users, manual selection does not become the best practice.

So, we suggest a system that has the capability to identify users who misuse or cause to decrease productivity by user intentions based on their and other users browsing behaviours and dynamically control

bandwidth and other restriction. To build that, first, need a method to predict user intention. So, this research is focused on identifying best methods for user intention prediction by using Web Usage Mining with access log local server.

## 2. Related work

In present internet is the most democratic of all the mass media [1]. So, millions of peoples access various kind websites for different kind intentions. When we surf through the internet that creates and left behind the innumerable amount of data in various places. As an example, when some people browse between various websites or various page within the website, they create historical data called browsing history, this history can save on different places such as users browser, connected local sever and some third-party location if the user used a search engine and web hosted servers. These spatiotemporal patterns of the category of websites visited by time of access can uniquely characterize and identify users [4]. Researchers suggested the various types of methods, algorithms and techniques to identify and predict user behaviours, user information and authenticate users using data which users created internet history. The findings shown that the long-term interest of the users could predict more accurately by the historical context of the user [5]. In 2007 group of Chinese researchers suggest that they can predict user's demographic data such as age, gender using weblog data by improved SVM method [2]. In 2016 group of researchers show out how to use browsing behaviours to identify criminal profiles using custom build method [3].

According to John S. Brees Bayesian Networks and Bayesian classifiers typically have smaller memory requirements and allow for faster prediction for collaborative filtering and recommendation systems [6]. Also, Tong Zang's research [7] recommend linear classifiers for recommendation systems which use historical data on user preference and collaborative filtering method which focused on using just history of user preference to make the recommendation.

In the work [8], the K-Nearest Neighbor algorithm was used to combine mining of web server logs and web contents for classifying users' navigation pattern and predict users' future request.

Within this research, researchers focus on techniques to categorize user interest using users' historical log data.

## 3. Data set

This historical behavioural information can be stored in mainly in 3 different ways as browsing history of a web browser, access log of a proxy server or local server and request log. The primary source of data for this study was

the logs of URLs visited by users which stored in local servers. Researchers used local sever's history log of the University of Kelaniya, Sri Lanka for this study. Those history log files contain log entries belong to the time duration between April to June 2017 and approximately containing more than 60 million records and it is spread over 340 GB hard space. Usually, these log entries contain 120 to 200 fields including username, timestamp, mac address, IP address, sent and received data amount, duration, app-id and protocols. "SmartViewTracker" in "checkpoint" firewall system was used to generate those weblogs. With its ability to categorize user requests users can easily add the category field to log files.

## 4. Methodology

### 4.1. Data Preprocessing

Dataset contained many irrelevant records for this research study such as record requests which are done by the operating system (Microsoft Windows), not by users. So, first removed those type of data. The researcher was able to remove those records since they were categorized as "Microsoft & Office365 Services" by "SmartView". Also, removed records that unable to categorize.

In this research when predicting, we consider what are the previous tasks, users did in a relevant hour. So, extract hour from the time field. The date field extracted as day, month, and year. Also, since the user's behaviours may depend on a weekday. So, the researcher extracted relevant weekday for each date. This assumption was confirmed by the study of Muhammad Rizwan [10]. Also, test experiment done by researcher shows that weekday number causes 0 to 2 percent (0-2%) average precision change.

Usernames in the data set generally contain faculty or department and academic year information. The researcher was able to extract 12 types of faculty or department information. But, there are some other usernames which do not follow the general format. These other users can be academic, non-academic staff and system. So, they are categorized under a different category.

In selected data set there were 72 categories which requests belong. Creating a model that can predict 72 categories is a much harder task and need more computational power. For that type of model need a large amount of data for each of 72 categories. Also, many of these categories can be re-categorized again. As an example, "VoIP" and "Instant Messaging" categories can be categorizing into "Communication" category. So, the researcher manually categorizes these 72 categories into 12 main categories. Those 12 categories were Information (1), Social networking (2), Communication (3), Entertainment (4), Sport (5), Software (6), Education (7),



Sex (8), Other (9), Business (10), File (11) and Questionable (12). Figure 1 indicates how records are distributed over each main category and Figure 2 indicates data usage percentages used for each main category.

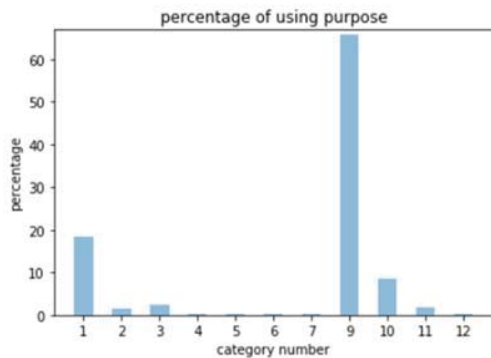


Figure 1. Percentage of using frequency vs main category.

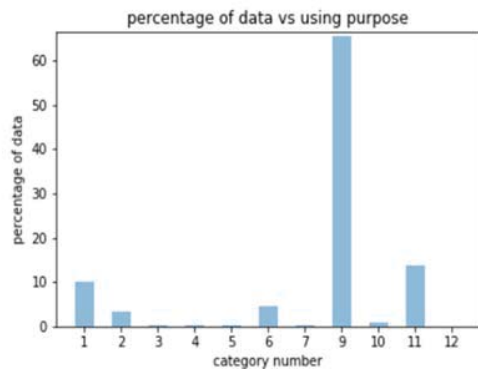


Figure 2. Percentage of data usage vs main category.

After the data preprocessing, split data by 0.7 to 0.3 train to test ratio. Also, selected UserId, Month, Day, Year, Hour, Weekday, Department or faculty id, Academic year and Main category id as data fields for analyzing.

#### 4.2. Experimental design

When creating experiments, we mainly concerned about How appropriate the selected method for the generalized model and single user model.

In the single user model, there is a single prediction model for each user and train each model for a specific user. In the generalized user model, there is a single model for every user and that model is trained for each user. By using the system with single user models, we can achieve high personalization while by System with a generalized model, we can achieve a high computational advantage over the single user model for lager user environment. Models are created using Multinomial Naïve Bayes.,

Gaussian Naïve Bayes, Support Vector Machine, Multi-layer Perception and K Nearest Neighbor.

After the parameter tuning for SVM selected “rbf” as a kernel, set probability as false and shrinking as true. For MLP used stochastic gradient-based optimizer, “Adam” as suitable weight optimization function. KNN used 2 different distance metric, hamming distance and Euclidean distance metric. For evaluating single user models, we used top 10 users who have most records on dataset and train and test each model for each user. Since the data set is more bias to category 9(other) researcher tested each model with and without (less bias) category 9. For evaluating the generalized model used a 30% of random data from data set.

As evaluation metrics, we considered accuracy, f1-score, precision, recall and time consumption of model for testing and training.

## 5. Results

### 5.1. Single user models

When considering single user models with original data set, the 9th and 1st category has shown the best f1-score, precision, recall for all models. Also, it shows that both KNN methods have similar performance and their performances are better than other models (Figure 3).

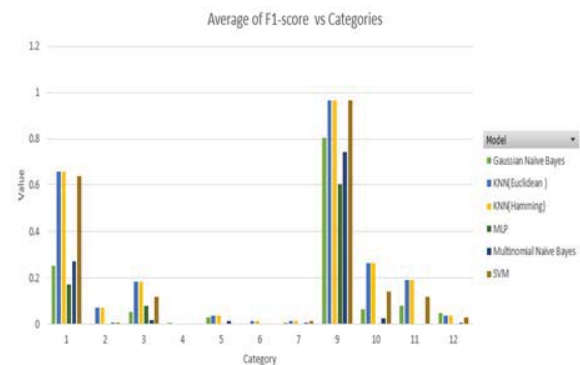


Figure 3. Average F1-score for categories between single user models.

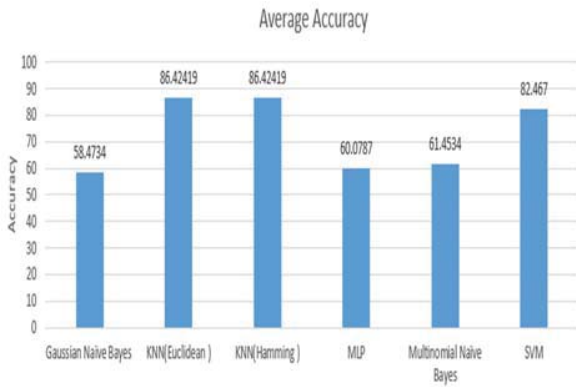


Figure 4. Average accuracy between Single user models for original dataset.

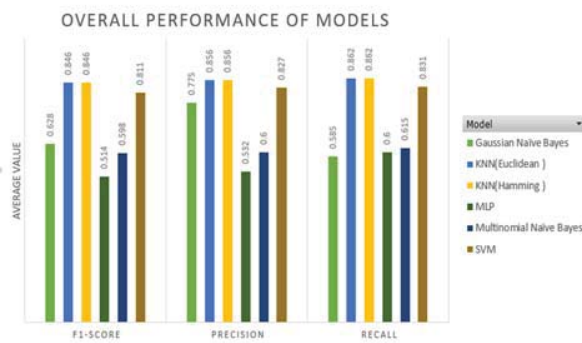


Figure 5. Overall performance between single user models.

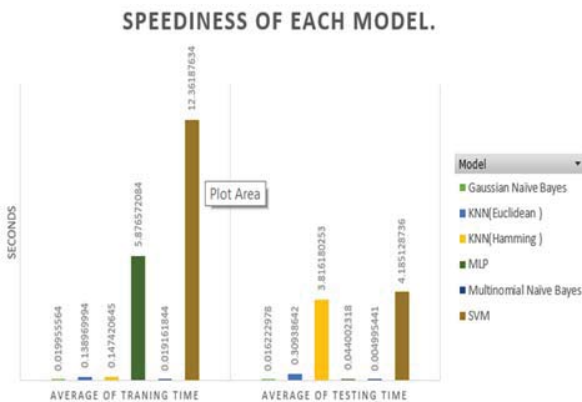


Figure 6. Average time consumption for training and testing.

Figure 4 shows how these models are fitting to the given data set. In here, both KNN models have shown a similar accuracy. By considering the overall performance of each model, both KNN models and SVM model have quite similar performance. But, KNN algorithms have provided better performance (Figure 5). When considering

speediness, the SVM model was slower than other models for both training and testing (predicting). Also, results indicate KNN (Hamming) needed more time than KNN (Euclidean) for predictions (Figure 6). Following figures are containing a comparison between results of Single User Models for data set which does not contain records belong to the “other” category.

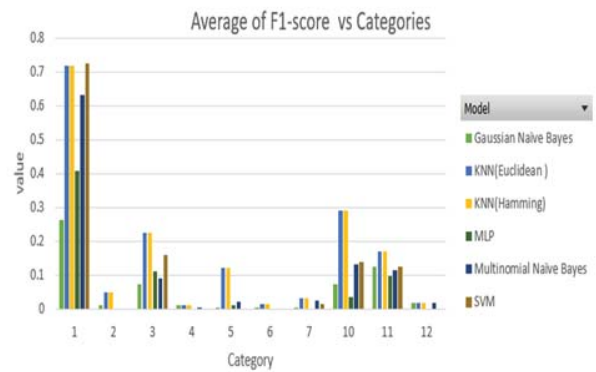


Figure 7. Average F1-score for categories between single user models with less biased dataset.

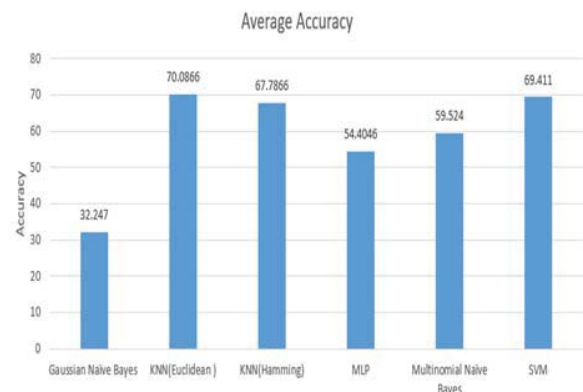


Figure 8. Average accuracy between Single user models for less biased dataset.

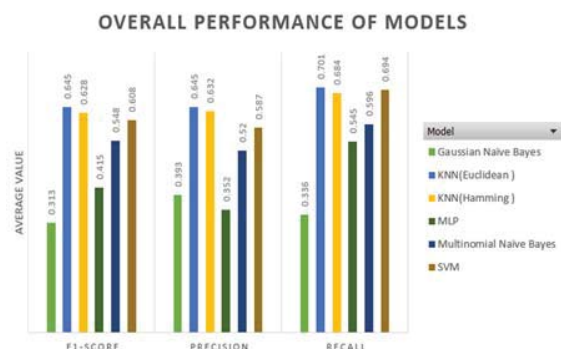


Figure 9. Overall performance between single user models with less biased dataset.

When considering single user models with fewer bias data set, the 1st category has shown the best f1- score, precision, recall for all models. Also, it shows that both KNN methods have similar performance and their performances are better than other models (Figure 7).

Figure 8 shows how these models are fitting to the given data set. In here, both KNN models have shown similar accuracy. By considering the overall performance of each model, both KNN models and SVM model has quite similar performances. But, KNN algorithms provide better performance (Figure 9.). When considering speediness, the SVM model is slower than other models for both training and testing (predicting). Also, it indicates KNN (Hamming) needs more time than the KNN (Euclidean) for predictions (Figure 10).

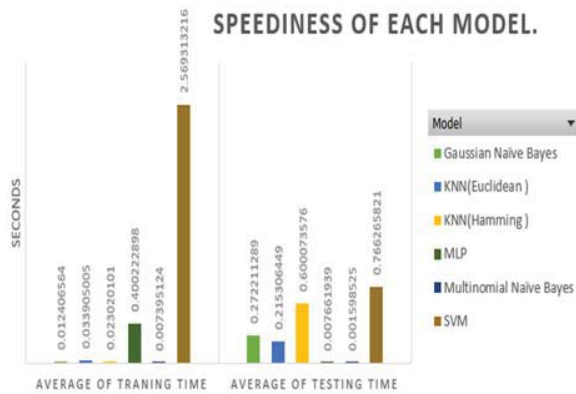


Figure 10. Average time consumption for training and testing with less biased dataset

## 5.2. For generalized model

For generalized user models with original data set, the 9<sup>th</sup> and 1<sup>st</sup> category have shown the best f1- score, precision, recall for all models. Also, it shows that both KNN methods have similar performance and their performances are better than other models (Figure 11).

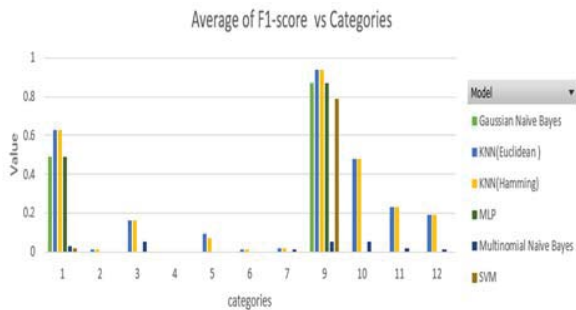


Figure 11. Average F1-score for categories between generalized models

Figure 12 shows how these models are fitting to the given data set. In here, both KNN models have shown similar accuracy. By considering overall performance of each models, both KNN models have performance. But, KNN with Euclidean provide better performance (Figure 13). When considering speediness, the SVM model is slower than other models for both training and testing (predicting) (Figure 14).

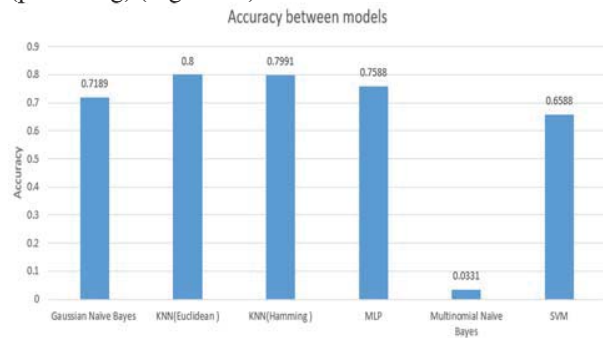


Figure 12. Accuracy between generalized models

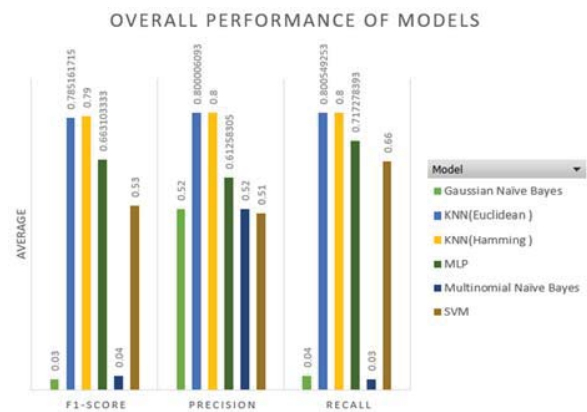
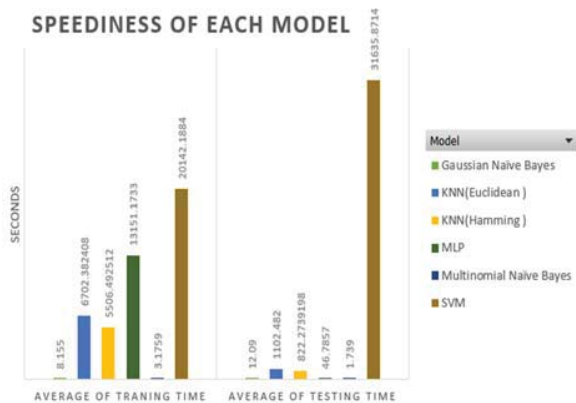


Figure 13. Overall performance between generalized models



**Figure 14. Average time consumption for training and testing**

## 6. Discussion

### 6.1. Multinomial Naïve Bayes models

Models based on Multinomial Naïve Bayes was able to achieve nearly 50% f1-scores for single user models according to Figure 5 and Figure 9. But, for generalized model F1-score value become 4%. When considered with other models, this model cannot recommend for user intention prediction based on past behaviours.

### 6.2. Gaussian Naïve Bayes models

Models based on, Gaussian Naïve Bayes achieved above 80% f1-scores. On most biased category, when the data set was less bias performance of these modes reduced according to Figure 5 and Figure 9. These models have not any ability to identify less biased data. When considered with other models, this model cannot recommend for user intention prediction based on past behaviours.

### 6.3. SVM

Models based on, SVM achieved above 70% f1-scores. On most biased category, when the data set was less bias performance of these models reduced. Since SVM is memory-based model, time consumption for training and testing relatively become more expansive.

### 6.4. MLP

Models based on, MLP achieved nearly 70% f1-scores when data set become larger. Time consumption for training can be large for these models; but, it is capable to predict faster than others. Since in this study researcher did not try to optimize this model. With better optimization, this model may produce greater results for large data sets.

## 6.5. KNN

Models based on KNN achieved the best performance above all other tested models. Every result in this study illustrates that these models could give better results than other models. In this study, researcher used 2 different KNN models and in every scenario, they produce nearly identical results. But, considering time consumption KNN (Euclidean) is giving better time consumption for single user models, or when data set is relatively small. While KNN (Hamming) gives better time consumption for generalized user models or relatively large data set.

All the above models were relatively performing better on the most biased class/category. This can be obvious when considering categorize f1-scores for each model. Since in study does not consider minutes and minutes grouped into relevant hours these tests should have less accuracy value, fewer performance values because of users may have different intentions during the one-hour duration.

Also, results in this study shown that there is some sort of behavioural patterns for peoples to use the internet according to time and group they belong. Otherwise, any of the used algorithms should not be able to give considerable results.

## 7. Limitation of the study

For this study researcher used 3 months of data at the University of Kelaniya. From that researcher extracted information from patterns which can identify in those 3 months. But, users may have behaviours pattern longer than 3 months. Also, this study did not use any demographical data about users. But, previous studies show that there is a relation between demographical data and browsing behaviours. So, this cause reduces performance.

## 8. Conclusion

Prediction of user intentions is an interesting and useful study and it can use for many applications. In this study, the researcher has tried to predict user intention from the user's past historical behaviours. Within this study, he used multiclass classification techniques for prediction and this study was done for the single user model and multi-user generalized model. In this study, KNN algorithms with Euclidean metric and KNN algorithms with Hamming metric able to give better results than other used algorithms. Also, results in this study shown that there is some sort of behavioural patterns for peoples to Using the internet according to time and group they belong or interact.

## 9. Future studies

Future studies can be done by using numerous data, by optimizing MLP, by adding demographic information, by considering MAC-Address as a parameter and optimizing KNN algorithm. Identify methods and architecture for a system that can identify user intentions based on their and other users browsing behaviours and suggest user who misuse or cause to decrease productivity. This study can be extended to some other user behaviours prediction study.

## 10. References

- [1] Singh, L., Gour, D., and Saluja, A., "Web Usage Mining Approaches for User's Request Prediction: A Survey", *International Journal of Computer Science and Information Technologies*, 6(3), 2015, pp. 2321-2325.
- [2] Hu, J., Zeng, H., Li, H., Niu, C., and Chen, Z., "Demographic prediction based on users browsing behavior", in *Proceedings of the 16th International Conference on World Wide Web - WWW 07*, 2007, doi:10.1145/1242572.1242594
- [3] Fan, X., Chow, K., and Xu, F., "Web user profiling based on behavior analysis", in *Proceedings of the 10th IFIP International Conference on Digital Forensics*, 2014, pp. 57-71. doi:10.5353/th\_b5731085
- [4] Abramson, M., and Gore, S., "Associative Patterns of Web Browsing Behavior (Internet)", in *Proceedings of the AAAI Fall Symposium Series*, 2013, Palo Alto, California. Retrieved from <http://www.aaai.org/ocs/index.php/FSS/FSS13/paper/viewPDFInterstitial/7587/7560>
- [5] White, R., Bailey, P., and Chen, L., "Predicting user interests from contextual information", in *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2009, doi: 10.1145/1571941.1572005
- [6] Breese, J. S., Heckerman, D., and Kadie, C., "Empirical analysis of predictive algorithms for collaborative filtering", in *Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence*, 1998, pp. 43-52.
- [7] Iyengar, V. S., and Zhang, T., "Recommender Systems Using Linear Classifiers", *Advances in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*, 2001, pp. 16-27. doi:10.1007/3-540-45357-1\_5
- [8] Liu, H., and Kešelj, V., "Combined mining of Web server logs and web contents for classifying user navigation patterns and predicting users' future requests", *Data & Knowledge Engineering*, 61(2), 2007, pp. 304-330. doi: 10.1016/j.datak.2006.06.001
- [9] Raschka, S., *Naive Bayes and Text Classification. Introduction and Theory*, 2014, Cornell University Library, Ithaca.

- [10] Rizwan, M., & Wan, W. (2018). Big Data Analysis to Observe Check-in Behavior Using Location-Based Social Media Data. *Information*,9(10), 257. doi:10.3390/info9100257

## Occupancy monitoring system for workplace washrooms

V. M. P. Godakandage, K. R. M. Kothalawala  
*Faculty of Graduate Studies and Research*  
*Sri Lanka Institute of Information Technology*  
*Sri Lanka*  
*mahirangi3@yahoo.com*

E. J. A. Chathumali, A. W. Madhubhashana  
*Faculty of Graduate Studies and Research*  
*Sri Lanka Institute of Information Technology*  
*Sri Lanka*

### Abstract

*With regard to rapid technological advancements majorly influencing our daily lives, Internet of Things (IoT) has been a topic of broad and current interest in the recent years. The capabilities of IoT can assist in revolutionizing the way people live and work, thereby improving quality of life. With the impact of IoT only continuing to propagate in the future, it can be used as a means of easing our day-to-day struggles. Therefore, with the assistance of IoT along with a few hardware, the proposed system, addresses the displeasing reality of queues and several visits for the washrooms due to them coming forth occupied. Thus, the focus of the intended system is on delivering a pleasant washroom experience for employees in an office environment providing them with an at-desk indication on the occupancy of the washroom cubicles reducing queues and disappointments.*

**Keywords:** *Internet of Things, Raspberry Pi, Message Queuing Telemetry Transport, Occupancy Detection, Publisher, Subscriber, Sensors*

### 1. Introduction

Nowadays, with highly active work environments, a customary issue faced by the employees is, having to make several trips to the washroom only to return back to their seats due to the cubicles being already occupied or wait around until a cubicle is free. Since, the occupancy of the washroom cubicles are not mostly visible at a glance once a person enters the washroom due to the inefficacy of the indicators attached the door locks, tenants often have to try each and every door to find an unoccupied cubicle. The issue is even more emphasized in case of the workplace having only a limited number of washrooms or the washrooms are being situated far from the employee desks. In a highly dynamic office environment where every minute counts, the matter could result in time wastage as well as bringing frustration to certain extremely pressurized employees.

As a solution for the aforementioned problem, a mechanism of informing the employees on the availability

of the washroom cubicles while they remain in their seats, would be ideal. Using the potential of hardware along with the Smart Building concept, the proposed system provides the resolution. The Internet of Things or IoT, which plays a major role in Smart Building systems, refers to a network of devices or things. These devices can collect and send information, receive and act on information or do both. IoT integrates intelligence to otherwise dumb devices, giving them the capability of interaction without the necessity of human engagement. Hence, the application of IoT can assist in solving a number of real-world complications encountered by people in their daily lives.

The proposed system decides on the occupancy of the cubicles by detecting the motions inside the cubicles and placement of the cubicle doors. In case of a particular cubicle being occupied, the system would light up the bulb inside the cubicle in conjunction with the relevant bulb outside the washroom indicating the occupancy. Simultaneously, the cubicle door would be automatically locked and would only be unlocked upon a button press action performed by the occupant. Once the door is unlocked and the occupant leaves the cubicle, both the light bulbs placed inside the cubicle and outside the washroom would be turned off in order to save electricity and indicate cubicle unavailability respectively. While the lights outside the washroom would indicate the washroom availability status to the employees nearby, employees all around the office can also check the availability status of the washroom cubicles, as they remain seated, with the assistance of the web interface provided by the system. In general, the proposed system incorporates smartness to the conventional washrooms thus saving time as well as energy.

### 2. Background study

A. R. A. Rudin et al. [1] proposed a system that monitors the occupancy of campus sport facilities by using the concept of internet of things (IOT). The aim of this proposed system is to facilitate staff members and students to find out the information about available sport facilities which are already provided by the university. It

has been firstly implemented for campus badminton court. Monitoring occupancy of the sport facilities, booking the sport facilities and providing information about sports are major functionalities implemented in the system. In the system, Raspberry Pi 2 has been used as the system's main microcontroller. This microcontroller is the one who find out the availability of sport facilities, by detecting the movements near the doors of each facility. Also, the system has been used Passive Infrared (PIR) sensors to identify movements near the door. In order to book the sport areas for events they have used a web-based application system which is implemented using HTML. If the area is already booked at a particular time, the door will be locked, and it will be only unlocked when identification number is read by the RFID scanner.

Jing He et al. [2] proposed a smart resource management system which is based on Internet of Things (IOT). Proposed system mainly focused on checking the occupancy status of the chairs. In the intelligent chair system, all the physical chairs with embedded sensors are connected to a network for collecting relevant information. This is conducted by the functionality of IOT. System has used a cloud architecture to upload that collected information. They have used the cloud architecture to access the data from anywhere when they need. Firstly, the system is implemented to a single chair and they are going to expand the implementation to connected chairs as an intelligent chair system. Proposed system has three main components. First one is the assembled Arduino system which has the major functionalities like scanning user ID, gaining chair occupancy status and to send all that information to the cloud server. Second component of the system is the cloud server which is the place they have been gathered all the collected data. Final component is the mobile application which has been developed using Android technology. Main functionality of this application is to display and monitor the chair occupancy states and user information. They have used RFID readers to identify users. It has been automatically done from the reader by using the user's identification (User ID). Then system passes the collected information to the Arduino microcontroller. Also, there is a pressure resistor to check the pressure levels. If the pressure level exceeds the pre-set threshold, timestamp is passed to the Atheros AR9331 microcontroller. This process has been done using a bridge library. Lastly, the bridge library provides a HTTP POST to the cloud server along with the User ID which has been scanned from the RFID reader.

Xiao Ling et al. [3] proposed system to identify vehicle parking spaces and detecting occupancy using vision based IOT devices. Aim of this project is to visible free parking spaces to the drivers. System has been developed with the concept of smart cities. Proposed

system developed using a camera-based space identification system. It has been used a single camera connected with IOT device. The purpose of having this camera and the edge device is to identify and monitor the status of available parking areas within the covering view. System contains with two major components. They are parking lot topology learning and parking lot occupation detection. In parking lot topology learning system, car-driven strategy has been used to identify the parking points. To do that they have used Haar-featured object identification algorithm which is included in OpenCV library. From the second component system identifies the occupied parking slots. This system has been developed using Raspberry Pi and by using an occupation detection pipeline. It identifies the available parking slots correctly.

### 3. Methodology

The main focus of the system is to detect and indicate the occupancy of washroom cubicles and automating the door locks and lights. A prototype model has been created in order to simulate the scenario.

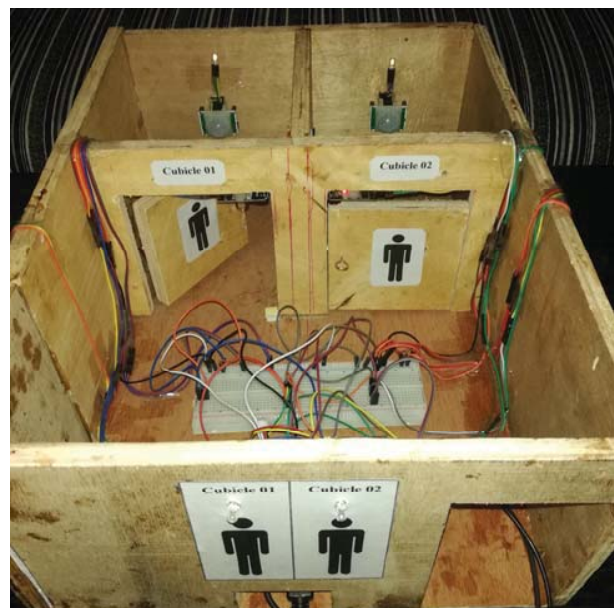


Figure 1. Prototype model.

Two types of sensors have been used in implementing the prototype. PIR (Passive Infrared) Motion sensors, as the name itself implies, are small, light, low-cost sensors capable of detecting human movement in or out of its range, with the assistance of variations in the infrared light, is used. IR (Infrared) Proximity sensor, is a multipurpose sensor capable of sensing obstacles by continuously transmitting infrared light and identifying

nearby objects via the means of reflected light, is also used. A PIR (Passive Infrared) Motion sensor is placed inside each cubicle which aids in detecting motions inside the cubicle. Furthermore, IR (Infrared) Proximity sensors are positioned at the top of the washroom cubicle doors, senses the status of the door identifying whether it is closed or opened. If both the PIR and IR sensors yield positive outputs, the cubicle is considered occupied [4]. The reason for checking positive outputs from both the sensors when deciding the occupancy is for better accuracy. For an instance, in case of the cubicle door being closed due to wind without a person inside, the IR Proximity sensor would detect positive inputs. Moreover, if a person goes inside the cubicle without closing the door perhaps with the intention of immediately returning, the PIR Motion sensor would yield positive inputs. In both the scenarios cubicle would be assumed occupied resulting in inaccurate outcomes provided that only one of the sensor data is taken into consideration. Hence, both the sensor inputs are taken into account when determining the conditions to check the occupancy. The sensor data would be utilized by the Raspberry Pi 3 which acts as the main microcontroller. Using the MQTT (Message Queuing Telemetry Transport) protocol which is a lightweight publish-subscribe based messaging protocol widely used in IoT applications, the Raspberry Pi acts as the publisher and publishes the data gathered through the PIR and IR sensors under a single sensor data topic for each cubicle [5], [6].

```
Motion_IR_status_1 =
str(GPIO.input(Motion_1)) + ';' +
str(GPIO.input(sensor_1))

mqttc.publish("topic/Motion_IR_status",
Motion_IR_status_1)
```

The publisher continuously checks and publishes the sensor data in order to process real time data. The MQTT broker, which is a separate computer, plays the role of the subscriber. Once data is published under the corresponding sensor data topic, it would be redirected to a callback function which processes the data published. If both the PIR and IR sensor results are positive, it means that there is motion detected inside the cubicle and the cubicle door is closed, implying that the particular cubicle is occupied. Since both the sensor data are checked in order to decide on the occupancy of the cubicle, both sensor data are published under a single topic. The occupancy of the cubicle suggests that the door lock should be automatically locked and the relevant light bulbs; both inside and outside the cubicle should be turned on simultaneously. Therefore, the callback

function publishes topics indicating what the status of the lock and lights should be.

```
if Motion_01 is "1" and IR_01 is "1":
mqttc.publish('topic/Light_1_status', 1)
mqttc.publish('topic/Door_Lock_1', "Lock")
```

The subscriber in the Raspberry Pi is subscribed to the corresponding topics and accordingly, it automatically locks the door with the assistance a Servo Motor connected to the door lock and in the meantime, turns on the light bulb inside the cubicle and one outside the washroom, representing the specific cubicle [7].

```
if "UnLock" in Lock_1_status:
p.ChangeDutyCycle(12.6)
else:
p.ChangeDutyCycle(6)
```

Once the occupant inside the cubicle wants to leave, he or she can press on a button placed inside the cubicle. The publisher in the Raspberry Pi publishes the status of the aforementioned button under a specific topic when publishing sensor data. The subscribed MQTT broker has a separate callback function to be executed in case of a message from the particular button related topic. The Raspberry Pi subscriber unlocks the cubicle door according to the messages published by the callback function. When the cubicle door is unlocked, the relevant light bulbs are also turned off indicating the availability of the cubicle. While the light outside the cubicle denotes the availability and unavailability status of the cubicles to those nearby, the same occupancy status can be monitored by tenants via the use of a web interface. The provided web interface exhibits a statistical view of occupancy state along with the time durations. The web application is implemented using Node-RED which is an open source visual tool developed by the IBM's Emerging Technology Services for wiring together hardware devices, APIs and online services [8]. Node-RED is widely used for IoT applications.

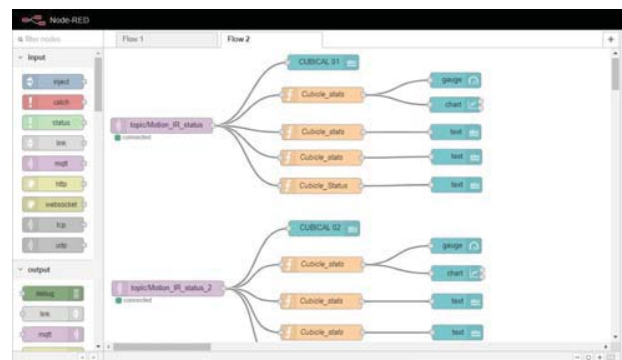


Figure 2. Node-RED flow.



A MQTT node is used in order to connect to the MQTT broker. Based on the topic data, the wired function nodes decide on the output projection in the web interface.

#### 4. Experimental results

The result of the system is an at-desk indication of the availability of the washroom cubicles. Initially, PIR and IR sensor data along with the button status are published from the Raspberry Pi. The MQTT broker is subscribed to these topics and therefore the messages published are received by it.

```
Motion 01 = 1

('Received message Button 01 =', '1')
('Received Message = ', '1;0')
```

Figure 3. Messages received by the Subscriber.

According to the data published by the topics, the relevant callback functions publish topics indicating whether to lock or unlock the door and turn on or turn off the lights. Then the Raspberry Pi subscriber performs the actions accordingly. For an instance, if cubicle 02 is considered occupied, as demonstrated in Figure 4, the door is automatically locked, the light bulb inside the cubicle is turned on and same happens with the light bulb outside the washroom to denote the occupancy.



Figure 4. Model showing occupancy of Cubicle 02.

In the meantime, with regards to the node-RED application, the occupancy status is illustrated along with the time durational statistics. As an example, in case of cubicle 01 being occupied, the unavailability is represented through the web interface using a gauge and a chart as shown in Figure 5.

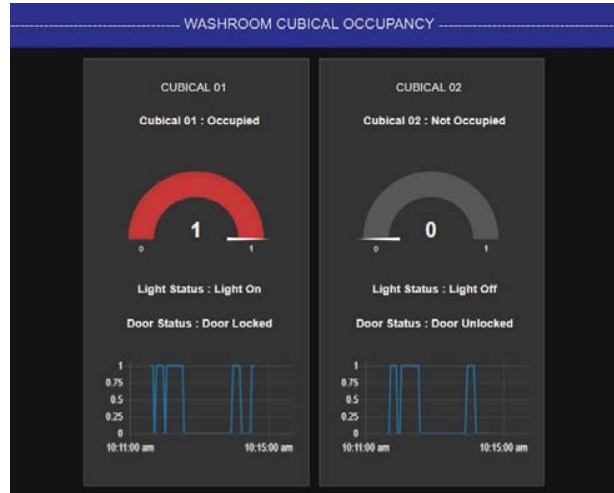


Figure 5. Web interface showing occupancy of Cubicle 01.

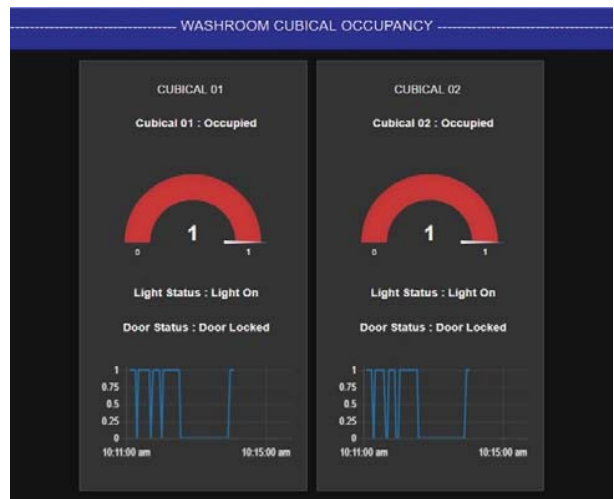


Figure 6. Web interface showing occupancy of both Cubicles.

#### 5. Conclusion

Present-day, with the emergence of Internet of Things or IoT, almost all the devices are interconnected via the internet which sends and receives data. Proposed washroom occupancy system was developed with the assistance of IoT concepts incorporated with a few hardware items. The system is mainly focused delivering an at-desk indication on the availability status of the washroom cubicles in the organizations, eliminating the need for the employees having to physically go there. In order to decide on the cubicle occupancy status, system

has considered two main conditions. The motions inside the cubicles are captured using PIR motion sensors and the cubicle door being closed is detected via the use of IR proximity sensors. Accordingly, occupancy is detected and door locking or unlocking and turning on or off lights are automated. The results of the system indicate that the system was able to achieve the expected goal. The system can be further expanded to be implemented in several other public places in the future.

## 6. References

- [1] A. Rudin, L. Audah, A. Jamil and J. Abdullah, "Occupancy Monitoring System for Campus Sport Facilities Using the Internet of Things(IoT)," *IEEE Conference on Wireless Sensors (ICWiSe)*, 2016.
- [2] J. He, A. Atabekov and H. M. Haddad, "Internet-of-Things Based Smart Resource Management System: A Case Study Intelligent Chair System," *IEEE*, 2016.
- [3] X. Ling, J. Sheng, O. Baiocchi, X. Liu and M. E. Tolentino, "Identifying Parking Spaces & Detecting Occupancy Using Vision-based IoT Devices," *IEEE*, 2017.
- [4] "Maker Pro," [Online]. Available: <https://maker.pro/raspberry-pi/tutorial/how-to-interface-a-pir-motion-sensor-with-raspberry-pi-gpio>. [Accessed 30 September 2018].
- [5] S. Cope, "Steves-internet-guide.com," [Online]. Available: <http://www.steves-internet-guide.com/install-mosquitto-broker/#windows>. [Accessed 1 October 2018].
- [6] "HiveMQ," [Online] Available: <https://www.hivemq.com/blog/how-to-get-started-with-mqtt>. [Accessed 28 September 2018].
- [7] "Components101.com," [Online]. Available: <https://components101.com/servo-motor-basics-pinout-datasheet>. [Accessed 18 October 2018].
- [8] "Node-RED," [Online]. Available: <https://nodered.org/>. [Accessed 2 October 2018].

## Candidate recruitment based on automatic answer evaluation using WordNet

J. Kasundi  
*Faculty of Information Technology*  
*University of Moratuwa*  
*Sri Lanka*  
*jeneeshakasundi94@gmail.com*

G.U. Ganegoda  
*Faculty of Information Technology*  
*University of Moratuwa*  
*Sri Lanka*  
*upekshag@uom.lk*

### Abstract

*Any organization's future survival in the industry mainly depends on the decisions they take today and recruiting a new employee is one of those critical decisions that an organization has to take. Recruitment process of an organization should have the capability to find the right person to the right job. Most of the organizations today, interview candidates to test their job skills but not their personality level. But recruiting the most skilled person does not work out in a good manner always. Especially for the positions like leaders, they have to consider about the personality of the employee. And also, in the industries like IT industry, employees are supposed to work in project teams but not individually. That is where personality matters the most. To work in a team, all the members should be able to cooperate with each other without facing any difficulty. Therefore candidate recommendation chatbot system is suggested and it is capable of evaluating both technical skills and personality traits of the candidate and providing a final recommendation based on the scores obtained to those two sections. The system mainly contains four modules; Question Generation and Dialogue Flow Maintaining Module, Technical Answers Evaluating Module, Vocabulary Based Personality Evaluating Module and Candidate Recommendation Module. This paper will be discussing about the answer evaluation module of the system which contains an ontology to store java related technical questions and answers and has used WordNet to measure the correctness of the answers.*

**Keywords:** *Automatic Answer Evaluation, Chatbot, WordNet, IT industry Recruitment*

### 1. Introduction

Recruitment process in any organization is a critical and challenging process since it affects the future survival of the organization. Having good quality employees is the number one goal for any employer. The efficient functioning of any organization is largely dependent upon

the employees that comprise that organization. For this reason, most of the organizations should not only be mindful of, but critically evaluate the methods they use to make selection decisions. Most of the companies test the job skills of the candidates at the interviews and then decide on recruiting them or not [1]. But having technical skills and job skills are not the only factors that determine the performance of a person inside the organization. No matter how skilled the employee is, s/he has to work well with the coworkers. This is where the personality starts to matter. Specially in industries like IT industry where people are working in teams to achieve organizational and project goals, the first thing is the team should be strong [2]. This means there should be good communication, mutual understanding and respect between the team members. However, even a very skilled and experienced employee may have poor personality when working with his/her coworkers. For example, some people are having more dominant personality and they try to dominate other people.

This is not always a good thing. Especially when a person is in a position of a leader, then they should be very flexible and should respect other peoples' ideas and communicate well. But the existing recruitment processes in many organizations do not evaluate the personality levels of the candidate [3]. This can be a critical issue because successful recruitment is a direct reflection of the validity and professionalism of any company. Employing the right people for a business is the most important part of the organization. It is essential to have a good recruitment process to attract the right kind of employees for a business. Therefore, the recruitment process should be cost effective as well as time effective. Recruiting and training can be expensive and time consuming. Therefore, when a company is recruiting, they have to make sure that they are making the right choices.

Hence, a solution is suggested to this problem where both technical skills and personality of the candidates are evaluated and the time and cost of the organization are saved. An automated chatbot system is then suggested and the most important thing is that, it is capable of evaluating both technical skills and personality traits of the

candidates. The suggested system is capable of analyzing the CVs of candidates and asking questions based on the technical areas mentioned in their CVs. In this case the system is limited to technical questions only from java. And also, the candidates are asked some HR related questions. Then based on the technical answers they have given, technical skill level is calculated and based on the vocabulary they have used to answer HR related questions, personality traits are calculated according to the big five modal personality traits.

This paper will discuss how the technical answer evaluation module works and other related information. The system will ask a question based on Java knowledge from the candidate and then candidate will provide the answer. The module has to analyze the correctness of the answer compared to the answers retrieved from the system. Our system stores Java knowledge in an ontology and retrieve answers from that ontology to evaluate the correctness of user given answers. An ontology is a formal representation of the knowledge by a set of concepts within a domain and the relationships between those concepts. An Ontology is used here because of the flexibility of it and it is great for combining information from various sources and it is capable of inferring new knowledge from the existing knowledge. More details about the ontology are discussed in the next chapters of this paper. Since the ontology is in RDF format, Sparql Protocol and RDF Query Language (SPARQL) queries have to be used to extract information from the ontology. When generating SPARQL queries, RDF triples format is needed. RDF triple format represents a sentence in the form of subject-predicate-object. Predicate shows how subject and object are related. In this case the SPARQL query is generated using the questions to match the triples patterns in the RDF graph and then the unknown answers are retrieved from the ontology. When evaluating answers one important point is that there can be multiple answers for a given question and one answer can be given in multiple ways. Therefore, the system will extract all the possible answers from the ontology and then evaluate user given answer based on that. Semantic similarity between sentences are used to evaluate the correctness of answers. To measure semantic similarity, there are various methods such as cosine similarity, Euclidian distance, word2vec, doc2vec, wordnet and so on. In this approach wordnet is used because it is a thesaurus which is structured according to sense relations. In wordnet words are represented by using the sense relations such as synonyms, antonyms, hypernyms and homonyms. Therefor it can compute the similarity based on the meanings of words. Then at the end of the process in this technical answer evaluation module, it will output a percentage value for the technical skill level of the candidate.

In this paper the section II will describe related work done in the similar research areas, and section III will

discuss about the ontology which is developed for the system. A description generating SPARQL queries will be given in section IV and then in the section V there will be a description on answer evaluation and semantic similarity. Methodology is described in section VI and section VII contains the results and discussion of the module. Finally, section VIII will provide a conclusion on the paper.

## 2. Related work

There are many other researches have been conducted in the areas related to this. For example, to evaluate the semantic similarity between sentences there are other methods such as cosine similarity, word2vec approach and doc2vec approach.

The cosine similarity is calculated based on the cosine of the angle between two vectors. This metric is not a measurement of magnitude but a measurement of orientation. Cosine similarity can be considered as a measurement of similarity between documents on a normalized space. The cosine similarity formula (1) is given below.

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|} \quad (1)$$

Cosine similarity can generate a metric that describes how related are two documents. But when using cosine similarity, we cannot get a measurement on how two text segments are similar, by their meaning.

And also, there are other methods such as vector space models which are used to measure the similarity between text. Vector space models can transform different length text such as word, sentence to a numeric vector. The most basic text vectorization method is term frequency-inverse document frequency (TFIDF) and it is capable of defining a space where each term in the vocabulary is represented by a separate and orthogonal dimension. But it can have some issues like ignoring n-gram phrases and having many numbers of dimensions [4].

One of the other existing approach to measure text similarity is by using Euclidian distance. Euclidian distance is used to calculate the semantic similarity between sentences. Euclidian distance is the ordinary straight-line distance between two points in Euclidian space [5]. Euclidian distance is calculated using the below formula (2).

$$\sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (2)$$

When used in machine learning, Euclidian distance can be used to measure the similarity between vectors. In here this will show how far the vectors are in Euclidian space and it means higher the vectors are less similarity they are having.

And also, there is an existing project named Platypus which is used to convert natural language questions into RDF triples and provide answers to the user [6]. Platypus is an open source question answering system which is based on a modular approach. The user input is passed to a core that distribute inputs between modules. Each module performs their operations to simplify the user input and solve it partially and the process is repeated as long as there are improvements in the results. Finally, results are displayed to the user in the user interface. Advanced natural language techniques are used in the system and it uses Wikidata. Platypus is capable of answering a lot of questions from the simple to the complicated. The system is capable of performing spell checking for the user while solving some simple mathematics formula given in natural language. The below figure 1 shows an overview of the developed system.

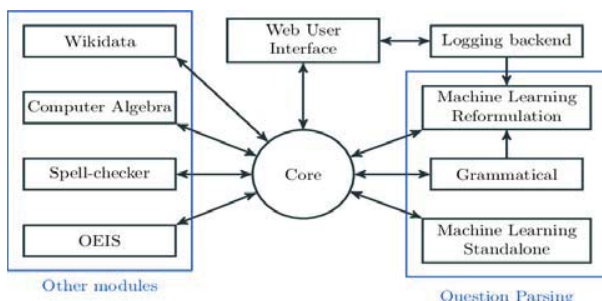


Figure 1 : Overview of Platypus.

### 3. Ontology

In our system an ontology is used to model questions and answers related to java technical knowledge. An ontology is a formal representation of the knowledge by a set of concepts within a domain and the relationships between those concepts. An ontology is used here because of the flexibility of it and it is great for combining information from various sources and it is capable of inferring new knowledge from the existing knowledge. And also, with compared to traditional knowledge base, an ontology is more suited for an evolutionary approach to the specification of requirements and domain knowledge [7].

Modeling ontologies is a difficult task which requires high cost. But it has many advantages of using ontologies to model domain knowledge rather than using a traditional database. The most important factor is the flexibility of ontologies compared to traditional databases and therefore ontologies can be used to combine variety of information from different sources and infer new facts using the existing information. And also due to the flexibility of ontologies it is very easy to extend and use for future work. Ontologies are more suitable for evolutionary approaches when compared with traditional databases.

Every ontology contains a dictionary with explanation of the terms and relations among the terms used in ontology. In other words, ontology is a specification of concepts and it shows the conceptual description of the specific content and help to identify appropriate terms and relationships in a particular domain. Ontologies can represent the hierarchical dependents between terms while providing explanations, descriptions and definitions for the terms. And also, it is possible to have a graphical representation of the ontology. When using ontologies, it is easy to capture knowledge about the problem domain.

The ontology is represented in RDF language and therefore SPARQL query templates are used to extract data from the ontology [8]. RDF is a model which is used to encode semantic relationships between data so that the relationships can be interpreted computationally. In RDF, triples are used to store data in the form of subject-predicate-object. And also, the ontology contains individuals, classes and properties. In this module the ontology represents java domain knowledge using classes, individuals and relationships that represent use of data type properties and object type properties [9]. The structure of the developed ontology is shown in Figure 2.

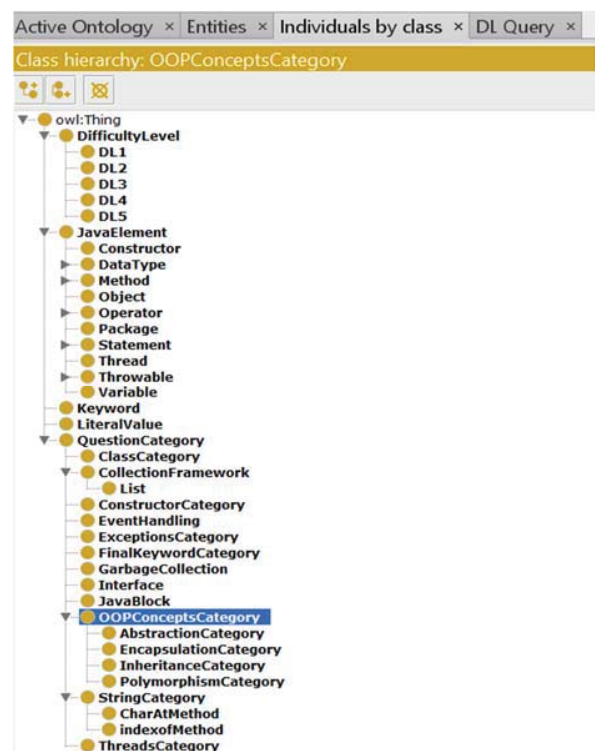


Figure 2 : Java ontology structure.

The basic elements of the ontology are classes, relations, instances, functions and axioms. Classes are used to represent concepts which are related to a specific knowledge domain. Instances are members of a class and

it represents specific entities. Axioms are statements which are always true, and relations are used to describe the types of interactions between entities in the ontology. Properties are used to represent relations and there are two types of properties. One is data type property and it is used to link individuals to data values. The other type is object type property and it is used to show relationships between individuals.

#### 4. Generating SPARQL queries

Sparql Protocol and RDF Query Language (SPARQL) is used to retrieve data which are stored in RDF format. SPARQL can extract data in different formats such as RDF subgraphs, typed and untyped literals, URIs with the help of aggregate functions, subqueries, property paths and complex joins. And also, SPARQL can transform data from one vocabulary into another and it has the power to explore data via query for relations which are unknown. SPARQL is based on RDF Turtle serialization and basic graph pattern matching. The basic graph patterns are also like RDF triples but in graph patterns there can be a variable for subject, predicate or object. A Basic Graph Pattern is a set of Triple Pattern. In SPARQL also the "SELECT" statement is used to retrieve values for variables and the result is a table. The graph patterns in the SPARQL is written in a way that it matches to a particular subgraph of the RDF data. In this module SPARQL is used to access the RDF document of the ontology and retrieve unknown data. And then the variables, used in the graph pattern is matched with the RDF terms in the subgraph. Then it will result the RDF graph equivalent to that subgraph with the values for our variables. An example for graph pattern matching is given in the below Figure 3.

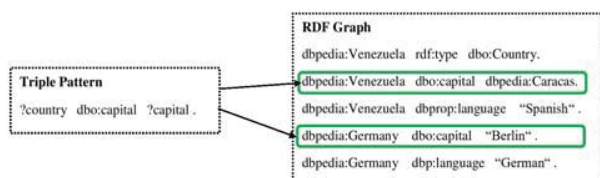


Figure 3: Graph pattern matching.

In the above Figure 3, ?country and ?capital are variable and the values for those variables are what we want to retrieve from the RDF file. The SPARQL query is then generated using the matching triple pattern to retrieve our data.

#### 5. Comparing answers given by the user

The module uses a corpus based and knowledge-based approach to measure the semantic similarity. The typical approach to measure similarity is usually done by using

lexical similarity. Lexical similarity is measured based on the degree to which the word sets of two texts are similar and it depends on the number of common words in the texts. But this is not always applicable because there can be same thing expressed in different words. Therefore, in this module semantic similarity approach is used. Semantic similarity is measured based on the meaning of words and it uses a metric based on the likeliness of their meanings. When two sentences are given, the module needs to calculate a score to indicate the semantic similarity between them. To measure semantic similarity of sentences, this module uses the semantic similarity of words in each sentence.

For that task the system uses the WordNet similarity measurement. WordNet is a large lexical database developed for the English language. It contains synsets. Synsets are sets of synonyms and it consists of groups of Nouns, Verbs, Adjectives and Adverbs. And also, WordNet is capable of providing usage examples and short definitions while recording relations in the synsets. It can also be called as a combination of dictionary and thesaurus and it is very useful for NLP and computational linguistics. By using WordNet, the system calculates the average similarity for a candidate given answer, for a question against the list of answers retrieved from the system. That simply means the same answer is evaluated against multiple correct answers stored in the system and calculate the similarity with each answer and then calculate the average similarity. Then the average similarity is accepted by the system as the user obtained marks for that particular question. Then the marks for each question is calculated in the same manner and at the end, the system gets a list of questions, candidate given answers for each question and the obtained marks for each question by a particular candidate.

When computing similarity between words, the algorithm finds the maximum semantic similarity only within the same POS (parts-of-speech) word classes. In word net, it can group only nouns, verbs, adjectives and adverbs into synsets. All those belong to open class words. Therefore, the algorithm first uses parts-of-speech to tag the sentence and then indicate what tags we are looking for, in WordNet. Then the algorithm will find the maximum similarity within the same POS tag classes.

#### POS tagging

POS tagging (parts of speech tagging) is used to read text in natural language and assign parts of speech tag to each word such as noun, adjective, verb and so on. POS tagging is not generic. A single word can have different parts-of-speech tags based on the sentences they are used in. For example, a simple sentence like "OOP concepts are encapsulation, abstraction, Polymorphism and inheritance" will be tagged in the below format.

[('OOP', 'NNP'), ('concepts', 'NNS'), ('are', 'VBP'), ('encapsulation', 'NN'), (';', ','), ('abstraction', 'NN'), (';', ','), ('Polymorphism', 'NNP'), ('and', 'CC'), ('inheritance', 'NN')]

## 6. Methodology

The below figure 4 explains the flow of the module developed for the system which consists of several steps. As the first step, the module has to get the question and the answer given by the user as an input. In the next step the SPARQL query is generated by using the question. Then the answers will be extracted from the ontology. And in the final stage of the module, the answer given by the user will be evaluated against the answers retrieved from the system. This process will be briefly explained further in this section.

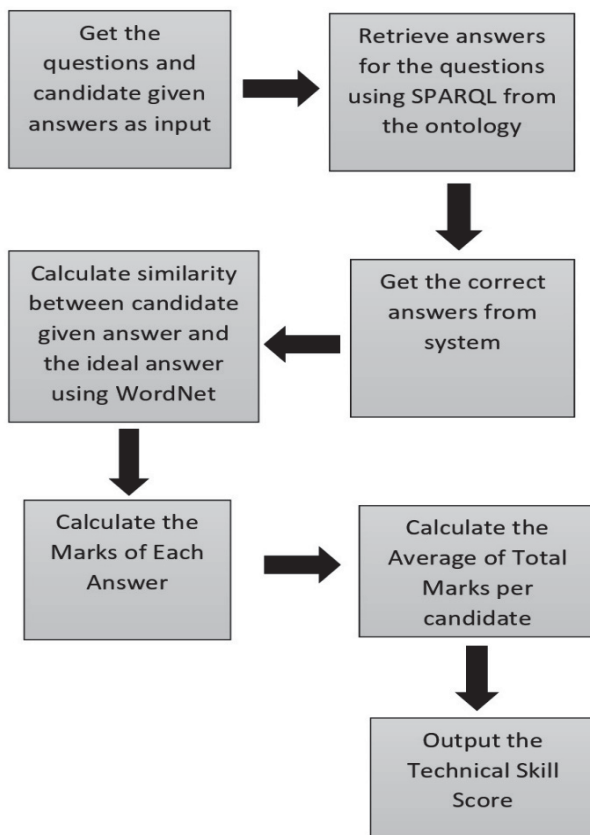


Figure 4: Flow of the module.

The first step in this module is to get the questions and the answers given by users to those particular questions. Then the system needs to extract answers from the ontology. For that purpose, SPARQL (Sparql Protocol and Query Language) is used. SPARQL can be used to retrieve data which are stored in RDF format. SPARQL is based on RDF Turtle serialization and basic graph pattern matching. The basic graph patterns are also like RDF triples but in

graph patterns there can be a variable for subject, predicate or object. A Basic Graph Pattern is a set of Triple Pattern and in SPARQL it is written in a way that it matches to a particular sub graph of the RDF data. The matching graph pattern should be generated to query the RDF document the generated SPARQL query is showed in the below Figure 5.

```

SELECT DISTINCT ?answer ?otherAnswers
WHERE {
  {JavaOntology:Q1 JavaOntology:hasAnswer ?answer. }
  UNION {
    JavaOntology:Q1 owl:sameAs ?sameQuestion.
    ?sameQuestion JavaOntology:hasAnswer ?otherAnswers.
  } UNION {
    ?sameQuestion owl:sameAs JavaOntology:Q1.
    ?sameQuestion JavaOntology:hasAnswer ?otherAnswers.
  }
}
  
```

Figure 5: SPARQL query format.

By using the above query, the answers are retrieved from the ontology and then the next step is to evaluate the answers. This is done by measuring semantic similarity of the candidate given answer against the system retrieved answers using wordnet. According to the developed algorithm if the module needs to measure semantic similarity of two sentences S1 and S2, the first step is finding the word in S2 which has the maximum semantic similarity with each word of S1. Then the same process is done for each word in S2 to find the most similar word in S1. And then the algorithm sums up each maximum word similarity and normalize it by dividing from the length of the sentence.

The similarity between two sentences S1 and S2 is calculated using the below formula. “w” represents a word in S1.

$$similarity(S1, S2) = \left( \frac{\sum_{(w \in S1)} maximumSimilarity(w, S2)}{No\ of\ words\ in\ S1} \right) (3)$$

When computing similarity between words, the algorithm finds the maximum semantic similarity only within the same POS (parts-of-speech) word classes. In word net, it can group only nouns, verbs, adjectives and adverbs into synsets. All those are belonging to open class words. Therefore, the algorithm first use parts-of-speech to tag the sentence and then indicate what tags we are looking for, in WordNet. Then the algorithm will find the maximum similarity within the same pos tag classes. For example, if w is a noun, then the similar word to w will be calculated only within the nouns in sentence 2. For other words which do not belong to the open word classes, then the semantic similarity will not be calculated for those words and if those two words are identical the algorithm will assign 1 as the similarity measure.

When measuring semantic similarity, the algorithm uses path\_similarity method in wordnet. Wordnet is a kind of thesaurus and it uses not only by using the word but also by using the word senses. When comparing the similarity of two sentences it is not enough to calculate only the word but also the sense of the words has to be taken in to account. Since wordnet is a thesaurus, the structure of the thesaurus can be used to define similarities between words. The synsets can be organized as graphs and therefore the similarity of sunsets is measured based on the shortest path between them. The path similarity used in wordnet calculates the shortest number of edges from one-word sense to another in a hierarchical structure. If two-word senses are having a longer path, then those words are considered less similar words and if the path is shorter between word senses then it is more similar. Path similarity return a score which is between 0 -1 based on the shortest path that connects the senses in the is-a taxonomy. It will return 1 if the two compared words are identical. Therefore, the higher the path similarity is, the more the similarity between two words. The path similarity is calculated using the below formula in wordnet

$$\text{path similarity} = \frac{1}{(\text{shortest\_path\_distance}(\text{synset1}, \text{synset2}) + 1)} \quad (4)$$

And also, when counting the similarity using the above-mentioned equation, the similarity (S1, S2) is not identical to similarity (S2, S1). Therefore, it is an asymmetric similarity. If we want to get the symmetric similarity equation can be updated as given below.

$$\text{similarity}(S1, S2) = \frac{1}{2} \times \left( \frac{\sum_{(w \in S1)} \text{maximumSimilarity}(w, S2)}{\text{No of words in S1}} + \frac{\sum_{(w \in S2)} \text{maximumSimilarity}(w, S1)}{\text{No of words in S2}} \right) \quad (5)$$

Now the above equation will return a symmetric similarity measure. But in this module asymmetric similarity is used because it has the maximum accuracy by compared to other two methods. Asymmetric similarity can have two measures of similarity which is obtained by similarity (S1, S2) and similarity (S2, S1). When S1 is equal to answer retrieved from the system and S2 is equal to candidate given answer, then the maximum accuracy was obtained by using similarity (S2, S1). Therefore, in this module the algorithm uses that similarity measurement to calculate the correctness of the answers given by candidates. Since the ontology contains a set of correct answers for a particular question, the overall correctness of the candidate given answer is calculated by taking the average as mentioned in the below equation. In the below equation n is the number of answers in the expected answer

list and  $A_i$  is the  $i^{\text{th}}$  answer in the expected answer list. S is the answer given by the candidate to that particular question and n is the number of retrieved answers.

$$\text{Correctness} = \frac{\sum_{i=0}^{n-1} \text{similarity}(A_i, S)}{n} \quad (6)$$

But the module needs to calculate one single percentage value as the technical skill level of the candidate. The technical score is calculated by taking the average of those marks by summing up all the M values and dividing by number of questions. The computed technical score is accepted as the technical skill level percentage of the candidate.

## 7. Result and discussion

The system uses an ontology to represent java questions and answers. Modeling java domain knowledge is not a simple and easy task. When modeling the java ontology, first step was to collect all knowledge concerning the java programming language and build a glossary of terms. The java oracle documentation was used as the main formal guide in this task. After extracting the answers from the ontology, the candidate given answer is evaluated against the system retrieved answers. For that task, semantic similarity measuring techniques are used. In this module wordnet approach is used to measure the similarity between answers.

An example technical question in the system and the expected answer (S1) is given in the below table.

**Table 1: Sample question and expected answer.**

Question	Expected Answer (S1)
“What is a class in java?”	“A class, in the context of Java, are templates that are used to create objects, and to define object data types and methods.”

For this question, different answers from different users were collected and all the user given answers can be evaluated against the same correct answer using wordnet and then the below similarity scores were obtained for each answer. The similarity scores can be measured in three ways; similarity (S1, S2), similarity (S2, S1) and symmetric Similarity (S1, S2). As mentioned earlier the similarity is not symmetric and therefore similarity (S1, S2) and similarity (S2, S1) is not having equal similarity scores. Similarity (S1, S2) is obtained by calculating maximum word similarities in S2 for each word in S1 and similarity (S2, S1) is obtained by calculating maximum word similarities from S1 for each word in S2. And the



symmetric Similarity (S1, S2) is obtained by summing up the two asymmetric similarities and dividing them by two.

The below Table 2 shows the obtained similarity measures for the user given answers for the above question. The similarity scores were calculated by evaluating each user given answer with the expected answer from the system which is given in Table 1. S1 is used to represent system expected answer and S2 is used to represent the user given answers by each user.

**Table 2 : Similarity scores for different answers.**

User Given Answer	similarity(S1, S2)	similarity(S2, S1)	symmetric Similarity(S1, S2)
'A class is a collection of objects'	48.9898989 899	87.5	68.2449494949
'Blueprints of object.'	23.2631257 631	57.1428571 429	40.202991453
'Class is a template'	46.8467643 468	100.0	73.4233821734
'class is a blueprint which defines the variables and the methods common to all particular objects'	58.8217338 217	77.0408163 265	67.9312750741
'A class means a category of objects that can generalize objects with similar behavior.'	35.1709401 709	63.0158730 159	49.0934065934
'Class can be identified as a mold or a template which helps to generate objects in a formatted manner'	55.8760683 761	58.5317460 317	57.2039072039

Then the answers were categorized in to 5 classes based on their similarity scores which are shown in the above Table 2. The criteria used to categorize classes is given in the below Table 3. If the score is between 0 and 20 then it is categorized to class C1, if the score is between 20 and 40 then it is categorized in to C2, if the score is between 40 and 60 then it is categorized as C4 and finally if the score is between 80 and 100 then it is categorized in to C5. Since each answer is having three similarity scores obtained from three different methods as given in Table 2, the categorizing process was applied to each similarity calculation method. That means at the end each user given answer was categorized in to 5 classes based on each of the similarity scores.

**Table 3 : Categorizing criteria.**

Marks out of 100 (M)	Category
$0 < M \leq 20.00$	C1
$20.00 < M \leq 40.00$	C2
$40.00 < M \leq 60.00$	C3
$60.00 < M \leq 80.00$	C4
$80.00 < M \leq 100.00$	C5

Then all the three similarity scores were evaluated using the same approach and calculated the accuracy. To evaluate this semantic similarity measures first the relevant data has been collected from final year undergraduate of University of Moratuwa. A question set consisting java related technical questions which are mostly asked at java interviews were given to the final year students of Faculty of Information Technology and collected their answers. The collected data shows how different candidates answer to the same question in different ways. Then, their questions and answers were mixed in a common set. The next step is to identify how these different answers will be evaluated by humans. To accomplish that task the question and answers were divided in to sets and then again given to research participants to categorize those answers in to five sections C1, C2, C3, C4 and C5 based on the correctness of their answers with compared to the system expected answers. That means the participants had to give a similarity score for a user given question, by comparing it with the expected answer of the system. Then based on that score the answers were categorized by using the same criteria as shown in Table 3. Then that human given category was considered as the true category for each answer. And then the confusion matrix related to those questions were generated and there are total three confusion matrices because three different similarity scores were used. But among the three different similarity

measuring methods similarity (S1, S2), similarity (S2, S1) and symmetric Similarity (S1, S2); the method which had the maximum accuracy was similarity (S1, S2). In here S1 is the system expected answer and S2 is the user given answer. The confusion matrix generated for similarity (S1, S2) is given below in the Table 4.

**Table 4: Confusion matrix for similarity (S1, S2).**

		Predicted				
		C1	C2	C3	C4	C5
Actual	C1	6	5	7	2	0
	C2	2	27	9	4	0
	C3	0	4	43	6	0
	C4	0	2	8	23	0
	C5	0	0	2	5	14

In the above confusion matrix, the diagonal values are the true positive values. True positive values are in which, the module has correctly predicted the category of the given answer. For example, true positive value for C1 is 6, and that simply means those 6 values actually belong to class C1 and the module also have successfully predicted it to be in C1.

Given below in the table are the other calculations done using the confusion matrix. In the below table FN stands for “False Negative” and that means the module has incorrectly predicted the value to be in a different class. For example, false negative of C4 is 10 and that means the module has predicted 10 values which actually belong to C4 in to other classes. FP stands for “False Positive” which indicates the values that are incorrectly predicted by the module to be in a class. According to the below table false positive of C2 is 11 and that means 11 values which actually do not belong to C2 have been incorrectly predicted by the system to be in C2. TN indicates “True Negative” and those are the values which are correctly predicted by the model to be in negative class. C5 has 148 true negatives and that means all those 148 values actually do not belong to C5 and the model has correctly predicted those values not to be in C5.

**Table 5 : Data obtained from confusion matrix.**

	C1	C2	C3	C4	C5
FN	14	15	10	10	7
FP	2	11	26	17	0
TN	147	116	90	119	148
TP	6	27	43	23	14

By using the above data accuracy was calculated with precession, recall and f-score for each class, and the obtained evaluation results are given in the following tables. C5 has the highest precision of 1 and that is because C5 actually has less data than the other classes and when

the prediction is C5 it always belongs to C5. C4 is having the least precision value 0.575 and that means when a prediction is done in to C5 the fraction of the correct prediction is equal to 0.575. Recall indicates the correct prediction of a class as a fraction of total number of actual data in the class. It can be obtained from the below table that even though the precision of C5 is 1, recall is 0.6666. That explains all the data points belong to C5 has not been actually predicted to C5 by the module. Next F-score is calculated by using both precision and recall. C5 has the highest f-score value and C1 has the lowest f-score values. That indicates the module is more suitable to predict values which actually belong to C5 than the values which actually belong to C1.

**Table 6 : Accuracy for different similarity types.**

	Similarity (S1,S2)	Similarity (S2,S1)	symmetric Similarity(S1,S2)
Accuracy	66.86%	41.4%	53.25%

**Table 7: Precision, recall and F-score for each class using similarity (S1, S2).**

	C1	C2	C3	C4	C5
Precision	0.75	0.71	0.62	0.58	1
Recall	0.30	0.64	0.81	0.69	0.66
F-Score	0.43	0.67	0.70	0.63	0.79

## 8. Conclusion

This paper describes about one of the four modules of a system that is used to recommend candidates based on their technical skills and personality traits for the position of java software engineer. A software engineer is someone who has to work cooperatively with his/her team. Therefore, it is very important to evaluate both their technical skills and personality traits. The paper describes about the answer evaluation module of the system.

The module takes two inputs; the question and answer and provides one output; technical skill level of the candidate. The questions are to test java technical skills and the system contains an ontology which includes java questions and answers. Therefore the correct answers are needed to be found by querying the ontology. In order to do that, SPARQL queries are used. SPARQL is used as a query language to retrieve data from ontologies and it contains RDF triples. In this case the triples pattern is generated using the questions. After generating triples, SPARQL query can be generated and the answers will be extracted from the ontology. Then the user given answer is evaluated against the system retrieved answers using wordnet similarity approach.

## 9. Acknowledgement

We would like to express my heartiest gratitude to the team members who supported me for this research and the Faculty of Information Technology for the guidance and strength given to us.

## 10. References

- [1] V. Vijay Anand, M. Shanthanlakshmi, G. Uppili Srinivasan, V. Arunkumar, G. Icewarya, S. Nandhu and S. Monisa Kamatchi, "A Study on Effectiveness of Recruitment Organizational Support in ITeS", *International Journal of Pure and Applied Mathematics*, Volume 119 No. 7 2018, 2755-2764
- [2] Djabatey, Edward Nartey, "Recruitment and Selection Practices of Organizations, A Case Study of HFC Bank (GH) LTD.", April 2012
- [3] N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," p. 25.
- [4] Omid Shahmirzadi, Adam Lugowski, and Kenneth Younge, "Text Similarity in Vector Space Models: A Comparative Study", September 2018.
- [5] Lakshmi Palaniappan, N. Sambasiva Rao, "An Ontology-based Question Answering Method with the use of Query Template", *International Journal of Computer Applications* (0975 – 8887), Volume 9– No.9, November 2010.
- [6] A. Kouneli, G. Solomou, C. Pierrakeas, and A. Kameas, "Modeling the Knowledge Domain of the Java Programming Language as an Ontology," in *Advances in Web-Based Learning - ICWL 2012*, vol. 7558, E. Popescu, Q. Li, R. Klamma, H. Leung, and M. Specht, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 152–159.
- [7] Maurizio Atzori, Giuseppe M. Mazzeo and Carlo Zaniolo, "A Natural Language Approach to Question Answering over RDF Data Cubes", *Semantic Web I* (2016) 1–5 1 IOS Press
- [8] Y. Hamoudi and T. Cornebize, "Extracting RDF triples using the Stanford Parser," p. 7.
- [9] Amal Zouaq, Roger Nkambou, and Claude Frasson.2007. Building domain ontologies from text for educational purposes. *In Proceedings of the Second European Conference on Technology Enhanced Learning: Creating New Learning Experiences on a Global Scale, EC-TEL'07*, pp. 393–407, Berlin, Heidelberg. Springer-Verlag.
- [10] Rowida Alfrjani, Taha Osman and Georgina Cosma, "A New Approach to Ontology-Based Semantic Modelling for Opinion Mining", 2016 *UKSim-AMSS 18th International Conference on Computer Modelling and Simulation*
- [11] Sapna Chauhan, Pridhi Arora, Pawan Bhadana, "Algorithm for Semantic Based Similarity Measure", *International Journal*

*of Engineering Science Invention* ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726 www.ijesi.org Volume 2 Issue 6 | June. 2013 | PP.75-78

## Implicit intention and activity recognition of a human using neural networks for a service robot eye

M.W.C.N. Moladande  
 Department of Mechanical Engineering  
 The Open University of Sri Lanka  
 Nugegoda, Sri Lanka  
 cmoladande@gmail.com

B.G.D.A. Madhusanka  
 Department of Mechanical Engineering  
 The Open University of Sri Lanka  
 Nugegoda, Sri Lanka  
 bgmad@ou.ac.lk

### Abstract

*Introduction of the assistive robot concept has created numerous ways to restore vital degrees of independence for the elderly and disabled people on their Activities of Daily Living (ADL). The most important aspect of an assistive robot is to understand the user's intentions with minimum number of interactions. Based on these facts, in this study we suggest a novel method to recognize the implicit intention of a human user, by using verbal communication, behavior recognition and motion recognition from the combination of machine learning, computer vision and voice recognition technologies. After recognizing the implicit intention of the user, the system will be able to identify the necessary objects from the domestic area that is going to help the human user and point them out to fulfil his/her intention. By far, this study is expected to simplify the human robot interaction (HRI) while consequently enhancing the adoption of assistive technologies and improving the user's independence in ADL. These findings will certainly help to guide future designs on implicit intention recognition and activity recognition to an accurate intention inference algorithm and intuitive HRI.*

**Keywords:** *Human-robot interactions, activities of daily livings, human friendly robotics, service robots, implicit intention*

### 1. Introduction

From the beginning of the 90's the robot caregiving technology has emerged in to the community to serve elderly and disabled people. These caregiving robots [1] perform a series of physical tasks for the well-being of a person with a disability. The task is embedded in the context of normal human ADLs and would otherwise have to be performed by an attendant. The emergence of assistive robots presents the possibility of restoring vital degrees of independence to the elderly and impaired in activities of daily living [2].

In the present day, aging is becoming a real issue on world's population. It's estimated that the number of older persons those aged 60 years or over is expected to more than double by 2050 and triple in 2100 [3]. This situation prompts the intelligent technologies such as assistive robots to play a crucial role in assisting the aging people with some daily activities. Robot assistance in ADL especially for the elderly and disabled people, should be executed in an intuitive and natural way. This requires a robot, which can proactively understand the user's intention and automatically provide the desired service [4-8]. Further, the robot should also be able to predict human intention and then choose the most reasonable actions to cooperate with humans.

There are two types of intentions, which can be identified in the human cognition. They are explicit-intentions and implicit-intentions. Explicit intentions are stated clearly and in detail, leaving no room for confusion or doubt. It is relatively easy to understand when someone communicates with explicit-intentions. Implicit-Intentions are suggested logically, according to the necessary consequences though not directly expressed. Implicit intentions are little harder to understand than explicit-intentions because the probability acts a major role in here.

Intention Recognition generally relies on voice recognition, biological motion recognition and activity or behavior recognition. These features could be extracted by using image processing, gesture detection, and motion calculation technologies and object affordances.

The biggest problem with the present robot caregiving technology is the lack of understanding of the user's intentions from minimal interactions with the user.

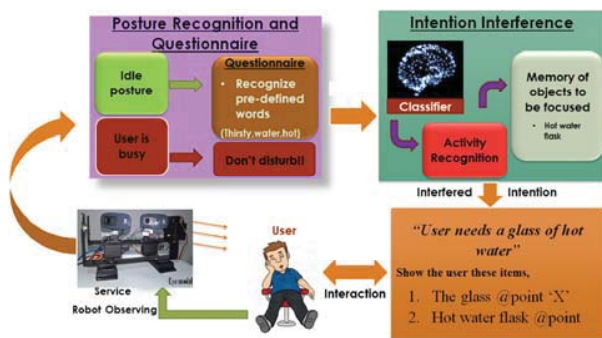
This research paper proposes an interactive robotic eye system to recognize the implicit intention and activity of an elderly or a disabled person. This robot eye unit was designed by mimicking the patterns on human eye. So, the way it functions is also identical to a human eye. Supervised learning techniques have used to identify the posture and human tracking. A questionnaire based on context features is used for the intention recognition.

Later on, necessary objects in the domestic are will be identified that will help to fulfil the user's intention by the system.

- The methodical approach is done by using six stages.
- Development of the robot eye with pan-tilt movement for a service robot.
  - Develop a human tracking algorithm to monitor the human user.
  - Develop the verbal implicit communication framework and interface to communicate with the human in the domestic environments.
  - Develop an intelligent system to recognize the implicit intention of the human.
  - Develop a system to recognize the predefined objects that fulfill user's intentions from the domestic environment.
  - Develop the human robot interaction and test to recognize the implicit intention and activity in domestic environment

## 2. System Overview

### 2.1. Overall system overview



**Figure 1. The implicit intention communication framework.**

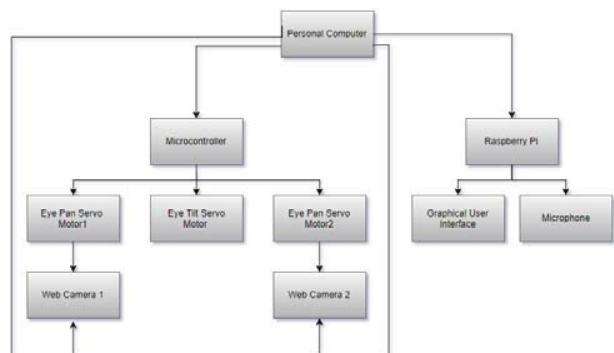
The complete overview of the implicit intention communication framework is illustrated on Figure 1. When the user is in the pre-defined domestic area, the robot eye system monitors the user during the entire period. While monitoring the user, the robot eye system also examines the user's body posture to find out a suitable moment to interact with the user. When the user is in a pre-defined relaxed posture, the system decides to interact with the user. Then the robot will start a conversation. Through the conversation, the system will ask a series of questions that help to recognize the implicit intention of the user. Then the robot eye system

will decide the user's implicit intention based on the user's response to the questionnaire.

### 2.2. Electrical system overview

Three direct current (DC) servo motors have been used to control all the degrees of freedom of the robot eye system. All these motors are controlled through the microcontroller, by using separate motor controllers.

The main computer and the microcontroller is connected with each other through a serial data link with the baud rate of 19200. This data link is a bi-directional link, which allows passing of data from main computer to the microcontroller to provide commands and position information calculated by image processing functions in some operational modes and for monitoring of different variables controlled via the microcontroller using the main computer interface. To control three servo motors, there is a requirement of generating 3 Pulse Width Modulation (PWM) signals. Therefore, the microcontroller should have a minimum of 3 PWM output channels. Since DC servo motors having potentiometer type encoders the microcontroller needs to have minimum of 3 analog to digital conversion channels [9]. The microcontroller should have high processing power also. Considering all the factors, we selected the Arduino as the microcontroller.



**Figure 2. Electrical system overview.**

Figure 2 shows the data acquisition from a human. The main computer controls two web camera movements via the microcontroller. It is necessary to develop a proper communication system to communicate with the human user. This part is not practically doable by using the main computer. Because the main computer is a bulky stable unit. Hence it is positioned somewhere inside the domestic area, it is unable to interact with the user when the user is away from it. For this purpose, another microprocessor is required to capture audio data from the microphone and then it has to encode, translate,

decode recorded data as a text. Later it has to transmit those translated data to the main computer through a wireless network for the intention recognition. At the same time, it should be capable to work as an external Graphical User Interface (GUI) too. To complete all these purposes, a Raspberry model 3B was selected. This microcontroller makes the interconnection between the main processor and the human user.

### 3. Implicit intention and activity of a human using Service Robot Eye

#### 3.1. Object tracking

Video tracking is the process of locating a moving object (or multiple objects) over time using a camera. In this project Kernelized Correlation Filters (KCF) based method is used [10] with face recognition for the human tracking purpose. This classifier is typically trained with translated and scaled sample patches. Such sets of samples are riddled with redundancies - any overlapping pixels are constrained to be the same. Based on this simple observation, an analytic model for datasets of thousands of translated patches was proposed. By showing that the resulting data matrix is circulant, we can diagonalize it with the Discrete Fourier Transform, reducing both storage and computation by several orders of magnitude. Interestingly, for linear regression this formulation is equivalent to a correlation filter, used by some of the fastest competitive trackers. There is an inbuilt KCF library, face recognition algorithm in OpenCv [11, 12] library so that we can directly use it by using its attributes. Both cameras are used to track the human to get a better accuracy.

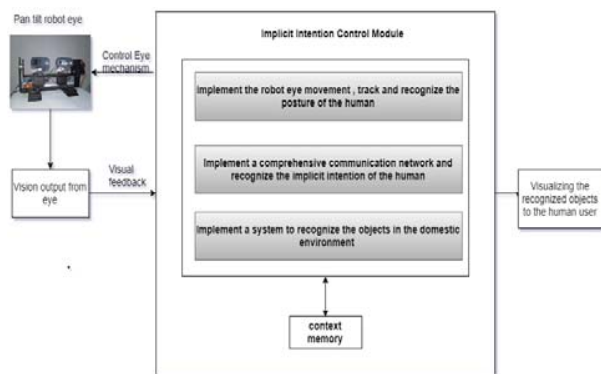


Figure 3. Implicit intention control module.

With the help of the face recognition and the object tracking libraries, we identified the human who is in the domestic area. After that the region of interest (ROI) was marked and the center of the ROI was calculated. The

cameras were coded to focus on to these center point of the ROI and keep it in line with the image center. When the ROI is changing, the cameras starts to rotate to keep a minimum distance between the ROI center and the image center.

#### 3.2. Posture recognition

While the system is tracking the human, it should also be able to identify the appropriate moment to interact with the human. If the human is busy with his work, he certainly would not like to get disturbed or distracted by the service robot. So that the system should be able to classify the difference between being busy and being relaxed. A pre-designed simple convolutional neural network [14, 15] was trained to classify these postures.



Figure 4. All the types of body postures that trained for relaxed mode.



Figure 5. All the types of body postures are trained for busy mode.

A TensorFlow based convolutional neural network is shown in Figure 6.  $128 \times 128 \times 3$  image is used as the input image. The objective of the training is to get the best possible values of the all these parameters which solve the problem reliably. TensorFlow™ [13] is an open source software library for high performance numerical computation. Its flexible architecture allows easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices.

In this training process the back-propagation method is used to calculate the correct weights in the neural network. Before training the neural network, a unique posture has pre-defined as the “Relaxed posture “As in the Figure 4. Five thousand images have been taken for

the positive data set and these photos have been taken while the human is in the pre-defined relaxed posture with different camera angles. Ambient lighting conditions were used while taking these training data. Another 5000 photos have taken for the negative dataset as shown in Figure 5. In the above data set, 70% of these photos have been used as the training data set and 20% of the data set has been used for the validation. Another 10% is selected for the testing dataset.

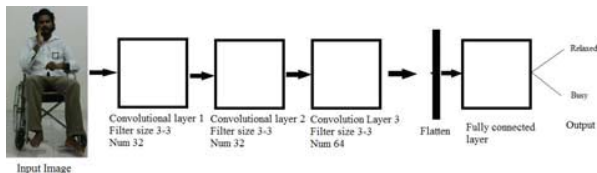


Figure 6. Tensorflow based convolutional neural network.

### 3.3. Verbal communication and implicit intention recognition

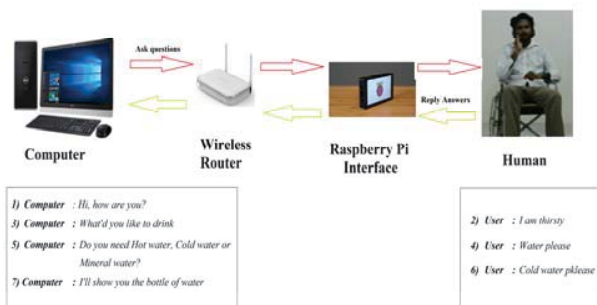


Figure 7. Verbal communication framework.

Figure 7 shows a series of questionnaire that have been developed to recognize the implicit thoughts of the human user. The questionnaire is presented through a GUI to the user and the vocal response of the user will be recognized by using the *google speech to text* API [16]. Then again, when the user has answered the questions, the system will recognize the intention of the human and move on to the next stage.

### 3.4. Object identification and visualizing

In the previous stage, the system has already recognized the implicit intention of the human. In this domestic environment, we have pre-defined objects that has a direct connection as in the Figure.8, with a specific intention. These objects are placed on constant and known angles with respect to the camera. For better understanding, a plan view of the domestic environment is shown in Figure 9 with the distances of the objects regarding to the robot eye system. When the relevant

object is triggered, the control system will rotate the cameras to the pre-defined angle and then the left camera will take a photo. This image will go through a TensorFlow [17] library based trained object detection network and then necessary object for the specific intention will be pointed out. Finally, the identified objects are sent to the GUI through wireless network.



Figure 8. Domestic environment.

Table 1. Predefined objects in the domestic environment.

Id	Details of the object		
	Object name	X - Angle( $^{\circ}$ )	Y - Angle( $^{\circ}$ )
A1	Ornament box	25	75
A2	Hot water flask	30	80
A3	Chair	50	90
A4	Picture	90	115
A5	Book rack	80	85
A6	Normal water flask	155	80

## 4. Result and discussion

### 4.1. Predefined domestic environment

The plan view of the pre-defined domestic environment is given in the Figure 8. As shown in the Table 1 each object has been placed in a static place while keeping a constant angle with the robot eye system. Figure 10 shows the setup of the entire robot eye system.

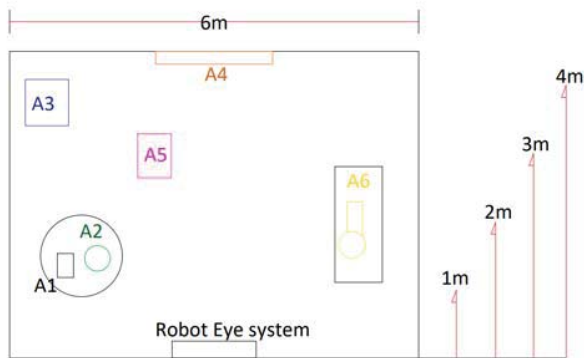


Figure 9. Domestic environment.

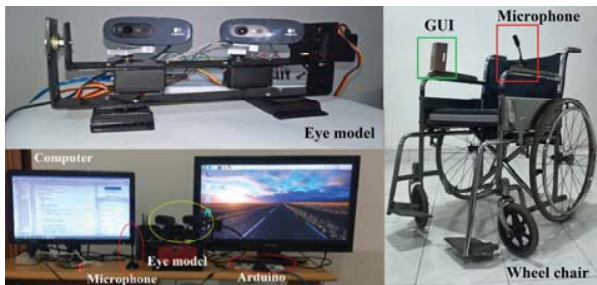


Figure 10. Complete setup.

#### 4.2. Rotation of the robot eye

Rotation speed of the robot eye system is an important factor that should have taken care of. With the higher motor speeds due to the vibration effects, human tracking is not very successful. So that it is necessary to reduce the motor speed as one degree per one step for better results. Figure 11 describes the pan and tilt motor angle variations while the user is moving in the pre-defined domestic area.

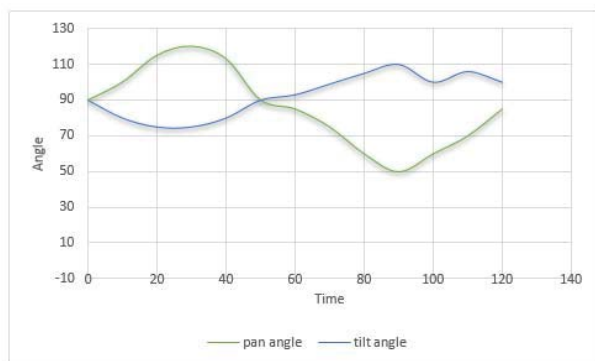


Figure 11. Camera angles vs time.

#### 4.3. Implicit intention and activity recognition

When a human is in a relaxed posture, generally he or she doesn't change their postures. So, when the person is not moving, the ROI is on the center point of the image and it is also not moving. Data has collected about the posture recognition accuracies. After the posture recognition, verbal communication process data has to be collected for every test scenario. Finally, the accuracies of the object recognition data is saved and all these different data sets are compared in Table 2 for further clarification.

Table 2. Implicit intention recognition system accuracy.

Implicit Intention	Known User			
	Relaxed Posture Recognition Accuracy		Voice Recognition Accuracy (%)	Object Recognition Accuracy (%)
	Distance to the Robot Eye (m)	Accuracy (%)		
Finding the ornament box	2	88	93	75
	3	76		
	4	65		
Need to drink hot water	2	90	95	85
	3	73		
	4	60		
Read a book	2	85	92	88
	3	70		
	4	69		
Need to drink normal water	2	80	89	78
	3	70		
	4	63		

#### 5. Conclusion

In here, a methodical approach has been followed to the design and developed an interactive robotic eye for recognizing the implicit intention and the activity of a human.

In the area of service robots, during the human-robot interaction, the human user may not use crisp commands to express their intentions with the robot. In such scenarios the robot should be able to identify the user's implicit intentions.

The process includes the mechanical and electrical designs. In the design process close attention has been paid to the human bio-mechanics to realize a design that reaches anthropomorphism to a closer degree. The proposed method is capable to mimic key visual functions of the human brain promises to robot eye



maneuver quickly and safely through adaptive vision field through the domestic environments.

For this project, different types of algorithms have been used in different stages. The KCF tracker shows superior performances when it comes to object tracking. A trained neural network has been used to recognize the postures and the classification made by neural networks, which are more accurate than the other cascades. Now days everything move towards the 'Internet of Things'. So, using the Google's speech recognition network will certainly change the approach of robot using the internet. Also, the proposed robotic eye is designed in such a way that it can be used as a platform for facilitating further developments in integrating more interactive features to robotic eye. The robot could be significantly improved with redesign.

After trials with robot eye for service robot to determine whether it is feasible to use the robot for its intended purpose, the overall design, materials, and manufacturing techniques should be evaluated and optimized for the quantity of robots that are expected to be made. The outward appearance of the robot eye needs to be improved upon. The eye is too big and somewhat unequal, and its appearance may keep it from being accepted by more human-like. The eye robot could become more interactive if it had more sensors in it. Sensors of some sort would allow it to sense physical interaction. Adding of a proper outer cover can be used to enhance the appearance of the interactive features. It is better to design the outer cover using molding techniques with low weight materials. And also, there are constraints in the head design due to movements of the links and actuators which should be considered before modeling the outer cover. The robotic eye can be used as a research platform for implementing various control strategies such as artificial intelligent systems and machine learning techniques. These can be implemented in further developments.

## 6. References

- [1] Ceccarelli, M., *Service robots and robotics*, 1st ed. Hershey, PA: Engineering Science Reference, 2012.
- [2] En.wikipedia.org, Service robot. [Online] Available at: [https://en.wikipedia.org/wiki/Service\\_robot](https://en.wikipedia.org/wiki/Service_robot) [Accessed 1 May 2017].
- [3] *World Population Prospects, 2017 Revision Key Findings and Advance Tables*, 1st Ed. New York: United Nations, Department of Economic and Social Affairs, Population Division, 2017.
- [4] Abowd, G., Dey, A., Brown, P., Davies, N., Smith, M. and Steggle, P., "Towards a Better Understanding of Context and Context-Awareness", *Handheld and Ubiquitous Computing*, 1999, pp. 304-307.
- [5] Madhusanka, B.G.D.A. and Jayasekara, A.G.B.P., "Design and development of adaptive vision attentive robot eye for service robot in domestic environment", in *Proceedings of the IEEE International Conference on Information and Automation for Sustainability*, 2016, pp. 1-6.
- [6] Jang, Y., Lee, S., Mallipeddi, R., Kwak, H. and Lee, M., "Recognition of Human's Implicit Intention Based on an Eyeball Movement Pattern Analysis", *Neural Information Processing*, 2011, pp. 138-145.
- [7] Li, S. and Zhang, X., Implicit Intention Communication in Human-Robot Interaction through Visual Behavior Studies. *IEEE Transactions on Human-Machine Systems*, 2017, pp. 1-12.
- [8] Liu, R., Zhang, X. and Li, S., "Use context to understand user's implicit intentions in Activities of Daily Living", in *Proceedings of the IEEE International Conference on Mechatronics and Automation*, 2014.
- [9] De Mel, W. and Siriwardana, P., *Sensors and Actuators*, 2nd Ed. Nugegoda: The Open University of Sri Lanka, 2011.
- [10] Henriques, J., Caseiro, R., Martins, P. and Batista, J., "High-Speed Tracking with Kernelized Correlation Filters", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3), 2015, pp. 583-596.
- [11] Docs.opencv.org, OpenCV: Face Detection using Haar Cascades. [Online] Available at: [https://docs.opencv.org/3.3.0/d7/d8b/tutorial\\_py\\_face\\_detection.html](https://docs.opencv.org/3.3.0/d7/d8b/tutorial_py_face_detection.html) [Accessed 25 Mar. 2017].
- [12] Docs.opencv.org, OpenCV: cv::TrackerKCF Class Reference. [Online] Available at: [https://docs.opencv.org/3.1.0/d2/dff/classcv\\_1\\_1TrackerKCF.html](https://docs.opencv.org/3.1.0/d2/dff/classcv_1_1TrackerKCF.html) [Accessed 25 Apr. 2017].
- [13] TensorFlow, TensorFlow. [Online] Available at: <https://www.tensorflow.org/> [Accessed 10 Apr. 2017].
- [14] YouTube, Deep Learning SIMPLIFIED: The Series Intro - Ep. 1. [Online] Available at: <https://www.youtube.com/watch?v=b99UVkWzYTQ> [Accessed 16 Apr. 2017].
- [15] Cv-tricks.com, TensorFlow Tutorial 2: image classifier using convolutional neural network - CV-Tricks.com. [online] Available at: <http://cv-tricks.com/tensorflow-tutorial/training-convolutional-neural-network-for-image-classification/> [Accessed 12 Apr. 2017].
- [16] Google Cloud Platform, Why Google's Cloud Infrastructure is Right for Your Business | Google Cloud Platform. [Online] Available at: <https://cloud.google.com/why-google/> [Accessed 12 May 2017].
- [17] En.wikipedia.org, TensorFlow. [Online] Available at: <https://en.wikipedia.org/wiki/TensorFlow> [Accessed 14 May 2017].

## Face and upper-body emotion recognition using service robot's eyes in a domestic environment

T. M. W. Vithanawasam  
Department of Mechanical Engineering  
The Open University of Sri Lanka  
Nugegoda, Sri Lanka  
tharind7a@gmail.com

B. G. D. A. Madhusanka  
Department of Mechanical Engineering  
The Open University of Sri Lanka  
Nugegoda, Sri Lanka  
bgmad@ou.ac.lk

### Abstract

*The population of the elderly/disabled people of the world is increasing rapidly. Taking care of these people has become a major issue since lack of professional caregivers or family members. Hence, the only feasible solution for this is using humanoid service robots. Available care-giving service robots lack the proper human emotion recognition. Hence, they cannot communicate with people as humans do. In addition, it is not preferable to people when robots are not androids. Therefore, this paper has proposed a method to recognize a face and upper-body emotions by using service robot's eyes. The service robot's eyes model is able to track a particular person in a domestic environment to mimic human eyes' behavior while providing visual feed for the system to recognize emotions. Face emotional expressions and upper-body gestures are recognized by using supervised learning methods. Finally, the results show that the trained system recognizes the emotions effectively in the domestic environment for a particular person.*

**Keywords:** Face Recognition, upper-body recognition, emotion recognition, humanoid service robots, robot eyes.

### 1. Introduction

The population of the world is aging rapidly. The population aged 60 or above, is growing at a rate of about 3 percent per year. In 2017, there were 962 million people aged 60 or over in the world, comprising 13 percent of the global population [1]. This aging population is needed to be taken care of. So that there has been vast job growth for the caregivers. But along with that, there is a shortage of caregivers. To solve this, many countries are desperately searching for a solution and the only feasible solution to it is use service robots to replace human caregivers.

Service robots are the robots that assist humans in many ways to perform a job. Caregiving robots are the type of robots which come under service robots to replace the human caregivers/nurses. But these service robots are at their infancy stage and need to be developed further. One of the most important features that a service robot should possess is the capability to recognize human emotions, in order to make human-robot interaction be more like human-human interaction. Then robots will be able to serve more likely a human caregiver. Usually, people like to be taken care by a humanoid robot rather than by any other shapes.

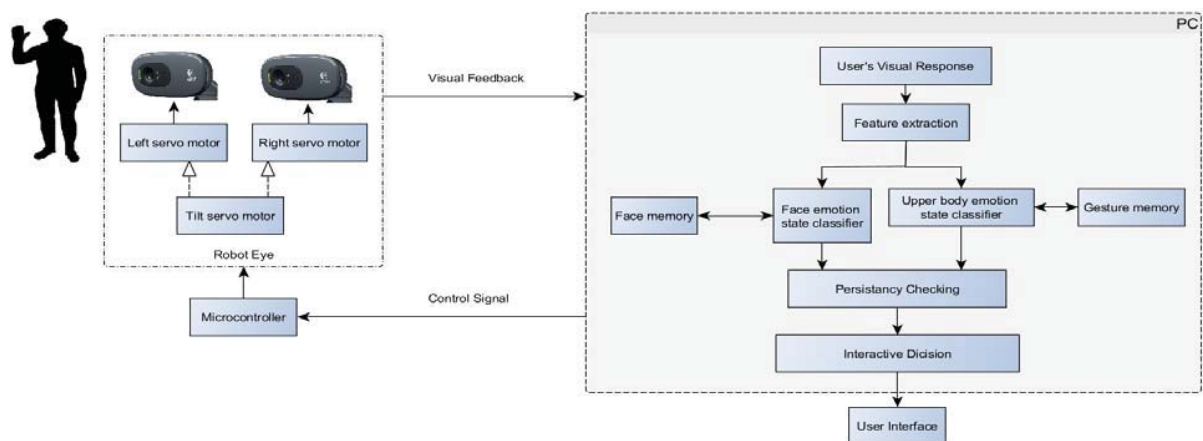


Figure 1. System overview.

There are service robots that can be used in caregiving situations with or without the capability of human emotion recognition. Such as robot Pepper. It is a human-shaped robot. It is the first humanoid robot capable of recognizing the principal human emotions joy, sadness, anger or surprise and adapting his behavior to the mood of its interlocutor. Based on the human voice, the expression on the face, body movements and the words used, Pepper will interpret emotion and offer appropriate content. It also responds personally to the mood of the moment, expressing himself through the color of his eyes, his tablet or his tone of voice [2]. Care-O-bot 4 is another robot supports humans in their daily life. It can be used for a variety of household tasks, for example, to deliver food and drinks, to assist with cooking or for cleaning. Its head cameras enable human-robot interaction based on a graphical user interface and gesture recognition, it also offers multi-modal use input such as touchscreen, microphones, and speaker [3]. A robotic bear named Robear was designed by scientists from Sumitomo Riko Company in Japan. The robot is capable of lifting patients out of bed and into wheelchairs or aid patients who need assistance walking [4].

According to Mehrabian (1982), 55% of communication is body language [5]. Also, an elderly or differently abled person usually communicate with their caregivers using the facial and upper-body expression to express their emotions, since they usually sit on a chair. Therefore, researching facial emotional expressions and upper-body emotional expression recognition was considered. Many types of vision sensors can be used for that. Available caregiving service robots do not have human-like eyes which focus to the caretaker, face and upper-body gesture recognition altogether. Therefore, this paper proposes a solution to that.

Interactive robot's eyes model was built [6]. It consists of two cameras to capture the emotions simultaneously. These cameras have pan-tilt motions

which enable them to move like human eyes, hence enabling them to track a person in the domestic environment. It also allows the system to recognize emotions while the person expresses their emotions naturally.

Supervised learning techniques are used to recognize the face and upper-body emotional expressions. Three emotional expressions are considered anger, boredom and fear. The two eyes are separately giving visual feeds to the intelligent system to recognize the face and upper-body emotions. Two classifiers are used to recognize the face and upper-body emotions respectively. Outputs of each classifier are compared and then gives the recognized emotion on Graphical User Interface (GUI). Since caretaking service robots are intended to serve a specific person, the system is developed to be trained and recognized emotions of one person.

## 2. System overview

The robot's eyes model's two cameras give the visual feedback for both face and upper-body emotion recognition. The cameras should provide images with sufficient resolution to recognize emotions. For that, total field of view and face recognition ability have considered. The shortest Inter-pupillary distance of the person's face and the minimum number of pixels between the eyes of captured image have been used for that. Hence a camera with 640×480 dynamic resolution was suitable for face detection in the selected domestic environment. So, for better recognition, the person should be at 1m away from the robot's eyes model.

The system initially searches for the person in the domestic environment by pan-tilt motion from the top right corner to bottom left corner. For that, three DC servo motors are used. Two for panning each camera individually and one for tilting the both simultaneously. Therefore, eyes of the model track the person while

recognizing the emotions, which enable the person to move freely in the given domestic environment. The model has a dynamic field of view of  $120^{\circ}$  vertically and horizontally. The robot's eyes model is shown in Figure 2.

When the person is identified, the emotional expression for a given video frame is recognized. Then if the same emotion is recognized in the next frame, the recognized emotion is displayed on GUI as expressed on the system overview Figure 1.

The domestic environment should be an indoor environment with an appropriate illumination of 300-500 Lux. Also, there should not be large objects equal to the size of the hand with the similar colour to the skin of the person inside the upper-body Region of Interest (ROI) for the proper emotion recognition.

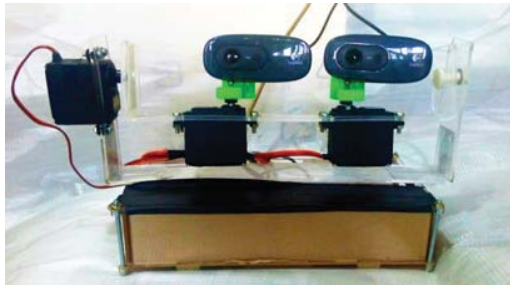


Figure 2. Robot's eyes model.

### 3. Recognition of modalities

#### 3.1. Face detection

Haar-cascade classifier is used for the face detection because of the robustness and light weightness [7], which enable to implement it in real time situations. For the face detection pre-trained face detector which comes with OpenCV [8] is used. It detects the face by developing an integral image for the detected face using Haar wavelet concept. For the detection, the person should be facing toward the robot's eyes model and it is able to detect the face with small rotations. In order to increase the accuracy, scaleFactor is set to 1.1, minNeighbors set to 25 and minSize is set to (10, 10).

#### 3.2. Upper-body detection

In order to recognize upper-body emotions, first of all, it must extract the upper-body from the visual feed by mean of an upper-body ROI. The human body has proportionality; which can be expressed using the head's height and head's width.

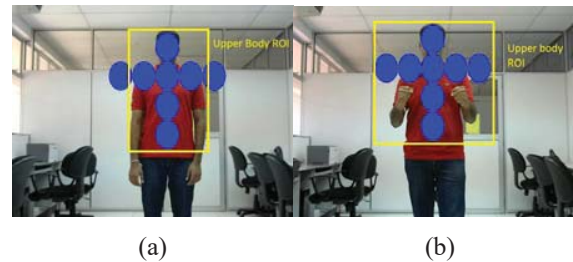


Figure 3. (a) Upper-body ROI when relax. (b) Upper-body ROI when expressing an emotion.

So human upper-body can be encompassed in a rectangle of five heads height and three heads width as shown in Figure 3 (a). However, this is suitable only for relaxed posture. As shown in Figure 3 (b), when expressing emotion using upper-body gesture, the upper-body ROI gets wider so that it is defined as five heads width and five heads height of the face ROI. By this method, the upper-body region can be quickly identified and eliminating noise generated by the background objects when emotion recognizing.

## 4. Recognition of emotions

### 4.1. Face emotion recognition

When it comes to compute-vision, the computers cannot understand the emotions as human do. It only understands the shapes and colours of an image. Since facial emotion is also can be identified by changes in the facial muscles, a computer should be able to differentiate the emotions with the help of pattern recognition algorithms.

Fisherfaces is a face recognition algorithm based on the linear discriminant analysis (LDA). Since it is needed to recognize different emotional expressions, Fisherfaces can be used because recognition of the different facial emotional expressions is a classification process. It separates the Eigenspace into classes, such that within-class variance is low while between-class variance is high. The advantage of using the Fisherfaces is, once the Eigenspace has been defined (or learned), the recognition can run in real-time. Even though Fisherface is complex and requires a lot of processing time, it gives a better classification result [10].

To recognize face emotion by Fisherface algorithm a data set is needed. The dataset should be included the images of the person with expressing the emotions such as anger, bored and fear.

Larger the data set more accurate it is, and longer the training time. Therefore, the dataset has been created with 2100 images, 700 images for each emotion. Initially, 15,000 images were taken while a person was expressing emotions naturally. Then these images were processed by

cropping out the face and converted them to grayscale since the color of the face do not have relevance to the emotions and grayscale images process faster than RGB images. Cropped out images were in various sizes due to person's distance to the camera was not fixed. Therefore, resized those to 250×250 pixels and in order to reduce the illumination variation effect, histogram equalization was done. Then, Fisherface emotions recognition classifier was trained. This process was done automatically. Therefore, in order to get the emotional expression from the person three videos of 15 minutes, which trigger each emotion, were shown.

Face emotional expression recognition classifier is able to recognize the three emotions, which are anger, fear and boredom. For the recognition, the person does not have to be in relax or neutral position before expressing the emotions. It can right away recognize the trained emotions. Before sending visual feed for the classifier, faces are cropped out and resized as similar to the data set images because otherwise, features would not be able to identify.

#### 4.2. Upper-body emotion recognition

Upper-body gestures are usually accompanied by various arms' and head's motions. When it comes to elderly people, upper-body emotions are mainly conveyed by arms, along with shoulder and head movements. Out of these, head and arm movements can be identified easily by skin detection. So, by identifying arms locations with respect to face in the body ROI, the upper-body emotions can be recognized. In this research, a person wearing either long sleeves or short sleeves T-shirt or shirt is considered. To recognize the arms locations, the color threshold is used with Hue Saturation Value (HSV) color space as shown in Figure 4. HSV thresholding values depend on the domestic environment and the person's skin tone. When compared to other color spaces HSV gives better color recognition in variance intensity level and it perceives colors as humans.

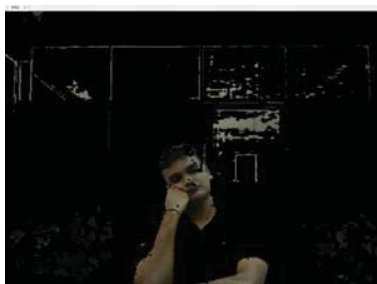


Figure 4. Skin detection.

After the color threshold, morphological operations are done to remove noises and optimize the blobs. Erode and dilate operations is done with a kernel with MORPH\_ELLIPSE with kernel size 3×3 in equation (1) with different iterations.

$$k = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (1)$$

Then, Gaussian blur is done to smooth the image by removing the Gaussian noise with kernel size with 5×5. Then, binary thresholding is done with Otsu's binarization to convert it to black and white image. It automatically calculates a threshold value from image histogram hence, it gives better thresholding.

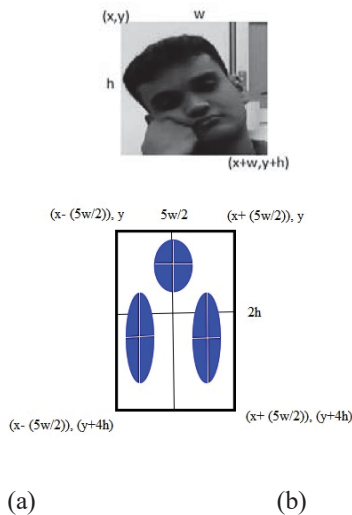
After that, black and white image obtains which represents the skin areas with black as shown in Fig. 5. Those areas are identified with drawing contours around them by calculating moment for the contours. Then centroid and area of each contour are obtained.



Figure 5. Skin extraction.

For the above recognition following, formulas are used to calculate center lines of the upper-body ROI, as shown in Figure 6. Blue ellipse represents the head and two arms from top to bottom and center lines of them are shown in white.

“x” and “y” gives the upper left-hand corner pixel value of the face ROI and “w” and “h” gives the width and the height of the ROI respectively as shown in Figure 6 (a). Then, Upper-body ROI rectangle's Vertices are defined as shown in Figure 6 (b).



**Figure 6. (a) Face ROI Vertices. (b) Upper-body ROI Vertices**

When the person is expressing the emotions, they are observed, and the arm locations for each emotion are found. When the person is in anger, usually hands are folded in front of the chest or arms are located around the chest area. Three or two contours are drawn for these blobs. Therefore, when the contour is drawn is two and their centroids are on both side of the horizontal center line “2h” of the upper-body ROI, then it expresses anger. When there are three contours and two contours are in the lower side of the horizontal center line “2h” and if these two are also in both side of the vertical center line “5w/2”, then it also expresses anger.

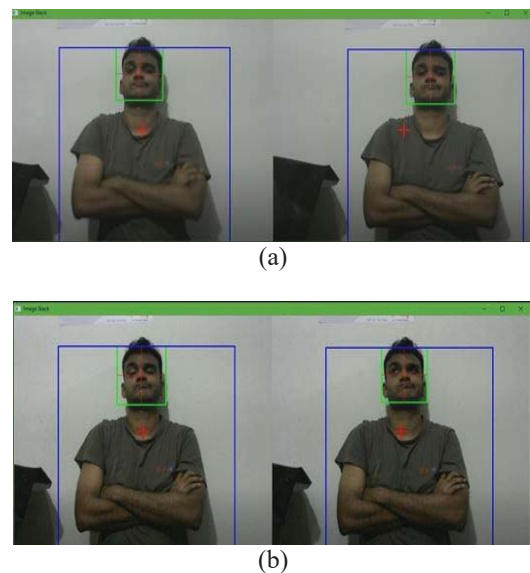
When the person is in boredom, usually arms are folded in front of the chest, supporting head with one arm while the other arm was folded in front of the chest or supporting the head with both arms. Therefore, contour drawn is one for boredom when both arms skin was visible. The first instance is similar to the anger and the same method is used to identify. The second instance is identified by removing face ROI to remove the face and neck area from the image. If the number of blobs is not two, then the emotion is boredom. But it is recognized as the bored-fear emotion because when supporting the head with either both arms or cover the mouth area by both arms are pretty much similar in geometry. First one expressed boredom while latter one expressed the fear emotion. However, if the person wears a long sleeve cloth, the number of blobs identified was one and deleting the Face ROI with neck area, reduces it to zero. If the person supporting head with one arm while the other arm is folded in front of the chest when wears a long sleeve cloth the number of blobs is two. If these two blobs are on both side of the “2h” line and one blob, other than face ROI was in on side of the “5w/2” line then it expresses bored emotion.

### 4.3. Emotions fusing

For the upper-body, individual and combined emotional classes have been defined as anger, anger\_bored, fear, fear\_bored, bored. For Face emotion, individual emotion classes have been defined as fear, anger and boredom. Emotion recognition conducts taking face modality as the primary modality while upper-body was secondary modality. So, recognized emotion is only displayed on the GUI when recognized emotional expression for both modalities and both eyes are the same.

## 5. Results and discussion

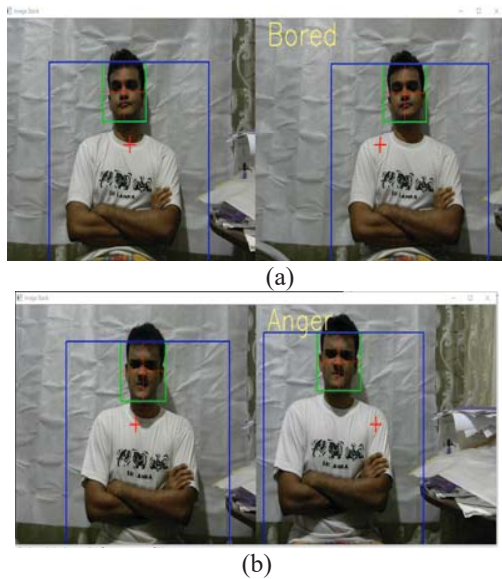
The interactive service robot’s eyes were tested for selected emotional expression in the domestic environments. It was able to recognize the anger, fear and bored emotions by both modalities while sitting, standing or waking facing towards it. The system was able to successfully track the person when expressing emotion as shown in Figure 7. As shown Figure 7 (a) right eye was centered to the neck area while left was centering. Figure 7 (b) shows the when both eyes are centered. This enabled person to move freely without giving much attention to the robot’s eyes model. Also, it gave the impression of robot’s eyes were more human-like.



**Figure 7. (a) Two eyes are not centered. (b) Two eyes are centered on the neck area.**

#### 4.1. Emotion recognition

Both upper-body and face emotional expression recognition were sensitive to the lighting condition of the domestic environment. Because the angle of the light came from the light source to the person had generated an unwanted silhouette which made emotion recognition algorithms to falsely identify features.



**Figure 8. Emotion recognition results of the system. (a) Bored emotion recognition. (b) Anger emotion recognition.**

Results of the emotion recognition are shown in Figure 8. As shown on images (a), and (b) the domestic environment was set messy in order to show that it does not affect the system's output as long as light conditions are good and not containing skin-colored objects.

The same set up which is used to create the dataset was used for calculating the accuracy of the system. Three videos each with 15-minutes of run-time were shown to the person to trigger each emotion. The system captured 22, 500 frames as instances for each emotion.

The result obtained are shown in Table 1 as a confusion matrix. The green colour diagonal cells represent the correctly recognized instances of each emotion while white colour cells represent the false recognition of each emotion out of the 22, 500 instances. The system correctly recognized 18, 225 anger instances which correspond to 81%. The system correctly recognized 14, 249 fear instances correspond to 63.33 %. The system correctly recognized 17,550 fear instances which correspond to 78 %. Out of the 22, 550 anger instances 2, 919 recognized as fear and 1, 356 recognized as bored. Out of the 22, 550 fear instances 2, 849

recognized as anger and 5, 402 recognized as bored. Out of the 22, 550 bored instances 1, 276 recognized as anger and 3, 674 recognized as fear. Out of 22, 351 anger instances 18, 225 were correctly predicted, which corresponds to 81.54 % and 2, 849 were predicted as fear and 1, 276 were predicted as bored. Out of the 20, 842 fear instances 14, 249 were correctly predicted, which corresponds to 68.37 % and 2, 919 were predicted as anger and 3, 674 were predicted as bored. Out of the 24, 307 bored instances 17, 550 were correctly predicted, which corresponds to 72.20 % and 1, 356 were predicted as anger and 5, 402 were predicted as fear.

**Table 1. The result of emotion recognition.**

		Actual class			Classification overall	Producer accuracy
		Anger	Fear	Bored		
Predicted class	Anger	18225	2849	1276	22350	81.54%
	Fear	2919	14249	3674	20842	68.37%
	Bored	1356	5402	17550	24308	72.20%
Truth overall		22500	22500	22500	67500	
User accuracy (Recall)		81%	63.33%	78%		

The overall accuracy of the system can be increased further by adding images to the dataset with different hairstyles, with more illumination variations so that only the features related to each emotion are extracted while others are discarded as noises. The error rate cases occurred mostly when the person moves faster, which causes motion blur and cover the face with a hand or hands. Such action cannot be controlled since the person was not advised to control his actions.

The system detects the face of the person whenever he turns his face towards the robot's eyes and then starts to recognize emotions. The system can be used to recognize those emotions of another person after retraining the system for that person. At the current stage, it can only be trained and used for one person at a time. By using the two cameras for emotion recognition, the system's accuracy can be increased, since it ensured the recognized emotion was the same for both sources. When one eye is blocked, not detecting a face, the system recognized the emotion with one eye. Therefore, it acts as a backup system, similar to human's eyes.

## 6. References

- [1] Ageing, Un.org, [Online]. Available: <http://www.un.org/en/sections/issues-depth/ageing/>. [Accessed: 12 Jan 2018].
- [2] Who is Pepper?, SoftBank Robotics, [Online]. Available: <https://www.ald.softbankrobotics.com/en/robots/pepper>. [Accessed: 2 Jan 2018].
- [3] Care-O-bot 4, Fraunhofer Institute for Manufacturing Engineering and Automation, [Online]. Available: <https://www.care-o-bot.de/en/care-o-bot-4.html>. [Accessed: 2 Jan 2018].
- [4] The strong robot with the gentle touch | RIKEN, Riken.jp, [Online]. Available: [http://www.riken.jp/en/pr/press/2015/20150223\\_2/](http://www.riken.jp/en/pr/press/2015/20150223_2/). [Accessed: 2 Jan 2018].
- [5] Mehrabian, A. Communication without words, *Psychology Today*, Volume 2, 1968, pp. 52-55.
- [6] Madhusanka, A. and Jayasekara, P., "Design and development of adaptive vision attentive robot eye for service robot in domestic environment", in *Proceedings of the IEEE International Conference on Information and Automation for Sustainability*, 2016, pp. 1-6.
- [7] Viola, P. and Jones, M., "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR 2001, 2001, pp. 1-511-I-518 vol. 1.
- [8] Cascade Classification - OpenCV 2.4.13.5 documentation, Docs.opencv.org, [Online]. Available: [https://docs.opencv.org/2.4/modules/objdetect/doc/cascade\\_classification.html?highlight=cascadeclassifier#cascadeclassifier](https://docs.opencv.org/2.4/modules/objdetect/doc/cascade_classification.html?highlight=cascadeclassifier#cascadeclassifier). [Accessed: 3 Jan 2018].
- [9] Face Recognition with OpenCV - OpenCV 2.4.13.5 documentation, Docs.opencv.org, [Online]. Available: [https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec\\_tutorial.html](https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html). [Accessed: 3 Jan 2018].



## MRI based Glioma segmentation using Deep Learning algorithms

H. N. T. K. Kaldera

*Department of Electrical  
and Electronic Engineering,  
Faculty of Engineering,  
University of Peradeniya,  
Sri Lanka  
nikalal.k@eng.pdn.ac.lk*

S. R. Gunasekara

*Department of Electrical  
and Electronic Engineering,  
Faculty of Engineering,  
University of Peradeniya,  
Sri Lanka  
shanagak@eng.pdn.ac.lk*

M. B. Dissanayake

*Department of Electrical  
and Electronic Engineering,  
Faculty of Engineering,  
University of Peradeniya,  
Sri Lanka  
maheshid@ee.pdn.ac.lk*

### Abstract

*Primary brain tumors can be malignant (cancerous) or benign (non-cancerous). Out of primary brain tumors, gliomas are the most common and, high grade gliomas carry a poor prognosis. In our paper, we present a technique to segment the glioma cells in Magnetic Resonance Imaging (MRI) using faster Region based Convolutional Neural Network (R-CNN) and edge detection techniques in image processing algorithms. This study identifies the region of interest that is glioma cells, with higher confidence level and localize the tumor on the MRI with the tumor mask. Further, analysis shows that with the proposed technique it is possible to achieve an average detection accuracy, sensitivity, Dice score and confidence level of 99.81%, 87.72%, 91.14% and 93.6% respectively.*

**Keywords:** *Magnetic Resonance Imaging, Region based Convolutional Neural Network, Glioma segmentation, deep Learning.*

### 1. Introduction

Brain tumors are abnormal growths found in the brain. There are more than 120 types of brain and central nervous system tumors. These tumors may arise from brain tissues or deposits from tumors elsewhere. The tumors which originates within the brain is called primary brain tumors, whereas tumors grown within the brain as a result of spreading of a malignant tumor elsewhere in the body are called as secondary metastatic brain tumors [1].

Among these kinds of tumors, the most common primary brain tumors are gliomas [1]. Gliomas are brain tumors starting in the glial cells. There are 3 types of gliomas. Astrocytoma, Oligodendroglioma, and Ependymoma. Glioma can be either low grade which exhibits a slow growth rate or high grade with a fast growth rate. The necessary and sufficient treatments are planned by medical practitioner, considering the tumor grades. Astrocytoma are the most common type of glioma

in children and adults. They develop from astrocytes. The astrocytes are supportive cells of neurons. As a treatment plan for the low grade glioma, tumor monitoring is practiced. Regular MRI scans are done to monitor the tumor and it is called, watchful waiting. After two to three years of monitoring, low grade tumors will be eventually removed through surgery.

Hence, the detection of the existence of a brain tumor, the approximate location of the brain tumor, the size and grade of the tumor is critical and its accuracy depends on the observation and the decision of the medical practitioner who analyze the MRI scan. Contrast-enhanced MRI is the gold standard imaging modality and provides highly sensitive anatomical information about the tumors. If the technology can be upgraded to localize the tumor within a MRI based on convolutional network, it may be more precious and accurate and it will aid the decisions making process of the observer.

According to the literature, devising an optimal model for tumor segmentation is much harder using discriminative and generative methods [2, 3]. The drawback of the selective search is, it groups each and every individual pixel. Therefore it needs lots of computational costs as well as it uses a huge time for training and testing. A review of brain tumor segmentation using the deep neural network is presented in [4]. It shows different CNN methods for the tumor segmentation with significant accuracy for BRATS dataset. In the system proposed in this paper, we were able achieve better performance in terms of Dice score, sensitivity, accuracy and confidence level compared to the architecture presented in [4].

One of the main insight of our study is, we use faster R-CNN, here it generates region proposals and from that, it proceeds to localize the tumor. The studies[5, 6] present a powerful object localization technique. The presented solution is fully autonomous in segmenting, and training and modeling time is relatively low than existing methods.

The rest of the paper is organized as follows: The section 2 gives a brief introduction to the faster R-CNN. In section 3, we describe the methodology adapted in

dataset preparation, proposed segmentation model architecture, and tumor boundary localization. In section 4 we present and discuss about the results obtained from the experiment and finally draws conclusions in Section 5.

## 1.2 Introduction to faster R-CNN

Region proposal network (RPN) and R-CNN are types of networks which are used generally in faster R-CNN. In the recent past, region proposal [7] and region based convolutional neural networks [5] were frequently used in object detection. Before these methods, exhaustive search was used in object detection [8]. The R-CNN in general acts as a classifier and its accuracy lies on the performance of region proposal algorithm [6].

The RPN is developed by adding additional convolutional layers that outputs objectness score at various locations of the image. Further, it also outputs region bounds of these regions of interest (ROI). Basically RPN receives a feature map and outputs object proposals of rectangular dimension with an objectness score. The object proposals are created by sliding small kernel over feature maps.

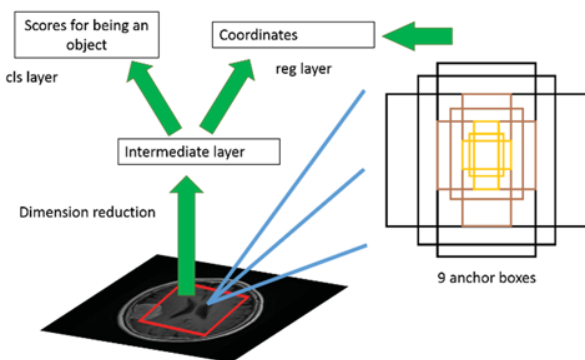


Figure 1. A RPN.

These sliding windows are mapped to a lower dimensional feature and feed to fully connected layers, box regression layer and classifier. The fully connected layer is shared with all spatial locations [6]. Figure 1 shows a region proposal network.

The intention of faster R-CNN is to locate ROI bounding boxes in the image. For this, anchors are being used and they are defined based on the convolutional feature maps. These anchors have rectangular shapes with different sizes and ratios. Figure 2 exhibits sample set of 9 anchors. For each anchor box the network should outputs corresponding coordinates as  $x_{min}$ ,  $x_{max}$ ,  $y_{min}$  and  $y_{max}$ .

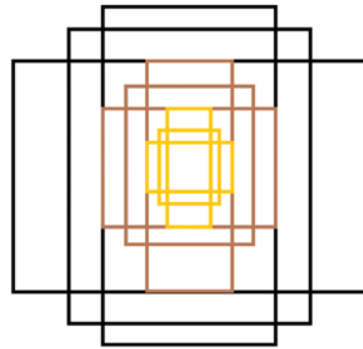


Figure 2. Nine anchors with different sizes and ratios.

Further if there are several objects in the image, it may outputs several anchor boxes with corresponding locations.

There are several moving parts in the network, therefore this architecture is complex. An image is represented as 3 dimensional array with height, width and depth. This array is passed through a pre trained CNN and obtains feature maps. This feature map is used as the feature extractor for the next stage. Since our dataset is a small dataset, we used faster R-CNN in our proposed method. This technique is very popular in terms of transfer learning because it can train a classifier on a smaller dataset. The main difficulty with the deep learning for object detection is making bounding boxes with variable lengths. This problem is solved by using RPN, with anchors. For every anchor, the RPN checks whether the anchor contains the relevant object and it adjusts the anchor dimensions to better fit the relevant object. Then the RPN applies Region of Interest (ROI) pooling and extracts the features which correspond to the relevant object.

Next the R-CNN module uses this information to classify the content in the bounding box. R-CNN module classifies the proposals into one of the existing classes, and remove bad proposals as background class. This module also adjusts the coordinates of the corresponding bounding box to fit in the ROI.

The loss function employed at RPN training, assigns a binary class label, to every anchor box to indicate whether it contains object or not. The loss functions employed at RPN contains classification loss as well as regression loss and it is defined as,

$$L(\{p_i\}, \{t_i\}) = \frac{\lambda}{n_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) + \frac{1}{n_{cls}} \sum_i L_{cls}(p_i, p_i^*), \quad (1)$$

where  $L$  is the loss function and  $L_{reg}$ ,  $L_{cls}$  are regression loss function and classification log loss over classes respectively. The  $i$  correspond to the index of an anchor and  $p_i$  is the prediction probability of anchor  $i$  being an object and  $p_i^*$  is 1 or 0 if anchor is positive or negative respectively. The vector with 4 normalized coordinates of the predicted rectangle is represented by  $t_i$  and  $t_i^*$  is that of ground truth box with a positive anchor and  $\lambda$  is a weighted parameter. The  $n_{cls}$  and  $n_{reg}$  refers to batch size and number of anchor locations respectively[6].

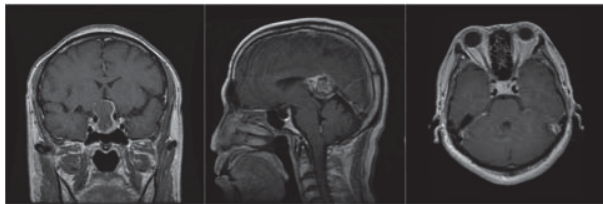


Figure 3. Sample images of dataset.

## 2. Methodology

### 2.1 Dataset preparation

In our research, we have used only MRI images with 2D slices. The image data, are generated at Guangzhou, China, and General Hospital, Tianjing Medical University, China, from 2005 to 2010 [9, 10]. This brain tumor dataset contains 3064 T1-weighted contrast-enhanced images from 233 patients with three types of brain tumors; namely meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices). All the MRIs are grayscale images. These images dataset are organized as MATLAB “.mat” files and each file was saved as struct type. This struct data type consists with tumor label (either meningioma, glioma, or pituitary tumor), patient ID, image data, tumor border coordinates, 5-fold cross validation indices, and tumor mask. The T1-enhanced MR images is one of the popular imaging modality acquired for treatment planning[11]. Since T1-enhanced images have a contrast agent (gadolinium), they can improve the contrast of the affected tumor regions of the MRIs. Hence, we analyze only T1 images in this research.

Further, this dataset consists of images of sagittal, axial and coronal planes of the brain. Figure 3 shows sample images of the dataset. For this experiment, we used only the axial plane dataset, because they have higher resolution and less noise compared to coronal and sagittal plane slices. We assume that, if tumor is identified at initial stage using axial images, doctors could proceed further by using sagittal and coronal planes. Therefore, for the glioma segmentation process, we selected 123 axial MRI brain tumors with the help of a neurologist. These selected

axial MRIs again separated for the training and testing purposes.

### 2.2 Architecture of the proposed model

The faster R-CNN algorithm is a speedup solution to integrate region proposals into the CNN model. Generally, this creates a single, unified model composed of RPN and

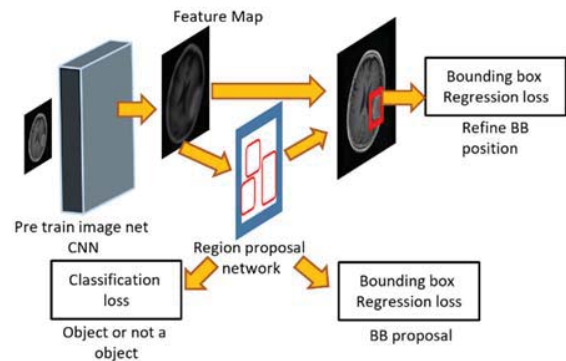


Figure 4. Faster R-CNN model.

R-CNN with convolutional feature layers. Figure 4 shows an illustration of a faster R-CNN model and every bounding box in it represents the probability (softmax) of encircling an object within the rectangle. If a particular anchor has an object score above a certain limit, it is moved forward as a region proposal. Next, the system executes a one full CNN over the complete image. This in turns extend the neural network for the prediction by softmax layer. The specialty of this approach is, rather than training many different models to classify each object class, it gives the class probability. Since model only aims to localize the tumor area bounding box rather than classifying, proposed model—performs much better in terms of speed in training as well as in testing. As the final output, this proposed method gives the bounding box and the confidence level indicating the probability of the region box being a glioma.

### 2.3 Tumor boundary location

The proposed model in section 3.2 outputs, the coordinates of the tumor bounding box and the confidence level of that decision. These coordinates are reference to low-resolution images because when we feed the MRIs to the proposed network we initially scale down the images from  $512 \times 512$  to  $128 \times 128$ . After taking the corresponding coordinates of the bounding box of the low resolution images, it was mapped to original resolution images.

In order to extract the glioma from the input MRI, it is necessary to generate a tumor mask with the estimated tumor boundary with a high precision. Although R-CNN detects the tumor region, its output is a rectangular box

enclosing the tumor with the confidence level of the selected ROI including a glioma. Yet, we were unable to extract the tumor boundary or the tumor mask, using R-CNN. To detect the tumor boundary, several edge detection techniques such as Prewitt and Sobel were tested. The detected edge contour lines were improved as in [12], to obtain the approximated tumor boundary. Finally, the binary mask generated using the improved edge detection process, is applied to the input image to extract the estimated tumor regions.

### 3. Results

The analyzing, training and testing were performed on a computer with Intel(R) 8 Series Chip set processor, 4.00GB RAM and 64-bit operating system, ×64-based processor. The analyzing process is continued with the dataset we selected as in section 3.1. That dataset consists of MRI images of glioma with a resolution of 512×512. Each image was converted to 16-bit integer type and 100 images were used to train the proposed method while 23 images were kept to test the trained model. In order to reduce the computational cost and training time, testing time, all MRIs were downscaled to 128×128. We observe that, downscaling to this level don't remove the important features of the images. The downscaled images were fed to the trained network.

The results obtained in this proposed system is tabulated in Table 1. The 1<sup>st</sup> row of the Table 1 shows the original T1 enhanced axial images from the dataset. These images were selected randomly from the validation dataset and there resolution is 512×512. When we input dataset to the proposed model it was down sampled to 128×128. Therefore the output images of the model is also in low resolution (128×128). The output for selected images from the validation dataset is shown under 2<sup>nd</sup> row of the Table 1. Each bounding box shows the confidence level of being a glioma tumor.

Then we mapped our low resolution images into 512×512 in order to improve the clarity of detection. After that edge detection techniques were applied to the bounding box of MRIs. The closed contour of the edge detected image is created using neighborhood search based simple contouring techniques. The extracted tumor regions are created using these contours and it is shown under the 3<sup>rd</sup> row. The predicted tumor regions are shown under white colour of binary image. The finalized comparison between original glioma and predicted is shown under “extracted tumor boarder vs original tumor boarder against original image”. In images, red colour contour is for the boundary of original tumor, and green colour is for the boundary of predicted tumor mask. This visual interpretation depicts the high performance of the proposed method in detecting glioma tumors.

The developed system sensitivity and accuracy are shown in final two rows of the Table 1. These results are obtained by comparing segmented images with reference images. These two parameters are evaluated by adopting pixel wise criteria. The predicted segmentation in true positive (TP) which are correctly detected as foreground when compare with actual tumor. The false positive (FP) is defined as when the pixels of the output is falsely detected as foreground and when pixels are correctly segmented as background they are taken as (TN). Finally when the segmentation process falsely recognized as background it is labeled as false negative (FN) [12].

Sensitivity gives an idea about the performance in the identification process and it is the ratio of actual positives to total [13]. Hence, the sensitivity is defined as,

$$sensitivity = \frac{TP}{(TP + FN)} \quad (2)$$

The accuracy highlights the correctly segmented area with reference to original tumor. It is defined as,

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (3)$$

The traditional or balanced Dice score is defined as the harmonic mean of precision and recall, and it is defined as [4],

$$Dice\ Score = \frac{2TP}{(2TP + FP + FN)} \quad (4)$$

The proposed model is perform very well under every axial T1 enhanced MRIs. When it is the 1<sup>st</sup> column images there are artifact due to the bone structure of the eye sockets. Due these artifacts most of the previous studies performances are not good or we have to insert huge dataset with MRIs of similar behavior. But our proposed model doesn't identify these artifacts as malignant. The column 1 result proves this situation solidly.

Further, the average confidence level is 93.6% for all the tested images. Then average accuracy, Dice score and sensitivity are 99.81%, 91.14% and 87.72% respectively. Hence, there is a high degree of confidence about the system design which is proposed in this research.

### 4. Conclusion

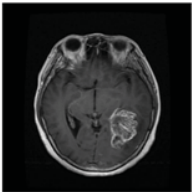
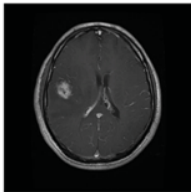
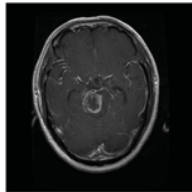
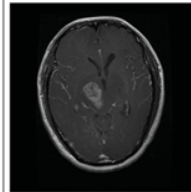
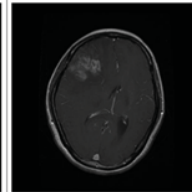
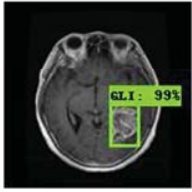
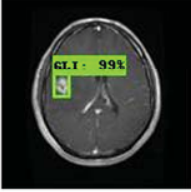
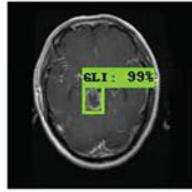
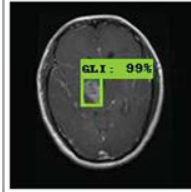
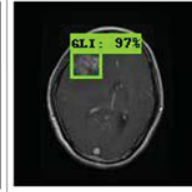

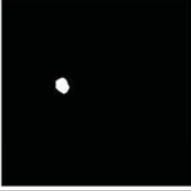
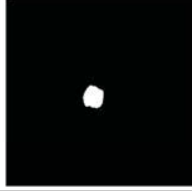
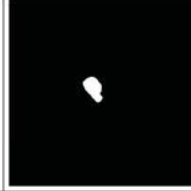
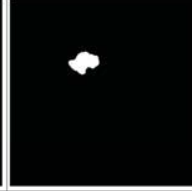
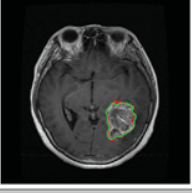
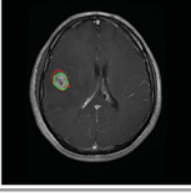
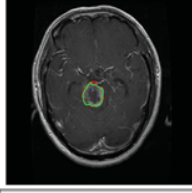
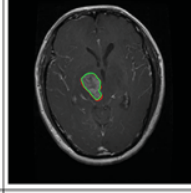
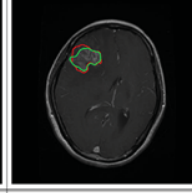
Automatic brain tumor localization and segmentation from MR image is a relatively difficult task. The paper presents a CNN based glioma tumor localization and segmentation process of T1 weighted MR images. This segmentation process is validated as accurate by means of ground truth labeling presented by a Neurologist. The

main insight of our presented method is it is fully autonomous in extraction of tumor region, as well as training and testing time and computation cost is significantly low.

## 5. References

- [1] KabirAnaraki, A., Ayati, M., and Kazemi, F., "Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," *Biocybernetics and Biomedical Engineering*, Vol. 39, pp. 63-74.
- [2] Angelini, E., Clatz, O., Mandonnet, E., Konukoglu, L. Capelle, and Duffau, H., "Glioma Dynamics and Computational Models: A Review of Segmentation, Registration, and In Silico Growth Algorithms and their Clinical Applications," vol. 3, no. 4, pp. 262–276, Nov. 2007.
- [3] Liu, J., Li, M., Wang, J., Wu, F., Liu, T. and Pan, Y., "A survey of MRI-based brain tumor segmentation methods," *Tsinghua Sci. Technol.*, vol. 19, no. 6, pp. 578–595, Dec. 2014.
- [4] Işın, A., Direkoğlu, C., and Şah, M., "Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods," *Procedia Comput. Sci.*, vol. 102, pp. 317–324, 2016.
- [5] Girshick, R., Donahue, J., Darrell, T., and Malik, T., "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, pp. 580-587, 2014.
- [6] Ren, S., He, K., Girshick R. and Sun, J., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.
- [7] Sande, K.E.A., Uijlings, J.R.R., Gevers, T., and Smeulders, A.W.M., "Segmentation as selective search for object recognition," in *Proceedings of 2011 International Conference on Computer Vision*, Barcelona, Spain, 2011, pp. 1879–1886.
- [8] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005, vol. 1, pp. 886–893.
- [9] J. Cheng, "brain tumor dataset." 02-Apr-2017. Available online: [https://figshare.com/articles/brain\\_tumor\\_dataset/1512427](https://figshare.com/articles/brain_tumor_dataset/1512427) (accessed on 4 November 2018).
- [10] Cheng J, Huang W, Cao S, Yang R, Yang W, Yun Z, et al. (2015) Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. *PLoS ONE* 10(10): e0140381. <https://doi.org/10.1371/journal.pone.0140381>
- [11] Liu, Y., "A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery," *PLOS ONE*, vol. 12, no. 10, p. e0185844, Oct. 2017.
- [12] Malkanthi, C., and Dissanayake, M. "Brain Tumor Boundary Segmentation of MR Imaging using Spatial Domain Image Processing", *International Journal for Innovation Education and Research*, 5(10), pp. 1-9, (2017). Retrieved from <http://ijer.net/index.php/ijer/article/view/621>
- [13] Mohane, S., and Borse, M., "Comparative study of brain tumor detection using morphological operators," *Int. J. Res. Eng. Technol.*, vol. 04, no. 03, pp. 422–428, Mar. 2015.

Table 1

T1 Images 512x512					
Bounding box of tumor with confident levels 128x128					
Extracted tumor region					
Extracted tumor boarder vs original tumor boarder against original image					
Sensitivity	0.949	0.792	0.9129	0.911	0.821
Accuracy	0.9976	0.9985	0.999	0.9991	0.9961
Dice score	0.948	0.8548	0.95	0.95	0.854

## Data mining approach for identifying suitable sport for beginners

P. T. Amarasena, B. T. G. S. Kumara,  
*Department of Computing & Information System*  
*Faculty of Applied Sciences*  
*Sabaragamuwa University of Sri Lanka*  
*Sri Lanka.*  
*pasindutharushika@gmail.com*

S Jointion  
*Department of Sport Sciences & Physical Education*  
*Faculty of Applied Sciences*  
*Sabaragamuwa University of Sri Lanka*  
*Sri Lanka.*

### Abstract

*Anthropometric measurements are generally used to determine and predict achievement in different sports. An athlete's anthropometric and physical characteristics may perform important preconditions for successful participation in any given sport. Further, anthropometric profiles indicate whether the player would be suitable for the competition at the highest level in a specific sport. Recently, more researches have been carried out on Sport Data mining. In this study, we propose an approach to identify the most suitable sport for beginners using data mining and anthropometric profiles. We propose clustering base approach. We apply a spatial clustering technique called the Spherical Associated Keyword Space which is projected clustering result from a three-dimensional sphere to a two dimensional (2D) spherical surface for 2D visualization. Empirical study of our approach has proved the effectiveness of clustering results.*

**Keywords:** *Anthropometric measurements, Data mining, Clustering, Euclidean distance*

### 1. Introduction

Anthropometric measurements are generally used to determine and predict achievement in different sports. Anthropometric measurements and morphological characteristics play significant roles in determining the success of a sports person. An athlete's anthropometric and physical characteristics may perform important preconditions for successful participation in any given sport. Indeed, it can be assumed that an athlete's anthropometric characteristics, can, in some way power his/her level of performance, at the same time helping to determine a suitable physique for a certain sport. It has been well established that specific physical characteristics or anthropometric profiles indicate whether the player would be suitable for the competition at the highest level in a specific sport [1].

The advent of data mining confined, sports organizations almost exclusively on human proficiency

[2]. It was trusted that domain professionals (coaches, managers and scouts) could effectively possess their collected data into valuable knowledge. As various types of data collected increase in scope, more practical methods are found to extract helpful knowledge using data mining techniques. Properly leveraged Sports Data Mining techniques can result in better team performance by matching players to certain situations, identifying individual player contribution, evaluating the tendencies of opposition, and exploiting any weakness.

The study was conducted for boys aged from 18 to 28 years who represented the 41st national sports festival competition. In this research study, we propose an approach to determine the suitable sport for the beginners. First, the different features from the players are collected by us. Then, the similarity between players is calculated. Finally, Spherical Associated Keyword Space (SASKS) algorithms [3] are applied to cluster the players. The algorithm aids to search players by visualization of the cluster data on a spherical surface and it is effective for noisy data. Finally, the beginners are assigned for suitable sport according their Anthropometric measurements using cluster.

The remainder of this paper is organized as follows. In section II we discuss related works. Section III discusses the proposed approach. In the next section we describe similarity calculation method, and the spatial clustering algorithm. Section IV discusses the experiment and evaluation. Finally section V concludes the study.

### 2. Related works

From a vast duration of time most researches have been carried out on Sport Data mining like basketball, baseball, cricket and etc. All those researches were done mostly in the view of either of coaches or of owners of the clubs. One reason for that is, some sports can't be predicted at the first stage of sportsman's skills, and therefore several tests need to be done before predicting the choice of sport accurately. So the researchers mostly stuck because they cannot predict anthropometric variables properly [4].

Research work [5] predicted a pattern for the basketball. In this analysis researcher has mostly used data from internet. Research work [6] has done Reference Values for Body Composition and Anthropometric Measurements in players. This research was done by using statistical approach. A research done by Valentina Cavedon, Carlo Zancanaro and Chiara Milanese [4] analyses Anthropometry, Body Composition, and Performance in Sport-Specific Field Test of Female Wheelchair Basketball Players. This research highlighted that physical and performance characteristics of female wheelchair basketball players are scarce. In most countries, female players train and contend with male players on mixed groups because of the restricted total population of players, which might otherwise cause large territorial spread for each team.

Marcel POMOHACI and Ioan Sabin SOPA have done research emphasizing the importance of anthropometry measurements in analyzing the impact of sports activities on students. The anthropometric tests included in the research were the measurement of height, weight, different length of superior and inferior segments and chest perimeter. [8]

Though there are many researches on sport data mining which can predict patterns of specific sports by utilizing players' performance data, there are a few dedicated patterns. There are less existing researches carried out to predict the mathematical module. Different types of machine learning approaches were used and the effectiveness of k-means clustering algorithm was analyzed for the pattern recognition in some studies.

### 3. Proposed clustering approach

The study was conducted for representatives of the 41st national sports festival competition 2015 in Sri Lanka. Hundred players, participated from nine provinces for national sports competition in Tennis, kabadi and koko were used as data set for this study. The Pushup, Sit-up, Cardiac respect endurance, flexibility, agility, Height, Weight, Upper Arm Relaxed Girth, Fore Arm Girth, Chest Girth, Wrist Girth, Waist Girth, Thigh Girth, Calf Girth, Angle Girth, Acromiale Radial Length, Radials lion Dactyl ion, Foot Length, body composition, fore arm length, hand length, Explosive power, Breath hold in time, resting heart rate, volume of oxygen, Force vital capacity, force explotaty volume in 1second, Upper arm radius and Leg Length, were selected as features and tested by standardized procedure. Figure 1 shows the steps of the proposed approach.

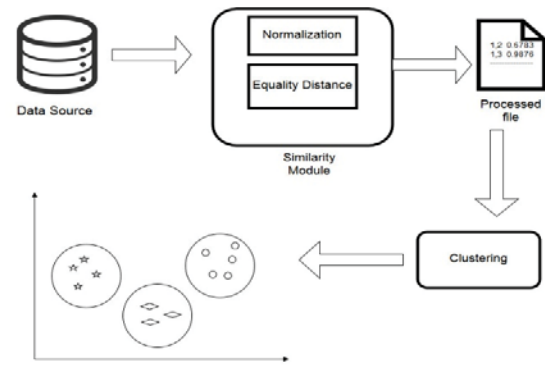


Figure 1. Steps of the proposed approach.

#### 3.1. Normalization of data set

Normalization is a scaling technique, a mapping technique or a preprocessing stage where, we can discover a new range from current range. It can be helpful for giving prediction or forecasting purpose. As we learn there are so many ways to predict or forecast but all can vary with each other a lot. So, to continue the large variation of prediction and forecasting the Normalization technique is required to make them close [7].

The techniques which supply linear transformation on original range of data are called Min-Mix Normalization. The technique which keeps relationship among original data is called Min-Mix Normalization. Here, we normalized values in all the features using following (1).

$$a_i = \frac{v_i - \min v_i}{\max v_i - \min v_i} \quad (1)$$

Here,  $a_i$  is normalizing value between 1 and 0 and  $v_i$  represents i-th attributes value of player.  $\min v_i$  is minimum value of the i-th attributes.  $\max v_i$  is maximum value of the i-th attributes.

#### 3.2. Euclidean distance based similarity calculation

Here, Euclidean distance is used based on similarity calculation method to calculate the similarity between players. The most famous distance used for numerical data is probably the Euclidean distance. This is a special case of the Murkowski distance. Euclidean distance performs well when deployed to datasets that include compact or isolated clusters. As Euclidean distance is very common in clustering following (2) was used to calculate the similarity.

$$Sim(S_1, S_2) = 1 - ED(P_i - P_j) \quad (2)$$



Where,

$$ED(P_i - P_j) = \sqrt{\sum_{r=1}^n [A_r(P_i) - A_r(P_j)]^2} \quad (3)$$

Here,  $P_i$  represents the  $i$ -th player.  $ED(P_i - P_j)$  is Euclidean Distance between two Players and  $A_r(P_i)$  represents the  $r$ -th attribute value of player  $P_i$ . Further  $n$  is the number of considered attributes of a player.

### 3.3. Clustering algorithm

In this paper, it was mentioned that we apply SASKS [3] to cluster the services. The SASKS algorithm is an extended version of ASKS algorithm. For the spherical visualization of the clustering space, it was needed to create a 3D space using affinity values for pairs of players. This is same as the nonmetric multidimensional scaling (MDS) problem. The ASKS algorithm is an extended MDS algorithm that is able to represent services in 3D space. In addition, ASKS is able to achieve high precision by noise isolation. In addition, ASKS algorithm is able to represent players in 3D space and we can get a visual output. However, this 3D form is difficult to visualize on a 2D screen. Therefore, the SASKS technique was applied as our clustering approach. The SASKS technique plots players onto a 2D spherical surface for easy visualization on a 2D screen. It is needed to give an affinity matrix as the input for the SASKS algorithm. The affinity matrix is created using the calculated similarity values of players using above (3).

We can represent the distance  $D_{ij}$  between two players using following (4):

$$D_{ij} = f(x_j^{(k)} - x_i^{(k)}), \quad (4)$$

where  $x_i$  and  $x_j$  are the locations of the players  $i$  and  $j$ , respectively in the space. The dimension of the space in which the players are located is represented by  $k$ .

Further, function  $f$  can be defined using (3):

$$f(\theta) = \begin{cases} |\theta|^2, & |\theta| < a \\ 2a|\theta| - a^2, & |\theta| \geq a \end{cases} \quad (5)$$

Here,  $\theta = (x_j^{(k)} - x_i^{(k)})$  and parameter  $a$  is the density-control parameter.

The criterion function of ASKS is given by;

$$\varphi(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}) = \sum_{k=1}^p \sum_{ij} \{M_{ij} f(x_j^{(k)} - x_i^{(k)})\} \longrightarrow \max$$

Here,  $M_{ij}$  is the affinity value between players  $i$  and  $j$ .

The partial derivative of  $\varphi$  with respect to  $x_i$  provides the formula for determining the values of  $x_i$  that maximize  $\varphi$  ;

$$\sum_{j=1}^n M_{ij} f'(x_j^{(k)} - x_i^{(k)}) \equiv 0$$

Finally after several steps, following iterative computation (6) is used to calculate the new player locations on the sphere.

$$x_i^{(k)}(t+1) = \frac{\sum_{j=1}^n M_{ij} \{D(x_j^{(k)}(t) - x_i^{(k)}(t)) (x_j^{(k)}(t))\}}{\sum_{j=1}^n M_{ij} D(x_j^{(k)}(t) - x_i^{(k)}(t))} \quad (6)$$

Here,  $x_i$ ;  $ij = 1, 2, \dots, n$ ,  $k = 1, 2, 3, \dots, p$  and  $t = 1, 2, \dots$ . Further,  $M_{ij}$  is the affinity value between players  $i$  and  $j$ .

SASKS algorithm cluster players, based on sports and players of the same sport, are arranged into the same region.

## 4. Experiments and evaluation

The experiments were conducted on a computer running Microsoft Windows 10, with an Intel core i7-4702MQ, a 2.20 GHz CPU and 4 GB RAM. Java was used as the programming language to implement the program for similarity calculations, and the SASKS algorithm was implemented using MATLAB.

We selected 101 players from three different sports. Sports were Tennis, Kabaddi and Koko. According to manual classification players from 0 to 50 were the most suitable for Tennis. Players from 51 to 75 were the most suitable for Kabaddi and players from 76 to 100 were most suitable for Koko. Next, vector with 30 elements was prepared for each player. Here we used 30 different features as we discussed in section 3. Table 1, shows the sample data set that we have considered. Here, we only display two players due to space limitation. Then, eq. (2) and (3) were used to calculate the player's similarity with the generated vectors. Table 2 shows sample similarity values between players.

According to the results of the table, similarity value between player0 and player1 is 1. So, the players are the most suitable for same game. According to their anthropometric measurements, they are the most suitable for Tennis. Thus, they should be in same cluster. When

We analyzed the clustering results, we show that they are in same cluster. According to the results of the table, similarity value between player56 and player77 is 0.29026. This is a low similarity value. Thus, they should be in different clusters. According to their anthropometric measurements, player56 is the most suitable for Kabaddi and the most suitable game for player77 is Koko. After calculating the pair similarity of players, we generated the affinity matrix using those similarity values. Finally, the matrix was used to generate the service clusters.

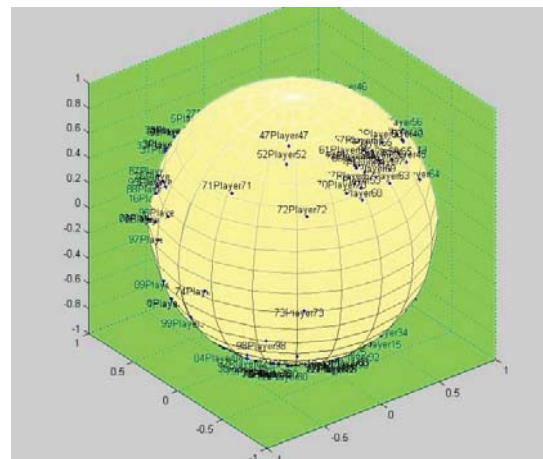
**Table 1. Sample data set.**

Variable	Player1	Player2
Push-up( in1 minute)	30	31
Sit-up	29	28
Cardiac respect endurance	65	68
Flexibility	28	30
Agility	26	29
Height(m)	1.63	1.61
Weight(kg)	53	50
HE	51	51
Upper arm radius(cm)	27	27
Upper arm length(cm)	31	33
Fore Arm girth(cm)	26	26
Chest Girth(cm)	80	79
Wrist Girth(cm)	17	16
Waist Girth(cm)	70	70
Thigh Girth(cm)	49	49
Calf Girth(cm)	35	35
Ankle Girth(cm)	19	18
Body composition	48	50
BEH	7	7
BEF	11	11
UA	27	27
Fore arm length(cm)	29	30
Hand length(cm)	19	16
Foot length(cm)	28	26
Explosive power	10.5	11.5
Heart Rate	80	79
Breath Holding Time	29	31
Force vital capacity	3.9	4.1
Force volume in 1second	3.5	3.6
Volume of Oxygen	49	48

**Table 2. Sample similarity values between players.**

Player	Player	Similarity
Player0	Player1	1.00000
Player24	Player77	0.35949
Player31	Player38	0.82677
Player54	Player60	0.89265
Player56	Player77	0.29026
Player85	Player94	1.00000
Player97	Player99	0.81355

Figure 2 shows result of spatial clustering approach. On the spherical surface the players are distributed according to their similarity. When analyzing the spherical surface, we observed three main regions where players are placed and it was evident that similar players in same sport are placed into one region. We observed clear separation of regions and density variation of players within the region and these regions can be considered as clusters.



**Figure 2. Clustering surface.**

Figure 3 shows a part of a cluster. We compared the cluster with the manual categorization and anthropometric measurements. We observed that most of the players in that cluster are suitable for Kabaddi, thus, we can consider this cluster as the kabaddi cluster. However there are some invalid members in this cluster. For example, Player 7, Player 40 and Player 33 are incorrectly placed in Kabaddi cluster (highlighted players in the cluster in fig. 3). According to their anthropometric measurements, they are the most suitable for Tennis and not for Kabaddi.

Figure 5 shows a part of Koko cluster. Here, we can observe players such as Player87, Player96 and Player 93 as correctly placed. However, in this cluster also we can observe some invalid members. Player32 is incorrectly placed in the cluster.

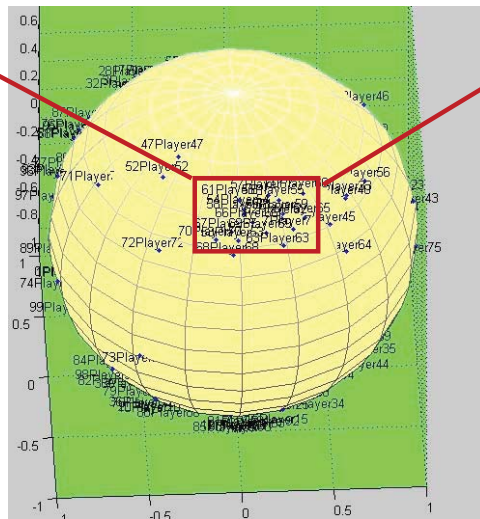
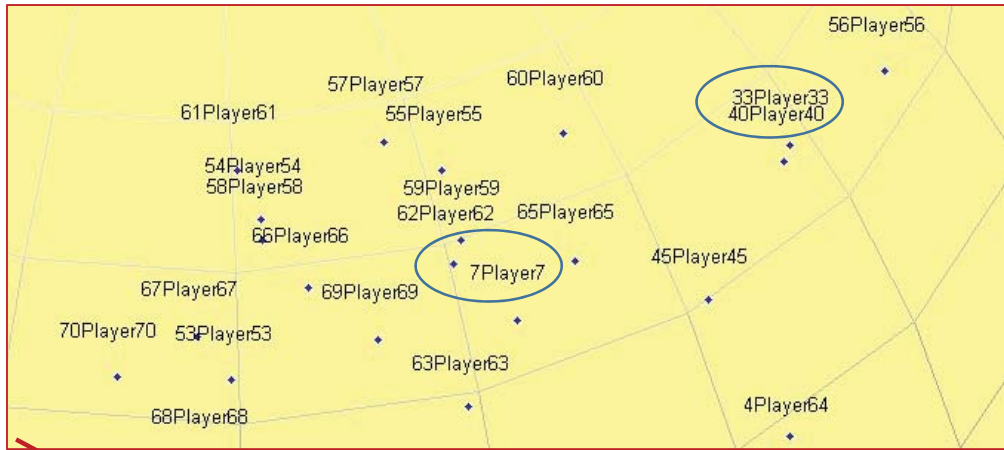


Figure 3. Output of kabaddi cluster.



Figure 4. Output of kabaddi cluster.

Figure 4 shows a part of Tennis cluster. Here, we can observe players such as Player27, Player5 and Player 17 as correctly placed. However, here also we can observe some invalid members.

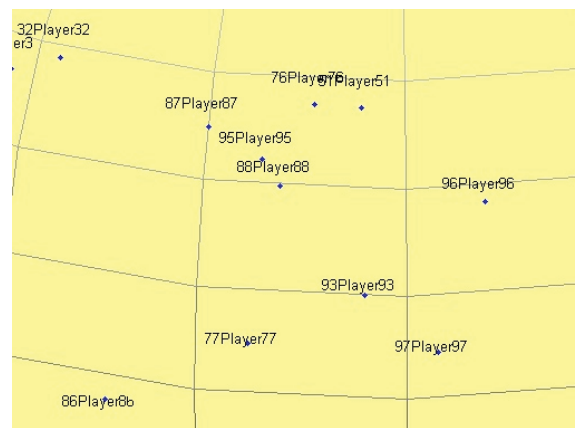


Figure 5. Part of Koko cluster.

## 5. Conclusion

Anthropometric measurements are generally used to determine and predict achievement in different sports. Further, anthropometric profiles indicate whether the player would be suitable for the competition at the highest level in a specific sport. In this paper we have presented an approach that uses anthropometric profiles and data mining techniques to identify the most suitable sport for the beginners. Here we used spatial clustering technique called ASKS projected from a 3D sphere to 2D spherical surface, for 2D visualization. We considered player similarity values in creating the affinity matrix for ASKS algorithm. Our objective was to plot the players on sphere to visualize the clusters. The algorithm clustered the players on the sphere based on similarity values.

However, there are some invalid members in clusters. So, in future works, we hope to increase the accuracy of the clusters by selecting only the most suitable features from the above 30 features. Further, in our future work, we plan to conduct a thorough evaluation of our proposed clustering approach to measure the performance.

## 6. References

- [1] S. Mining, R. Schumaker, O. Solieman, H. Chen and S. US, "Sports Data Mining | Robert P. Schumaker | Springer", Springer.com, 2018. [Online]. Available: <https://www.springer.com/us/book/9781441967299>. [Accessed: 17- Oct- 2018].
- [2] Sun Ya Xue, "Research on the Influence of Big Data on School Physical Education Development," *Sports frontier*, vol. 8, no. 6, 2016, pp. 1-3.
- [3] Banage T. G. S. Kumara, Yuichi Yaguchi, Incheon Paik and Wuhui Chen, "Clustering and spherical visualization of Web services," in *Proceedings of the IEEE International Conference on Services Computing*, 2013.
- [4] V.Cavedon, C.Zancanaro and C. Milanese, "Anthropometry, Body Composition, and Performance in Sport-Specific Field Test in Female Wheelchair Basketball Players", 2018 [Online]. Available:<https://www.frontiersin.org/articles/10.3389/fphys.2018.00568/full> [Accessed: 17 Oct 2018].
- [5] J. Solarte, "A Proposed Data Mining Methodology and its Application to Industrial Engineering", Trace: Tennessee Research and Creative Exchange, 2018.
- [6] D. Santos et al., "Reference Values for Body Composition and Anthropometric Measurements in Athletes", 2019. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0097846>. [Accessed: 14- Oct- 2018].
- [7] S. Patro and K. Sahu, "Normalization: A Preprocessing Stage", arXiv.org, 2019. [Online]. Available: <https://arxiv.org/abs/1503.06462>. [Accessed: 12-Oct- 2018].
- [8] S. Mining, R. Schumaker, O. Solieman, H. Chen and S. US, "Sports Data Mining | Robert P. Schumaker | Springer", Springer.com, 2018. [Online]. Available: <https://www.springer.com/us/book/9781441967299>. [Accessed: 17- Oct- 2018].
- [9] Journalofsports.com, 2019. [Online]. Available: <http://www.journalofsports.com/pdf/2017/vol2issue2/PartA/2-1-109-955.pdf>. [Accessed: 12- Oct- 2018].

## Distribution cost optimization using Big Data Analytics, Machine Learning and Computer Simulation for FMCG Sector

A. M. C. Adikari

Department of Transport and Logistics  
Management

University of Moratuwa, Sri Lanka  
amchandimal@gmail.com

T. P. Amalan

Department of Transport and Logistics  
Management

University of Moratuwa, Sri Lanka  
thivya@uom.lk

### Abstract

*Developments in information and communication technology has made significant impact on every sector. Unfortunately, limited research exists regarding information systems for the distribution networks in Supply Chain. This study made an effort to investigate the linkage between information systems and transportation cost optimization in FMCG (Fast Moving Consumer Goods) sector. Information systems should support the management at operational and strategic level. The study focused on the operational level implementation of information system with machine learning and big data analytics. Factors, variables and constraints affecting the cost of transportation were identified from industry experts and literature. Then a case study approach applied by analyzing the distribution network data of a Sri Lankan FMCG company. A quantitative model was developed to reflect the transport cost structure and a software model was developed considering the constraints and the cost structure, to reduce the cost of transportation by big data analytics, machine learning and computer simulation. Developed model has been compared with the existing model of transportation in the FMCG manufacturer to benchmark the optimization. In proposed model, the usage of vehicles are reduced, thereby minimizing the transportation cost by increasing the consolidation possibilities, route planning and stacking models.*

**Keywords:** Big Data Analytics, Distribution logistics, Fast-moving Consumer Goods (FMCG), Machine learning, Supply chain management

### 1. Introduction

In the business world, Information Systems (IS) are used to improve efficiency and effectiveness of business processes and throughput. MIS (Management Information Systems), DSS (Decision Support Systems) and ERP (Enterprise Resource Planning Systems) are few examples for use of Information Systems in business world.

Various information systems such as SAP are also used as supply chain process support systems. However use of information systems for network planning, route planning and for the cost optimization of the delivery is limited in Sri Lankan context. Most of the FMCG companies in Sri Lanka do not get the maximum use of information systems to minimize their transportation cost.

FMCG companies handle various types of SKUs (Stock Keeping Units) in large quantities. Comparing to a vehicle load with single items, utilization of vehicles depend on various variables such as load balancing, stacking norms and delivery constraints to delivery points. In a case study of a FMCG company in the UK, was found that the overall vehicle effectiveness was reported as 54% but it can be increased to 70% [1]. Considering the number of vehicle movements of a FMCG which is high, it can be reduced by increasing the vehicle utilization which will directly address the transportation cost.

Reducing distribution cost can be done mainly by two methods, network planning and transport operation planning [2]. Network planning is a strategic level process which determines the number of manufacturer's distribution centers and to determine the inventory policy [3]. Strategic management practices are defined in early stages of business, where the changes of the practices will have significant increase in management cost and capital investment.

Without information systems, most of the planning is done with human intuition and it has an effect on optimum output of distribution operation. However, using

human-information system collaboration with machine learning, both benefits of human intuition and computational power of information systems can be achieved. The collaboration will lead to automation of industry which is also known as Industry 4.0 [4].

Accurate decisions are primarily based on availability of information. The “Big Data” within and outside the organization can be used to make more accurate decisions [5]. Enterprise supply chain networks which are embedded with sensors, computer systems, smart phones and computerized devices generates a lot of unstructured data[6]. Because of the volume, the speed of the data generation and variety of the data, the traditional statistical approach cannot process the generated data[7]. Therefore, there was a need to develop advanced computing methods to use this vast amount of data efficiently[8]. Machine learning will enable pattern recognitions from big data and enable organizations to make accurate, data driven and fast decisions[9]. With accurate and effective decision making the organization can reduce its operational cost.

The rest of the paper is organized as follows; Section 2 presents the methodology of the research. Evaluation area and data analysis from pilot study is presented in Section 3. Techniques used on software model is presented in Section 4, Section 5 consists the approach of formulating the software model, Section 6 covers model validation, Section 7 addresses the validation of the model and the results while Section presents the limitations of the study.

## 2. Methodology

The study was carried out in three phases. In the initial phase, an extensive literature survey was conducted in the areas of information systems and distribution operations. Interviews were carried out from experienced personnel in the industry and it was based on shortlisting the factors that affect distribution performance and cost in FMCG Sector. The objectives of the initial phase is to identify the processes where information systems can be used in the distribution operation in supply chain.

In the second phase, through a meta-analysis the most prominent factors which affect the FMCG distribution cost were shortlisted and categorized followed by a pilot study at a leading FMCG company to observe and understand the FMCG distribution process, cost factors of distribution and possible areas to implement information systems in FMCG distribution. Factors affecting implementation of information systems to distribution operation optimization was determined by the meta-analysis and pilot study.

In third phase the FMCG company in which the pilot survey was carried out, was selected to check the applicability of information system and implementation.

The distribution operation data of the FMCG company were gathered by a pilot study and analyzed with operation research and big data analytics. From the analysis, possible processes where information systems can be implemented in the FMCG distribution operation is identified. A software was also developed and implemented for the FMCG firm to determine and to evaluate the expected outcome. Internal benchmarking was done to compare the distribution operation with and without information system in terms of distribution cost. The applicability of software in Sri Lankan or wider context was evaluated with the data from first phase and second phase. It was aspired to evaluate the areas and processes in distribution operation that information systems can be implemented.

The final output was a software model with big data analytics, machine learning and computer simulation which is functional with given constraints of distribution to obtain optimum transportation mix to achieve optimum cost of FMCG distribution.

## 3. Data analysis

### 3.1 Evaluation of area

Data was analyzed in two phases. In the first phase data collected from the industry experts and from the literature survey was analyzed in order to shortlist a set of factors of FMCG distribution cost and determine the processes where information systems can be implemented.

**Table 1. Factors affecting cost of FMCG distribution.**

Strategic Decision Area		Operational Decision Area	
Factor	%	Factor	%
Geographical Location	3.3%	Flexibility of Customer	16.7%
Network Plan	6.7%	Technical Barriers	23.3%
Organizational Cultural Issues	13.3%	Volume Fluctuations	10%
Lack of Investment	6.7%	Safety Issues	6.7%
		Practical Issues in Operation	13.3%
Total	<u>30%</u>	Total	<u>70%</u>

From the results, it is evident that the factors affecting the FMCG distribution cost and processes where information systems can be implemented in FMCG distribution can be categorized into two categories based on the management decision areas (Table 1).

Processes where information systems can be implemented in FMCG distribution can be also divided into strategic decisions or operational decisions with prominence given by the literature review and respondents (Table 2).

**Table 2: Process where IS can be implemented.**

Strategic Decision Area		Operational Decision Area	
Factor	%	Factor	%
Network Planning	3.3%	Route Planning	23.3%
		Vehicle Allocation	26.7%
		Stacking Planning	6.7%
		Vehicle Tracking	20%
		KPI monitoring	20%
Total	3.3%	Total	96.7%

The meta-analysis suggests that operational decision area has significant impact on implementing of information systems to FMCG distribution. Therefore, further analysis was done by eliminating strategic decisions for distribution cost reduction and only focusing on the operational aspect, as meta-analysis is ensured.

The narrowed down areas from the meta-analysis show direct linkage to the distribution. In operation aspect the shorter the distance vehicle runs, lower the cost of distribution. Dijkstra algorithm is widely used to obtain the shortest path of distribution in most instances[10]. Apart from Dijkstra algorithm, algorithms like Floyd-Warshall algorithm are used to obtain least cost routes and paths of distribution[11].

However, these algorithms are focused on connecting all the nodes of distribution, in a way that minimize the cost of distribution. The practical constraints like, distribution points accessibility constraints were not considered in those algorithms. Therefore, past vehicle distribution data as big data will reflect more accurate route planning, because the new routes were derived from actual distribution data, rather than hypothetical, theoretical suggestions. Further evaluation was conducted to focus on optimizing distribution based on past data. Therefore, a case study firm with past distribution data was selected for evaluation.

### 3.2 Existing model

In the company where the pilot study was done, Microsoft Excel was used as a tool to aid the route planning process. The process followed was not scientific, but was a more trial and error method, with manual

arrangement of ordering rows with manual calculations until the operator gets the route plan output. Since it was a manual method, and it did not have a standard process, the method for route planning was different from user to user.

In existing system, problems of obtaining the best consolidation mix, lack of vehicle distance utilization, and getting optimum product mix for stacking were identified which can be improved from implementation of a data backed information system.

## 4. Problem formulation

The problem focuses on allocating the vehicles to the orders while minimizing the overall cost.

Let  $\mathbf{O}$  be the set of orders, where each order  $o \in \mathbf{O}$  has a delivery location ( $A_L$ ). An order consists of only one type of item  $\mathbf{I}$ . These orders are to be processed by a set of trucks  $\mathbf{V}$ . There are five types of trucks (A, B, C, D, E) based on the capacity of the truck. Each delivery location has set of truck types where the loading and unloading operation can be handled.

**Table 3. List of symbols.**

Symbol	Description
$N$	No of orders Received
$T_i$	Time of $i^{\text{th}}$ order received
$T_m$	Time of $m^{\text{th}}$ shipment sent
$W_i$	Weight of $i^{\text{th}}$ order received
$V_i$	Volume of $i^{\text{th}}$ order received
$W_k$	Maximum weight the type $k$ vehicle can be carried
$V_k$	Maximum volume the type $k$ vehicle can be carried
$C_{lk}$	Cost of distribution to $l$ distribution point by type $k$ vehicle
$X_l$	{Vehicle Types that distribution point $l$ can be catered by}
$A_m$	{delivery points in $m^{\text{th}}$ shipment}
$O_m$	{Orders in $m^{\text{th}}$ shipment}
$N$	{Consolidated orders}
$N_c$	Number of Consolidated orders
$C_T$	Total Cost of Distribution
$S_l$	No stacks that $l$ is stacked

Our objective is to cover all orders  $\mathbf{O}$  with trucks  $\mathbf{V}$ . It is assumed that the number of trucks available for each truck type is unlimited.

### 4.1 Constraints and conditions

#### Vehicle capacity constraints

Each vehicle type has its own loading capacities by weight and volume. Let  $W_A, W_B, W_C, W_D, W_E$  be the

maximum weights by kilograms (kg), A,B,C,D,E be the type of vehicles that can be loaded and let  $V_A, V_B, V_C, V_D, V_E$  be the maximum volumes by cubic meters ( $m^3$ ) type A, B, C, D, E vehicles can be loaded where,

$$V_A < V_B < V_C < V_D < V_E \text{ and } W_A < W_B < W_C < W_D < W_E$$

Each delivery location ( $A_L$ ) has set of vehicle types that can be handled  $X_{A_L}$ . In the shipment there cannot be delivery locations which violate that rule.

$$\forall m \in N \sum_{l \in O_m} v_l < \max(V_k) ; k \in X_l ; l \in A_m$$

$$\forall m \in N \sum_{l \in O_m} W_l < \max(W_k) ; k \in X_l ; l \in A_m$$

#### Stacking of items

Each item type ( $I$ ) has maximum number of stacks that the item can be stacked ( $S_{I\_MAX}$ ). No of stackings ( $S_I$ ) per item in a shipment should not exceed  $S_{I\_MAX}$ . Therefore, let  $M$  be an arbitrary shipment,

$$\forall o \in M S_I \leq S_{I\_MAX} ; I \in o$$

#### Time of delivery

All order received should be shipped within 24 hours.

$$\forall m \in N T_m - T_l (\forall l \in O_m) \leq 24 \text{ hours}$$

All the symbols in the notations are mentioned in Table 3.

#### Route of distribution point

All distribution points in consolidated order group should be in same route. Let  $R$  be the set of routes where a route  $r$  ( $r \in R$ ) consists set of delivery locations.  $A_L \in O$  and  $M$  is an arbitrary shipment,

$$\forall o \in M \rightarrow A_L \in r ; A_L \in o$$

#### Payment on vehicle conditions

The third party transporters are paid by the type of the vehicle used for the shipment and payment clusters ( $P$ ) and every delivery location is categorized its own payment cluster ( $A_L \in P$ ). Each payment cluster has different rates for different type of vehicle used. Let  $p_A, p_B, p_C, p_D, p_E$  be the rates for the delivery to payment cluster  $p$  where  $p \in P$  for the vehicle types A, B, C, D and E where,

$$p_A < p_B < p_C < p_D < p_E$$

Let us assume shipment  $m$ , consists only with orders of  $A_Z$  where  $A_Z \in p$ .

If  $\sum_{l \in O_m} W_l < W_A$  OR  $\sum_{l \in O_m} V_l < V_A$  then, vehicle is paid by  $p_A$ . Likewise let  $k$  be the set of vehicle types. Then, if  $\sum_{l \in O_m} W_l < W_k$  OR  $\sum_{l \in O_m} V_l < V_k$  then, vehicle owner is paid by  $p_k$ .

If there are several payment clusters per order, the payment is done for the highest payment cluster.

## 4.2 Objectives

Given the orders, vehicles and cost, the primary objective is to cover all the orders and minimize the cost. To minimize the cost it is necessary to find optimum order mix to a vehicle under the constraints and conditions.

$$\text{Min} \sum_{m=1}^{N_c} (\max(C_{lk}) k \in R_l ; l \in A_m)$$

where  $l$  is delivery location and  $k$  is type of the vehicle used to make the delivery and  $C_{lk}$ , the cost of delivery to delivery point  $l$  by  $k$  type vehicle as per the payment clusters.

## 5. Solution approach

Once the orders are received vehicle allocation to minimize the cost of distribution is the expected output of the solution. It was assumed that there weren't any orders which were not allocated into vehicles. We assumed that the vehicle allocation for next day's delivery orders will be determined on the previous midnight's confirmed orders.

Rule checker was enforced in the proposed solution to enforce the constraints and conditions. Rule checker evaluates the constraints which were defined in Section 4. To define the routes for distribution and identify stacking constraints for vehicles for each type of item, big data analytics and machine learning was used.

### 5.1 Route definition and delivery point constraints

Two years of past distribution data was analyzed to define the routes for the delivery. The distribution points which were served in a single shipment was identified from Big Data Analytics. The vehicle accessibility constraints were also identified from past years data using Big Data Analytics. Machine learning, will be used for



future, redefinition of routes based on changes of distribution points.

### 5.2 Route definition and delivery point constraints

In the order request document, relating to the item only the volume and weight is available. From the pilot study it was found that weight to volume ratio is unique for an item and each row of order request document contains only a single item. From the list, the items with weight to volume ratio can be created. From the past distribution data, maximum number of items which can be stacked in each type of vehicle, can also be found out. With item information from the FMCG Company, the weight and volume for one case of a certain item, were gathered.

From the distribution data of two years, maximum loads that were carried out for each vehicle type for each item were calculated/generated. Further, the item which does not have maximum load details, the data were filled with items with same weight, volume and stacking attributes. For others it was calculated by proportion.

### 5.3 Consolidating orders

The process is to find all the ways a set of order can be sent through the partitioning and simulation. After that, sets will be filtered with predefined routes and what orders can be sent in a same vehicle, found. From all the feasible results, the best solution can be selected which has the least cost.

As an example, from Set Partitioning Theory for 3 orders that were received for route Planning Order No 1, 2 & 3, partitions can be arranged in

$\{\{1,2,3\}\}, \{1,\{2,3\}\}, \{\{1,2\},3\}, \{2,\{1,3\}\}, \{\{1\},\{2\},\{3\}\}$

In the next step, the partitions which have combinations and which does not comply with constraints, will be removed. Let us assume only order 2 and order 3 are situated in connecting city by a same predefined route, considering that all the combinations with  $\{1,2\}$  and  $\{1,3\}$  and  $\{1,2,3\}$  can be removed.

Then existing partitions are,  $\{1,\{2,3\}\}, \{\{1\},\{2\},\{3\}\}$  In this step, there is also a need to check whether these partitions does not violate constraints for distribution points. If there are violations, these should also be removed. Let us assume in the existing sets of partitions that there are no partitions that violate given constraints. Then the costs for route plans  $\{1, \{2, 3\}\}$  and  $\{\{1\}, \{2\}, \{3\}\}$  should be estimated using the algorithm. Then the most cost effective route plan can be selected.

However, the problem is, when the number of orders are increasing number of partition sets that can be arranged also get increased exponentially. The number of partitions of a set can be derived from Bell No, as demonstrated in Table 4.

**Table 4. No of orders vs No of sets of partitions.**

No of Orders	No of Sets of Partitions
1	1
2	2
3	5
4	15
5	52
6	203
7	877
8	4140
9	21147
10	115975
11	678570
12	4213597

As observed in pilot study there are about 60-100 orders in a day and 300-400 in a peak day. Therefore, processing speed may affect the operation when the number of orders get higher.

From simulation, the initial algorithm to get an output of 12 orders takes 10 mins (used laptop with 12GB RAM and Intel Core i5 6200U processor). If number of orders are less, the output will be generated in a shorter time.

In order to reduce the computational time, the number of orders should be split into few segments, if the number of orders were greater than 10 orders. However, there were many vehicles used to fulfill all the delivery orders splits, to reduce the workload for each collection. Among the other methods, Tabu search performed well in splitting and clustering orders [12]. Therefore, to obtain the output with speed the Tabu Search approach was taken for creation of algorithm by dividing orders into clusters.

The results from simulation clustering with Tabu Search algorithm was, 95% - 100%, similar to the most cost effective theoretically available route plan.

Clustering and Tabu Search approach is implemented by following steps.

- Clustering orders by distribution point
- Find all the ways of orders can be sent within the same cluster by set partitioning theory
- Remove all the partitions which does not comply with constraints
- Estimating and comparing existing sets for transportation cost
- Selecting minimum cost partitions set of distribution points
- Clustering orders by cities

- g) Find all the ways of orders can be sent within the same cluster
- h) Remove all the partitions which does not comply with routes and constraints
- i) Estimating and comparing the existing set of partitions for transportation cost
- j) Selecting the minimum cost partitions for each city.
- k) Sorting by utilization highest to lowest and iterating from highest utilized consolidated groups to lowest and consolidating orders.

The segmentation was done based on the distribution point, city and route. The segmentation can be defined differently, based on the distribution point mapping as per the organization's requirement and default usage. In order to get the cities that can be consolidated, the routes should be defined. To define routes Big Data Analytics were used from previous distribution data. Further, machine learning was used with big data analytics to identify the constraints for the distribution points. In the software model, constraints and routes will be updated with the user inputs and through machine learning. Backward mapping is used to map the consolidated orders to the original order to the traceability.

## 6. Validation

### 6.1 Benchmarking with past delivery costs

Parameters like cost of distribution, types of vehicles used, number of vehicles used, vehicle utilization by volume, vehicle utilization by weight were considered for the distribution data of random month for the manual model and also for the software model implemented.

The actual cost of distribution in selected month was reduced by 7.8% using software simulation model, compared to manual method. Further, average vehicle utilization by weight for the selected month is increased from 72% to 76%, and vehicle utilization increased by volume from 58% to 60%. The vehicle turns will be also reduced from 1195 vehicle turns per month to 1067 vehicle turns per month.

The vehicle type shift from smaller to larger vehicle can be observed in system aided distribution planning. For the selected month, the information system aided distribution plan has 26% Type D vehicles compared to 15% Type D vehicles in the actual distribution. Further, the usage of Type C was reduced by 13% with information system aided route planning. The computational power to consolidate by considering most of the combination is the reason to have this type shift between larger capacity vehicle types.

## 6.2 Results after implementation

After the implementation of the distribution planning information system, the month of June was recorded as the month with lowest cost per case delivered (LKR 33.11 per case) and highest vehicle utilization (74%) by weight for the year 2018 (Figure 1). Ten human hours for the planning process was saved by implementation of the system. The cost per case delivered, is indicator for minimization of vehicles compared to orders received. Since, the demand for the items are different from month to month, the number of vehicles were not taken to measure and compare the effectiveness of the implemented system. As a result cost per case delivered is taken as a key measure to compare the effectiveness of the system. The cost per case was reduced from the average of LKR 36.17 per case delivered to LKR 33.11 per case delivered by 8.5% (Figure 1).

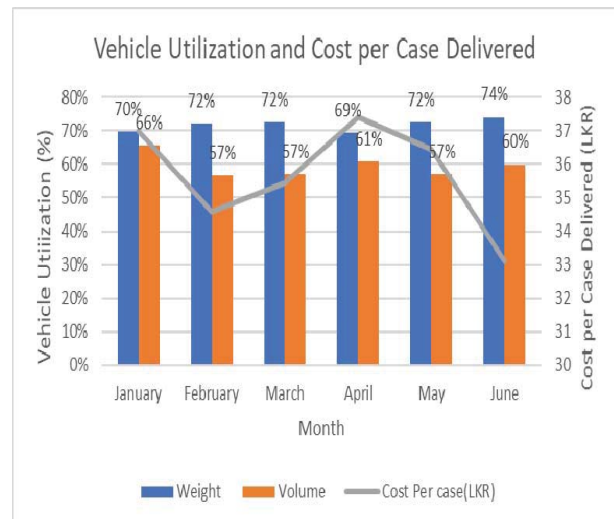


Figure 1. Vehicle utilization and cost per case delivered.

## 7. Limitations

From the developed model the cases and instances that should be evaluated with human intuition were not considered as overriding constraints. The software model can be applied only for the distribution of unitized items, such as FMCG items. There can be complications in results if it is applied for bulk cargo transportation or project cargo transportation. The model is designed for centralized distribution model and the distribution operation from multiple distribution points is not considered in the study.

Previous distribution data should be available to determine the routes and constraints for the distribution points. Otherwise the data should be input manually and it

can reduce the accuracy of the model. The stacking information should be defined for each item for the accurate output. The pricing method considered in the study for the distribution was fixed matrix based pricing based on type of vehicle. The use of vehicle was by the third party transporters. The transporters were paid to the point where the highest amount is paid according to the price matrix for the delivery made. The software model does not comply with other pricing models.

However, through further studies the derived software model can be further developed, to the other distribution sectors as well. From exploration of constraints affecting the cost of distribution for various industries worldwide, the software can be updated to encompass a wider scope. Other pricing models such as per kilometer pricing, ton kilometer, and time based pricing models can be considered in future research. For stacking optimizing purposes other parameters such as area and height of cargo loading area can be considered.

## 8. References

- [1] Simons, D., Mason, R. and Gardner, B., "Overall vehicle effectiveness," *Int. J. Logist. Res. Appl.*, vol. 7, no. 2, pp. 119–135, Jun. 2004.
- [2] Patidar, R., Khan, I. and Ghosh, M., "Linking the Production Planning and Supply Chain using Fuzzy Logic: an Integrated Model for FMCG Products," *academia.edu*.
- [3] Tang, C., "Robust strategies for mitigating supply chain disruptions," *Int. J. Logist.*, vol. 9, no. 1, pp. 33–45, Mar. 2006.
- [4] Ansari, F., Erol, S. and Sihn, W., "Rethinking Human-Machine Learning in Industry 4.0: How Does the Paradigm Shift Treat the Role of Human Learning?," *Procedia Manuf.*, vol. 23, pp. 117–122, Jan. 2018.
- [5] Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L., & Cahnoto, A. (2018). Big Data, Big Decisions: The Impact of Big Data on Board Level Decision-Making. *Journal of Business Research*, 93, 67-78, Dec. 2018.
- [6] Addo-Tenkorang, R. and Helo, P. T., "Big data applications in operations/supply-chain management: A literature review," *Comput. Ind. Eng.*, vol. 101, pp. 528–543, 2016.
- [7] Tiwari, S., Wee, H. M. and Daryanto, Y., "Big data analytics in supply chain management between 2010 and 2016: Insights to industries," *Comput. Ind. Eng.*, vol. 115, pp. 319–330, 2018.
- [8] Paaajenen, S., "Opportunities of big data analytics in supply market intelligence to reinforce supply management," 2017.
- [9] Zhou, L., Pan, S., Wang, J. and Vasilakos, A. V., "Machine learning on big data: Opportunities and challenges," *Neurocomputing*, vol. 237, pp. 350–361, May 2017.
- [10] Xu, M. H., Liu, Y. Q., Huang, Q. L., Zhang, Y. X. and Luan, G. F., "An improved Dijkstra's shortest path algorithm for sparse network," *Appl. Math. Comput.*, vol. 185, no. 1, pp. 247–254, Feb. 2007.
- [11] Aini A., and Salehipour, A., "Speeding up the Floyd–Warshall algorithm for the cycled shortest path problem," *Appl. Math. Lett.*, vol. 25, no. 1, pp. 1–5, Jan. 2012.
- [12] Qiu, M. , Fu, Z., Eglese, R. and Tang, Q., "A Tabu Search algorithm for the vehicle routing problem with discrete split deliveries and pickups," *Comput. Oper. Res.*, vol. 100, pp. 102–116, Dec. 2018.

## A scalable bioinformatics analysis platform based on Microservices Architecture

S. Rajapaksa, A. Wickramarachchi, V.  
Mallawaarachchi  
*Department of Computer Science and  
Engineering  
University of Moratuwa  
Sri Lanka  
sandunip@cse.mrt.ac.lk*

W. Rasanjana, I. Perera, D. Meedeniya  
*Department of Computer Science and  
Engineering  
University of Moratuwa  
Sri Lanka*

### Abstract

*With the advancement of technologies, web services play a significant role to maintain infrastructure in healthcare domain due to the increasing demand of performance. In such systems, adoption of novel technologies is necessary to increase the productivity and reduce the burden of maintenance associated with legacy systems. Microservices architecture has become prominent in deploying server-side enterprise applications by allowing maintainable functionalities. However, it is challenging to utilize microservices in the domain of bioinformatics, although it enables independent process execution and maintenance. This paper introduces the utilization of microservices architecture to build an optimized platform for bioinformatics analyses. We present a hybrid architecture that consists of different hardware platforms to execute accelerated computational services, independently. The core communication is based on an Application Programming Interface (API) gateway. Furthermore, the paper presents the evaluation of results related to the performance of the proposed solution under varying biological sequences as inputs and algorithms.*

**Keywords:** *Bioinformatics analysis, microservices architecture, enterprise applications*

### 1. Introduction

Bioinformatics is an emerging discipline that draws upon the strengths of computer sciences and mathematics to analyze information flow in biological systems. This consists of diverse areas such as genomics and proteomics, where the subject experts utilize different technologies for the advances in each field [1]. Many bioinformatics computations are accessed via web services, while the underlying technologies differ vastly. This enables the provision of robust Application Programming Interfaces (APIs), which can be used by experts to perform the required computations [2]. Current

research focuses on web technologies, cloud computing, distributed and parallel computing infrastructures, to enable responsive and efficient services, considering the heterogeneous nature of the hardware and software technologies.

Bioinformatics domain requires software solutions with high performance and computational capacities. The developers are motivated to split applications into small, easily maintainable functional units that can be replicated on demand [3, 4]. The microservices architecture abstracts the underlying technologies by splitting a complete solution to a set of self-contained services. These can be developed, deployed and scaled independently, thus enables a language agnostic communication with services using JavaScript Object Notation (JSON) [5]. Thus, microservices architecture has become popular, emphasizing the design and development of highly maintainable and scalable software [6]. Adoption of microservices in bioinformatics analyses enables the development of a platform that supports to use heterogeneous technologies and can be made available for external users using an API gateway [2].

Although there are existing computerized bioinformatics analysis tools, and web services, the efforts made to improve their architecture and performance using the state of the art microservice architecture are limited [3]. Thus, the services provided by vendors such as NCBI [7], Taverna [8] and Galaxy [9] remain to use traditional software architectural approaches and deployment methodologies. Additionally, the use of optimized algorithms in the existing bioinformatics computational services is limited, although they provide high throughputs, due to the integration of incompatibilities in hardware and software levels of the technologies [8]. Thus, it is challenging to overcome the heterogeneity of the executables and the nature of computations, when providing a unified platform for algorithmic execution.

This paper presents an approach to create a bioinformatics platform to support the execution of different bioinformatics computations with the use of microservices architecture. The solution introduces a novel architecture, which utilizes the API gateway pattern that exposes microservices to the external users via a single API masking the underlying discrepancies. Section II elaborates the literature on the usage of microservices in bioinformatics computations and existing services. Section III describes the architecture of the proposed system and Section IV explains the methodology. The experimental and evaluation results are presented in Section V. Finally, Section VI concludes the paper with the inferences obtained from the results and possible future extensions.

## 2. Literature review

### 2.1 Microservices

Microservices are self-contained units of functionality with loosely coupled dependencies on other services and are designed, developed, tested and released independently [10]. They can be reused across many different solutions and can scale appropriately. Although the microservices technology was introduced into the UNIX kernel as early as the 1970s, it has emerged in web technologies only recently. Shadija *et al.* [4] has analysed computational systems for their granularity and performance and suggested that microservices can increase the system granularity, thus increasing the scalability and maintainability. The microservices deployed over many server nodes cause a delay due to network overheads but can perform better by optimizing the execution environments.

At present, bioinformatics analyses are used for simulation and testing of different domains such as genetic studies and drug discovery. The need for microservices based solutions has been highlighted, with the increasing need for efficient, maintainable and scalable computational techniques [3]. The growing need for microservices based solutions to address bioinformatics computations is shown by Williams *et al.* [3]. Literature has highlighted the developments in the related field due to the increased agility of the microservices architecture with heterogeneous analysis techniques.

### 2.2 Containers

A *container* is a lightweight, stand-alone, executable package of software and a widely adopted method to develop microservices. A container includes all the required artefacts such as the code, runtime, system tools, libraries and settings [11]. Docker [12] is a well-known container platform provider. It provides the ability to

package applications along with their dependencies into lightweight containers that can be easily moved between different distros, start up quickly and independently. Many related works that emphasize the importance of using microservices in the domain of bioinformatics are provided as deployable containers.

CodonGenie [13] web-based tool is an ambiguous codons design tool that supports protein mutagenesis applications. It is designed as a microservices so that it can be integrated to applications by a RESTful web service API as well as Docker container. ChEMBL API allows integrating data of bioactive molecules with drug-like properties to other applications [14]. Moreover, Khoonsari *et al.* [15], have introduced a generic method using the microservices architecture, where software tools are encapsulated as Docker containers that can be connected to scientific workflows and executed in parallel. The approach utilizes containers to provide microservices and the containers are managed using Kubernetes [16]. The service caters for functions in workflows related to the domain of metabolomics data analysis.

### 2.3 Microservices in Biology, Medicine and Bioinformatics

A Reaction balancing web service for computational systems Biology has described by Dobson *et al.* [17]. They have implemented a RESTful web service that offers a language-agnostic way of binding services together. This platform is built using many technologies in JAVA and Python. This is a web-based solution, where message passing is based on JSON communication. The underlying logic has used many technologies with different forms of input types. Unlike in a service-oriented architectural design, the system is developed as a web service, which connects the individual components with more maintainability.

Fjukstad *et al.* [18] have presented a data exploration application for systems biology using microservices. They have used microservices due to the language-agnostic means of communication between the building components. Although this gives a solution for data exploration using visualization components such as heat maps, the other functionalities related to genome analysis have not considered. A similar research by Hill *et al.* [19], have performed medical data processing by incorporating IoT devices to gather data and build a community healthcare system. However, they have not addressed the bioinformatics related use cases and complex computational requirements.

Arkas-Quantification and Arkas-Analysis [20] are cloud-scale RNAseq pipelines, which are versioned into Docker containers and publicly deployed in Illumina's BaseSpace platform [21]. PAPAyA [1] is a cloud-based

framework that provides genomic processing services for bespoke therapy guidance. It provides diverse pipelines and tasks (detecting variants, mutations, copy number variation, differential gene expression and DNA methylation) as Linux-based containers.

According to the literature, most of the research has been done using cloud services that provide RESTful services. Many solutions are presented as individual services with isolated APIs and should be integrated through a single API gateway. Thus, an aggregation of a set of microservices is required for a language agnostic execution of subroutines in workflow modelling. This enables the workflow modelling tool [22], to operate seamlessly by calling an API gateway that communicates with relevant service to complete workflow. Although the related work has identified importance of microservices to provide robust analytical capabilities, there has been limited emphasis on the usage of microservices architecture in bioinformatics workflow design.

### 3. System architecture

#### 3.1 System components

We describe the architecture of the proposed microservices based platform to support optimized bioinformatics workflow design and modelling [22, 23]. It utilizes both central processing units (CPU) and graphics processing units (GPU) powered service containers to support optimized execution of algorithms. The proposed solution provides the end users with an API, which exposes algorithmic functionality in a language-agnostic manner using JSON message passing. The proposed solution consists of three major components; *Web Server*, *API Gateway* and *Microservices*, as shown in Figure 1. The service layer contains highly decoupled service instances that are wired to the external API using the gateway. The Web Server enables the functions to be accessed in the form of a REST API. The requests to the web server are forwarded to the API gateway for service invocation.

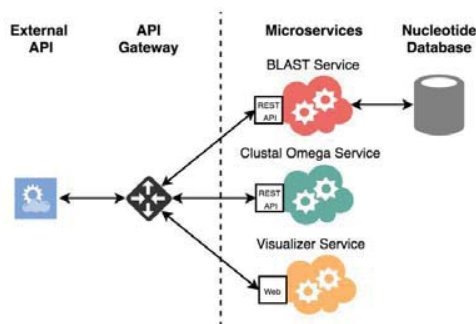


Figure 1. System components of the microservices based solution.

Since the services are running their own REST APIs to handle and process requests, an API gateway is required to connect each service and control its access. This component also manages to add, remove and migrate services to support further functionality. Addition of new services to the API gateway is indicated by the addition of API resources along with resource methods (GET or POST). The new resources can be configured with path parameters other than query parameters.

Microservices is the collection of the actual executable components and deploy in several environments that support CPU and GPU computing to optimal execution of algorithms. The services with data requirements are provided with independent databases to retrieve data.

The microservices architecture is adopted due to its scalability with the addition of services. This allows distributed execution of bioinformatics related analyses using a cluster of computing nodes that support both GPUs and CPUs. In the scenarios, where databases are used to perform queries, separate database instances can be maintained alongside the microservice in its own container space. Depending on the resource utilization and the growing demand of an instance, the number of computing nodes can be increased. Further, extra nodes to support different technologies and computing capabilities can be added. The API gateway manages the addition of services and routing requests to those services.

The deployment of the new services involves the steps of spinning up a microservice container or a process, assigning an API resource along with a method and wiring up the API gateway resource endpoint with the desired microservice endpoint and port. Services such as BLAST [24] that require data to be fetched are provided with independent databases. The services which require specific hardware configurations run on clusters that possess such capabilities. The system consists of a web application that renders visualized components and requires sending data once they are processed.

#### 3.2 Methodology

Our solution processes data using two types of clusters; GPU and CPU clusters. GPU clusters execute algorithms that exploit GPU level parallelism such as GPU BLAST [25], SW-CUDA [26], while the CPU clusters execute algorithms that are not optimized for GPUs such as Clustal Omega [27] and T-Coffee [28]. The use of GPUs is highlighted due to the capability of executing GPU powered containers, which can cater to many implementations that exploit GPU parallelism. CPU clusters are configured with different specifications as demanded by the process. The services that run I/O intensive operations such as BLAST require higher disk space allocation and a RAM of nearly 16GB to perform optimally. However, multiple sequence alignment

algorithms do not require more disk space, but they can be made run faster by having a RAM of reasonable capacity. Section V describes the evaluation of the services for memory footprint in detail.

### 3.3 Interaction model

Majority of the data processing in our solution happens in an asynchronous manner due to the longer times consumed for heavy computations. Hence, the end results are provided for users in three ways, (1) *HTTP requests and WebSocket push*, (2) *Webhook calls* and (3) provision of a *Results endpoint*. Different services are configured to use different forms of response methods depending on the time complexity of the operations. For example, results of BLAST processes are made available through a results endpoint and a WebSocket push, whereas Clustal Omega results are directly sent as HTTP responses. This is because BLAST results take longer times than a normal HTTP timeout and Clustal Omega provides results much faster. This eliminates the resources being used up by connection waiting times.

#### 1) HTTP requests and WebSocket push

This provides REST call access via HTTP requests and WebSocket access to the services. WebSocket methodology keeps alive WebSocket throughout the service execution and data is pushed back to the user. All the services are designed to support this functionality since this can be used by other web-based applications that use the microservices to obtain a greater user experience.

#### 2) Webhook Call

Webhook Call enables the provision of another API endpoint to receive the execution results. This is only used for heavy execution services such as BLAST. The results will be sent to the provided URL by means of a POST request once the computations are completed.

#### 3) Results Endpoint

This generates a new endpoint in each microservice, providing direct access to the results. The results will be available for a limited time, after which a scheduled task will clear the results to save the storage space to serve future demands.

## 4 Methodology

### 4.1 Hardware arrangement

The system comprises two types of computations that require CPUs and GPUs. Therefore, two clusters are utilized in the system to perform relevant executions. Figure 2 shows the processing units in our solution. The solution is implemented on top of the two clusters to

perform GPU computations and CPU computations separately.

The two clusters are exposed to the external API via the API gateway.

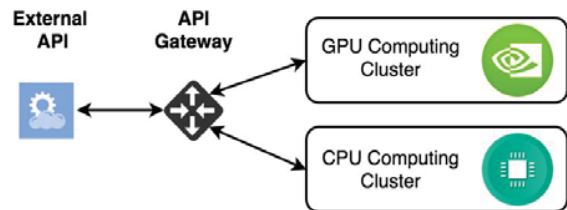


Figure 2. Arrangement of processing units.

Table 1 states the specifications of the hardware used in each cluster. The system consists of a CUDA powered GPU and a CPU that can run up to 8 threads using hyperthreading.

Table 1. Hardware specifications.

Specification	CPU	GPU
Model	i7 4770 3.4GHz	GeForce GTX 480
Physical Cores	4	384 (CUDA)
Logical Cores	8	384 (CUDA)
Memory lane size	8GB	4GB

### 4.2 Service access endpoints

The provided services are accessed via the API gateway. The API gateway can be configured to forward the incoming requests as new services, are being added to the system. Since there are two types of services as a web application and REST services, they are identified through the API URL.

Service endpoints are accessed via the following address.

```
http://192.248.8.242/api/services/<service-name>
```

Web endpoints are access via the following address

```
http://192.248.8.242/api/web/<web-application-name>
```

Figure 3 illustrates the JSON formatted request object for the BLAST service. The users can send all flags along with the request to perform a BLAST search as demonstrated.

```
{
  "data": "<SEQUENCE DATA IN FASTA
  FORMAT>",
  "gpu": "true",
  "threads": "4"
}
```

Figure 3. Request body to BLAST service.

Figure 4 shows the response object for an HTTP REST invocation of the BLAST service. The users are provided with a results endpoint as a results file or a message indicating the status of the service as *queued*, *in progress* or *failed*.

```
{
  "resultsEndpoint":
    "<DOMAIN>/api/gpu-blast/results/id_1bas98xA"
}
```

Figure 4. Response JSON for BLAST service.

### 4.3 Service execution and process management

The system consists of several services running in each cluster as shown in Figure 2. Thus, a given service only allows the execution of a single process at a time. Figure 5 illustrates the execution of requests given to a service. Initially, all the requests are queued and followed by sequential execution within the computing unit in a First-In-First-Out (FIFO) manner. A job ID is assigned at the beginning of execution to provide a result uniquely for each job. This job ID is used to generate the results link. This approach ensures that the resources of the system will not be overwhelmed due to concurrent requests. Additionally, the methodology makes available the resources for all the services running on a given computing unit. Further, the queue is used to manage the concurrency to achieve the optimum throughput of each of the services by exploiting the maximum concurrency.

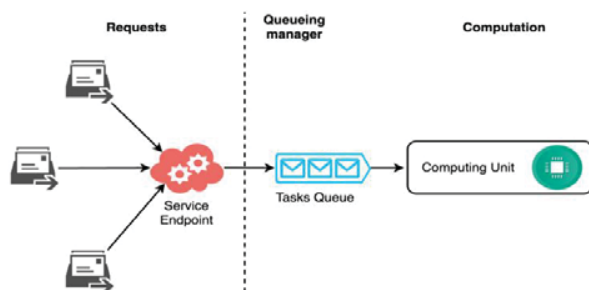


Figure 5. Request processing.

Execution of algorithms such as BLAST requires a separate database to perform the search. Therefore, a database is maintained independently for that particular service. Services can be configured to fetch data via FTP access. For example, BLAST service can access the NCBI FTP Site [29].

## 5. Results and discussion

The system is tested for response times and compared the features to obtain the outstanding metrics. We have evaluated the system to compare the performance of

different services running in heterogeneous environments while being subjected to varying concurrency levels during high demands. Figure 6, shows the time taken by the BLAST search service to get results by executing on CPUs and GPUs for varying lengths of input sequences. The service searches through the *env\_nr* protein database from NCBI BLAST FTP Server [29], which consists of 7,007,470 protein sequences with 1,397,713,333 characters.

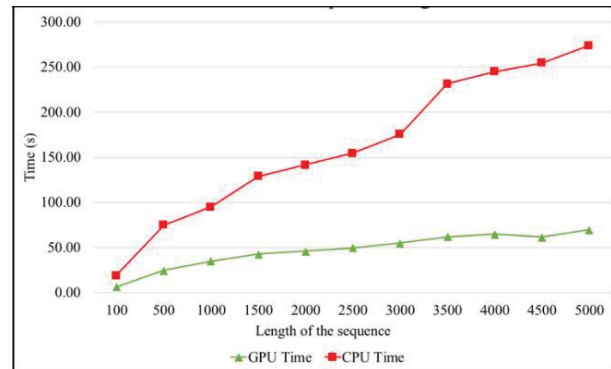


Figure 6. Execution time of BLAST in CPU and GPU vs sequence length.

It is evident that the times taken for CPU computations have increased rapidly with the increasing length of the input sequence, whereas times taken for the GPU computations have a slight increase with the increase of the sequence length. Although in sequential execution, the time was increasing rapidly, the time increase for GPU execution was not so steep, leading to an increase in speedups. Therefore, the GPU powered microservices based solution provides better performance while saving resources and time for more computational requests from users.

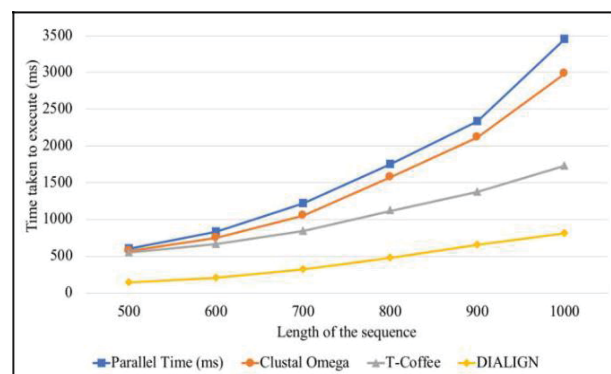


Figure 7. Service execution time in parallel and individually vs length of the sequence.

Additionally, the system is tested for the performance of three Multiple Sequence Alignment (MSA) tools;



Clustal Omega [27], DIALIGN [30] and T-Coffee [28]. Figure 7, shows the times taken for invocation of MSA services that execute in sequentially and concurrently for different lengths of the sequences. The results were obtained to ensure that the execution of multiple services on the same computing unit does not cause to delay operations due to the process of scheduling scheme of operating system. The experiment was conducted using 5 sequences from the *env\_nr* protein database of lengths varying from 500 to 1000 characters as inputs for each MSA tool. According to the performance measures, the parallel execution required nearly the same time as that of the process that consumed the most time. Thus, the parallel execution of independent services does not hinder the performance of other services running in the same computing unit.

The proposed solution was further evaluated based on the ability to handle concurrency. Exhaustive tests were conducted for the three MSA tools Clustal Omega, DIALIGN and T-Coffee using up to 10 concurrent service executions (CSE) for each. We have obtained the measurements related to the usage of system memory and consumption of time. This information can be used to decide on the amount of allowable concurrency per microservice container and the resources to be allocated for each of them. The test was conducted using 5 sequences from the *env\_nr* protein database, where each sequence was 500 characters long.

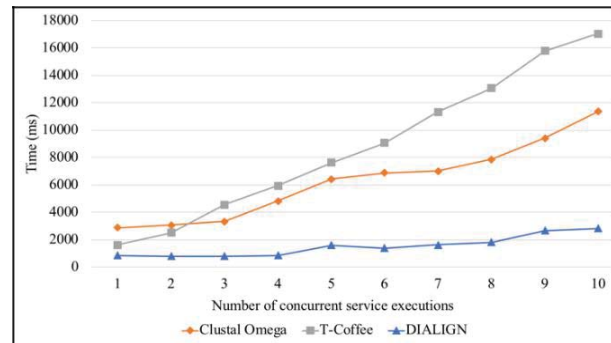
**Table 2. Feature comparison between the implemented solution and existing systems.**

Feature	Standalone tools	Web applications	Microservices based solution
Technology	JAVA/C++, CLI programs	Web based	Microservices based
Fault Isolation	Entire system can fail	Entire system can fail	Improved fault isolation
Scalability	Not scalable	Horizontally scalable	Individual services can be scaled at a more granular level.
Real-time processing	No support	No support	Publish-subscribe framework enables real-time data processing
Technology stack support	Whole system is developed using a single technology	Whole system is developed using a single technology	Services can be developed using different technologies suitable for the application.
GPU support	Limited by machine specs	Expensive and inefficient to run all on a single GPU cluster	Can have GPU clusters deployed for specific services

Figure 8 shows the times taken by each of the services Clustal Omega, T-Coffee and DIALIGN to complete execution under varying levels of concurrency. According to the results, the increase of concurrency has degraded the overall performance of the service. DIALIGN shows a low gradient whereas the times observed for Clustal Omega and T-Coffee have a steep increase along with the increasing number of CSE. Clustal Omega has a 5-fold time increase of 10 CSE. However, up to a concurrency of 4 service executions can be permitted as the increase of the time is nearly twice that of a single execution. Similarly, for T-Coffee, a concurrency of 1 to 3 service executions can be allowed. For DIALIGN the maximum concurrency of 10 service executions can be allowed without degrading the

performance significantly. Thus, the implementation of the service must use a concurrent queue that limits the number of active executions at any given time.

Furthermore, the system is tested for the use of memory along with the increase of the concurrency, which determines the nature of the physical resources demanded by the services. Figure 9, shows the maximum memory used by each service with the increase of the concurrency. The memory footprint of DIALIGN is not affected much by the concurrency.

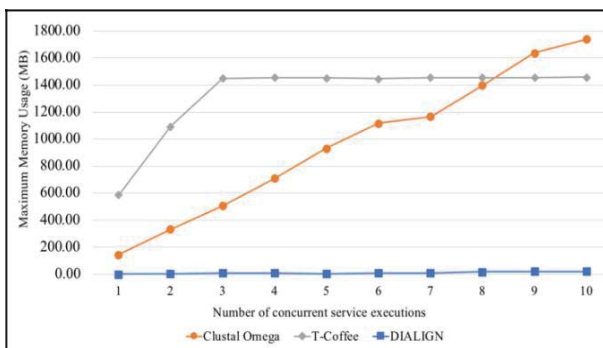


**Figure 8. Times taken by each MSA service vs the number of CSE.**

However, Clustal Omega has linearly increasing memory usage with concurrency depicting a positive correlation. According to the results, the memory usage for T-Coffee has plateaued after a concurrency of 3 CSE. Although the memory demand has increased along with the growth of the CSE, the memory requirement for DIALIGN is not high, which agrees with the observed times. Thus, a maximum concurrency of 10 service executions can be allowed for this service. In contrast, Clustal Omega has increased its memory usage more than 10 times as the concurrency has increased. Thus, a concurrency level of 1 to 5 service executions can be safely allowed and limited the maximum memory under 1GB. However, considering the time, an optimum concurrency of 1 to 4 service executions can be considered as ideal. Similarly, for T-Coffee, an ideal

concurrency level of 1 to 2 service executions can be agreed, given the maximum memory, demand lies around 1GB.

According to the results, for an optimal service operation, Clustal Omega should be limited to a concurrency of 4 service executions, T-Coffee to 2 service executions and DIALIGN with a maximum concurrency of 10 service executions. Here, all the services can be expected to run, each consuming a maximum memory of 1GB and provide results in less than 5 seconds of waiting time, which is reasonable for an HTTP request to complete in an analytical use case. The evaluation is based on a single operating environment without using containers, which may result in a slightly lower performance due to the virtualization overheads of the Docker platform.



**Figure 9. Maximum memory utilized vs level of concurrency (CSE).**

We have compared the existing bioinformatics analysis platform architectures and the proposed microservices solution as shown in Table 2. Since the microservices architecture has become popular in addressing many issues and bottlenecks in the cloud computing domain, this can be used in bioinformatics domain. Further, the microservices based solution outstands among other standalone and web-based solutions in terms of fault tolerance, scalability, real-time processing and technological flexibility [4].

## 6. Conclusion and future work

The microservices architecture has become one of the most trending architectures for cloud computing and web services. This paper has used microservices based architectural pattern for algorithm execution in bioinformatics platforms. The adopted architecture of the proposed solution allows the integration of heterogeneous techniques and provides a hybrid platform with both GPU and CPU computing units. This enabled a language-agnostic means of building a robust bioinformatics platform for computations with the integration of

different technologies while using message passing for communication.

This work can be extended by integrating more services and evaluating their performance in different environments with the use of cloud computing services. Furthermore, we intend to integrate this solution architecture for our proposed work on workflow modelling tool to execute bioinformatics workflows enabling the seamless integration of function in the form of independent services.

## 7. References

- [1] F. Andry, N. Dimitrova, A. Mankovic, V. Agrawal, A. Bder and A. David, "PAPAY: A Highly Scalable Cloud-based Framework for Genomic Processing," in *Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2016)*, Rome, Italy, 2016, pp. 198-206.
- [2] C. Richardson, "API gateway pattern," [Online]. Available: <http://microservices.io/patterns/apigateway.html> [Accessed 25 February 2018].
- [3] C. L. Williams, J. C. Sica, R. T. Killen and U. G. J. Balis, "The growing need for microservices in bioinformatics," *Journal of Pathology Informatics*, vol. 7, pp. 45-48, 2016.
- [4] D. Shadija, M. Rezai and R. Hill, "Microservices: Granularity vs. Performance," in *Proceedings of the 10th International Conference on Utility and Cloud Computing*, Austin, Texas, USA, 2017.
- [5] J. Thönes, "Microservices," *IEEE Software*, vol. 32, no. 1, 2015, pp. 113 – 116.
- [6] N. Dragoni, S. Giallorenzo, A. L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin and L. Safina, "Microservices: Yesterday, Today, and Tomorrow," in *Present and Ulterior Software Engineering*, Cham, Springer, 2017.
- [7] "National Center for Biotechnology Information," National Center for Biotechnology Information, U.S. National Library of Medicine, [Online]. Available: <https://www.ncbi.nlm.nih.gov>. [Accessed 4 January 2018].
- [8] T. Oinn et al., "Taverna: a tool for the composition and enactment of bioinformatics workflows," *Bioinformatics*, vol. 20, no. 17, 2004, pp. 3045–3054.
- [9] J. Goecks, A. Nekrutenko, J. Taylor and The Galaxy Team, "Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences," *Genome Biology*, vol. 11, no. 8, p. R86, 2010.
- [10] B. Familiar, *Microservices, IoT and Azure: Leveraging DevOps and Microservice Architecture to deliver SaaS Solutions*, New York, United States: Apress, 2015.
- [11] "What is a Container | Docker," Docker, [Online]. Available: <https://www.docker.com/what-container>. [Accessed 22 March 2017].

- [12] D. Merkel, "Docker: lightweight Linux containers for consistent development and deployment," *Linux Journal*, no. 239, p. Article No. 2, 2014.
- [13] N. Swainston, A. Currin, L. Green, R. Breitling, P. J. Day and D. B. Kell, "CodonGenie: optimized ambiguous codon design tools," *PeerJ Computer Science*, vol. 3, no. e120, 2017.
- [14] M. M. Nowotka, A. Gaulton, D. Mendez, A. P. Bento, A. Hersey and A. Leach, "Using ChEMBL web services for building applications and data processing workflows relevant to drug discovery," *Expert Opinion on Drug Discovery*, vol. 12, no. 8, 2017, pp. 757-767.
- [15] P. E. Khoonsari et al., "Interoperable and scalable metabolomics data analysis with microservices," *bioRxiv* 213603, 2017.
- [16] "Kubernetes | Production-Grade Container Orchestration," [Online]. Available: <https://kubernetes.io>. [Accessed 25 February 2018].
- [17] P. D. Dobson, P. Mendes, D. B. Kell and N. Swainston, "A Metabolic Reaction Balancing Web Service for Computational Systems Biology," *bioRxiv* 187328, 2017.
- [18] B. Fjukstad, V. Dumeaux, K. S. Olsen, E. Lund, M. Hallett and L. A. Bongo, "Building applications for interactive data exploration in systems biology," in *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*, Boston, Massachusetts, USA, 2017, pp. 556-561.
- [19] R. Hill, D. Shadija and M. Rezai, "Enabling Community Health Care with Microservices," in *Proceedings of the 16th IEEE International Conference on Ubiquitous Computing and Communications*, Guangzhou, China, 2017.
- [20] A. R. Colombo, T. J. Triche Jr and G. Ramsingh, "Arkas: Rapid reproducible RNAseq analysis," *F1000Research*, vol. 6, no. 586, 2017.
- [21] "Welcome - BaseSpace Sequence Hub," Illumina, [Online]. Available: <https://basespace.illumina.com/home/index>. [Accessed 16 January 2018].
- [22] A. Welivita, I. Perera, D. Meedeniya, A. Wickramarachchi and V. Mallawaarachchi, "Managing Complex Workflows in Bioinformatics - An Interactive Toolkit with GPU Acceleration," *IEEE Transactions on Nano Bioscience*, 2018, (to appear).
- [23] A. Welivita, I. Perera and D. Meedeniya, "An Interactive Workflow Generator to Support Bioinformatics Analysis through GPU Acceleration," in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine*, Kansas City, MO, USA, pp. 457-462, 2017.
- [24] "BLAST: Basic Local Alignment Search Tool," NCBI, [Online]. Available: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. [Accessed 6 January 2018].
- [25] "GPU-Blast | Sahindis," [Online]. Available: <http://archimedes.cheme.cmu.edu/?q=gpublast>. [Accessed 3 January 2018].
- [26] S. A. Manavski and G. Valle, "CUDA compatible GPU cards as efficient hardware accelerators for Smith-Waterman sequence alignment," *BMC Bioinformatics*, vol. 9, no. 2, 2008.
- [27] "Clustal Omega - fast, accurate, scalable multiple sequence alignment for proteins," [Online]. Available: <http://www.clustal.org/omega/>. [Accessed 12 January 2018].
- [28] "T-Coffee Home Page," [Online]. Available: <http://www.tcoffee.org/Projects/tcoffee/>. [Accessed 12 January 2018].
- [29] "BLAST FTP Site - BLAST Help - NCBI Bookshelf," National Center for Biotechnology Information, U.S. National Library of Medicine, [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK62345/>. [Accessed 4 January 2018].
- [30] L. Al Ait, Z. Yamak and B. Morgenstern, "DIALIGN at GOBICS - multiple sequence alignment using various sources of external information," *Nucleic Acids Research*, vol. 41, no. W1, p. W3-W7, 2013.

## Novel machine learning ensemble approach for landslide prediction

C. N. Madawala

Department of Computing and  
Information Systems,  
Sabaragamuwa University, Sri Lanka  
cnmadawala@std.appsc.sab.ac.lk

B. T. G. S. Kumara

Department of Computing and  
Information Systems,  
Sabaragamuwa University, Sri Lanka  
btgsk2000@gmail.com

L. Indrathilaka

National Research Building  
Organization, Sri Lanka.  
laksiriindrathilaka@gmail.com

### Abstract

*Haphazard development activities on mountain slopes and inadequate attention to construction aspects have led to the increase of landslides and consequently sustaining damage to lives and infrastructure. Nearly 3275 sq.km of area spread over the Ratnapura District, seems to be highly prone to landslides and mass wastage of 2178 sq.km. Landslides occur in many regions of Ratnapura district and nearly 90 deaths have been reported according to National Research Building Organization (NBRO) in 2017. Most landslides or potential failures could be predicted fairly accurately if proper investigations were performed in time. The primary objective of this study is landslide-hazard mapping and risk evaluation to determine the real extent, timing, and severity of landslide processes in Ratnapura district. Such knowledge will provide the most significant benefit to government officials, consulting engineering firms, and the general public in avoiding the landslide hazard or in mitigating the losses. Hybrid Machine Learning techniques can be used to develop prediction models using existing data. Ensemble approach based on Support Vector Machine (SVM), Naïve Bayes model were combined and implemented for the final prediction. This study possesses a strong capability to predict landslides by causative factors, slope, land use, elevation, geology, soil materials and triggering factor; rainfall was extracted and applied to the machine learning algorithms. This research introduces a novel architecture to produce a more relevant and accurate prediction of the landslide vulnerability within the study area. Moreover, it was revealed that all of the factors had relatively positive relationship with occurrence of landslides. An improvement in hazard monitoring, accuracy of early warning and disaster mitigation is documented.*

**Keywords:** Landslide, Support Vector Machine (SVM), Naïve Bayes, Hydrological, Rainfall

### 1. Introduction

Landslides are the geological incident which includes widely spread land movement resulting meticulous damages to the people and their belongings. Fundamentally the landslide transpired when a part of a natural slope is not capable of bearing its weight. The gravity is the fundamental driving force of the landslide refuse flow relying on the slope of the area. Landslide happens when the stability of the slope changes from a stable to an unstable state. In the last decades, there was a considerable increase in landslide frequency, in accord with the climatic changes, improper land uses and the expansion of urbanized areas in the world. Thus, landslide detection is a crucial requirement in pre and post-disaster, hazard analysis processes.

The recognition of landslide susceptibility is important to get some preventive and control actions and give some early warnings to reduce or mitigate hazards impacts. Most developed countries in the world apply the latest tools for landslide prediction. But Sri Lanka being an underdeveloped country cannot meet the expenses for such technologies. So, it is necessary to bring a research study for landslide predictions to give early warnings which lead save the lives of naive people.

Machine learning is a sophisticated fusion of applied mathematics and computational intelligence. It focuses on ‘training’ an algorithm to probe for and learn from data structure robust enough to make predictions; even without predecessor knowledge of the structure. In Machine Learning perspective, a majority of studies discussed aspect related to Artificial Neural Network (ANN), Support Vector Machine (SVM), Decision tree, Classifying and Clustering, Bayesian, WebGIS and General data mining techniques.

Even though these models have been applied successfully and efficiently in landslide susceptibility assessment, no model is perfect. Therefore, the improvement in these models is needed to achieve desirable results. The performance of landslide models can be enhanced by using feature selection and ensemble. Whereas, the Ensemble frameworks that combine multiple classifiers to improve the performance of

individual classifiers based on characteristics of the diversity [1].

Ensemble frameworks started in the 1990s but received significant attention of researchers in recent years. Ensemble techniques such as Bagging; AdaBoost; Naïve Bayes and SVM have been applied efficiently in an improvement of the performance of individual classifiers for different problems. Therefore, the main objective of present study is to attempt a novel classifier Ensemble data mining approach for landslide susceptibility assessment at Ratnapura District, Sri Lanka. This method is a combination of Naive Bayes classifier and SVM ensemble [2].

Recently, many algorithms were proposed for landslide prediction. Thought in the landslide studies, machine learning algorithms are more accurate than conventional statistical techniques. However, Ensemble methods are known as proper machine learning techniques in combining statistical methods for better landslide prediction.

The formation and occurrence of landslides is a complicated evolution process, which is caused by the interaction of multiple instability factors. However, most of the methods consider only the current value of the instability factors while ignoring the factors' evolution feature over time. This study proposes an Architecture related to Ensemble approach of Machine Learning algorithms, SVM, Naïve Bayes model. A variation of spatial data, including landslides, geology, topography, slope, soil, and land cover, and triggering factor rainfall were identified and collected in the study areas.

## 2. Research problem

The detection of natural disasters is a significant and non-trivial problem. A traditional method relies on dedicated physical sensors to detect specific disasters. With the advances in information communication technologies (ICT), and Machine Learning, it is critical to improving the efficiency and accuracy of disaster management systems through modern data processing techniques. They help the decision-makers understand near real-time possibilities during an event.

If the data related to a landslide can be collected and then analyzed using Machine Learning techniques, it may provide valuable insights into the disaster management. Furthermore, if we develop a prediction model which can be embedded in to develop a user friendly and efficient computer program which is used by any ordinary person who is living in a landslide-prone area to determine "am I safe in the current place with regards to current geological and weather condition or not?" by dealing with data of current situation rather than living blindly until National Building Research Organization (NBRO) issue disaster warnings [3].

In the past, disaster information extraction and prediction were mainly based on artificial visual interpretation. Apart from being time-consuming and strenuous, the traditional method also has a limitation in that the measurement process lacks accuracy and depends heavily on experts' experience [4]. With the development of the computer vision and pattern recognition technologies, it is possible to make the hazard assessment automatic. The other type of monitoring method is to embed different kinds of sensors related to slope, rainfall, water table level, and other factors into the landslide and sense the dynamic change of signals. Wireless sensor networks are therefore being used to achieve large-scale data collection and transmission.

During recent decades, a number of different methods were proposed for landslide modelling including heuristic, deterministic, and probabilistic methods [5]. But that methods have some limitations [2]. However, recent approaches to landslide modelling show that the prediction of landslide susceptibility could be enhanced with the use of hybrid machine learning techniques. Therefore, exploration of new hybrid machine learning methods for landslide susceptibility modelling should be further carried out.

The reason behind this research is to use Ensemble approach and propose a suitable landslide prediction Architecture. Meanwhile, identify, an expected location that would cause fatalities, damages or disruptions to existing standards of safety in Ratnapura District, Sri Lanka. Hence, the study helps to predict the most triggering factors of a landslide and changes can be expected in the activity of massive landslides in the future under the impact of environmental changes.

## 3. Literature review

When concerning the effect of hydrological characteristics and slope failures in different locations in the world, major landslide disasters triggered by rainfall are reported every year in different countries [6].

The objective of landslide-hazard mapping and risk evaluation is to determine the real extent, timing, and severity of landslide processes in selected high-priority areas of the Sri Lanka, where such knowledge will provide the most significant benefit to government officials, consulting engineering firms, and the general public in avoiding the landslide hazard or in mitigating the losses.

More specifically hydrological triggering is generally known as one of the major natural landslide beginning mechanisms. Hydrological triggering can be defined as a reduction in shear strength due to an enhance in pore-water pressure on the potential failure surface which eventually causes to the slope failure [7]. In Sri Lanka, landslides are mostly triggered due to massive and

continued rainfall. During the last few decades, landslides have occurred with increasing frequency and intensity. The intense rainfalls with shorter durations could trigger more destructive landslides[3].

Currently in Sri Lanka NBRO and many researchers (i.e. Karunanayake, K.B.A.A.M.; Wijayanayake, W.M.J.I. e tal) issue landslide early warning based on historical data collected using the contours, map of land use, the map of overburden and map of landslides[8]. The decision tree algorithm and the Neural Network technique will be used to develop prediction models out of the predictive analysis of Machine Learning techniques [9]. But the aforementioned literature review shows that SVM are widely accepted to be an effective and robust method for landslide modelling compared to the other mentioned methods and techniques[10]. Therefore, in this study, aimed at filling this gap by developing an SVM model and evaluating its performance for the prediction of landslides at the Sabaragamuwa province area of Sri Lanka.

In Machine Learning perspective, a majority of studies discussed aspect related to ANN, SVM, Decision Tree (DT), Classifying and Clustering, Naïve Bayesian, WebGIS and General data mining techniques from Machine learning.

During recent decades, some different methods were proposed for landslide modelling including heuristic, deterministic (engineering approach), and probabilistic (non-deterministic or data-driven) methods. Due to limitations of these methods, exploration of new hybrid machine learning methods for landslide susceptibility modelling should be further carried out.

In the present study, this gap in the literature is partially filled by proposing a new hybrid machine learning approach for landslide susceptibility modelling with a study of the Ratnapura District, Sri Lanka. The proposed approach relies on an integration of the Support Vector Machine and Naïve Bayes.

#### 4. Objectives

The natural hazards are beyond human control, but their destruction can be reduced if prediction mechanisms are carried out in advance. Researchers worldwide are having a great pace to develop early prediction mechanisms for such natural hazards. It was hard to use traditional mathematical methods for analyzing.

The paper aims in presenting an architecture using Ensemble approach relying on an integration of SVM and Naïve Bayes which possess a strong capability to predict landslides and analyze the climate variability (rainfall and temperature).

### 5. Data and methodology / materials and methods

#### 5.1. Study area

Nearly 3275 sq.km of area spread over the Ratnapura District, 1097 sq.km forestry areas and seems to be highly prone to land sliding of 2178 sq.km. 473 out of 575 GM divisional areas reported as Landslide-prone areas. A severe landslide occurred in Eheliyagoda, Nivithigala, Ayagama, kalawana, Dolapallehena, Kiribathgala, Alupathgala, Hortonwatta, Girapagama and approximately 14 of Landslide-prone and Embilipitiya, Godakawela, Kolonna 03 non-landslide-prone AGM divisional areas around Ratnapura District, Sabaragamuwa province in Sri Lanka[11].

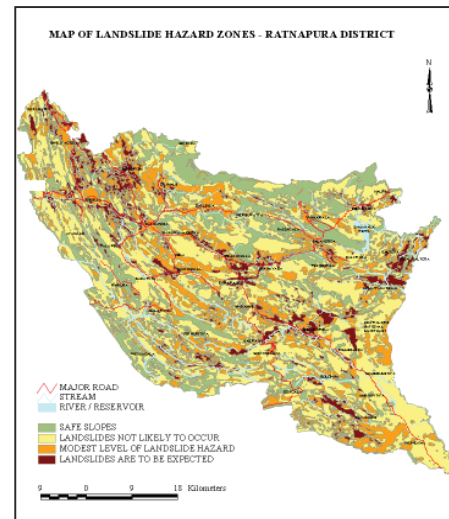


Figure 1. The study areas, the Ratnapura District, Sabaragamuwa Province.



Figure 2. (a) Kiribathgala, in Ratnapura district, Sri Lanka, Monday, May 29, 2017, (b) Ratnapura, Kalawana and Ayagama areas (c) Colombo Hatton main road due to rock falls

## 5.2. Materials

### 5.2.1. Required data

To determine the tributary characteristics of the study area a) Digital topographic maps of the study area (1:10000) and b) Digital elevation model (DEM) - satellite images of specific areas were obtained.

**Table 1. Data used and sources.**

Data use	Sources
Rainfall	Metrology Department
Soil Materials	Irrigation Department
Geology	Survey Department
Land Use	LUPP Department
Landslide & non Landslide	NBRO
Soil Texture	Survey Department
Distance from road, river	NBRO
Influence of constructions	NBRO

To analyze the climate variability (rainfall and temperature), monthly rainfall and temperature data in the past, recent five years (2012-2017)

To study the slope aspect, angle within the landslide are National Building Research Organization | Hazard Zonation Map was referred.

### 5.2.2. Software tools and models

- Arc GIS 10.4 version
- Digital Elevation Model (DEM) Classifier Ensemble model

### 5.2.3. Programming Language

- Java

## 5.3. Methodology of ensemble model

The fundamental data required for producing the landslide susceptibility model of the study area was obtained from Standard Topographic Maps at a scale of 1/50000. Six major contributory factors can be identified for landslide events. That is the bedrock geology with the degree of weathering and nature and intensity of defects, Slope angle, overstrain soil cover, Landform, drainage pattern and land use pattern. But the critical triggering reason for a landslide is the high-intensity rainfall.

### 5.3.1. Naïve Bayes Classifier

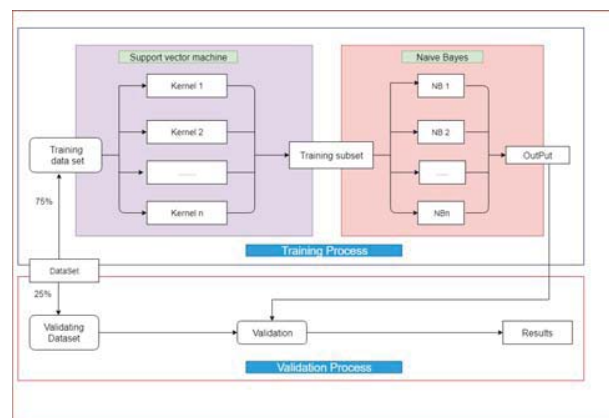
Naive Bayes classifier is one of the simplest soft computing methods which is based on the Bayesian theory and the maximum posterior hypothesis [12]. Naive Bayes classifier uses a statistical hypothesis that all values of numeric attributes are independent and normally distributed in each class. Naive Bayes classifier has been applied effectively in many fields such as medical diagnosis, and management. However, its application is still limited to landslide problems [13].

### 5.3.2. Support Vector Machine

Support Vector Machine is primary a classifier method that performs the classification tasks by constructing hyperplanes in a multi-dimensional space that separates cases of different class labels. SVM supports both regression method and classification technique tasks and can handle multiple continuous and categorical variables [14].

### 5.3.3. The novel classifier ensemble model

In this study, the novel ensemble classifier model was generated by the combination of Naive Bayes classifier and SVM ensemble. SVM ensemble was first applied to create the subsets of training[15]. After that, Naive Bayes classifier was used to construct base classifiers from these subsets for classification. Methodological flow chart of the novel classifier ensemble model is shown in Figure 3.



**Figure 3. Methodological flow chart of the novel classifier ensemble model.**

The advantage of the novel classifier ensemble model is that the training subsets are being optimized using SVM ensemble, and then these training subsets are utilized for training a base classifier of Naive Bayes. Therefore, the

novel classifier ensemble model could improve the predictive capability of a base classifier of Naïve Bayes.

#### 5.4. A methodology for reference architecture

Based on the classification results, proposed an Architecture which is meant to possess a strong capability to predict landslides by factors of landslide dataset using Machine Learning concept.

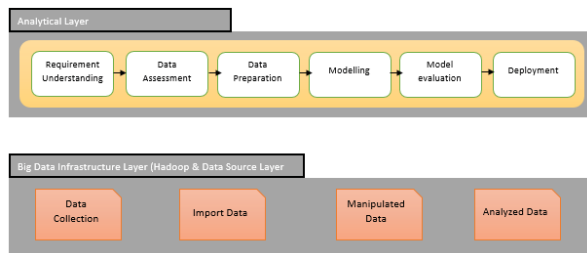


Figure 4. Reference architecture of prediction model.

### 6. Results and discussion

The natural disasters are beyond human control, but their destruction can be reduced if prediction mechanisms are carried out in advance. Researchers worldwide are having a great place to develop an application of the landslide prediction model an exploration always, and it includes many things to be researched further. The SVM apply, and the results use to produce landslide prediction model of the study areas.

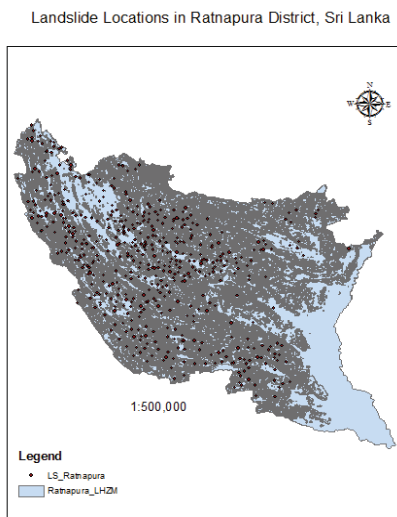


Figure 5. Landslide susceptibility map, Ratnapura District.

Furthermore, researcher aim to separate instances belongs to two classes: (i) areas where landslides occurred or likely to happen in the future; and (ii) areas where landslides did not occur or are not expected to occur in the future. The index was classified into a class based on an area for visual and straightforward interpretation of Landslides are most likely to occur. The proposed architecture is meant to possess a strong capability to predict landslides by factors of landslide dataset using Machine Learning concept.

#### 6.1. Model performance and validation

The performance capability of the novel classifier ensemble model has evaluating with three main Ensemble Learning techniques.

##### 6.1.1. Ensemble learning techniques

###### Bagging

Bagging uses to implement similar learners on small sample populations and then takes a mean of all the predictions.

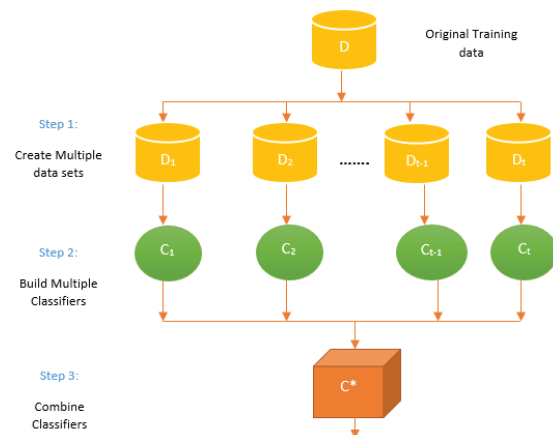


Figure 6. Bagging process, to reduce variance error.

###### Boosting

This technique uses to adjust the weight of an observation based on the last combined classification. If an observation was classified incorrectly, it tries to increase the weight of these observations.

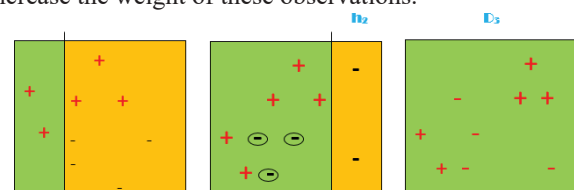
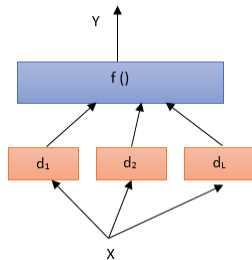


Figure 7. Boosting process, to decrease the bias error.



## Stacking

We use a learner to combine output from different learners. This leads to a decrease in either bias or variance error depending on the combined learner we use.



**Figure 8. Stacking process, to decrease in either bias or variance error depending on the combining learner.**

Ensemble approach produces more accurate solutions than a single model would. When we are trying to predict the target variable using any machine learning Ensemble technique, the leading causes of difference in actual and predicted values are noise, variance, and bias. Ensemble approach has helped to reduce these factors, except noise, which is an irreducible error.

A landslide prediction model should be able to make an adequate prediction of possible landslide areas to improve the hazard monitoring, accuracy of early warning and disaster mitigation by collecting data of the triggering factors of the disaster. Data collected for this research may provide a suitable database for hazard forecasting and future studies.

## 7. Conclusion

Landslide susceptibility assessment has been done in the Ratnapura District, Sri Lanka using the novel ensemble classifier model which is a combination of Naïve Bayes classifier and SVM. Naïve Bayes is an effective classifier. Analysis results show the novel classifier ensemble model has the best degree of fit to landslide susceptibility assessment compared to other models. This study identified factors that may be involved in landslides, and the results and methods that can be used for landslide predicting in other regions beyond the study areas. Landslide prediction models can help implement a guide for planning the mass evacuation of residents in the case of a landslide, and also to prevent or reduce the disruptive impacts of a natural disaster on surrounding communities.

Therefore, the present study proposed an Architecture for predicting landslide occurrences in the areas susceptible to these phenomena in the Ratnapura

District, where landslides are expected to continue to strike in the future.

Moreover, soon, a landslide warning system may be established by forecasting landslides induced by rainfall. And as a recommendation conducting programs in hilly or mountainous areas about the control of run-off and erosion through the water and soil conservation can have extremely beneficial effects to control the hazards.

## 8. Acknowledgement

The access to Hazard Zonation map of Ratnapura District from the Sri Lanka National Research Building Organization (NBRO), Director General, Department of Metrology, Dr. Mr. R. M. S. Bandara has been acknowledged, Mr. A. L. K. Wijemanna, Director (Computer, Research, Climate change, International Affairs) was supported to collect Rainfall Data and also Director General, Department of Land Use & Policy Planning, Mrs. A. S. Illangange helped to collect Land Use data and my sincere gratitude to the Sabaragamuwa University of Sri Lanka for the encouragement.

## 9. References

- [1] Pham, B. T., Bui, D.T., Dholakia, M.B., Prakash, I., Pham, H.V., Mehmood, K. and Le, H.Q., "A novel ensemble classifier of rotation forest and Naïve Bayes for landslide susceptibility assessment at the Luc Yen district, Yen Bai Province (Viet Nam) using GIS," *Geomatics, Nat. Hazards Risk*, vol. 8, no. 2, pp. 649–671, 2017.
- [2] Pham, B. T., Shirzadi, A., Tien Bui, D., Prakash, I. and Dholakia, M. B., "A hybrid machine learning ensemble approach based on a Radial Basis Function neural network and Rotation Forest for landslide susceptibility modeling: A case study in the Himalayan area, India," *Int. J. Sediment Res.*, vol. 33, no. 2, pp. 157–170, 2018.
- [3] Government of Sri Lanka, World Bank, UN Sri Lanka, and Global Facility for Disaster Reduction and Recovery, Sri Lanka Rapid Post Disaster Needs Assessment: Floods and Landslides, May 2017, no. May. 2017.
- [4] Kadavi, P. R., Lee, C. W., and S. Lee, "Application of ensemble-based machine learning models to landslide susceptibility mapping," *Remote Sens.*, vol. 10, no. 8, pp. 1–18, 2018.
- [5] Gariano, S. L. and Guzzetti, F., "Landslides in a changing climate," *Earth-Science Rev.*, vol. 162, no. August 2016, pp. 227–252, 2016.
- [6] Pham, B. T. and Prakash, I., "A novel hybrid model of Bagging-based Naïve Bayes Trees for landslide susceptibility assessment," *Bull. Eng. Geol. Environ.*, pp. 1–15, 2017.
- [7] Terlien, M. T. J., "The determination of statistical and deterministic hydrological landslide-triggering thresholds," *Environ. Geol.*, vol. 35, no. 2–3, pp. 124–130, 1998.

- [8] “Predicting landslides in hill country of Sri Lanka using data mining techniques,” pp. 2016, 2016.
- [9] Lee, S., Lee, M.-J. and Jung, H.-S., “Data Mining Approaches for Landslide Susceptibility Mapping in Umyeonsan, Seoul, South Korea,” *Appl. Sci.*, vol. 7, no. 7, p. 683, 2017.
- [10] International Consortium on Landslides; University of Ljubljana; Geological Survey of Slovenia, Landslide Research and Risk Reduction for Advancing Culture of Living with Natural Hazards, *Local Proceedings with Programme*, no. June. 2017.
- [11] Baba, H., “Introductory study on Disaster Risk Assessment and Area Business Continuity Planning in industry agglomerated areas in the ASEAN,” *J. Integr. Disaster Risk Management*, vol. 3, no. 2, pp. 184–195, 2013.
- [12] Hellerstein, J. L., Jayram, T. S. and Rish, I., “Recognizing End-User Transactions in Performance Management,” New York, vol. 1, 2000.
- [13] Zhou, C., Yin, K., Cao, Y. and Ahmed, B., “Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China,” *Eng. Geol.*, vol. 204, no. November 2017, pp. 108–120, 2016.
- [14] Korup, O. and Stolle, A., “Landslide prediction from machine learning,” *Geol. Today*, vol. 30, no. 1, pp. 26–33, 2014.
- [15] Goetz, J. N., Brenning, A., Petschko, H. and Leopold, P., “Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling,” *Comput. Geosci*, vol. 81, pp. 1–11, 2015.

## Communication framework for vehicular ad-hoc networks using Blockchain: A case study of Metro Manila Electric Shuttle automation project

A. S. Kulathunge

Department of Statistics and Computer Science  
University of Peradeniya, Sri Lanka  
askkulathunge@gmail.com

H. R. O. E. Dayarathna

Department of Statistics and Computer Science  
University of Peradeniya, Sri Lanka  
erunikad@pdn.ac.lk

### Abstract

*Vehicular Networks or the Vehicular Ad-hoc Networks (VANET) are experiencing revolutionary growth in research and industry, but still suffers from various breaches such as poor decentralization of communication, security vulnerabilities, scalability and trust issues in its communication as well as in its design. Major issues identified in VANET, are trust, data accuracy, and reliability of communication in the communication channel. Blockchain technology is a technology adopted by crypto currency, namely Bitcoin, which is recently used to build trust and reliability in peer-to-peer networks. This study proposes a communication framework for VANET exploring capabilities of blockchaining. Metro Manila Electric Shuttle Automation Project is used as a case study to verify the communication framework. It fulfills Vehicle-to-Vehicle (V2V) as well as Vehicle-to-Infrastructure (V2I) communication requirements of the considered project. It includes an Intelligent Toll Payment (ITP) system (V2I communication) and automated vehicular following, known as goose tracking (V2V and V2I communication), and those covering main communication requirements. Further, a simulator named SimulatorZ is implemented to model Goose Tracking which should support multi-vehicle simulation to understand data requirements of master and slave vehicles and the timing of communication. Communication framework provides trustworthiness for vehicles behavior, cashless secure transaction between two untrusted party as well as rewards and penalties for vehicles' actions. Data communication in goose tracking is done with 10ms latency between two vehicles and  $1-10^{-5}$  reliability. Slave vehicles' movements depends on master vehicle's speed, position, angle and the timestamp.*

**Keywords:** BlockChain, Communication framework, Metro Manila, VANET

### 1. Introduction

One of the most crucial issues faced by many metro cities in the world due to the dense urbanization, is automobile traffic. One of the main reasons for this is inefficient infrastructure, which might take years to resolve. However, utilization of existing infrastructure more efficiently may reduce the traffic condition to a great extent. This may require highly efficient real-time communication and understanding between vehicles as well as infrastructure. Inspired by this thought, Vehicular Ad-hoc network (VANET) [1] was introduced. But, it has various data communication issues such as security vulnerabilities, scalability issues, trust issues, etc. These are extremely hard to be solved due to the nature of participants, namely privately owned vehicles.



**Figure 1. Typical traffic condition of Metro Manila**  
Source: Google Maps [2]

The project of concern, i.e. Metro Manila Electric Shuttle Automation Project, focuses on traffic as well as environment pollution within the Metro Manila area. Manila was identified to have the worst traffic on earth in year 2015 by Waze [3]. Figure 1 shows the typical traffic condition of Metro Manila during the day time.

The mentioned project tries to introduce an Electric Vehicle (EV) as a shuttle service to manage the transportation within Metro Manila. It is mainly expected to replace “Jeepney” (Figure 2) which is the most common public transportation means adopted in Manila. An eight-passenger prototype of the proposed EV (Figure 3) has

been already tested and the mass production is expected to commence in January, 2019.



**Figure 2. Jeepney - Makati, Philippines.**



**Figure 3. Prototype of the EV - Intramuros, Metro Manila Philippines.**

The EVs are expected to be operated by licensed drivers, though journey details, any payments, rewards, penalties, etc. are attributed to the driver. However, this EV is expected to be partially automated. During rush hours, due to the high demand, multiple EVs are logically joined together (not physically), thus, the master EV is expected to command the slave EVs to allow them to follow the master, known as goose tracking.

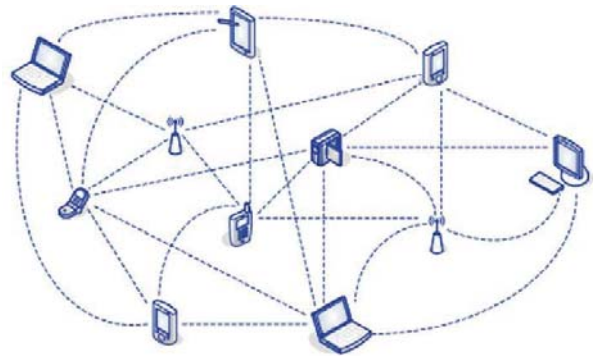
This study focuses on addressing the above requirements by establishing a reliable and secured communication framework which enables both V2V and V2I communication between the EVs. Further, as existing vehicle simulators do not provide multiple vehicle simulation for scenarios like goose tracking, a simulator is implemented to identify timing for message passing, parameters that should be passed and the required calculations at each peer.

As outcomes, a vehicle simulator for multiple vehicles is built with Java technology and a Blockchain-based data communication framework is designed for the VANET of Metro Manila Electric Shuttle Project. It is based on a Peer-to-Peer (P2P) Blockchain network with clients and dashboards. P2P Blockchain network is built using distributed ledger and Spring REST is used for the

communication. Finally, a tolling application is implemented and used to test the applicability of the framework.

## 2. Literature review

Vehicular ad-hoc network (VANET) and vehicular automation have been in the focus of research for quite a long time. There are various studies in the literature considering blockchain for Intelligent Transportation Systems (ITS). Most systems focus on Mobile Ad-hoc Blockchain (MONET) [4] on demand, which is formed by small, localized, groups that are involved in an activity common to the group. Usually, the participants are mobile nodes. No trust is required between nodes where each node is identified by its address. Its functionality is briefed using Figure 4.



**Figure 4. Mobile ad-hoc network architecture.**

*Source: Research Gate*

Temporary disconnection of node groups from the overall vehicular ad-hoc network is supported where it allows them to enact and transact as local groups while being disconnected.

Even though the above mentioned approaches exist in literature, they are yet to address issues such as, decentralization of communication, security vulnerabilities, scalability and trust issues in its communication.

Blockchain is the underlying technology of the Bitcoin protocol [6] that emerged in 2008. It is a distributed public ledger and has a consensus mechanism based on a proof of work algorithm (Figure 5). These significant features of Blockchain make it potential for constructing the desirable trust model in VANETs. All the broadcast messages and actions of vehicles will be written into the immutable and unforgettable record, which can be verified and audited by every entity in the network. However, the transparency of Blockchain means privacy is not considered naturally.

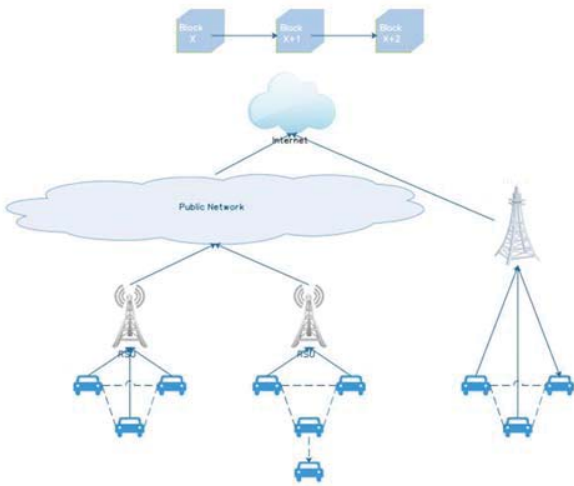


Figure 5. Mobile ad-hoc Blockchain.

By reviewing the ledger, the actions made with any public key is traceable to a real identity.

### 3. Methodology

Considering the integration of blockchain for both V2V and V2I communication in the target VANET environment, the following method is proposed.

This research work has three phases (Figure 6). First, implementation and development of the simulator: SimulatorZ for Automated Vehicular Following (AVF) known as goose tracking. This phase allows the simulation of EV's behavior and identifies data necessities while semi-automated vehicular following.

Then, the Blockchain-based data communication framework for VANET is developed. Finally, implementation and testing of framework applications using the simulator and the blockchain-based data communication framework takes place. Final phase verifies the research work.



Figure 6. Overview of the research.

All communication takes place through Dedicated Short Range Communication (DSRC) [7] radio where EVs exchange messages with nearby EVs using V2V paradigm and communicate directly with roadside units (RSUs) using V2I communication paradigm.

### 3.1 Construction of the simulator: SimulatorZ

The simulator was created by Java Swing Graphics libraries with a Physics engine. SimulatorZ creates car objects and can control them using keyboard inputs. Car has two vector parameters, *Position* parameter is used to get the position of the car and *Velocity* is used to get the velocity of the absolute vehicle. Both position and velocity has two values through the x-axis and y-axis. Also a car has angle, angular velocity, steer angle, throttle, brake, front slip, rear slip and car number, parameters. Each vector is defined by,

$$\text{Vector}(A) = \{x\text{-axis value for } A, y\text{-axis value for } A\}$$

Table 1. Parameters of car objective.

Parameter	Definition
Position	Position of car
Velocity	Velocity vector
Angle	Angle of car body orientation
Angular velocity	Angular velocity of the car
Steer angle	Angle of steering/wheels
Throttle	Amount of throttle
Brake	Amount of braking
Car number	A unique number for car

### 3.2 Establishment of the communication framework: Block-Net

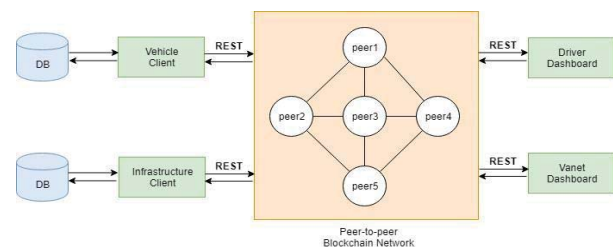
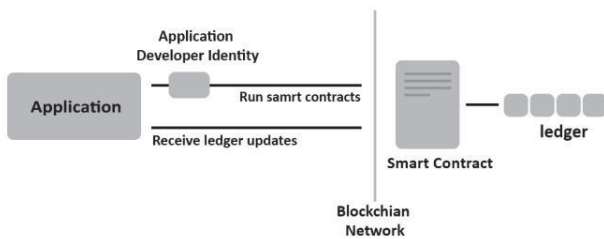


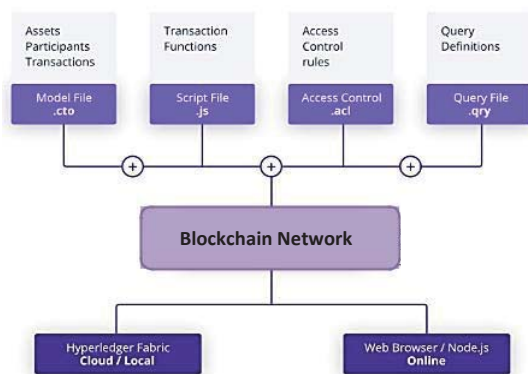
Figure 7: Framework architecture.

Framework has a peer-to-peer Blockchain network that is built using Hyperledger Fabric and it has vehicle clients, infrastructure clients, VANET-dashboard and driver-dashboard that are communicated using REST APIs (Figure 7).

Network is built using the Hyperledger Fabric network Configurations [8]. It establishes a P2P network with distributed ledger. Also, it configures the smart contracts and update protocols by default.



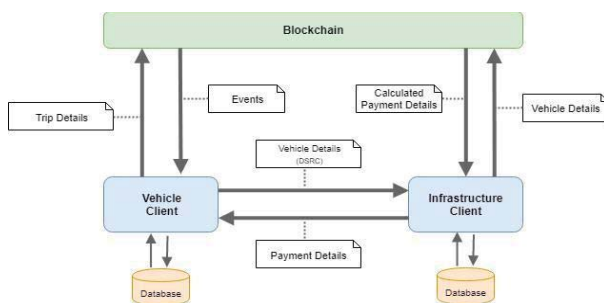
**Figure 8. Development environment of peer-to-peer Blockchain network.**



**Figure 9. Hyperledger Blockchain network.**

### 3.3 Implementation of applications

Two applications are considered to verify the applicability of the simulator and the framework, which are based on V2I and V2V paradigms. Goose tracking mainly utilizes V2V communication paradigm whereas Tolling mainly utilizes V2I communication paradigm. These two are selected for the purpose of validation of the framework.

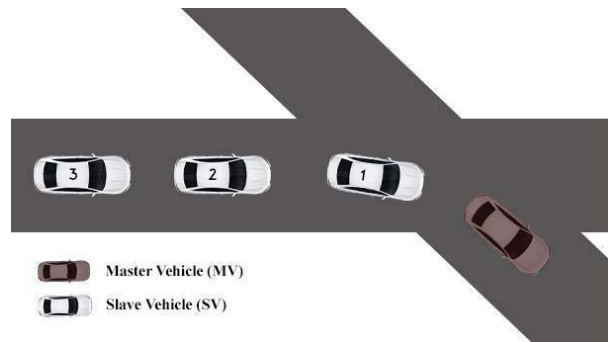


**Figure 10. Modeling assets, access control and transactions on the framework.**

In general, the client is implemented using Spring RESTful web services as REST client. Clients use Historian Client API. It records the history of actions taken in Blockchain using Composer. It is a registry that stores *HistorianRecords* and each record is created in response to a transaction being executed as depicted in Figure 8.

Vehicle client communicates with both Blockchain network and infrastructure clients. It transfers vehicle details to the Infrastructure client. It transfers driver ID, vehicle type, timestamp parameters. In fact it transfers all data that is needed in goose tracking to the Blockchain. Infrastructure client transfers the details provided by the vehicle client and the infrastructure details to the Blockchain. It transfers calculated and analyzed data that comes from Blockchain to the vehicle clients as illustrated in Figure 9. Blockchain records all the transactions of vehicle clients as well as infrastructure clients and transfers analyzed and calculated data and published events using *AssetRegistry*, *IdentityRegistry*, *ParticipantRegistry*, *Query*, *TransactionRegistry* APIs in Hyperledger Composer (Figure 10).

#### 3.2.1 Goose tracking



**Figure 11. Goose tracking.**

In Goose Tracking, it considers all EVs as time synchronized vehicles that belong to one authority. Simulator can add multiple vehicles by creating car objects of same type. But, the user needs to control the Master Vehicle (MV), which is currently chosen manually by the car agent and it will be the master for the shuttle round. Other vehicles that are known as Slave Vehicles (SV) are simulated according to the movements of the MV. MV can be changed at any instance and roles can be interchanged and any other SV can be a MV with the traffic situation (Figure 11). SV's will choose to join the MV as they arrive at the MV location, known as selective model. However, in later stages, full autonomous decision making is expected to be implemented in selecting MV and SVs, which is out of scope for this study. In goose tracking, MV needs to pass its current data to other SVs that are connected with MV, such that they follow the MV. MV has the data structure given by Table 2. Main purpose is to

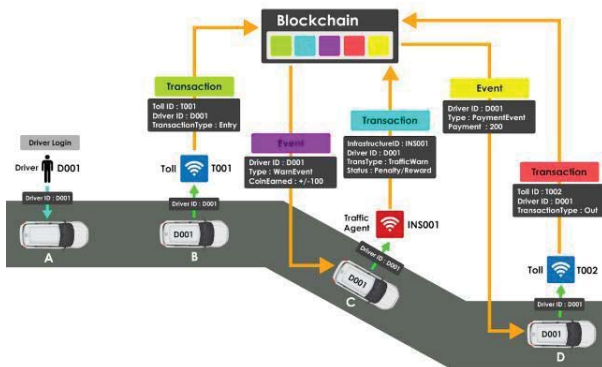
identify data transfer instances and the communication model in order to control the SVs efficiently.

**Table 2. Parameters of the communication data structure.**

Parameter	Definition
Speed	Current speed of MV
Angle	Current angle of MV
X-axis position	Position of the MV (x-axis)
Y-axis position	Position of the MV (y-axis)
Time stamp	Current time of MV

### 3.2.1 Toll application

In Toll Payment System, the vehicle client first needs to log into the vehicle with his/her driver identification number as EVs are distinguished by the driver. A vehicle client participant is created using the Hyperledger Composer and that is submitted to the Blockchain immediately. Similarly, infrastructure clients can be added to the network. When a driver passes an infrastructure client, vehicle transfer its driver ID and a time stamp to the infrastructure client through a Dedicated Short Range Communication (DSRC) protocol. The infrastructure client then makes a transaction using the driver details and infrastructure ID and sends it to the blockchain. Then, the blockchain accepts that record and stores it in its distributed ledger. Proposed Toll Payment System has several types of infrastructure clients like Entry/Exit Tolls and Traffic agents (Figure 12).



**Figure 12. Toll application flow overview.**

## 4. Results and discussion

Data communication in goose tracking is achieved with 10ms latency between two vehicles and  $1-10^{-5}$  reliability. SVs follow MV as depicted in Fig.12 based on MVs speed, position, angle and the current timestamp. The SimulatorZ supports multiple vehicles at the same time.



**Figure 13. SimulatorZ.**

Performance of framework is tested using the SimulatorZ and the implemented dashboard (Figures 13 – 15).

Date, Time	Entry Type	Participant
2018-10-10, 21:47:...	Toll Entry	driver001
2018-10-10, 08:14:...	AddParticipant	driver001

**Figure 14. VANET dashboard.**

From	To	Amount	Description
Anuradha	Toll	200	Toll Payment
Anuradha	Toll	200	Toll Payment
Anuradha	Toll	200	Toll Payment
Traffic Agent	Anuradha	20	Reward
Anuradha	Toll	200	Toll Payment
Anuradha	Toll	200	Toll Payment
Anuradha	Toll	200	Toll Payment
Traffic Agent	Anuradha	20	Reward
Anuradha	Toll	200	Toll Payment

**Figure 15. Driver dashboard.**

## 5. Conclusions

In this project, the main outcome is the implementation of a communication medium to achieve automated vehicular following known as goose tracking of EVs in the Metro Manila electric shuttle automation project. This work simulated the vehicular following scenario introducing a simulator known as SimulatorZ that supports multiple vehicles. It is used to identify the necessary information to be passed along with the timing of data passing. A standard optimal data structure is identified to be passed between stations and proposed a Blockchain based data communication to use in V2V data communication. Further, V2I communication model is also implemented and tested by considering tolling system. Data communicate is done with 10ms latency and 1-10-5 reliability in both V2V and V2I scenarios.

The support for goose tracking and tolling is implemented and tested in this project scope, however, it can be extended for any other application which needs V2V and V2I data communication. The framework guarantees decentralization, extended security, scalability and privacy of the data communication. It, also introduced penalty/reward and cashless secure transactions between two untrusted parties. This can be used for many application areas in vehicle transportation and can be adopted in any metro vehicular project thus, not limited to the Metro Manila Electric shuttle project. Further, all the data collected during the transportation are stored in the Blockchain. So, the stored Blockchain data can be used for further analysis and can support in Artificial Intelligent (AI) module training. AI modules can be used for trip data analysis, traffic prediction, autonomous car training, etc. As future work, SimulatorZ is expected to be enhanced as a three dimensional (3D) simulator and support all conditions in real world environment. Data communication is also expected to be enhanced by adopting faster communication technologies.

## 6. Acknowledgement

Authors would like to acknowledge AssemblePoint (Japan) Co. Ltd, specially, the CEO Shozo Kurihara, design and engineering teams of AssemblePoint (Japan) Co, and SBX project team in Makati, Intramuros, Philippines for the opportunity and the support.

## 7. References

- [1] Qian, Y., Lu, K., and Moayeri, N. (2008). A Secure VANET MAC Protocol for DSRC Applications.
- [2] Google Maps (<https://www.google.com/maps/>) [Accessed November 2018]
- [3] WAZE (<https://www.waze.com/>) [Accessed November 2018]
- [4] Arrivets, M. (2018). MONET: Mobile Ad Hoc
- [5] Blockchains. [online] Available at: <https://www.mosaicnetworks.io/> [Accessed 06 July 2018].
- [6] Benjamin Leiding and William V. Vorobev (2018). Tezos-based Vehicular Ad Hoc Blockchains
- [7] Wikipedia, (2018). Bitcoin. [online] Available at: Bitcoin: <https://en.wikipedia.org/wiki/Bitcoin> [Accessed 20 Apr. 2018].
- [8] DSRC [online] Available at: [http://www.arib.or.jp/english/html/overview/doc/5-STD-T75v1\\_0-E2.pdf](http://www.arib.or.jp/english/html/overview/doc/5-STD-T75v1_0-E2.pdf) [Accessed 15 May 2018]
- [9] Hyperledger fabric, (2018). Hyperledger fabric configuration. [online] Available at: [https://hyperledger-fabric.readthedocs.io/en/latest/write\\_first\\_app.html](https://hyperledger-fabric.readthedocs.io/en/latest/write_first_app.html) [Accessed 14 Mar. 2018].
- [10] [9] E-Trike- Intramuros [online] Available at: <https://www.clemenisle-ev.com/>



## A novel computational method to capture FPGA technology trends from Patent information

Masashi Shibata  
*Graduate School of Sciences and  
 Technology for Innovation  
 Yamaguchi University  
 Yamaguchi, Japan  
 g501wc@yamaguchi-u.ac.jp*

Masakazu Takahashi  
*Graduate School of Innovation and  
 Technology Management  
 Yamaguchi University  
 Yamaguchi, Japan  
 masakazu@yamaguchi-u.ac.jp*

### Abstract

*This paper provides a novel trend analysis of FPGA development with Machine Learning. Recently, demands for the computing power are expanding due to reform of industrial structure such as the Industry 4.0 and the explosive expansion of AI. In this paper, we reveal the technical development trend of the leading FPGA companies from the patent information with Machine Learning. We focus on the classification codes in the patent and employ Link Mining method as the analytical method. Link Mining is a conventional method to analyze the structural features of things. It simplifies the objects and the relations as the nodes and the edges. With the proposed method, we succeed in revealing the companies' focused technology fields, the transition of focusing areas, and their differences and common points from the results of extracting the graphs' features.*

**Keywords:** *Field Programmable Gate Array, Machine Learning, System-On-a-Chip*

### 1. Introduction

The demand for implementing various kinds of sensors, collecting a large amount of data from them, transmitting the collected data, and processing the data in real time is rapidly increasing. It is caused by the computerization of manufacturing industry represented by the Industry 4.0, popularization of the mobile devices such as smartphones, installation of communication functions to various devices by IoT, and machine learning methods' explosive high performance and adoption in various fields. For satisfying these requirements, data processing performance is improved on both the edge side and the cloud side.

Depending on the applications, various types of sensors are implemented in edge devices. In addition to them, edge devices often have various modules such as communication modules; display modules depending on the purposes. Furthermore, edge devices need to process user application, power control, and so forth. Thus,

regarding HW (hardware), edge devices need high-performance processors, and regarding SW (software), implementing OS (Operating System) is required.

Edge devices have various installation locations and usages. Sometimes they are installed in severe environments such as underwater and motor-vehicle engine room. Thus, they often require severe restrictions on shapes, heat generation amount, and power consumption. In order to achieve both required performance and constraint conditions, it is common to use the system LSI called SoC (System-On-a-Chip). It satisfies reducing power consumption, improving reliability, reducing footprint by combining necessary functions into one chip.

A large amount of data is transmitted from a large number of edge devices connected to the cloud to servers installed on the cloud side. The servers are required to have the ability to process the data to be transmitted in real time and reliability that processing does not halt. Also, in many of the main applications running on the server in recent years, such as image processing and machine learning require a large amount of arithmetic processing. Therefore, in addition to high-performance CPUs, it is common to use high-performance GPUs, and large-capacity memories by connecting them using broadband interconnect to perform parallel calculations.

A plurality of such systems is prepared and clustered to ensure reliability. Regarding SW, it is an important task to utilize HW assets efficiently using languages that can perform parallel processing and compilers that can efficiently distribute the processes to computing units. In general, data centers for clouds are often large-scale facilities. Thus, high energy efficiency is also an essential factor.

The SoC which is the core of the edge device is often realized as the LSI. However, it is difficult to apply to the edge device of small production because the manufacturing period is long, and the manufacturing cost is high. Also, servers on the cloud side mainly use x86 series CPUs and NVIDIA GPUs, but there are drawbacks such as lack of flexibility to support new protocols.

As a means to solve these drawbacks, adoption of an FPGA (Field Programmable Gate Array) is increasing. FPGA is a reconfigurable integrated circuit. FPGA's chip unit price is higher than ASIC, but it does not need the manufacturing process. Therefore the development period is significantly shorter. Also, since it can be obtained from a small amount, it is suitable for products of small amount and large varieties. Also, it is possible to implement protocols designed by users, and there is no limit on the degree of parallelism. Thus, it is appropriate for using as CPU accelerators and implementing user protocols.

Altera (acquired by Intel in 2016) and Xilinx are the leading suppliers of FPGA. In the past, FPGAs were products only packed with logic circuits, but in recent years, they have released products that can be used as SoCs with ARM's CPU installed. In the cloud, it has already been adopted by Microsoft Azure and Amazon AWS. Intel has a series of chipsets which consist of CPU and FPGA for the server and reference design for fog computing. Also in edge devices, it is adopted in various products.

From the above, it is considered that FPGA plays an essential role as a key device in the computing domain. Therefore, capturing the technology trends of FPGA suppliers is important for conducting computing-related projects.

In this paper, we reveal the technological trends of Altera and Xilinx by analyzing patent information filed by the companies. Knowledge extraction from the patent information has been done so far, but existing patent analysis methods rely on individual heuristic knowledge. It makes it difficult to extract the technological structure.

The link mining method is used as the analysis method. It is a method of finding structural features by creating a graph by abstracting the elements and their relationships to nodes and edges respectively. There are several types of graphs, such as directed/undirected graphs and weighted/unweighted graphs. Moreover, there are various kinds of indicators which express structural features such as density, transitivity, cluster coefficient, centrality, order, and so forth.

We are focusing on classification codes in the patent. The graphs which consist of the classification codes of the patents of every six months are made - and revealing the companies' technology trends by extracting the structural feature of graphs.

The remainder of this paper is organized as follows: Section 2 explains the research backgrounds and related work; Section 3 briefly summarizes the collected data on the target patent sector; Section 4 describes the analysis of the data and presents analysis results, and Section 5 explains some conclusions and future work.

## 2. Related work

This chapter describes related work on our proposed method. First, we investigate the patent analysis method. One of the methods of patent information analysis is citation analysis. Lee et al. suggested a method for forecasting technological changes using patent citation analysis and text mining [1]. Kim et al. analyzed the pattern of technological convergence using a network analysis of the patent citation by taking printed circuit board technology as an example [2]. Sato et al. proposed the importance calculation method of the patent document based on the citation information [3]. Ogawa et al. proposed a primary patent extraction based on citation information [4]. Albert conducted a validation of citation for essential patent among the industry [5]. Muguruma showed the validity of the patent citation analysis to propose FCA (Forward Citation Applicant) map [6]. Patent mapping is also a standard method. Kiriya made content analyses with this method [7]. Shide et al. performed finding the change of the positioning for a customer of research and development activities of the company using the patents analyses [8]. Kimura proposed a technology evaluation method based on patent analysis for technology strategy planning [9].

In place of that which were performed heuristically so far, as for the patent classification, such as the category of invention and problem, Tanaka proposed a method of extracting the feature automatically [10]. Yamashita proposed a method of surveillance technology and the specific method of patent classification with text mining [11]. Yamamoto et al., proposed a method to enhance the compatibility of the search by applying the information of related patent documents in search of academic papers. Yamamoto proposed a method to find scientific papers with a variety of further information [12, 13]. Kleinberg extracted the topic and description of the relationship with graph theory [14]. Eto proposed a measure of co-citation based on structural units of the paper [15]. Ueda proposed the technical analysis with an active mining method that focuses on the cognitive processes of the patent examiner, utilizing patent classification such as IPC (International Patent Classification), FI (File Index), and F-term (File Forming Term) [16]. Thus, with the application of the technology of intelligent informatics, knowledge extraction is performed to the patent information.

For a business solution, TRIZ (Teoriya Resheniya Izobretatelskikh Zadatch) aims at performing technical development based on the structure of problem-solving which repeatedly appears in the patent [17, 18]. Thus, many methods were employed based on patent information as a primary business solution.

Then, we look down about mining technology, which is the one technique of knowledge extraction [19 - 21]. The relationship of analytical methods and technologies,

technology analysis method is described using patent information. A structured technique using graph theory has been applied in various fields. For example, chemical formula, WWW (World Wide Web), social network, statements with grammatical structure and dependency.

A Graph is made for expressing the structure of the object and make it useful for the analysis and the visualization of the relation of things. [22 - 23].

From the result of the related work study, the relationship can be explained using link mining. This paper employs it for patent analysis.

### 3. Experimental configuration

This chapter explains the configuration of this experiment. Patent information consists of 3 types of data, such as textual information, graphics information, and metadata. Metadata includes application date, applicant, classification codes, and so on. There are classification codes that are used internationally, such as IPC, CPC, and others used domestically, such as ECLA, USPC, FI, F-term, and so on. IPC and CPC indicate the technological fields of the main topics of the patent's claims.

CPC is enacted in 2013 [24, 25]. It is assigned retroactively to the patents filed before 2013. Like other classifiers, it has a hierarchical structure. It is separated into five fields such as Section (Alphabet 1 character) - Class (2 digits) - Subclass (Alphabet 1 character) - Main group - Subgroup. Figure 1 shows an example of a CPC notation. There are about 250 thousand groups in total.

There is a firm order to grant CPC to patents. First, a CPC which points to the primary technical field of the patent is given. It is called Head-CPC. Next, CPCs of the technical field to which the patent relates are granted. They are called Sub-CPC. The number of CPCs granted to a patent is arbitrary. It depends on the relevant technical field of the patent. For example, Xilinx's patent: US9, 831, 218B1 has 3 CPCs such as H01L 25/0657, H01L 25/50 and H01L 2225/06513, and its Head-CPC is H01L 25/0657.

We focus on the patents owned by Altera and Xilinx in the U.S. The analysis period is 16 years from 2001 to 2016. The full-text database of the U.S Patent and Trademark Office is used as an experimental database [26]. Applicant: Altera and Xilinx are used as search keywords. As a result,

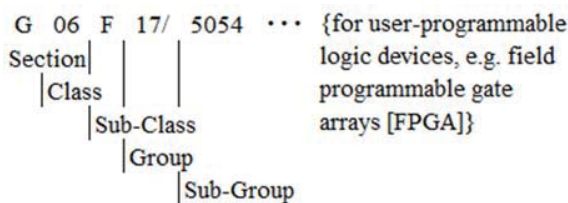


Figure 1. CPC notation.

3,885 and 3,704 patents are collected respectively. Among them, 3,302 and 3,113 patents are filed in the period.

First of all, patents are separated every six months. Then, a directed graph is created for each patent to express the technological structure of it. The CPCs granted to the patent are used as nodes, and edges are drawn between Head CPC and Sub CPCs. The direction of the edges is Head CPC to Sub CPC. There is no edge between Sub CPCs. The graph is an unweighted graph. Figure 2 shows an example of a graph of a patent.

Then, the semi-annual graphs of each supplier are created by superimposing the graphs every six months. The number of link occurrences is accumulated for each occurrence of an edge. Therefore, the graphs are directed and weighted graph. Figure 3 shows an example of a semi-annual graph.

In this experiment, we focus on the shape of the graphs and the inward degree of the nodes. The inward degree represents the number of times the CPC become Sub-CPC. In general, the nodes connected with a large number of edges, are considered to be core technologies. However, the Head CPC of the patent which has a large number of CPCs has a large number of outward degree, but it is not necessarily core technologies.

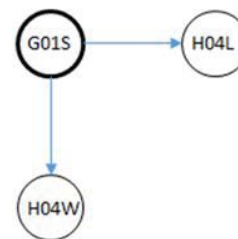


Figure 2. Patent graph.

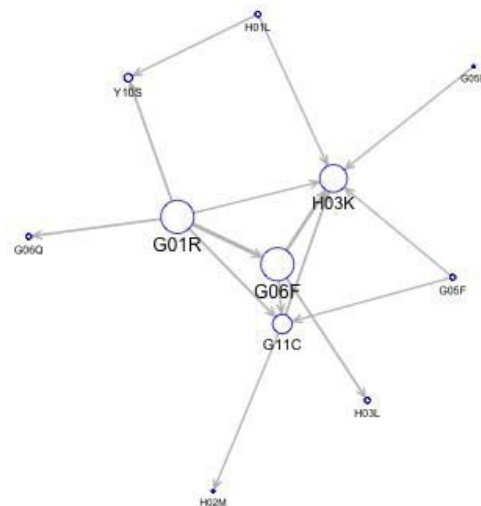


Figure 3. Semi-annual graph.

Therefore, we utilize the number of times used as the Head CPC rather than outward degree. In Fig.3, the size of nodes indicates the number of the inward degree plus the number of times the node become Head-CPC. The thickness of an edge indicates the number of times the nodes have the relation as the Head CPC and the sub CPC. Four CPCs out of 11 have a high degree. Therefore, in this case, these four CPCs are considered core technologies.

#### 4. Experimental results

This chapter explains the result of the experiment. Figure 4. represents the number of filed patents and granted CPCs every six months. It depicts that both companies continually acquire patents. Moreover, the number of CPCs increases throughout the period, and the increase is much more than the increase in the number of patents. It means a patent becomes related to a wider technological field. It is caused by the field of FPGA adaptation is expanded, and both companies expand their research and development fields.

From Figure 5 to 8 show the graphs of 2001 first half and 2016 second half from Altera and Xilinx respectively. They depict the research and development fields and their relations of both companies, and their expansions by increasing the number of nodes.

The graphs of 2001 first half of both companies consist of a large connected graph and a small amount of tiny connected graphs. There are some hubs in the large graphs such as H03K 19/17736, H03K 19/17728, H03L 19/1776 and H03K 19/1778 in Altera, and H03K 19/17704, H03K 19/17736, H03K 19/1737 and H03K 19/17744 in Xilinx. It means they are the main focus areas of the companies.

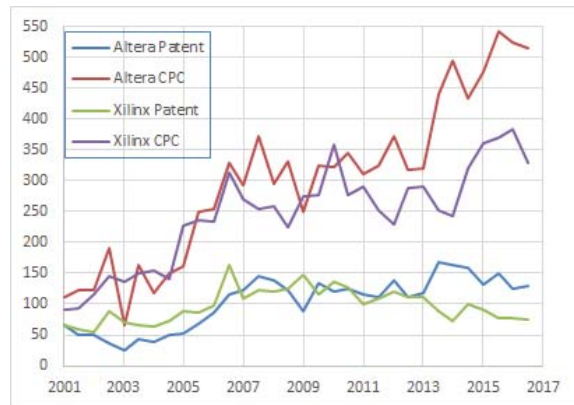


Figure 4. Time-series filed patents in CPC

On the other hand, in the second half of 2016, the graphs of both companies have more nodes and small connected components than the 2001 first half. It depicts their research and development fields are diversely expanded.

Table 1 represents the top ten CPCs which have the highest degree for each supplier every six months. It represents the transition of the focused area of their research and development. The color of cells: red, yellow, and green represent CPCs: H03K, H01L, and G06F respectively.

Comparing Altera and Xilinx, the CPCs that appear in both companies' lists are almost common. It shows that both companies' directions of technology development are very similar. Both G06F and H01L have a large number of the degree all over the period. It represents that the fields of "Electric digital data processing" and "Semiconductor

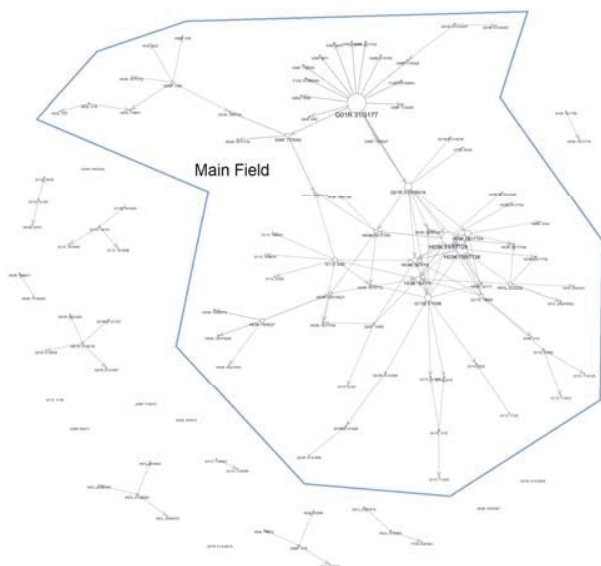


Figure 5. Altera graph in 2001 first half.

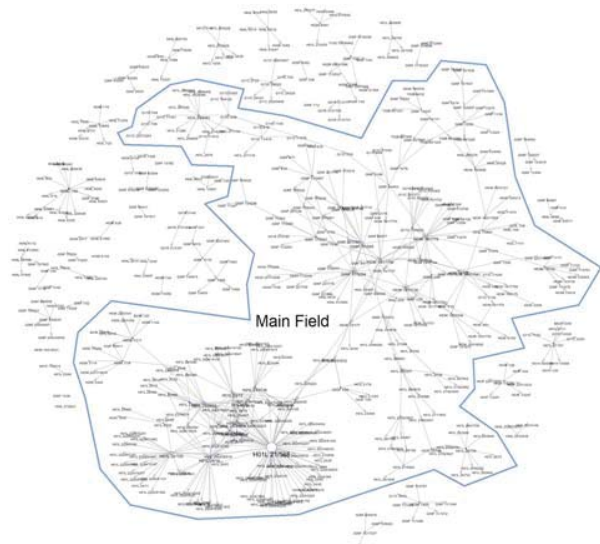


Figure 6. Altera graph in 2016 second half.

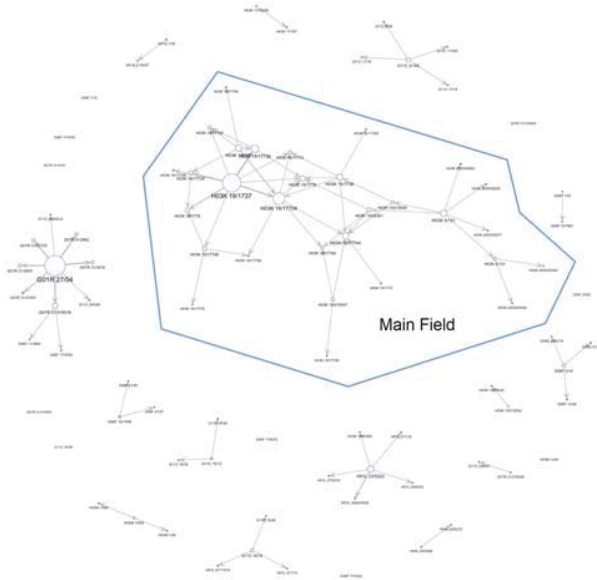


Figure 7. Xilinx graph in 2001 first half.

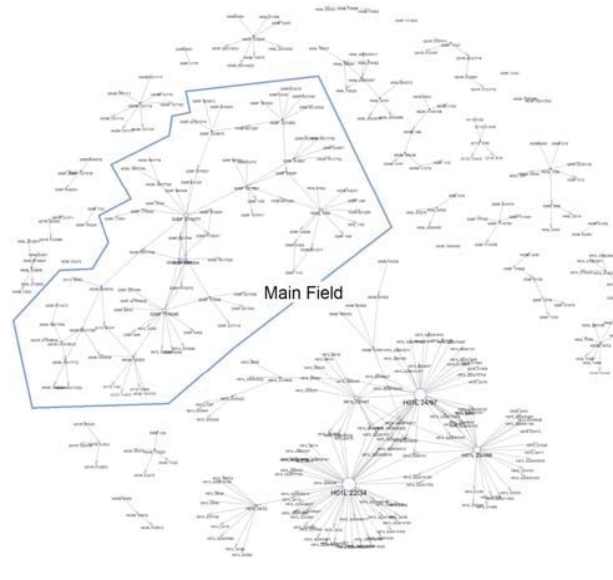


Figure 8. Xilinx graph in 2016 second half.

Table 1. The transition of Focus Areas by High degree CPCs.

Altera		Xilinx																				
2001	fh	H03K	H03K	H03K	H03K	G06F	H03K	H03K	G01R	G06F	G06F	G06F	H03K	H03K	H03K	H03K	H03K	H03K	H03K	H03K	G01R	G01R
	sh	H03K	H03K	H03K	H03K	H01L	H03K	G01R	G01R	G06F	H03K	G06F	H03K	H03K	G11C	H03K	H03K	H03K	H03K	H03K	H03K	H03L
2002	fh	H03K	H03K	H03K	H03K	H03K	H01L	H03K	H03K	G06F	G06F	G06F	H03K	H03K	H03K	G06F	G06F	G06F	G06F	G06F	G06F	G06F
	sh	H01L	H01L	H03K	H01L	H01L	H01L	H01L	H03K	H01L	H01L	G06F	H03K	G06F	H03K	H03L	G06F	H03L	G06F	H03K	G01R	G01R
2003	fh	H03K	H03K	H03K	H03K	H03K	H03K	H03K	H03K	H03K	G06F	G06F	H03K	H01L	H01L	H01L	H03K	H03K	G06F	H03K	H03K	
	sh	H03K	G06F	H01L	H03K	H03K	H01L	H01L	H03L	H01L	H01L	H03K	G01R	G06F	H01L	G01R	G06F	H03K	H03K	H03L	G01R	
2004	fh	H01L	H03K	H03K	H03K	H03K	G01R	H01L	H01L	H01L	H01L	G06F	H01L	G01R	G06F	H01L	G06F	H01L	H03K	G06F	G06F	
	sh	H01L	H03K	H01L	H03K	H01L	H01L	H01L	H01L	H01L	H01L	H03K	H01L	G06F	H03K	G06F	H01L	H03K	H03K	G06F	G06F	
2005	fh	H01L	H03K	H01L	H01L	H03K	H01L	H03K	H01L	H01L	H01L	H01L	G06F	H01L	G06F	G06F	G01R	H01L	H01L	H01L	G01R	
	sh	H01L	H01L	H01L	H01L	H01L	H01L	H03K	H01L	H01L	H01L	G06F	H01L	H03K	H01L	H03K	G01R	G06F	G06F	H03K	G06F	
2006	fh	G06F	H03K	H01L	H03K	H01L	H01L	H03K	H03K	H03K	H03K	G06F	H03K	H03K	H01L	H01L	H03K	H03K	G01R	H01L	H03K	
	sh	H01L	H03K	G06F	H03K	G06F	H01L	H01L	H03K	H03K	H03K	G06F	G06F	G06F	H01L	H03K	G01R	G06F	G06F	H03K	H03K	
2007	fh	G06F	H01L	H03K	H03K	H03L	G11C	G11C	H01L	H03K	H03K	H03K	G01R	H01L	H03K	H03K	G06F	G06F	G06F	H03K	G06F	
	sh	G06F	H03K	H01L	H01L	H03K	G06F	G06F	H01L	H03K	H03K	G06F	H03K	H03K	H03K	G01R	H01L	G06F	H03K	G06F	H03K	
2008	fh	G06F	H03K	H03K	G06F	H03K	H01L	H03K	H03K	H03K	G06F	H01L	G06F	H01L	H01L	H01L	G06F	H01L	H01L	H01L	H01L	
	sh	G06F	H01L	H01L	H01L	H01L	H01L	H03K	H03K	H01L	H01L	G06F	G06F	G06F	G01R	G06F	H03K	G06F	G06F	G06F	G06F	
2009	fh	G06F	H03K	H01L	G06F	H01L	H01L	H03K	H03K	G01R	G06F	G06F	H01L	H03K	G06F	H03K	G06F	H01L	H03K	H03K	H03K	
	sh	H03K	H03K	G06F	H03K	H03K	G06F	H01L	H01L	H01L	H01L	H01L	G06F	G06F	H03K	H03K	H03K	G01R	G06F	G06F	G06F	
2010	fh	G06F	H03K	H03K	H01L	G06F	H03L	G06F	G06F	H03K	H03K	H01L	H03K	H01L	H01L	H03K	H01L	G06F	H01L	H01L	H01L	
	sh	G06F	H01L	H01L	H03K	H03K	G06F	H01L	H03K	H03K	G06F	H01L	G06F	G06F	G06F	G06F	G06F	G06F	G06F	G06K	H03K	
2011	fh	H01L	H01L	G06F	H01L	H01L	H03K	G11C	H01L	G06F	H01L	H01L	G06F	H01L	H01L	G06F	H01L	H01L	H01L	H01L	H01L	
	sh	G06F	H01L	H01L	G06F	H01L	G06F	G11C	H01L	H01L	H04L	G06F	H01L	H01L	H01L	G06F	H01L	H01L	H03K	H01L	H01L	
2012	fh	G06F	H01L	G06F	H01L	H01L	H01L	H01L	H01L	H01L	G06F	G06F	H01L	H01L	H01L	H01L	H01L	G06F	H01L	H01L	G06F	
	sh	G06F	H01L	H01L	H01L	H01L	H01L	H01L	G06F	H01L	G06F	H01L	H01L	G06F	H01L	H01L	G06F	H03K	H03K	G06F	G06F	
2013	fh	G06F	G06F	G06F	G06F	H01L	H03M	H04L	G06F	G06F	G06F	H01L	G06F	H01L	H01L	H01L	G06F	G06F	G06F	G06F	G06F	
	sh	G06F	H01L	H01L	H01L	G06F	H01L	H01L	G06F	G06F	G11C	H01L	G06F	H01L	G06F	H01L	H01L	H01L	H01L	H03K	H01L	H01L
2014	fh	G06F	H01L	H01L	G06F	G06F	H03K	G06F	H01L	H01L	H04L	H01L	H01L	H01L	G06F	H01L	G06F	G06F	H01L	H01L	H01L	
	sh	H01L	H01L	G06F	H01L	G06F	H03K	G06F	H01L	H01L	H01L	H01L	H01L	H01L	H01L	H01L	G06F	H01L	H01L	G06F	G06F	
2015	fh	G06F	H01L	A61K	G06F	H03K	G06F	G06F	G06F	G06F	G06F	H01L	H01L	G06F	G06F	H01L	H01L	H01L	H03M	G06F	H01L	
	sh	H01L	H01L	G06F	G06F	G06F	G06F	H01L	H01L	H03K	H03K	H01L	H01L	G06F	H01L	H01L	H01L	H01L	G06F	H01L	H01L	
2016	fh	H01L	G06F	H01L	H01L	H01L	H01L	H01L	H01L	G06F	H01L	H01L	H01L	H01L	H01L	H01L	H04L	H01L	H01L	H01L	H01L	
	sh	G06F	G06F	H01L	H01L	H03K	H01L	H01L	H01L	H01L	H03K	G06F	H01L	H01L	G06F	H01L	H01L	H01L	G06F	G06F	G06F	

devices: Electric solid state devices not otherwise provided for” continue to be important for the companies throughout the period. At the beginning of the period, the degree of H03K is high. It means they focused on “Pulse technique” at that time.

## 5. Concluding remarks

In this paper, we propose a sufficient method for revealing corporate technology development trends with link mining. We gathered the US patent data filed by FPGA suppliers. We utilized the patent classifier for making the graph of technology structure. With the

proposed method, we succeed in revealing corporate focusing technology area with the technological transition.

Our future work includes 1) examining another machine learning method such as xgboost [27, 28], 2) generate new patterns with meta-heuristic algorithms, and 3) predicting the future technological structure of FPGA by utilizing the research results. This work will require practical experiments and further survey studies.

## References

- [1] C. Lee, Y. Cho, H. Seol, Y. Park, "A stochastic patent citation analysis approach to assessing future technological impacts," *Technol. Forecast. Soc. Change*, vol. 79, 2012m pp. 16-29.
- [2] D. H. Kim, B. K. Lee, and S. Y. Sohn, "Quantifying technology-industry spillover effects based on patent citation network analysis of unmanned aerial vehicle (UAV)," *Technol. Forecast. Soc. Change*, vol. 105, 2016, pp. 140-157.
- [3] Y. Sato, M. Iwayama, "A study of patent document score using patent-specific attributes in citation analysis," *Journal of Information Processing and Management*, vol. 51(5), 2008, pp 334-344.
- [4] T. Ogawa, I. Watanabe, "Finding Basic Patents using Patent Citations," *IP SJ SIG Notes, Information Processing Society of Japan*, vol. 35, 2005, pp.41-48.
- [5] M. B. Albert, D. Avery, F. Narin, P. McAllister, "Direct Validation of Citation Counts as Indicators of Industrially Important Patents," *Research Policy*, vol. 20(3), June 1991, pp.251-259.
- [6] M. Muguruma, "The Usefulness of Patent Forward Citation Analysis and its Practical Examples," *The Journal of Information Science and Technology Association*, vol. 56(3), 2006, pp.114-119.
- [7] T. Kiriya, "IP Information Analysis (<Special feature>Patent Information: Analysis and Effective Utilization)," *The Journal of Information Science and Technology Association*, vol. 60(8), 2010, pp.306-312.
- [8] K. Shide, M. Ando, "The Shift of Positioning of Japanese General Contractors' R&D Activities," *Journal of Architecture and Planning*, vol. 76(668), 2011, pp.1929-1935.
- [9] H. Kimura, "One Approach of Technology Stocktaking and Evaluation for Corporate Technology Strategies: Emphasizing Future Intentions and Quantification through Patent Analysis," *The Journal of Science Policy and Research Management*, vol. 26(1/2), 2012, pp.52-61.
- [10] K. Tanaka, "Multi-viewpoint Clustering of Patent Documents," *IP SJ SIG Notes*, vol. 4, 2008, pp.9-14.
- [11] Y. Yamashita, "Text mining Technology for Patent Analysis and Patent Search: Patent Search and Patent Analysis Service Patent Integration," *Journal of Information Processing and Management*, vol. 52(10), 2010pp.581-591.
- [12] M. Yamamoto, H. Maze, T. Yajima, H. Kinugawa, "A Journal Paper Filtering Using the Profile Revised by Patent Document Information," *IEEJ Transactions on Electronics, Information and Systems*, vol. 130(2), 2010, pp.358-366.
- [13] M. Yamamoto, H. Kinugawa, "A Journal Paper Filtering Using the Multiple Information," *IEEJ Transactions on Electronics, Information and Systems*, vol. 131(6), 2013, pp.1250-1259.
- [14] J. M. Kleinberg, "Authoritative Sources in a Hyperlinked Environment," *Journal of the ACM*, vol. 46(5), September, 1999, pp. 604-632.
- [15] M. Eto, "A New Co-citation Measure Based on Structures of Citing Papers," *Database, 49(SIG 7(TOD 37)), Information Processing Society of Japan*, 2008, pp.1-15.
- [16] I. Ueda, "Active Mining Utilizing the Patent Classification IPC, F1, F Term on the Basis of the Cognitive Processes of the Examiner," in *Proceedings of the SIG-FAI, Japanese Society for Artificial Intelligence*, vol. 46, 2001, pp.13-21.
- [17] G. Altshuller, "The Innovation Algorithm: TRIZ, systematic innovation, and technical creativity," Worcester, MA: Technical Innovation Center, 1999.
- [18] G. Altshuller, "40 Principles: Extended Edition," Worcester, MA: Technical Innovation Center, 2005.
- [19] J. Karamon, Y. Matsuo, M. Ishizuka, "Link Mining from Networks of Academic Papers," *Technical report of IEICE. KBSE*, vol. 106(473), 2007, pp.73-78.
- [20] H. Kashima, "Mining Graphs and Networks," *The Journal of the Institute of Electronics, Information, and Communication Engineers*, vol. 93(9), 2010, pp.797-802.
- [21] Y. Kajikawa, "Utilization of citation information by link mining," *The Journal of Information Science and Technology Association*, vol. 60(6), 2010, pp.224-229.
- [22] L. Gettor, "Link mining: a new data mining challenge," *ACM SIGKDD Explorations Newsletter 5.1*, 2003, pp.84-89.
- [23] R. Diestel "Graph Theory 5th Edition", Springer 2017
- [24] Y. Ota, "About CPC," Tokugikon, no.270, pp.34-47, 2013
- [25] Cooperative Patent Classification <https://www.cooperativepatentclassification.org//index.html> (accessed on Dec 20th, 2018)
- [26] USPTO Patent Full-Text Databases <http://appft.uspto.gov/netahtml/PTO/index.html> (accessed on Dec 2nd, 2018)
- [27] xgboost <https://github.com/dmlc/xgboost> (accessed on Dec 2nd, 2018)
- [28] xgboost <https://xgboost.readthedocs.io/> (accessed on Dec 2nd, 2018)

## Use of LIME for human interpretability in Sinhala document classification

P. K. S. Kumari

*Faculty of Graduate Studies and Research  
Sri Lanka Institute of Information Technology,  
Sri Lanka  
suriyaa.k@sliit.lk*

P. S. Haddela

*Department of Information Technology  
Sri Lanka Institute of Information Technology,  
Sri Lanka  
prasanna.s@sliit.lk*

### Abstract

*With advancement of technology in Sri Lanka, use of Sinhala text usage has grown rapidly over the time where automatic categorization is helpful for efficient content management. As a result, experts tend to use machine learning application to categorize this large volume of data in an efficient and accurate manner. Most of these learning models are operating in a black-box where there is no way to understand how the model has decided which category an instance is assigned. Understanding the reason behind why learning model makes these predictions is very important to trust such models and to provide reasonable justifications in real world application. Intention of this research is to present the work carried on related to document classification model prediction interpretation where a set of text classifiers has been studied with use of SinNG5, freely available Sinhala Document corpus.*

**Keywords:** *Machine learning, Sinhala text, document classification, human interpretability*

### 1. Introduction

Wide availability of electronic data has led to the enormous interest in text analysis, information retrieval and content classification methods where Natural Language Processing (NLP) techniques are widely engaged. Text documents are one of the richest sources of data for businesses: whether an email, technical documents, user reviews or news articles, they all contain valuable information that can be used to automate slow manual processes, better understand users, or find valuable insights [1]. Most Natural Language Processing (NLP) applications follows supervised learning techniques where a machine learning algorithm is trained to handle given content according to its domain/subject relevance where documents are assigned into set of predefined categories. This makes it easier to find the relevant information at the right time and for filtering and routing documents directly to correct users. But the problem is how anyone is going to trust the predictions made by such a model for them to apply in world real applications. Even the model verified during the

development stage with use of test and validation data set after being deployed may handle a data set very different from what was used to build the model. So, predictions made by machine learning model can be wrong in such situation and there is no way to note such problems because most of complex machine learning algorithms operates in a black-box. The classifier provides set of model evaluation metrics, but it is hard to depict how learning model has processed data internally to come up with the results. To address this issue, recently some research work [6, 10] has proposed an explanation technique that explain predictions of any classifier in an interpretable and faithful manner which simulates how learning models analyze the content of text and images related to the category assigned. The proposed LIME library [6, 12] learns an interpretable model locally around its prediction and visualize the reasons behind why model made came up with its classification result.

During this study, a set of text classification algorithms has been studied and the predictions made by these models further construed with use of prediction interpretable libraries to produce descriptive model representations to visualize the reason why learning model has made such classification.

The research paper is organized as follows. Section 2 contains a review on Sinhala related text classification work and a few other text categorization research studies including application prediction interpretation techniques; In section 3, a text classification process is discussed; in section 4, set of experimented classifiers are discussed; in section 5, Use of LIME explainer for Sinhalese is discussed; in the section 6, the result and research outcomes are outlined and in section 7 and 8, the paper is concluded with research challenges and future work.

### 2. Review: Prediction, interpretation and text classification

To achieve high accuracy level with learning model, practitioners use complex algorithms like Random Forest, deep learning network which are highly complex and there is no understanding on the inner working of such models. The objective of interpretation is to give

ability to explain or provide reasons to justify outcomes of learning model in understandable terms to a general practitioner.

Many of the state-of-the-art machine learning models are functionally black boxes, as it is nearly impossible to get a feeling of its inner workings. While computers beating professionals in games like AlphaGo [23], many people have started think of ways to Interpret Machine Learning models. This brings us to a question of trust: “do I trust prediction coming from the model? or do I even trust that the model is making reasonable predictions in general? [6]”.

During literature survey, it was observed that only limited number of research works [1-4] has been done relevant to Sinhala text classification work where none of these have attempted to study on prediction interpretation to make these black box classifiers more understandable to its users.

As mentioned in “Challenges of enabling IT in the Sinhala Language” [1] the main reason for low use of Sinhala in supportive application development was the lack of a single standard. A study outlined the approaches to categorize non-English documents to provide better service for Sri Lankan users [2]. The approach makes use of a transformation based on the term frequency and inverse document frequency (tf-idf), which enhances the clustering performance. Another study proposed an automated system to do document classification which will minimize the language ambiguities when categorizing digital documents [3] while another carried out an experiment to study about use of rule-based classification on Sinhala Text Documents [5]. The studies have used SinNG5 corpus [4] and SinSWL [4] for the experimental purposes. There was also a study which documented about an algorithm implemented to generate domain-specific stopword list from data set SinNG5 news data set [4]. Stopwords are set of words contained in a document that frequently occurs which add noise to data. Mostly in classification work, stopword removal applies during preprocessing stage which cleans data, removing frequently found unrelated terms for classification.

These days there are different kind of recommendation systems e.g. Patient symptoms diagnosis and treatment recommendation, identify victim of a court case, where these systems provide evidences and facts can be used to support decision making process, so expertise on subject domain can use such services to make effective use of them. But the problem is how developers are going to convince its users to get more out of such services because behavior of such complicated algorithms cannot be audited. To overcome such issues, recently researchers have proposed some solutions which let them to interpret the predictions made by these

models. The following are some of these related studies which address the interpretability of prediction models.

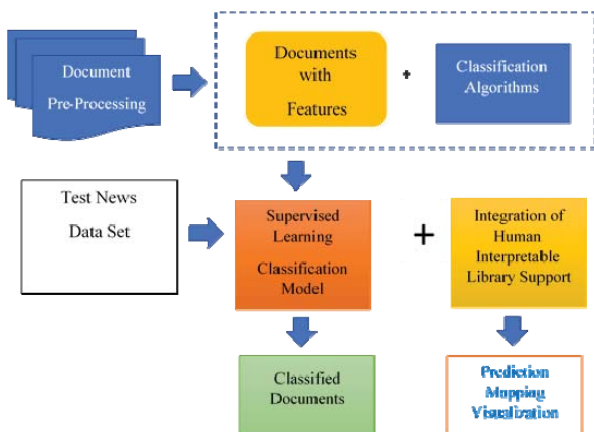
A study proposed a library known as Lime (Local Interpretable Agnostic Explanations) where they have studies about how predictions made by a classification algorithm can be interpreted in a more understandable manner for its users. LIME [12] produce visual explanations where looking at them you can understand the logic behind how the internally model has analyzed the data. The authors of LIME library have studied model prediction explanations with text and image classification tasks and shows how complicated model results can be explained in a more precise and clear manner. Another study has analyzed ICU patient treatment records to predict ICU mortality level. During research study authors has applied LIME library [12] to interpret predictions given by complex Random Forest Classifier. The authors were successfully able to predict the model predictions which clearly visualized what factors are contributing to predictions made by learning model [7]. Authors also proposed a unified framework known as SHAP (Sharpley Additive Explanations) for viewing explanation of a model’s predictions as a model itself, where predictions are interpreted with use of a class indicating important feature properties related to decision taken [10].

Many research papers have discussed classification algorithms like Naïve Bayes, K-NN and SVM, Random Forest where those experiments have mostly based on text data related to English Data Set. There was however no evidence available to indicate where these algorithms were studies with Sinhala data set. This study focuses on experimental evaluation of above-mentioned text classifiers to identify the most effective classifier for the Sinhala text classification problem. Later with use of these models classifier prediction results are explained in an interpretable manner to make them more understandable to users of applications.

### **3. Text classification and prediction interpretation process**

Figure 1 shows the stages of the text categorization process which has been followed during experimental studies where text classifiers were integrated with prediction interpretable libraries. The aim of this process was to do classification of documents into a fixed number of pre-defined categories and obtained the prediction results for the sample set of data. Next, integrated libraries which support prediction visualization, map input features are used for the classification into class label probabilities.





**Fig 1: Text classifier prediction result visualization.**

Text classification task mainly goes with few important steps. During pre-processing task, collected set of documents need to be converted into the form where can identify input features the classifiers going to work with. Next, selected set of classification algorithms can be trained with labeled text documents to get a supervised learning text classification model. Afterwards, providing test data set model can be tested to see how accurate it makes the text categorization task. The model predicts how many instances it has correctly been classified from entire data set it has worked. Finally, for all the selected algorithms a performance evaluation can be done where each classifier gets evaluated by evaluation measures like precision, recall, f-measure, accuracy, error rate to select best classification model can have applied for categorization task.

The most important task during this experimental work is to interpret learning model prediction record to its user. This will explain how classifier has worked with input features of test data while makes the classification, so prediction result can be evaluated. For that, selected text classifier is integrated with LIME python library [11] which support learning model prediction interpretation.

#### 4. Text classification methods

Text classification models [7, 10, 12] are pre-trained on a labeled data set to know about text patterns related to its categories. Later, to evaluate how well the model performs, it is tested unlabeled, which a subset of collected dataset. Among many supervised learning text classification algorithms available, the following set of classification algorithms has been considered for the study.

Naïve Bayes - Naïve Bayes classifier [13-16, 23] is a machine learning model which uses Bayes' theorem. It is based on probability models that incorporate strong

independence assumptions. Using Bayes theorem, it finds out the probability of **A** happening, given that **B** has occurred. When it comes to classification, it can handle binary classification or multiclass classification problems.

SVM - A support vector Machine (SVM) [8] is a classifier which tries to classify the data based on a hyper plane used to separate data points in a higher dimension. When deciding hyper plane, it select the separation line which makes the highest margin for the support vectors. Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. To increase the accuracy of learning model it uses kernel functions or regularization parameters to fine tune the model.

kNN - In pattern recognition, the k-nearest neighbors' algorithm (k-NN)[13-14, 16] is a non-parametric method used for classification and regression. In classification, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k-nearest neighbors.

Random Forest - Random Forest [17-21] is ensemble type of machine learning algorithm, which produces results more accurately compared to other type of classification algorithms without even hyper-parameter tuning. It is one of the most commonly used classification algorithms due to its accuracy and simplicity compared with deep learning networks. It builds collection of trees which make up the forest where each tree is trained with sub set of features related to the data model.

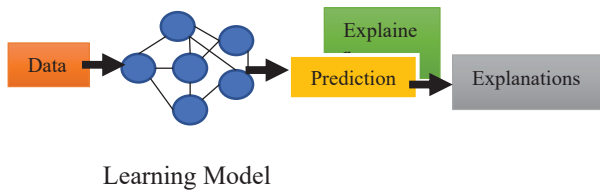
Machine learning models like linear regression, decision trees are interpretable due to their simplicity but produce less accurate results [10]. When it tries to increase level of accuracy of complex models like SVM, Random Forest, deep neural networks are more preferred [6]. But they are functionally black-box models, providing 'good' explanations and will increase understandability of inner work, which in return increase level of trust [6, 10].

#### 5. Use of LIME explainer for Sinhalese

Among few libraries which support prediction interpretation, LIME python library has used during this study to work out prediction interpretation. The library defines an explainer object which takes a text classifier instance as its parameter to learn about how classifier has analyzed its input features when making text categorization. Then for a randomly selected instance it can ask to explain the way how classifier has assigned class probabilities into each input vector. While classifier explains it, which class label assigned with maximum probability among the rest, can be seen

To trust a learning model prediction, it is very important to know how internal decision making happened for classification process. Through “Explaining a prediction” [6-7, 10], it’s expected to present textual or visual explanation that provides clear relationship between the predictions and input features or components has been used.

Figure 2 illustrated the process of explaining individual prediction making by a text classification model. The explainer visualizes terms or features the learning model has taken into consideration when it predicts the class the instance is belongs into.



**Fig 2. Use of LIME explainer to visualize individual prediction.**

Usually people who works in the application domain have prior knowledge about facts needed to be considered for classification. By looking at illustrated interpretations, any users can get a good understanding about the way classifiers handle the work and decide whether they are going to accept (trust) or reject the predictions if they can understand the reason behind it.

During this study, to apply LIME explainer for a text classifier used Sinhala news data set SinNG5 [4, 5] which contained articles related to five news categories named as, crime, politics, business, religion and sports. Text classification algorithms considered for the study were, SVM, KNN, Naïve Bayes and Random Forest. These text classifiers were trained using the training set. Next each classifier was tested against test set obtained measures related to evaluation metrics.

Intent of studied text classification experiments is to analyze how learning model prediction result can be interpreted in an understandable manner to a general practitioner [6, 7]. Understanding the reasons behind learning model predictions is relatively important in evaluating trust when trying to propose such automated models in real application. As experimental environments used Anaconda Jupiter note book, an integrated development environment which support natural language processing tasks where python is been used to code machine learning classifiers and prediction interpretation. By “explaining a prediction”, we mean presenting textual or visual artifacts that provide qualitative understanding of the relationship between the instance’s components and the model’s prediction [6]. LIME learns interpretable model locally around its

predictions [6] where LIME explainer uses different color code to visualize assigned class label probabilities related to each category learning model it works with. Interpretation of selected prediction results can be explained with LIME where explanations are visualized according to supportability of each input feature related to model prediction.

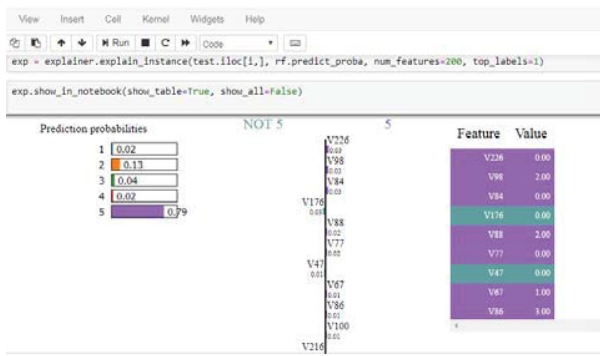
## 6. Results and research outcomes

In most cases machine learning classification models are using a Black-Box approach where internally how learning model interpret the data cannot be seen in default situation. Referring to the prediction made on training and test data, based on measures produced at the end of classifier execution a justification can be made if accuracy is in higher level (more than 90%) to be sure the learning model works well.

The problem with this approach is, during the development of any kind of learning model, when it uses supervised learning algorithms, the model is trained and tested with collected historical data related to the domain the model is developed to. But in real operation it’s possible that the data feed into the classifier to make predictions can be totally different from the data which has been used for model development. So, predictions made by classifier can be totally wrong with what is really expected. Machine learning models can only be debugged and audited when they can be interpreted.

The experimental worked which has been done with use of lime library during this research was focused on how LIME library support can be used to explain predictions made by classification model. The aim was to visually represent reasons behind prediction behavior of selected classification models.

All implementations have been studied with use of anaconda python note book application. The news article used during the experiment were labeled from 1-5 where 1 for crime, 2 for business, 3 for politics, 4 for religion and 5 for sports respectively. For each test instance the learning model predicted the class it labeled by the classifier with its’ respective number. Lime explainer prediction visualization can be obtained as shown by following Figure 3.



**Fig 3: Explaining Random Forest text classification prediction using LIME explainer.**

For a selected record from test data set, prediction result interpretation can be summarized as shown in Table 1. The probability assigned for each class label is indicated with its respective class label. The document classifier has assigned highest probability for class label Sports where respective input features model was taken into consideration and also represented clearly with color code used to represent each class label.

**Table 1: Summarizing text prediction with assigned probabilities.**

Class No	Class Name	Lime Explainer Colour	Probability assigned By Lime explainer
1	Crime	Blue	0.02
2	Business	Orange	0.13
3	Politics	Green	0.04
4	Religion	Red	0.02

Observing the explanations related to randomly selected instances, each learning model further can be audited to understand how well each classification model works with its data. For certain problems or tasks, it is not enough to get the prediction (the what). The model must also explain how it came to the prediction (the why). E.g.: Patient treatment recommendation system, Real time weather prediction system, Stock market exchange rate variation prediction system.

The process of integrating machines and algorithms into our daily lives requires interpretability to increase social acceptance where understanding can increase trust level.

## 7. Research challenges and future work

The work related to this study used Sinhala data set SinNG5 [4, 5] for classification purpose. Commonly available open-source tools have character encoding issues relates to them. These tools cannot represent Sinhala characters due to default encoding issues. Further, insufficient literature relates to application of interpretable libraries in non-English context and is a major challenge which limit applicability of such techniques directly to Sinhala text. This solution mainly helps future studies related to Information Retrieval and classification work based with Sinhala Text.

Future directions for research are,

- Investigate other approaches can be applied to Sinhala text corpuses where default encodings issues can be overcome.
- Investigate other libraries which support prediction interpretations as well as learning model interpretations.

## 8. Conclusion

In this paper, research work applied for document classification algorithms mainly aimed to analyze learning model prediction interpretation. Gaining trust is very crucial when planning to introduce these machine learning systems in real world applications. Explaining individual predictions is very important in assessing understandability and trustworthiness of such learning models. So, evaluating the prediction results will provide an assurance to users whether they can trust these models in their real world operations. With use of LIME library, it was able demonstrate how individual predictions results can be evaluated, independent to what learning model has been used for classification work.

## 9. References

- G. V. Dias, "Challenges of enabling IT in the Sinhala Language," in *27th Internationalization and Unicode Conference*, Berlin, Germany, 2005.
- D. A. Meedeniya and A. S. Perera, "Evaluation of Partition-Based Text Clustering Techniques to Categorize Indic Language Documents," in *2009 IEEE International Advance Computing Conference*, Patiala, India, 2009.
- C. M. Wijewickrema and R. Gamage, "An Ontology Based Fully Automatic Document Classification System Using an Existing Semi-Automatic System," in *79th IFLA General Conference and Assembly*, Suntec City, Singapore, 2013.
- S.V.S. Gunasekara and P. S. Haddela, "Context aware stopwords for Sinhala Text Classification," in *2017*

*National Information Technology Conference (NITC)*, Colombo, Sri Lanka, 2018.

- [5] K. B. N. Lakmali and P. S. Haddela, "Effectiveness of Rule-based Classifiers in Sinhala Text Categorization," in *2017 National Information Technology Conference (NITC)*, Colombo, Sri Lanka, 2017.
- [6] M. T. Ribeiro, C. Guestrin and S. Singh, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, KDD 2016 San Francisco, CA, USA, 2016.
- [7] G. J. Katuwal and R. Cen, "Machine Learning Model Interpretability for Precision Medicine".
- [8] "Problem-solving with ML: automatic document classification," Google, [Online]. Available: <https://cloud.google.com/blog/big-data/2018/01/problem-solving-with-ml-automatic-document-classification>.
- [9] I. Dilrukshi, K. D. Zoysa and A. Caldera, "Twitter News Classification Using SVM," in *The 8th International Conference on Computer Science & Education*, Colombo, Sri Lanka, 2013.
- [10] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *31st Conference on Neural Information Processing Systems*, CA, USA, 2017.
- [11] M. Makrehchi and M. S. Kamel, "A Text Classification Framework with a Local Feature Ranking for Learning Social Networks," in *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, Omaha, NE, USA, 2007.
- [12] M. T. C. Ribeiro, "Lime: Explaining the predictions of any machine learning classifier," 2016. [Online]. Available: <https://github.com/marcotcr/lime>.
- [13] Wongso, R., Luwinda, F. A., Trisnajaya, B. C., & Rusli, O. (2017). News Article Text Classification in Indonesian Language. *2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI*. Bali, Indonesia: Elsevier.
- [14] S. J. Sohrawardi, I. Azam and S. Hosain, "A Comparative Study of Text Classification Algorithms on User Submitted Bug Report," in *Ninth International Conference on Digital Information Management (ICDIM 2014)*, Phitsanulok, Thailand, 2014.
- [15] R. S. Zulfany Erlisa Rasjid, "Performance Comparison and Optimization of Text Document Classification using k-NN and Naïve Bayes Classification Techniques," in *2nd International Conference on Computer Science and Computational Intelligence 2017*, Bali, Indonesia, 2017.
- [16] "Naive Bayes Classifier," [Online]. Available: <https://machinelearningmastery.com/naive-bayes-tutorial-for-machine-learning/>.
- [17] Rasjid, Z. E., & Setiawan, R. (2017). Performance Comparison and Optimization of Text Document Classification using k-NN and Naïve Bayes Classification Techniques. *2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI 2017* (pp. 107-112). Bali, Indonesia: Elsevier.
- [18] A. Alsudais, G. Leroy and A. Corso, "We Know Where You Are Tweeting from: Assigning a Type of Place to Tweets," in *2014 IEEE International Congress on Big Data*, Anchorage, AK, USA, 2014.
- [19] Y. Hegde and S.K. Padma, "Sentiment Analysis Using Random Forest Ensemble for Mobile Product Reviews in Kannada," in *2017 IEEE 7th International Advance Computing Conference (IACC)*, Hyderabad, India, 2017.
- [20] Y. Zhang, C. Wang, B. Xiao and C. Shi, "A New Method for Text Verification based on Random Forests," in *2012 International Conference on Frontiers in Handwriting Recognition*, Bari, Italy, 2012.
- [21] A. John and M. Wilscy, "RANDOM FOREST CLASSIFIER BASED MULTI-DOCUMENT SUMMARIZATION SYSTEM," in *2013 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, Trivandrum, India, 2013.
- [22] M. N. Elagamy, C. Stanier and B. Sharp, "Stock market random forest-text mining system mining critical indicators of stock market movements," in *2018 2nd International Conference on Natural Language and Speech Processing (ICNLSP)*, Algiers, Algeria, 2018.
- [23] "AlphaGo," DeepMind, 2010. [Online]. Available: <https://deepmind.com/research/alphago/>.
- [24] "Naive Bayes Classifier," [Online]. Available: <https://machinelearningmastery.com/naive-bayes-tutorial-for-machine-learning/>.

## A Blockchain-based decentralized system to ensure the transparency of organic food supply chain

B. M. A .L. Basnayake

Department of Industrial Management,  
University of Kelaniya, Sri Lanka  
basnayake.bmal@gmail.com

C. Rajapakse

Department of Industrial Management,  
University of Kelaniya, Sri Lanka  
chathura@kln.ac.lk

### Abstract

Low quality agricultural products are added to the market daily. Over usage of chemicals in the production process, use of uncertified chemicals and mechanisms for preservation and ripening processes, are the major issues that impact on agricultural product's quality as well as overall health of the consumers. Mechanisms to identify the quality of the agricultural products are highly demanded due to the lack of transparency in the current process. Blockchain technology is emerging as a decentralized and secure infrastructure which can replace involvement of a third party to verify the transactions within the system. The purpose of the research was to implement a Blockchain based solution to verify the food quality and the origin of the agricultural supply chain. A public Blockchain concept was selected instead of a private Blockchain in this study to ensure transparency by allowing any person to access the network. Instances of the smart contract were created for each physical product and deployed to Blockchain network. A Quick Response code which contained the address of the instance, was a reference to the virtual product. All the actors who are involved in the supply chain must be able to interact with the system to achieve the transparency. Each transaction and events related to a product is validated by peers of the Blockchain system. Product ownership was changed for each relevant transaction. A token-based mechanism was used to indicate the farmers' reputation with their products. Farmers could place a certification request regarding their products and, they can gain reputation tokens for each certification done by peers. A unique Quick Response code was used to identify each product within the supply chain. The proposed system has been implemented as a prototype and validated within the study.

**Keywords:** Agriculture, Supply chains, Food security

### 1. Introduction

The applicability of Blockchain concept to the agricultural supply chain to verify the origin and quality of the agricultural products was considered for this study. The Customers are not satisfied with the quality of the agricultural products that they are consuming due to the presence of synthetic agricultural chemicals in the agro products.

Production of the organic foods is adequate to fulfill market demand [1] and a mechanism for product certification is required. Thus demand for the carbonic foods and products with health certifications are increasing due to the current global health issues. Due to lack of transparency of existing certification mechanisms, consumers are reluctant to trust these mechanisms.

Hence preparing a transparent and efficient architecture for food certification was the purpose of the study, by combining trusted certification system utilizing Blockchain concept and other technologies. A proof-of-concept has been constructed and tested in this paper.

#### 1.1. Organic agriculture industry

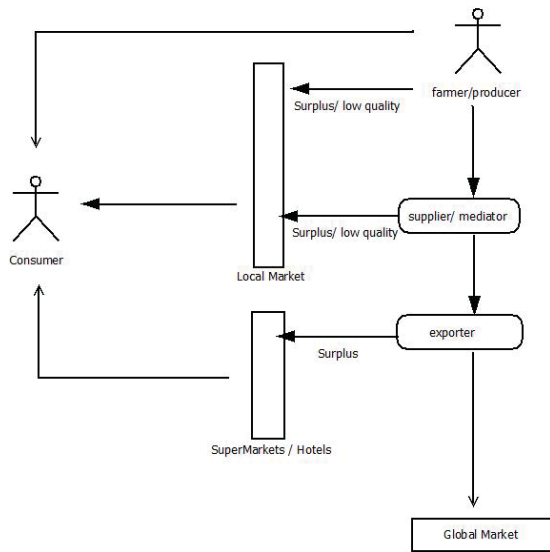
With the drastic environmental pollution due to agrochemicals and their effect on human health, demand for the organic products is increased. Incensement of global demand and new market, premium price in local and export market and increment of farmers' income by 30-100% are the inspiring factors for the farmers to involve in the organic food industry [2].

Sri Lanka has introduced a vast variety of organic products such as organically certified tea, spices, essential oils, herbs, desiccated coconut, oil seeds, pulses, cashew, rubber, tropical fruits and vegetables [3].

#### 1.2. Agricultural supply chain

Agricultural supply chain varies for each product. Each product/crop has a unique flow within the process. Value addition or formation processes may be taken place within the supply chain. Lot of waste, damage,

being unable to verify origin can be occurred when products reach the customer end point



**Figure 1. Current Food supply chain in Sri Lanka.**

Figure 1 describes the agricultural supply chain in brief. Damaged or low quality products are issued to the local market at low price or same price, in relation to non-organic products [4].

### 1.3. Food certification

Organic farmers are certified under several certifications within Sri Lanka. SLGAP (Agricultural Department, Sri Lanka), Control Union (Netherlands), IMO (Switzerland), NASAA (Australia), EcoCert (Germany), etc.[5]

Agro land details, soil, water supply, seeds, history of the land, crop, disease and control etc. are considered when issuing the certification [6]. These certifications are updated seasonally. Certifications can be misused to sell products from third parties or low quality products because only the farmers are certified, not the products. Most of the regular relationships between farmers and consumers, farmers and mediators depended on pre-experience and trust.

## 2. Literature review

In 2008, [7] published a white-paper on use of Bitcoin as a peer to peer Electronic Cash System. Initially, the Blockchain concept was applied in financial services.

With the invention of first digital currency Bitcoin as the first Blockchain platform, requirement of trusted authority or central server to avoid double spending<sup>1</sup> of

<sup>1</sup> Double-spending is a problem unique to digital currencies because digital information can be reproduced relatively easily [10]

currency was solved. Later on, Blockchain concept derived to other areas more than implementation as a currency in Bitcoin. With the emergence of second generation, Blockchain applications were derived to many fields. Supply chain is a key area where Blockchain concept is widely applied.

The Danish shipping company Maersk the world’s largest container carrier, tested the use of Blockchain applications in the international logistics to reduce the paper work in shipping and container tracking [8].

### 2.1. Flaws in centralized systems

Consumer trust, supply chain transparency, product quality, logistic issues, environmental impact, personal consumer data, fraud, food safety, etc. are the reliability issues which are faced by current supply chains. Consumers request greater transparency and information but current systems cannot provide such data. Most of the cases, information is audited by third parties, and records are stored in paper or centralized databases. This approach causes some problems mentioned below.

- High cost and inefficiency of paper based processes.
- Frauds, security errors in IT systems
- Integrity of digital records
- Misuse of paper certifications

These problems result in low transparency and trust. With the usage of Blockchain based platform, mechanisms such as smart contracts, decentralized database and proof-of-works ensured the security and transparency of supply chain [9].

### 2.2. Food tracking

Tracking mechanisms are dependent on the products/items which were exchanged within the supply chain. Customers were willing to pay additional costs to attach IoT devices to track the location of the product when they are high cost products. Heavy vehicles and high cost machines were attached with IoT devices to track locations. Product types like pharmaceutical items, foods, etc. were tracked in bulk. In most cases, containers were tracked instead of each item. In South Africa, Blockchain was applied to grape supply chains to interact between farmers, certifiers, auditors and grape traders through the assistance of digital certified grape boxes. Each box was identified by a unique barcode number and the ownership of the box was changed for each transaction [9].

### 3. Methodology

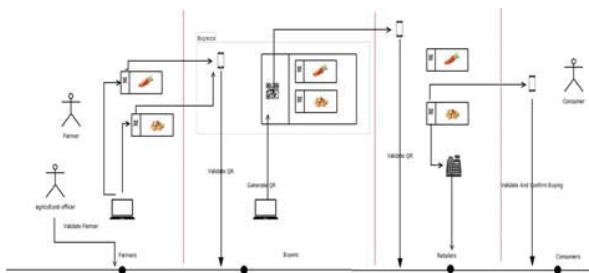


Figure 2. Overview of user interaction.

In the development process of the DApp described in Figure 2, Token model, Solidity based smart transaction model and QR code generation model to validate products were developed. Based on the developed DApp, applicability of Blockchain concept in the agricultural supply chain was tested. QR codes were used as the link layer between digital and physical world.

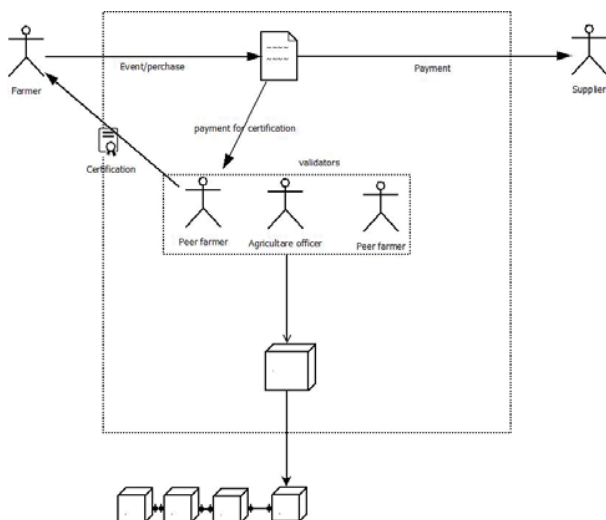


Figure 3. Production process validation.

Food production process and events such as water supply source, land history, pest control, fertilizer usage, packaging, etc. were identified in data collection and systematic review of literature. Furthermore, persons who can validate each process/event were identified. For each event/process, farmer can place a request, and peer farmers, agricultural officers or related persons should validate the request to gain reputation. Each event was recorded in Blockchain after the validation process. Figure 3 describes the certification/validation process for each event.

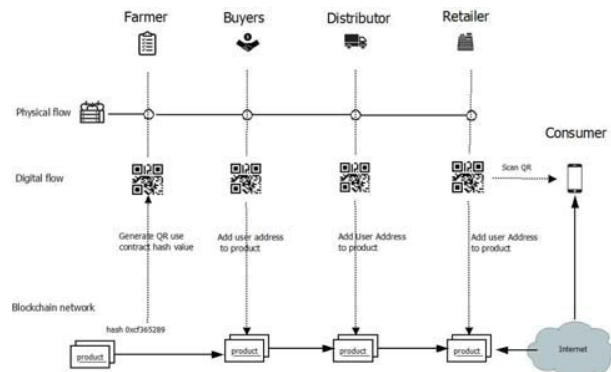


Figure 4. Overview architecture for Blockchain based supply chain.

Ethereum smart contracts were used to develop DApp which was connected to the Blockchain system. Only farmers were eligible to deploy “product” contract to indicate physical product. Farmers should have sufficient amount of reputation level to issue a product contract to control the quality of each product. For each deployment of the “product” contract, it will return an address which was used to generate the QR code to identify the physical product. Figure 4 depicts how virtual product was validated with physical product using QR code. For each transaction within the supply chain, user’s address has been added to the “product” solidity contract as product owners. Consumers could scan the QR code using a mobile application that enables access to the Blockchain system. Consumers are also eligible to rate the product quality to ensure the trust.

### 4. Implementation

Solidity classes were compiled into two parts as Application Binary Interface (ABI) and Byte Code by solidity compiler. Encoding and decoding solidity smart contract calls into data and machine code were used in Application Binary Interface. Compiled smart contracts were deployed to Blockchain network using web3.js and a hexadecimal address was returned when deployment occurs.

Business logic and the user actions were governed by solidity smart contracts. A new season was created by farmers to the network and they were governed by ‘Campaign’ class. Issuing of products was governed by ‘Products’ Solidity class and products were deployed only by the farmers. Only a product owner was given the authority to end the transaction at any point of the supply chain. Products would not be sold which contain its transaction state as ‘true’.

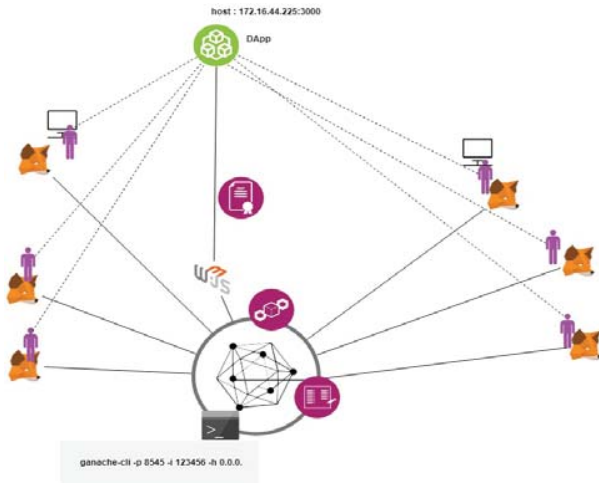


Figure 5. Implemented system.

Ganache test client was run on port 8545. MetaMask extensions were connected to Ganache test client using Remote Procedure Calls (RPC). Application was run on port 3000 and users could connect to application on <http://172.16.44.225:3000>. Figure 5 describes the implemented system architecture.

## 5. Testing and validation

Product history could be identified by owner's data which was added to smart contracts within each transaction. Smart contract was updated with new owners and its state whether it was sold out or not. Smart contracts were publicly available in the network and data was available to everyone. User interface to view product details is displayed in Figure 6.

0x5306d6a07DcCaa32B5e0257272F5C719aDC773a9  
Owners - 1

0x5306d6a07DcCaa32B5e0257272F5C719aDC773a9  
Owners - 2

0x99A445bc88Dd3ed126c7166972A2DB15557E32ba  
Owners - 3

0xaAc8b2AFc253EeAD9d5a6236bB77B6461Ee90d0F  
Owners - 4

Products



End Transaction

Figure 6. Product description.

System generated Quick Response (QR) code contains the address of the product, returned from the network when a product was deployed to the system by the farmer.

**Process Summary**

<p>0x90b451f31032ADc36eB3f0d A54C5b573928681B</p> <p>Address of Farmer</p> <p>The manager created this campaign and can create requests to withdraw money</p>	<p>0</p> <p>Minimum Contribution (wei)</p> <p>You must contribute at least this much wei to become an approver</p>
<p>4</p> <p>Number of Requests</p> <p>A request tries to withdraw money from the contract. Requests must be approved by approvers</p>	<p>2</p> <p>Number of Approvers</p> <p>Number of people who have already donated to this campaign</p>
<p>3</p> <p>Campaign Balance (ether)</p> <p>The balance is how much money this campaign has left to spend.</p>	<p>0/100</p> <p>max product limit</p> <p>Product limit</p>

[View Events](#)
[Generate QR code](#)
[View QR code](#)

Figure 7. Summary of the season.

Summary of a particular season has been generated from data which was fetched from deployed smart contracts as shown in Figure 7.

Requests

ID	Description	Amount	Recipient	Approval Count	Approve	Finalize
0	8ffcc9ff 0x1f 852ea	0	0x90b451f31032ADc36eB3f0dA54C5b573928681B	0/2	<a href="#">Approve</a>	<a href="#">Finalize</a>
1	8ffcc9ff 0x1f 852ea	0	0x90b451f31032ADc36eB3f0dA54C5b573928681B	0/2	<a href="#">Approve</a>	<a href="#">Finalize</a>
2	event outflow	0.01	0xEf089E9E8c7aCA2d098279F7CE887AC45a647	0/2	<a href="#">Approve</a>	<a href="#">Finalize</a>
3	event outflow	0.01	0xEf089E9E8c7aCA2d098279F7CE887AC45a647	0/2	<a href="#">Approve</a>	<a href="#">Finalize</a>

Found 4 requests.

Figure 8. Event / production process validation.

Events/processes were publicly available to the peers for approval. Events/processes would be finalized when accepted by peers. Event/process was open to anyone to be accepted, till it was finalized by the farmer. These functionalities were shown in Figure 8. Block creation can be identified in Figure 9 and Figure 10.

```
Transaction: 0xb99fd081df52c41640a7432382fe870d475b3f0292e62a717f817818c884f264
Contract created: 0xdd5f7245e691536017d118300349f5728767b8fb
Gas usage: 1522522
Block Number: 1
```

Figure 9. Block creation.

New transaction has been deployed to the network and these data were added to new block.

```
Transaction: 0x502c5aa88af812e9389c8a8f2cee6e9c1e1dba17ff3e2f4673b6539670f40a56
Gas usage: 1101828
Block Number: 2
```

Figure 10. Block creation.



## 6. Conclusion

This study is based on the applicability of Blockchain concept to improving transparency and validity of agricultural supply chain and its process. Since recent past, there has been a rapid change in the production of food and its raw materials. An efficient method to bridge the gap between the farmer producing commodities in the market and the end customer was studied. Blockchain based architecture and its concepts were taken for implanting trustworthiness and transparency within the users and their transactions.

Identified business logic in literature review were implemented using smart contracts. It was sufficient to control business logic within the transactions and other events which were related to food production. Quick Response (QR) code was identified as a cost effective and practical solution to identify each food product within the supply chain from farm to end consumer. Quick Response codes were generated including the address of the product, returned by the Blockchain network when products were deployed to the network. Smart contract's address can be given by scanning QR code. Address could be used to fetch data directly from the Blockchain network without depending on the application layer. High availability of data was ensured using this approach.

In a typical transaction, trust was built by involving a third party to regulate the rules between users. Trust was decentralized among all peers in the network by Blockchain network. Money was received to smart contracts by farmers when farmers deployed a new season. Approvers were given money when events and processes were being validated. Each process and event were validated by a participant of the network. All the transactions were tested within Ganache test network and data were stored in a Blockchain. Thus, trust and transparency were successfully ensured by the proposed mechanism.

### 6.1 Recommendation

Ethereum Geth client should be used for implementing a local private network and establishing a peer-to-peer network with each nodes. It is possible to connect Geth client to online Ethereum network such as ropsten, kovan Rinkeby main network etc. but with low computational power and storage limitations, it is not recommended. Connecting web3.js to third party provider like Infura is recommended to avoid these limitations.

## 7. References

- [1] Rosairo, H. S. R., "ARE ORGANIC VEGETABLES A MARKETING FAILURE IN SRI LANKA?," no. September, pp. 15–25, 2006.
- [2] OTA, "Organic Industry Survey in U.S.," *May 19, 2016*, 2016. [Online]. Available: <https://www.ota.com/news/press-releases/19031>.
- [3] "Organic Products in Sri Lanka|Organic Agriculture." [Online]. Available: <http://www.srilankabusiness.com/organic/>. [Accessed: 09-Jul-2018].
- [4] "එළවළු අපනයනය ආශ්‍රිත ස පසපම් ද ම විශ් ලේඛණය."
- [5] "Sri Lankan Organic Products Certification|Organic Agricultural Product Standards." [Online]. Available: <http://www.srilankabusiness.com/organic/organic-certification-in-sri-lanka.html>. [Accessed: 09-Jul-2018].
- [6] "GAP." [Online]. Available: <http://slgap.lk/>. [Accessed: 09-Jul-2018].
- [7] Nakamoto, S., "Bitcoin: A Peer-to-Peer Electronic Cash System."
- [8] Groenfeldt, T., "IBM And Maersk Apply Blockchain To Container Shipping," *Forbes*, 2017. [Online]. Available: <https://www.forbes.com/sites/tomgroenfeldt/2017/03/05/ibm-and-maersk-apply-blockchain-to-container-shipping/#7652eccb3f05>. [Accessed: 13-Jun-2018].
- [9] Ge, L., Brewster, C., Spek, J., Smeenk, A., and Top, J., *Findings from the pilot study Blockchain for Agriculture and Food*. 2017.
- [10] "Double-Spending Definition | Investopedia." [Online]. Available: <https://www.investopedia.com/terms/d/doublespending.asp>. [Accessed: 13-Jun-2018].

## Part of speech tagging for Twitter conversations using Conditional Random Fields model

A. R. F. S. Fanoon  
 Postgraduate Institute of Science,  
 University of Peradeniya, Sri Lanka  
 fanoonarfs@gmail.com

G. A. I. Uwanthika  
 Postgraduate Institute of Science,  
 University of Peradeniya, Sri Lanka  
 isuriuwanthika@gmail.com

### Abstract

*Part-of-Speech Tagging is the technology of assigning the appropriate parts-of-speech to a word. Part-of-speech tagging is very useful in information retrieval, information extraction, and speech processing. This research presents a part-of-speech tagging, especially for twitter text data. The process of part-of-speech tagging for twitter conversation is a difficult task. Several approaches have been made to develop an accurate tagging system but most of them are relevant to news text data and web contents. Therefore, this research intends to develop a part-of-speech tagger model for twitter speech. using CRF toolkit. The system was developed for nearly 1000 twitter conversations employing Conditional Random Field stochastic model. The data for twitter speech was downloaded from the internet. A POS-tagged text corpus, template file and CoNLL file for both training and testing database were prepared accordingly. The training was carried out for both the unigram model as well as the bigram model. The performance of the system over these models was obtained through this examination which showed a significant efficiency, calculated from the number of correctly tagged words and the total number of words.*

**Keywords:** *Bigram model, Conditional Random Field, Tagging, Unigram model*

### 1. Introduction

Enormous quantities of text data in electronic form are generated as a result of the usage of social media and user-created web content by the growing population. The social media web blogs contain mainly online conversational text. This online conversational text, which is in the forms of microblogs, chats and text messages is a very challenging task in Natural Language Processing.

Part-of-Speech (POS), also known as word classes or syntactic categories provide a piece of large information about a word and its neighbors. POS features are useful in finding named entities and other information extraction related tasks. A part-of-speech is very important for

pronunciations in speech synthesis and speech recognition.

Part of Speech (POS) Tagging is the process of assigning appropriate POS Tags to each word in a sentence. e.g., noun, verb, adverb, pronoun, depending on the context and definition. Simply, POS Tagging is labelling words with their appropriate Part-of-Speech. Tagging is a disambiguation task while words are ambiguous, and therefore, the main goal is to find the correct tag for the word.

POS Tagging is very essential in order to understand the meaning of any sentence or to extract relationships and build a knowledge graph.

The paper is organized as follows: Section 1 briefs on the introduction to POS Tagging, and POS Tagging techniques available. It also brings a theory on HMM and CRF models. A review of related work is briefly described in Section 2. In Section 3 we explain the methodology as to how the research was carried out for Twitter speech. In Section 4, the performance of the models is compared. Finally, we draw our conclusion based on the performance evaluated and discuss how we can further improve the research.

#### 1.1. Twitter speech

Twitter is an application which has been developed for the purpose of online news and social networking service for users where they can interact with messages, known as Tweets. Registered twitter users are allowed to post, like and retweet tweets while unregistered users are allowed only to read them. Twitter is one of the most visited websites and it has been described as “the SMS of the Internet”.

The conversational nature on Twitter is vastly different from the general text data such as news text. Lack of conventional orthography and the character limitation of a tweet are the main reasons for this. Therefore, tagging is a challenging task for Twitter speeches [3]. Hence the application of the standard modelling assumptions is inappropriate for Twitter conversations. An example is shown in the figure below.

ikr	smh	he	asked	fir	yo	last
!	G	O	V	P	D	A
name	so	he	can	add	u	on
N	P	O	V	V	O	P
fb	lololol					
^	!					

**Figure 1. POS tag example.**

In order to tackle the challenges faced in tagging the twitter conversation as well as Internet Relay Chat (IRC), [3] developed a system using HMM (Hidden Markov Model) at an accuracy of 82%. This study aims to use the CRF model for part-of-speech tagging for twitter speech

## 1.2. POS tagging techniques

There are mainly four different techniques available for POS Tagging [7]. They are as follows:

- a) Lexical based methods  
A lexicon-based approach is a simple, feasible and practical approach that does not need training for sentiment analysis of text data. It assigns the POS Tag for a word in the training corpus depending on the frequency of the occurrences of the word.
- b) Rule based methods  
Rule-based approach assigns POS Tag to a word depending on the rules pre-defined by the user. For example, we can define a rule that says words ending with '-ed' and '-ing' are verbs. Lexicon approach also could be combined with Rule-based method to tag words which are not in the training corpus but in the test data set. This combined approach has two stages, the first stage assigns POS tag using a lexicon and the second stage assigns POS tag using a large list of manually prepared rules.
- c) Probabilistic methods (Stochastic approach)  
POS Tags are assigned based on the probability of the occurrences of a particular tag sequence. Hidden Markov Model (HMM), Conditional Random Field (CRF) and Maximum Entropy Markov Model (MEMM) are probabilistic POS Tagging approaches. Need for a tagged corpus to obtain probabilities about the ambiguous word is the main idea behind this approach.
- d) Deep learning methods  
Deep Learning approach involves the application of Recurrent Neural Networks (RNN) for POS Tagging.

## 1.3. Conditional Random Fields

Conditional Random Fields (CRF) is the probabilistic models for labelling and segmenting structured data such

as sequences, trees, and lattices. It is a standard model to predict the most likely sequence of labels corresponding to a sequence of inputs [6]. CRF is a discriminative classifier, meaning that they try to model conditional probability distribution ( $P(y|x)$ ), whereas HMM is a generative model trying to model a joint probability distribution ( $P(x,y)$ ).

CRF takes a set of features which are derived from the sequence of input using feature functions, their respective weight and the previous label, as input. The task is to predict the correct label for the current word. The feature functions, in CRF, extract a set of features, known as State Features, for each word in a sentence. The extracted features may be, what the previous word is, whether the particular word is the first or the last word, what the next word is, and etc. The label of the previous, as well as the next word, is input in CRF model in order to determine the weight and the weight determination of different functions maximizes the probability of the tags in the training data. The feature function that is dependent on the previous tag is known as the Transient Feature.

## 1.4. Conditional Random Fields Vs other models

HMM, MEMM and CRF are the popularly used statistic models for pattern recognition and machine learning problems.

### 1.4.1. Hidden Markov Model (HMM)

The word 'Hidden' means that only the symbols from the system are visible and users are unaware of the underlying state changes in the system. Therefore, HMM is recognized as an infinite state machine.

HMM is an efficient algorithm allowing consistent insertion and deletion of penalties and also able to handle variables of varying lengths. But it is only dependent on every state and its corresponding observed object. HMM also causes a mismatch between the target function and the predicted function.

### 1.4.2. Maximum Entropy Markov Model (MEMM)

MEMM functions based on the dependencies between the neighboring states and the complete observed sequence, thus resulting in a better expression ability. MEMM does not consider the  $P(x)$ , as it reduces modelling workload and learns the target and estimated function consistency. The major drawback of MEMM is that it causes labelling bias.

### 1.4.3. Conditional Random Fields (CRF)

CRF model, in comparison to other models discussed above, has overcome the labelling bias issue in MEMM and eliminated two unreasonable hypotheses in HMM. As

a result of this, the CRF model is complicated than the other two. Further, the CRF model adopts global variance while MEMM adopts local variance.

CRF advantages are:

- Compared to HMM, CRF is not strictly dependent on assumptions like HMM model. CRF is capable to accommodate any context information and feature design is flexible.
- CRF computes the conditional probability by overcoming the label bias issue on MEMM.
- In the availability of observation sequence intended for labelling, CRF computes the joint probability distribution of the entire label sequence.

CRF does not only assume the features that are dependent on each other but also it considers future observations while it learns the pattern. Therefore, CRF combines both HMM and MEMM models.

CRF Disadvantages are:

- Highly complex when training the data.
- Difficult to re-train when new data are added to the corpus.

This study is carried out using CRF model and therefore, the comparisons among the three models, CRF, HMM and MEMM is [7]:

- HMM directly models the transition and the phenotype probability and finally calculates the co-occurrence probability.
- MEMM calculate the co-occurrence probability from the transition and phenotype probability thus, calculating the conditional probability. MEMM applies local variance normalization technique so that it will be easy to make it fall into a local optimum.
- In contrast to HMM and MEMM, CRF calculates the normalization probability using global variance, hence providing an optimal global solution and resolving the labelling bias issue in MEMM.
- HMM and MEMM are directed graph while CRF is an undirected graph.

## 2. Related work

POS tagging involves the task of assigning the part-of-speech marker to an input text. Tokenization is a process that involves the separation of punctuations such as commas, quotation marks, etc. before tagging since the tags are even applied for punctuations. A sequence of words and a tag set is input into the tagger which in return outputs a sequence of tags, a single best tag for each word.

Part-of-Speech tagging is one of the most fundamental parts of the linguistic pipeline. This is a basic form of

syntactic analysis and has infinite applications on NLP. Standard pipeline tools have been applied on news text data and have shown a richer performance for linguistic analysis while they show poor performance on Twitter data [4].

Social media part-of-speech tagging has a preliminary work [3], named entity recognition [5,9], and parsing [2], but however, the accuracy is slightly significantly lesser than that for traditional news text data. Most of the part-of-speech tagging system for the English language has been carried out using the HMM model.

## 3. Methodology

Part-of-Speech tagging systems mainly involve the training of text data corpus. There are several approaches to part-of-speech tagging process such as rule-based approach, stochastic models and morphology driven approach. In this study, we have adopted the CRF modelling technique as it is a much sophisticated and probabilistic computation technique for searching the most likely tag, tagged by the system.

An open source toolkit known as CRF which comprises of CRF++ library has been used here for the purpose of speech training and model development for Twitter speech. All research activities were carried on in Ubuntu platform as no features will be lost compared to the Windows platform, while training the data. The main steps involved in part-of-speech is shown in the diagram below.

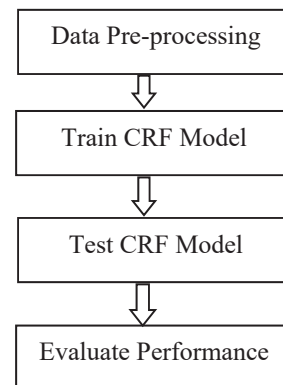


Figure 2. POS tagging approach.

### 3.1. Data preparation

The goal of this research is to tag the accurate part-of-speech to twitter conversations and for this purpose, it is required to prepare a text corpus containing twitter speech and tag them. The data for this research was downloaded from the internet, which has been prepared by [3]. The data set used for this research is the oct27 data set. When the

text corpus has been prepared, the following tasks were carried out on the data set.

### 3.1.1. Prepare a POS tagged corpus

From the literature review, it is very clear that the traditional annotations could not be used to tag the twitter speech and therefore [3] have come up with a tag set specifically for online conversations, which was carried out manually.

### 3.1.2. Convert to ConLL Format

CoNLL is a data format, in which each line represents a single word with a series of tab-separated fields. ‘\_’s indicates the empty values. The output would be oct27.conll.

N	common noun
O	pronoun (personal/WH; not possessive)
^	proper noun
S	nominal + possessive
Z	proper noun + possessive
V	verb including copula, auxiliaries
L	nominal + verbal (e.g. <i>i'm</i> ), verbal + nominal ( <i>let's</i> )
M	proper noun + verbal
A	adjective
R	adverb
!	interjection
D	determiner
P	pre- or postposition, or subordinating conjunction
&	coordinating conjunction
T	verb particle
X	existential <i>there</i> , predeterminers
Y	X + verbal
#	hashtag (indicates topic/category for tweet)
@	at-mention (indicates a user as a recipient of a tweet)
~	discourse marker, indications of continuation across multiple tweets
U	URL or email address
E	emoticon
\$	numeral
,	punctuation
G	other abbreviations, foreign words, possessive endings, symbols, garbage

Figure 3. POS tags for Twitter speech.

### 3.2. Train CRF model

Training a model requires a template file and the training data. A template file is where the features are defined for training the model. The features considered for this research are the previous word, the current word and the word that comes next.

Once the template is prepared, the POS tagged corpus has to be trained. For training purpose, CRF++ library on

Ubuntu was installed and then the training data set was trained.

### 3.3. Test CRF model

Testing the system is important in order to check whether the proposed model is working finely. The testing is performed on the test data set.

## 4. Results and discussion

The part-of-speech tags were tested on the test data corpus using the crf\_test command on the Ubuntu terminal for both unigram and bigram models. Later the accuracy was checked for both the models and analyzed.

According to the test results, the unigram model showed an accuracy of 60% and the bigram model showed an accuracy of 72% for twitter speech. The calculation is based on the formula that:

$$A = (CTW/W) * 100,$$

where

A = Accuracy

CTW = Number of correctly tagged words

W = Total number of words

The result of POS Tagging for both Unigram and Bigram Models is as given below.

Table 1. Accuracy of the POS tag.

	Unigram Model	Bigram Model
Correctly Tagged Words	4356	5215
Incorrectly Tagged Words	2796	1937
Total Words	7152	7152
Total Sentences	500	500
Accuracy (%)	60	72

Twitter is one of the main social media which is being widely used worldwide presently. People post media posts of their memories, like them and retweet tweets daily. Almost all the countries political news is mainly broadcasted through twitter in just a very few seconds. Unlike the traditional genre of news text, twitter messages have a different orthography and a creative language and therefore, it is a challenging task.

Very few researches have been carried out for online conversational text, and almost all of them have been carried out using Hidden Markov Model (HMM). The accuracy obtained from these researches is nearly 82% [3], which is comparatively higher than the accuracy obtained from the research we have carried out using CRF model.

## 5. Conclusion

The aim of this research was to develop a part-of-speech tagging for Twitter speech. The POS tagged corpus was prepared and converted to CoNLL format for training. Using the CRF++ library, CRF models were trained and tested in different ratios. Finally, after the training and testing of the Twitter conversational data, the accuracy of the tagged text was calculated for unigram and bigram models. In comparison with the previous works, the newly generated tagger model shows an accuracy of 72%, which is significantly lesser than the older versions. The database size used for this research work is approximately 1000 words. As the data size is small, the system gives a lesser accuracy than the system with large size vocabulary. Finally, the system shows the potential for further improvement and promises to be a nucleus for further work and research in part-of-speech tagging for twitter texts.

## 6. Recommendation

In the future, this project can be used in a large scale with very little modifications. During the experimental work medium size part-of-speech tagging system was implemented. The system can be extended to longer sentences with more emoticons and abbreviations, using the HMM/CRF Technique or using other growing techniques like Artificial Neural Network. At present we considered about only 1000 words for training and testing. The data size could be increased even more in order to get a part-of speech tagged with maximum accuracy.

Further, the study can be further extended to increase the accuracy by parallelizing part-of-speech tagging tasks, through means of the search algorithm using GPGPU processors. In the future, by leveraging both the CPU and GPU processors during the part-of-speech tagging process, the tagging can be performed for large, and in some cases multiple models, which could yield a high accuracy.

## 7. References

- [1] O'Connor, B., Krieger, M. and Ahn, D., "TweetMotif: Exploratory search and topic summarization for Twitter", in Proceedings of ICWSM.
- [2] Foster, J., Cetinoglu, O., Wagner, J., Roux, L., Hogan, S., Nivre, J., Hogan, D., and Genabith, J., " #hardtoparse: POS tagging and parsing the Twitterverse", in *Proceedings of AAAI-11 Workshop on Analysing Microtext*.
- [3] Gimpel, K., Schneider, N., O'Connor, B., Das, D., Mills, D., Eisenstein, J., Heilman, M., Yogatama, J., Flanigan, J., and Smith, N.A., "Part-of-speech tagging for Twitter: Annotation, features, and experiments. In *Proc. of ACL*",2011.
- [4] Lafferty, J., McCallum, A., and Pereira, F., "Conditional random fields: Probabilistic models for segmenting and labeling sequence data", in *Proceedings of ICML*, 2011.
- [5] Ritter, A., Clark, S., Mausam, and Etzioni, O., " Named entity recognition in tweets: An experimental study" in *Proceedings of EMNLP*, 2011.
- [6] Chen E., "Introduction to Conditional random Fields [online], Available at: <http://blog.echen.me/2012/01/03/introduction-to-conditional-random-fields/>
- [7] Clouder, A., "HMM, MEMM, and CRF: A Comparative Analysis of Statistical Modelling Methods" [online], Available at: [https://www.alibabacloud.com/blog/hmm%2C-memm%2C-and-crf%3A-a-comparative-analysis-of-statistical-modeling-methods\\_592049](https://www.alibabacloud.com/blog/hmm%2C-memm%2C-and-crf%3A-a-comparative-analysis-of-statistical-modeling-methods_592049)
- [8] Li S., "Part of Speech Tagging with Hidden Markov Chain Models" [online], Available at: <https://towardsdatascience.com/part-of-speech-tagging-with-hidden-markov-chain-models-e9fcc835c0e>
- [9] Liu, X., Zhang, S., Wei, F., and Zhou, M., Recognizing named entities in tweets. In *Proc. of ACL*, 2011.
- [10] Collister, L., "Meaning variation of the iconic deictic and in an online community. In *New Ways of Analyzing Variation*", 2011.
- [11] McCallum, A., "Efficiently inducing features of conditional random fields", in *Proceedings of the 19th Conference in Uncertainty in Artificial Intelligence (UAI-2003)*, 2003.
- [12] Analytics Vidhya, Complete Tutorial on Text Classification using Conditional Random Fields Model (in Python) [online], Available at: <https://www.analyticsvidhya.com/blog/2018/08/nlp-guide-conditional-random-fields-text-classification/>

## Language identification at word level in Sinhala-English code-mixed social media text

Kasthuri Shanmugalingam  
*Department of Computational Mathematics*  
*University of Moratuwa, Sri Lanka*  
*s.shanshiya@gmail.com*

Sagara Sumathipala  
*Department of Computational Mathematics*  
*University of Moratuwa, Sri Lanka*  
*sagaras@uom.lk*

### Abstract

*Automatic analyzing and extracting useful information from the noisy social media content are currently getting attention from the research community. It is common to find people easily mixing their native language along with the English language to express their thoughts in social media, using Unicode characters or the Unicode characters written in Roman Scripts. Thus these types of noisy code-mixed text are characterized by a high percentage of spelling mistakes with phonetic typing, wordplay, creative spelling, abbreviations, Meta tags, and so on. Identification of languages at word level become a necessary part for analyzing the noisy content in social media. It would be used as an intimidate language identifier for chatbot application by using the native languages. For this study we used Sinhala-English code-mixed text from social media. Natural Language Processing (NLP) and Machine Learning (ML) technologies are used to identify the language tags at the word level. A novel approach proposed for this system implemented is machine learning classifier based on features such as Sinhala Unicode characters written in Roman scripts, dictionaries, and term frequency. Different machine learning classifiers such as Support Vector Machines (SVM), Naive Bayes, Logistic Regression, Random Forest and Decision Trees were used in the evaluation process. Among them, the highest accuracy of 90.5% was obtained when using Random Forest classifier.*

**Keywords:** *Code-mixing, Language identification, Machine learning, Natural Language Processing (NLP)*

### 1. Introduction

Automatic analyzing and extracting useful information from noisy social media content are currently getting attention from the research community for for NLP [1]. The type of noisy social media text is characterized by a high percentage of spelling mistakes with phonetic typing ('good job putha'; since putha is a Sinhala word), wordplay ('helloooo' for 'hello'), creative spelling ('Gud 9t' for 'good night'), abbreviations ('TC'

for 'Take Care'), Meta tags (URLs), and so on. Currently, bilingual speakers use Unicode characters or Unicode characters written in Roman scripts to write in their own language and use phonetic typing, frequently adding English elements through the combination of multiple languages to express their thoughts. These type of text is called code-mixed text.

Make automatic language identification a precondition for the complete process of text analysis on social media. Although language identification is considered an almost solved problem in different applications [2], language detectors fail within the context of social media such as code-mixing, phonetic typing and lexical borrowing [3]. For this reason, the complexity of the analysis and understanding of information will increase within the context of social media. The reason for this limitation is due to the correct corpus acquisition. Automatic detection of the boundaries of the language in a code-mixed social media text, for English-Bengali, English-Hindi and Tamil-English has been proposed [3-5].

This study focuses on effectively detecting the language boundaries at word level of Sinhala-English code-mixed noisy text. Natural language processing and machine learning techniques are used to provide the solution. This study proposed a novel approach by adding Sinhala Unicode characters written in Roman scripts as new features including dictionaries, and term frequency as other features for machine learning classifiers.

Sinhala-English code-mixed data from popular social media posts and comments were taken as input for this purpose. An example is Sinhala - English code-mixed comment collected from Facebook and its corresponding annotations are shown in Table 1. This study includes (a) Feature identification and feature selection of language identification system for Sinhala-English code mixed data at word level; (b) Creation of annotated data set for collected Sinhala-English code-mixed data; (c) implementing the language identification system for the Sinhala-English code-mixed data, and (d) Evaluation of language identification with Sinhala-English code-mixed data.

**Table 1. Language-tagged sentence at the word level.**

Example code-mixed comment							
<b>Text</b>	screen	eke	pena	okkoma	photo	ekata	wadinawa
<b>Tags</b>	eng	sin	sin	sin	eng	sin	sin

This paper is divided into six sections. Section 1 gives an overall introduction to the paper. Section 2 critically reviews the work done on code-mixing and language identification with background information. Section 3 contains the details of the dataset used in this study. Section 4 presents a detailed methodology of a language identification system for Sinhala-English code mixing text. Section 5 reports on evaluation strategy with experimental results. Finally, section 6 concludes the paper with a note on possible further work.

## 2. Literature review

Code-mixing is a new phenomenon which has only attracted the attention of researchers in the last two decades. New complications have been added to the language Identification process under the context of code-mixed social media data.

### 2.1. Code-mixing

The types of traditional spoken and written languages have undergone great changes due to new forms of communication [6]. Internet and social media (particularly Facebook, Twitter, etc.) are the cause for these new forms. The sociological and conversational needs behind the code-mixing and its linguistic nature were mainly focused on the linguistic efforts in the field [7].

Mixing of two languages or more than two languages is referred to as Code-mixing. Two types of code-mixing can be categorized under this, namely Intra-sentential code-mixing and Inter-sentential code-mixing [4]. The code-mixing that takes place within sentence with no superficial change in topic is Intra-sentential code-mixing. The code-mixing that takes place when switching between native and second language to explain an incident is known as Inter-sentential code-mixing [8]. The users of Facebook have a tendency to use inter-sentential code mixing over intra-sentential code-mixing, and it is reported that real lexical needs was initiated 45% of the code mixing, 40% was to talk about a particular topic, and 5% to clarify the content [9]. Problems of storing and retrieving information from all types of sources, including social media, which is very difficult for tokenization and text processing are dealt by information retrievals [8]. Generally, it is difficult to understand and analyze texts written in multiple languages. An evaluation metrics is proposed to determine the complexity that occurs due to code-mixed

social media texts which gets developed rapidly due to multilingual interference [10].

### 2.2. Language identification

Joshi was the first one to work on code-mixing processing, more than 30 years ago, in 1982 [11], while efforts to develop automatic language identification tools began earlier [12]. Nevertheless, the solving of the problem of applying these language identification process to texts mixed in several languages, has only recently begun.

Automatic identification of word-level boundaries of different languages used in social media texts, illustrated by mixed English-Bengali and English-Hindi Facebook messages with standard techniques such as n-gram characters, dictionaries and, SVM classifiers [1]. Some different techniques were also used, including a simple unsupervised dictionary-based approach, supervised word-level classification with and without contextual clues, and sequence labeling using Conditional Random Fields. The dictionaries-based approach is surpassed by supervised classification and sequence labeling, and it is important to consider contextual clues [3]. Classification of Hindi-English code-mixed data was performed to categorize the data into English, Hindi, Mixed, Named Entity, Acronym, Universal, and undefined tags. Two types of embedding features were considered; character-based embedding features and word-based embedding features with the addition of context information. Support Vector Machine was used to train and test the system [13].

A word-based language identification system on mixed Turkish-Dutch messages randomly sampled from an online forum by comparing dictionary-based methods with language models and with logistic regression and conditional random fields with linear chain. This system achieved a high level of accuracy at the word level (97.6%), but with significantly lower accuracy at the post level (89.5%), although 83% of the messages were actually monolingual [14]. Similarly, using a bilingual case using Spanish-English Twitter messages, it uses only the ratio of the probability of words as a source of information and obtains good results, with a 96.9% accuracy at the word-level. However, the corpora are almost monolingual, so the result was obtained with a baseline of up to 92.3% [15].

The use of the most frequent word dictionaries is an established method in language identification. In this method, efficient and automatic segmentation of the input text in individual language blocks, in the case of multi-language documents [16]. However, this method has a number of challenges, especially for social media, as the text is full of noise.



Word-level language detection from code-mixed text can be defined as a classification problem. Support vector machines (SVMs) are one of the most popular methods for text classification, largely because of the automatic weighing of a large number of functions [17]. SVM is currently the most successful machine learning technique across multiple domains [18].

### 3. Dataset description

In the proposed system, Facebook comments and posts with Sinhala-English code mixed text were taken as input. All the posts and comments were broken down into sentences. Among them 1,500 Sinhala-English code mixed sentences were taken to the dataset. In this dataset 1,000 sentences were taken as training set and 500 sentences were taken as testing set. The statistics of the dataset used for training and testing are shown in Table 2.

**Table 2. Statistics of dataset.**

Sinhala-English code-mixed data	Train Data	Test Data
Sentences	1,000	500
Tokens	6,640	3,157
Words	2,896	1,579

The Sinhala-English code-mixed dataset annotated with language tags, by three human annotators using Tagtog [19] annotation tool. There are three main language tags, 'sin' for Sinhala words, 'eng' for English words and 'rest' for all other words. The 'rest' tag includes Named Entities, Acronyms, Universal, mixed and other language tags.

## 4. Proposed methodology

This section describes the language identification system for the Sinhala-English code-mixed data using feature-based embedding model.

### 4.1. Feature identification

Feature selection plays a significant part of this study. This section analyzed the pattern of the characters appears in the words from the Sinhala-English code-mixed dataset and identified relevant features among it. The below features were identified for language identification system.

#### 4.1.1. Sinhala Unicode characters in Roman scripts

Sinhala characters written in Roman scripts were taken as important features to identify Sinhala words. For example “ග” in Sinhala Unicode character written as “ga” in Roman scripts. The most common 284 Unicode characters used in writing, convert to Roman script. The Sinhala Unicode characters like “අ” and “ඇ” written same as “a” in Roman scripts in lower case. Further, the “න” and “ඬ” characters written same as “n” in Roman scripts. So among 284 Sinhala Unicode characters written in Roman script, the repeating 51 characters with same Roman script were removed and 233 Sinhala Unicode characters written in Roman scripts were taken as features. The features taken from Sinhala Unicode characters are shown in Table 3.

#### 4.1.2. Language-Specific Dictionaries

In order to identify the English words, two corpora were used as dictionaries. The corpora are British National Corpus (BNC) and LEXNORM corpus. Existence of a word in these two corpora were taken as two features.

BNC: A computer corpus of 100 million words of British English, written and spoken [20].

LEXNORM: A lexical normalization dataset released by Han et al. (2012) [21]. This dataset used to identify the spelling variations are expected in social media data.

#### 4.1.3. Term frequency

The frequency of each unique word was taken as a feature. Eventually, 236 features were taken for language identification system.

## 4.2. Proposed solution

In the proposed solution firstly, raw Sinhala-English code-mixed sentences were led to pre-processing steps. Then the pre-process data were taken as input for feature identification module for extracting identified above features from words. After that the feature-based embedding model was created according to selected classifiers in Weka. Ten-fold cross-validation was used to evaluate the performance based on language tags at word level. The output was recorded based on accuracy and the F-Measure value of each tag.

The top-level architecture of the proposed system is shown in Figure 1.

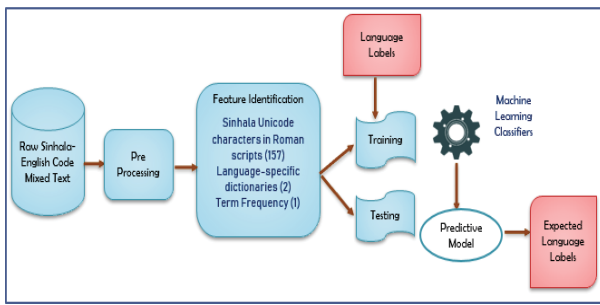


Figure 1. Top level architecture diagram of the proposed system.

Table 3. Sinhala Unicode characters.

Sinhala Alphabets															
a	aa	A	ae	i	ii	u	uu	e	ea	l	o	oe	au		
ආ	ආා	ආ	ආේ	ආී	ආීී	ආු	ආුු	ආේ	ආේා	ආී	ආී	ආේ	ආේා		
ka	kaa	kA	kae	ki	kii	ku	kuu	ke	kea	kl	ko	koe	kau	k	
ක	කා	කා	කේ	කී	කීී	කු	කුු	කේ	කේා	කී	කී	කේ	කේා	ක	
ba	baa	bA	bae	bi	bii	bu	buu	be	bea	bl	bo	boe	bau	b	
බ	බා	බා	බේ	බී	බීී	බු	බුු	බේ	බේා	බී	බී	බේ	බේා	බ	
ga	gaa	gA	gae	gi	gii	gu	guu	ge	gea	gl	go	goe	gau	g	
ග	ගා	ගා	ගේ	ගී	ගීී	ගු	ගුු	ගේ	ගේා	ගී	ගී	ගේ	ගේා	ග	
ma	maa	mA	mae	mi	mii	mu	muu	me	mea	ml	mo	moe	mau	m	
ම	මා	මා	මේ	මී	මීී	මු	මුු	මේ	මේා	මී	මී	මේ	මේා	ම	
cha	chaa	chA	chae	chi	chii	chu	chuu	che	chea	chl	cho	choe	chau	ch	
ච	චා	චා	චේ	චී	චීී	චු	චුු	චේ	චේා	චී	චී	චේ	චේා	ච	
ya	yaa	yA	yae	yi	yii	yu	yuu	ye	yea	yl	yo	yoe	yau	y	
ය	යා	යා	යේ	යී	යීී	යු	යුු	යේ	යේා	යී	යී	යේ	යේා	ය	
ja	jaa	jA	jae	ji	jii	ju	juu	je	jea	jl	jo	joe	jau	j	
ජ	ජා	ජා	ජේ	ජී	ජීී	ජු	ජුු	ජේ	ජේා	ජී	ජී	ජේ	ජේා	ජ	
ra	raa	rA	rae	ri	rii	ru	ruu	re	rea	rl	ro	roe	rau	r	
ර	රා	රා	රේ	රී	රීී	රු	රුු	රේ	රේා	රී	රී	රේ	රේා	ර	
la	laa	lA	lae	li	lii	lu	luu	le	lea	ll	lo	loe	lau	l	
ල	ලා	ලා	ලේ	ලී	ලීී	ලු	ලුු	ලේ	ලේා	ලී	ලී	ලේ	ලේා	ල	
da	daa	dA	dae	di	dii	du	duu	de	dea	dl	do	doe	dau	d	
ද	දා	දා	දේ	දී	දීී	දු	දුු	දේ	දේා	දී	දී	දේ	දේා	ද	
wa	waa	wA	wae	wi	wii	wu	wuu	we	wea	wl	wo	woe	wau	w	
ව	වා	වා	වේ	වී	වීී	වු	වුු	වේ	වේා	වී	වී	වේ	වේා	ව	
tha	thaa	thA	thae	thi	thii	thu	thuu	the	thea	thl	tho	thoe	thau	th	
ථ	ථා	ථා	ථේ	ථී	ථීී	ථු	ථුු	ථේ	ථේා	ථී	ථී	ථේ	ථේා	ථ	
sa	saa	sA	sae	si	sii	su	suu	se	sea	sl	so	soe	sau	s	
ස	සා	සා	සේ	සී	සීී	සු	සුු	සේ	සේා	සී	සී	සේ	සේා	ස	
dha	dhaa	dhA	dhae	dhi	dhii	dhu	dhuu	dhe	dhea	dhl	dho	dhoe	d hau	dh	
ඳ	ඳා	ඳා	ඳේ	ඳී	ඳීී	ඳු	ඳුු	ඳේ	ඳේා	ඳී	ඳී	ඳේ	ඳේා	ඳ	
ha	haa	hA	hae	hi	hii	hu	huu	he	hea	hl	ho	hoe	hau	h	
හ	හා	හා	හේ	හී	හීී	හු	හුු	හේ	හේා	හී	හී	හේ	හේා	හ	
na	naa	nA	nae	ni	nii	nu	nuu	ne	nea	nl	no	noe	nau	n	
න	නා	නා	නේ	නී	නීී	නු	නුු	නේ	නේා	නී	නී	නේ	නේා	න	
Na	Naa	NA	Nae	Ni	Nii	Nu	Nuu	Ne	Nea	Nl	No	Noe	Nau	N	
ඤ	ඤා	ඤා	ඤේ	ඤී	ඤීී	ඤු	ඤුු	ඤේ	ඤේා	ඤී	ඤී	ඤේ	ඤේා	ඤ	

The preprocessing steps include tokenization, convert the word in lowercase, removing wordplay characters and calculating the term frequency of each word.

Tokenizer in NLTK (Natural Language Tool Kit) was used to tokenize input data. The lower() function was used to convert the words in lowercase. The regular expression was used to remove the wordplay characters. The characters appear more than two times were eliminated. For example in the word “chooooooty” the character ‘o’ appears more than two times. So that, the word consider as wordplay word. So the regular expression used to replace the word “chooooooty” as “chooty”. After that, the term frequency of each unique word was taken using count() function.

Preprocessed data was embedded with language tags and identified features. The embedded data was taken as input for feature identification model, and the below feature identification was done in this process.

- **Sinhala Unicode characters in Roman scripts:** The presence of 233 features in the words were identified. This is a Boolean feature. Presence of each feature recorded as 1 or 0 otherwise.
- **English language-specific dictionaries:** Used to identify the presence of a word in dictionaries. This is a Boolean feature. Presence of word will be 1 or 0 otherwise. The presence of words in BNC and LEXNORM dictionaries were identified.

Now the trained data consists of words with embedded features such as Sinhala Unicode characters in Roman scripts, language-specific dictionaries, and term frequency.

Among this embedded features, some identified features had the mean value as zero. The 76 feature’s mean value identified as zero. These features did not play a significant role in the algorithm. Therefore, these features were removed from the training dataset.

The training dataset consists of 160 features. In addition to corresponding language tags, a total of 161 attributes were included in training dataset.

### 5. Evaluation

For the evaluation process, the Weka 3.9.2 tool was used. Different machine learning classifiers such as Support Vector Machine, Naïve Bayes, Logistic Regression, Random Forest and Decision Tree were used to evaluate the performance.

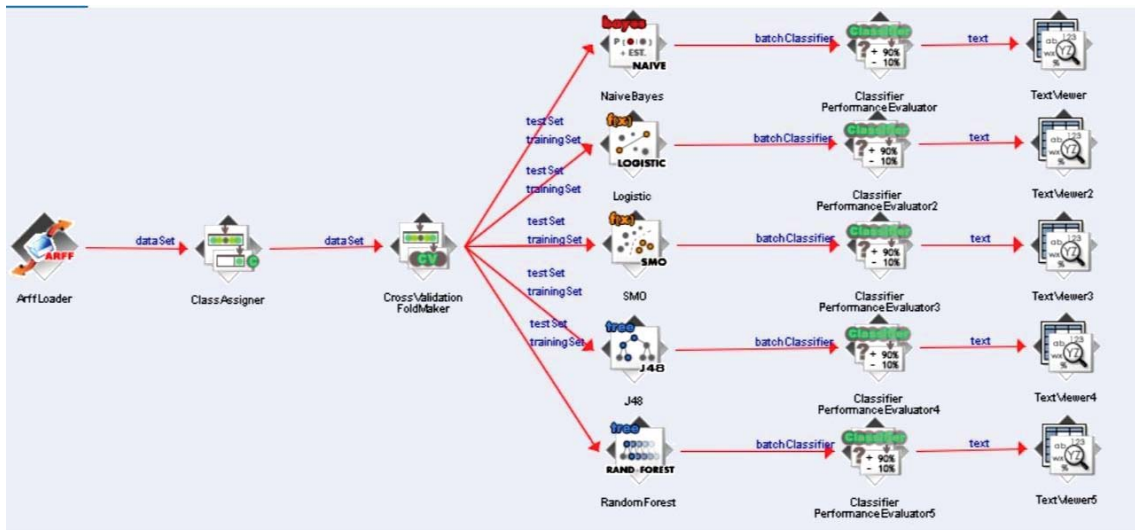


Figure 2. The knowledge flow of experiment process.

The feature-based model was created based on identified features and the dataset subjected to 10 fold cross-validation. The overall accuracy and F-Measure of different classifiers were recorded. The knowledge flow of experiment process is shown in Figure 2.

In order to identify the important features from feature set, feature evaluation was done with different classifiers by adding and removing the features. The accuracy obtained from different classifiers for different feature set was recorded. The features evaluation chart was illustrated in Figure 3. By observing the results most important features set were identified as Sinhala Unicode characters in Roman Script and English language-specific dictionaries.

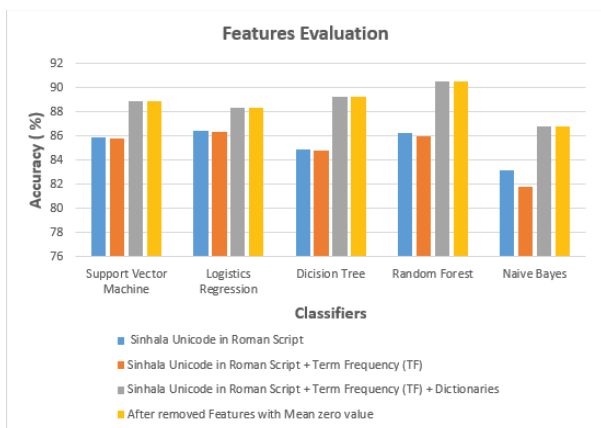


Figure 3. Evaluation of features importance with different classifiers.

The confusion matrix, overall accuracy and F-Measures of 'sin', 'eng', and 'rest' tags obtained from different classifiers are shown in Table 4.

Table 4. Overall results obtained from different classifiers.

Classifiers	Confusion Matrix	Accuracy	F-Measure (a=sin)	F-Measure (b=eng)	F-Measure (c=rest)
Support Vector Machine (Linear Kernel, C=1)	a 2274 44 6	88.92%	0.942	0.749	0.174
	b 106 283 8				
	c 125 32 18				
Logistics Regression (Ridge = 1.0E-8)	a 2241 55 28	88.33%	0.943	0.727	0.273
	b 86 280 31				
	c 100 38 37				
Decision Tree (Confident Factor=0.25, Num of objects=2)	a 2250 50 24	89.30%	0.943	0.731	0.474
	b 109 268 20				
	c 89 18 68				
Naive Bayes (Use Kernel Estimator=True)	a 2185 105 34	86.81%	0.936	0.715	0.087
	b 66 319 12				
	c 94 71 10				
Random Forest (Number of iteration = 100)	a 2280 24 20	90.50%	0.949	0.758	0.513
	b 118 271 8				
	c 82 23 70				

Among them, Random forest classifier gave 90.5% accuracy for language identification system for Sinhala-English code-mixed text at the word level. This model is good for identifying Sinhala and English language tags. Because the F-Measure for 'sin' and 'eng' tags were 0.949 and 0.758 respectively. However, Random forest model did not identify 'rest' tags accurately. The F-measure of 'rest' for Random forest classification model were 0.513. This was seen in that most words belonging to 'rest' tag were incorrectly classified in 'sin' and 'eng' tags. This happened because some Sinhala and English words were mixed with numbers. For example in "container1ka" word "container" belongs to English language, "ka" belongs to the Sinhala language and '1' belongs to number. Further, the name entities were identified wrongly in 'sin' and 'eng' tags.

## 6. Conclusion and further work

This paper discusses problems with code mixed data and proposed a feature-based embedded methodology to automatic language identification of Sinhala-English code mixed data. The methodology used for this system is a novel approach implemented as machine learning classifier based on features such as Sinhala Unicode characters in Roman scripts, dictionaries, and term frequency. Different machine learning classifiers such as SVM, Random Forest, Naive Bayes, Logistic Regression, and Decision Tree were used to evaluate the performance. Random Forest classifier achieved the highest accuracy of 90.5%.

In order to enhance the performance of the system, more features can be added. Pruning some features by parameter tuning with different classifiers may possibly improve the performance of the model. Further, the performance of the model can be evaluated with neural network techniques like multi-layer perception, recurrent neural networks and modern Deep Learning approaches.

## 7. References

- [1] Das, A. and Gambäck, B., “Code-mixing in social media text: the last language identification frontier?,” 2015.
- [2] McNamee, P., “Language Identification: A Solved Problem Suitable for Undergraduate Instruction,” *J Comput Sci Coll*, vol. 20, no. 3, pp. 94–101, Feb. 2005.
- [3] Barman, U., Das, A., Wagner, J., and Foster, J., “Code mixing: A challenge for language identification in the language of social media,” in *Proceedings of the first workshop on computational approaches to code switching*, 2014, pp. 13–23.
- [4] Das, A. and Gambäck, B., “Word Level Language Identification of Code Mixing Text in Social Media using NLP,” 2014.
- [5] Kasthuri, S. and Sagara, S., “Identifying languages at the word level in code-mixed indian social media text,” in *Proceedings of the Third International conference on Advances in Computing and Technology*, 2018.
- [6] Crystal, D., *Language and the Internet*. Cambridge [u.a.: Cambridge Univ. Press, 2008.
- [7] Muysken, P., “code-switching and grammatical theory,” p. 23.
- [8] Ajita, S. and Amit, K., “POS Tagging of Hindi-English Code Mixed Text from Social Media.”
- [9] Hidayat, T., *An analysis of code switching used by facebookers. Sekolah Tinggi Keguruan dan Ilmu Pendidikan (STKIP) Siliwangi Bandung*, 2008.
- [10] Ghosh, S., Ghosh, S., and Das, D., “Complexity Metric for Code-Mixed Social Media Text,” *ArXiv170701183 Cs*, Jul. 2017.
- [11] Joshi, A. K., “processing of sentences with intra-sentential code-switching,” in *Coling 1982: Proceedings of the Ninth International Conference on Computational Linguistics*, 1982.
- [12] Gold, E. M. “Language Identification in the Limit,” p. 28.
- [13] Veena, P. V., Kumar, M. A., and Soman, K. P., “Character Embedding for Language Identification in Hindi-English Code-mixed Social Media Text,” *Comput. Syst*, vol. 22, no. 1, Mar. 2018.
- [14] Nguyen, D. and Doğruöz, A. S., “Word Level Language Identification in Online Multilingual Communication,” in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, 2013, pp. 857–862.
- [15] Lignos, C. and Al, B. E., *Toward Web-scale Analysis of Codeswitching*. .
- [16] Řehůřek, R. and Kolkus, M., “Language identification on the web: Extending the dictionary method,” in *International Conference on Intelligent Text Processing and Computational Linguistics*, 2009, pp. 357–368.
- [17] Joachims, T., “Svmlight: Support vector machine,” *SVM-Light Support Vector Mach. Httpsvmlight Joachims Org Univ. Dortmund*, vol. 19, no. 4, 1999.
- [18] Baldwin, T. and Lui, M., “Language identification: The long and the short of the matter,” in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 2010, pp. 229–237.
- [19] Cejuela, J. M. et al., “tagtog: interactive and text-mining-assisted annotation of gene mentions in PLOS full-text articles,” *Database*, vol. 2014, no. 0, pp. bau033–bau033, Apr. 2014.
- [20] G. Nelson, “Guy Aston and Lou Burnard, *The BNC handbook: exploring the British National Corpus with SARA*. Edinburgh Textbooks in Empirical Linguistics. Edinburgh: Edinburgh University Press, 1998. Pp. 256. Hardback £43.50, ISBN 0 7486 1054 5; paperback £16.50, ISBN 0 7486 1055 3,” *Engl. Lang. Linguist.*, vol. 6, no. 1, pp. 197–221, May 2002.
- [21] B. Han, P. Cook, and T. Baldwin, “Automatically constructing a normalisation dictionary for microblogs,” in *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*, 2012, pp. 421–432.

## VoiceNote: An intelligent tool for monetary transactions with integrated voice support

Buddhima Hashan, Yoshani Abeyrathna,  
Mathisha Kaluaratchi,  
Arosha Piyalath  
Faculty of Computing, Sri Lanka Institute of  
Information Technology, Sri Lanka  
buddhimahashan@gmail.com,  
yoshanikavindya94@gmail.com,  
mathisha22@gmail.com,  
udam.arosha@gmail.com

Samantha Thelijjagoda  
Faculty of Business, Sri Lanka Institute of  
Information Technology, Sri Lanka  
samantha.t@sliit.lk

### Abstract

The smart phone has become a critical aspect of our daily lives, with many applications providing solutions to different problems that we encounter. Fintech have become a major priority in the present world. The use of petty cash has run its course throughout the world. Credit cards have proven to be successful in creating a cashless environment. Combining the smartphone with financial planning ability will be a powerful tool. In the meantime, customers in the modern society expect more convenient methods to make financial transaction with enhanced security measures. This paper describes how a QR based payment system which can be controlled by voice commands simplifies business operations between customer and vendor. QR code based mobile payments represent a relatively new technology that can be used in the new business arena of payments using smartphones. It provides a flexible and cashless environment for users. It is more secure and easier to; pay in supermarkets, shopping malls, as well as the user can share money with their friends and families, even can be manipulated by voice commands. The use of such devices in the payment process has lots of advantages in terms of speed and comfort when accessing services, and also improves the security and integrity of transactions, and the simplicity with user friendly interactions. To evaluate the potential of QR code-based voice support payment system, in the context of business transactions, VoiceNote updates monetary transaction in real-time and uses encryption and decryption methods for security. The transaction details can be analyzed and forecast through a web dashboard. VoiceNote research findings shows that transaction time is reduced by an average of 64 seconds when compared with a cash transaction and an average reduction of 13.5 seconds when compared with card transactions.

**Keywords:** Mobile payments, QR code payments, VoiceNote, Voice payments

### 1. Introduction

Today the world is moving towards mobile computing. Tasks done through mobile devices are increasing the easiness of doing everyday tasks. Day-to-day transactions can be done in cash or other payment methods. To make this process easy for consumers and vendors, several institutes are working on mobile payment systems. The major concerns on these payment systems are security and reliability.

Mobile phones have revolutionized human lives by bringing mobility to tasks that used to consume a lot of effort. The use of mobile phones has increased drastically with the passage of time. The mobile phone is the new gateway for all transactions [1]. At first, it was only used for small transactions and the medium was IVR or SMS, but today mobile phones can be trusted with bigger transactions and other mediums like QR Code, WLAN and NFC. The dark side of this new advances are security and integrity since there are ways of stealing valuable data. Therefore, users tend to lean toward technologies that make their transaction more secure and reliable. The information such as passwords and account numbers of users must only be confined to the relevant parties. This study was conducted to address the above concerns. The mobile phones are currently being used for the purchasing the music, pictures, and software applications. The mobile phone is entering to the new areas including the finance sector, payments through mobile phones is replacing credit cards. The rapid increase and growth of mobile payments has given permission to several rises like the breaching of security, privacy and reliability. Another factor that has to be looked into is user experience in these products. Transaction data is a valuable source which can

be utilized to give analytical insights into one's financial life. Therefore, in order to have a successful product, it has to be equipped with tight security measures and beneficial user experience features.

The main objective of VoiceNote is to develop a proper monetary management system to control the transaction between vendors and customers. Therefore, VoiceNote system is designed as a combination of mobile application, web application and web dashboard. VoiceNote mobile application reduces the time duration of transactions because VoiceNote is a cashless and cardless monetary management system. It also provides a solution for the shortage of Sri Lankan coins (loose change). VoiceNote budget plan feature will encourage the customer to have a proper financial and budget plan. Users can view and analyze their monetary behavior of a given period of time by the use of VoiceNote web dashboard. With the use of higher user-friendly graphical interfaces and good human computer interaction, VoiceNote provides a great user experience. VoiceNote ensures a high security measurement throughout the system.

## 2. Literature survey

There are many studies that have explained the essence of mobile payment through QR codes on smartphones. Some of the prominent issues include security, supply and demand problems, and user experience. Software agents will also gather the relevant information about the product and present it to the user, minimizing the errors.

According to a literature review of mobile payment research done by Dahlberg et al., most of the papers published covered technical issues of those payment systems and consumer-centric studies. This rather limited scope could be partly explained by the recent emergence of mobile payment research [2].

In light of the many past mobile payment system failures, there is a real need to analyze and understand what requirements are needed to succeed in this uncertain market. In fact, the technological trends are hard to predict as mobile technologies tend to behave as disruptive technologies. Interestingly enough, mobile payments services are currently under performing, though they are already deployed in niche markets.

It is a clear evidence that some technologies (e.g., RFID, NFC, QR Codes) are bringing better performance (e.g., speed) than traditional payment cards [3]. Therefore, it is important to replace traditional card-based systems with QR Code based technology for high performance.

Most of these traditional processes of payments and settlements involve a buyer-to-seller physical transfer of cash or a payment information exchange. There are no precautions to ensure security, controllability and

accuracy. This could also open an opportunity for QR Code payment method. However, the current phone-based payment systems are still in the development stage. There is still progress to be made in terms of ease of use, cost, reliability, and user/market acceptance. Therefore, QR Code technology becomes a value-added advantage for this domain to achieve these above targets.

In a study regarding consumer adoption of mobile wallets, Doan [4] explained that 'Mobile wallet is formed when your Smartphone functions as a leather wallet: it can have digital coupons, digital money (transactions), digital cards, and digital receipts'. Mobile wallet service allows the user to install an application from online stores in their smartphones and use them to pay for their online and offline purchases. Using latest technologies that connect smartphones to the physical world such as NFC (Near Field Communication), sound waves, and QR codes, cloud-based solutions, mobile wallets are believed to provide more convenient payment solutions to the customers in future [5].

According to Hoofnagle, et al. [6], payments made through wireless devices like mobile phones and smartphones are thought to provide more convenience, reduce the fee for the transaction, and increase the security of electronic payment. This payment system has also made it easier for businesses to collect useful information about their customers and their purchases. Paunov and Vickery [7] found the applicability of mobile payment systems to be quite wide due to the remarkable growth and greater penetration of mobile devices as compared to other telecommunication infrastructure.

According to the researchers' Mobile payment methods are suitable for offline micropayments as well as for online purchases. This method is a potential attraction for online traders due to an enormous user base of mobile phones. The use of mobile payment service does not only reduce the overall cost of a transaction but also offer better payment security.

## 3. Methodology

As discussed previously, there are many difficulties of traditional payment systems. Therefore, the study focused on developing a QR based voice support payment system to overcome difficulties such as time consumption, lack of security, unavailability of loose change and lack of user friendliness.

Our research mainly focuses on QR technology. This approach involved a structured process to build the QR based payment model. The system model is based on two major user access levels namely the Vender and Consumer.

The proposed system will assist in monetary transactions, budget planning and manage expenses.

The high-level architecture of the payment model is as shown in Figure 2. It shows how each module interacts with other modules in the payment model.

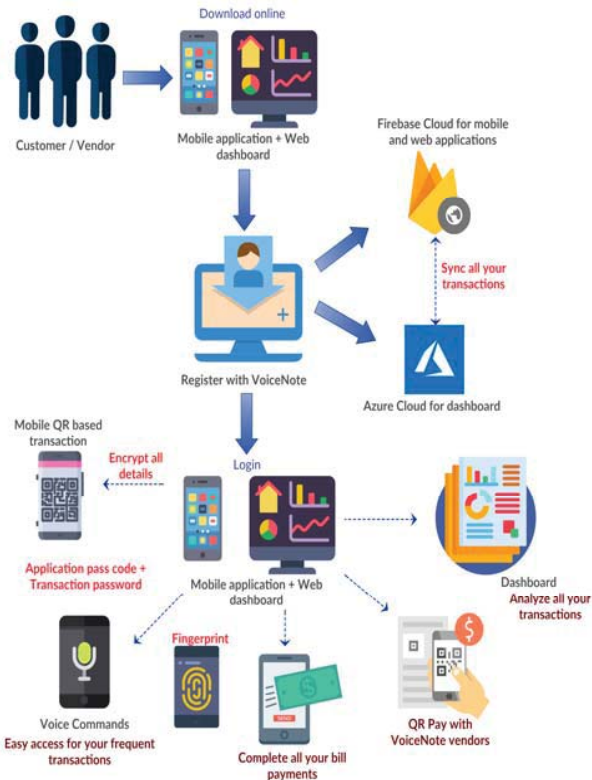


Figure 1. High-level architecture.

In order to initially assess the hypotheses, questionnaires were distributed to randomly selected vendors and customers. Based on the survey, the outcome of the following factors was determined.

### 3.1. Vendor and customer module

This module simplifies the vendor and the customer's role in the transaction. When the customer purchases the product, the vendor will prepare the bill by using their POS (Point of Sale) system with QR code. After the vendor prints the bill, the customer can easily settle the bill by scanning that QR code. After the bill details are received, the application uses a web service to verify the received data. The web server authenticates the data and sends back, then automatically the total bill value will be settled by the customer's credit or debit card through the "VoiceNote" application. In order to be secure, all the transaction data is encrypted.

### 3.2. Voice module

A user can give simple voice commands to get details from "VoiceNote" application. Such as, if a user wants to know the available balance in the Sampath bank card, user just needs to provide card nickname (which is easy to pronounce) and the system will identify the given voice input by using some algorithms and the relevant response will be generated. The fingerprint authentication is used to secure the voice module.

### 3.3. Security module

For the security purpose of the application, user can login to application by using fingerprint recognition. Other than this, VoiceNote provides special security features by using a backup password. User can use the backup password if the user is being robbed and if the user uses the backup password, the smartphone sends user's current location's longitude and latitude coordinates to the emergency contacts of the user or to the nearest police station.

### 3.4. Budget planning module

This module allows users to plan their monthly expenses at the beginning of each month. This is called as the budget feature. This option helps the user to divide their income under several main sections. As the user does transactions, the application will keep a track of records and notifies the user when they reach the allocated amount. At the end of each month, the user can generate a full report regarding his expenses.

### 3.5. Web module

Web application provides the basic transaction methods of mobile application such as mutual money transaction and bill payments. The user can generate reports of their monetary behavior of the day, week or month by using the web application. There will be a dashboard to view the activities of each day, weeks and months. The user can use some prediction methods to analyze monthly expenses and view it in the dashboard.

### 3.6. Databases

This application uses two main types of databases. NOSQL databases for mobile transactions handling and SQL RDMS for budget planning web application. SQL database uses to create Business Intelligence model inside budget planning module. When user is using the application, both databases are updated parallelly. Both

databases are hosted in the cloud to provide real time solutions to user.

VoiceNote payment system is unique and different from the existing payment mechanisms. It reduces the transaction time and provides more security than other payment systems.

#### 4. Results and discussion

In line with our assessment, many trials of the VoiceNote system showed greater results. Our focus is to illustrate the better performance of QR based voice support mobile payment system. For illustration purposes several trials with set of vendors and customers have been performed by us. The survey results are as follows;

*Benchmark – Vendor customer transactions:*

*Scenario 1: If transaction performs with cash*

**Table 1. Result when transaction perform with cash.**

Task	Related Time Frame
Get the wallet	3s
Choose money	3s – 5s
If the vendor doesn't have the change	1min – 3min
If the vendor has to provide change	6s – 10s
Get back the change	5s

Total time taken to complete the transaction (average) = 76s

*Scenario 2: If transaction performs with electronic card*

**Table 2. Result when transaction perform with electronic card.**

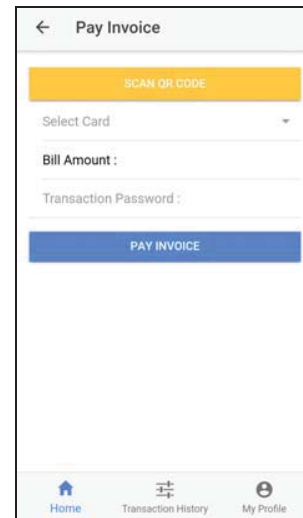
Task	Related Time Frame
Get the wallet	3s
Choose credit/debit card	3s – 5s
Virtual transaction	10s – 15s
Sign the receipt	3s
Put credit card back	3s

Average time taken to complete the transaction = 25.5s

*Scenario 3: If transaction performs with VoiceNote system (with QR code Option)*

The vendor will generate the QR code according to the bill details, scanning generates QR code through the

'VoiceNote' application customer can complete the payment.



**Figure 2. QR payment.**

**Table 3. Result when transaction performed with Voicenote.**

Task	Related Time Frame
Get the phone	3s
Login to the app	3s – 5s
Generate the QR code	1s
Perform the transaction (Scan QR code)	1s – 3s
Put the phone back	2s

Total time taken to complete the transaction (average) = 12s

*Benchmark – Mutual money transaction:*

*Scenario 1: If money deposited through the bank*

**Table 4. Result when money deposited through the bank.**

Task	Related Time Frame
Select the slip	10s – 12s
Fill the slip	40s – 1min
Wait in the queue	2min – 5min
Deposit the money in cashier	1min
Get the copy of the receipt	5s

Average time taken to complete the transaction = 4m 6s

*Scenario 2: If money deposited through the VoiceNote system (with QR code Option)*



The person who sends the money will generate the QR code using VoiceNote application. Then the receiver scans the QR code through the VoiceNote and completes the transaction.

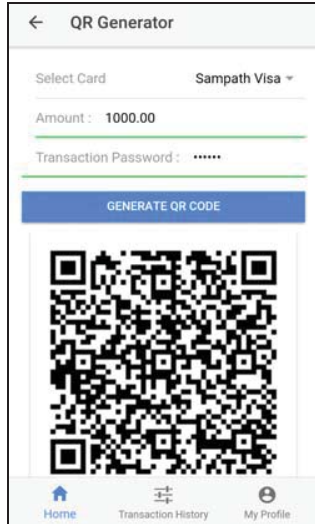


Figure 3. Generate QR code.

Table 5. Result when money deposit through the Voicenote.

Task	Related Time Frame
Get the phone	3s
Login to the app	3s – 5s
Generate the QR code	1s
Perform the transaction (Scan QR code)	1s- 3s
Put the phone back	2s

Average time taken to complete transaction = 12s

Scenario 3: If money deposited through the VoiceNote system (with Voice Option)

By using simple voice command, VoiceNote user can do the transaction. Other than this user can check balance by using voice commands.

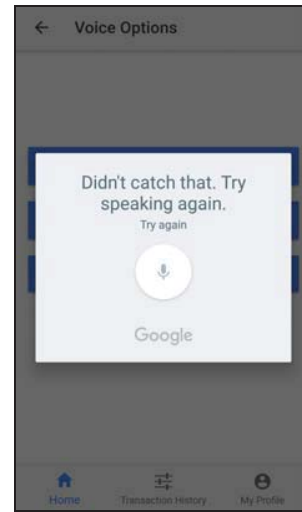


Figure 4. Voice payment.

Table 6. Result when money deposit through the Voicenote.

Task	Related Time Frame
Get the phone	3s
Login to the app	3s – 5s
Select the voice option and provide account nickname	3s
Provide amount	2s
Transaction complete	1s
Put the phone back	2s

Total time taken to complete the transaction (average) = 15s

According to above results it clearly shows how better QR and Voice payment works in the context of mobile payment systems.

We have seen many trials, system designs, and architectures developed for mobile payments. However it clearly shows that QR technology is the one that finally makes the mobile payments efficient and effective. The test figures prove that this is true.

According to the research results, in the Table 7, we have compared and contrast the gap between the existing systems and the VoiceNote.

We took Google Wallet, Apple Passbook and Lemon's Wallet as similar systems, with the VoiceNote system containing many features when compared with similar systems.

Only the VoiceNote system is capable of scanning and generating QR code for transactions within payments. Google Wallet and Lemon's Wallet only supports Android devices and Apple's Passbook only supports iOS devices yet VoiceNote system is support any operation system such as Android, iOS or Windows.

**Table 7. Research gap comparison.**

	<b>Google Wallet</b>	<b>Apple's Passbook</b>	<b>Lemon's Wallet</b>	<b>VoiceNote</b>
<b>QR Code</b>	Used NFC		Use PIN authorization	✓
<b>Device Support</b>	Only android mobile devices	Only devices based on IOS	Only android mobile phones	Any mobile device
<b>Users</b>	Only from US	Only Apple users	Only from US	For any organization
<b>Budget planning Features</b>	x	x	x	✓
<b>Security</b>	No in-built security	Integrated with Apple's security	Password security	Password, Fingerprint authentication, Backup passwords
<b>Monitor transaction on then and there</b>	x	x	x	✓
<b>Rules and regulation</b>	More	More	Less	Less
<b>Voice support</b>	x	x	x	✓
<b>Provide E-receipt</b>	x	x	x	✓
<b>Mutual money transaction</b>	x	✓	x	✓
<b>Transaction within banks</b>	x	x	x	✓

VoiceNote can also be used by any person or by any organization. Other than any similar monetary management systems, VoiceNote is capable with budget planning feature and forecasting and prediction feature. VoiceNote also has only less number of rules and regulations other than similar payment systems, therefore it is more convenient than using other monetary management applications. VoiceNote is mainly capable of integrated voice support and mutual money transferring feature. Therefore, user can check their bank account

balance and transfer money within VoiceNote application by giving voice commands.

## 5. Conclusion

Developing a reliable, secured, well performing monetary management system for the research context by considering existing monetary management applications in the current market was the main objective of the authors of this comprehensive report. Globally, researchers have created many E-wallets by using different types of technologies. Even though they provide solutions to users when doing monetary transactions, there were some limitations as well.

VoiceNote mobile and web application "Quick Response Code (QR Code) based payment solution with integrated voice support" is developed to overcome the existing monetary management issues.

After analyzing the data and information that was gathered from the observations and survey questionnaires as well as the inspiration which were mentioned in the global research papers and studies; the authors/research team were able to develop a reliable, secured, well performing monetary management application for both vendors and customers.

The research team has developed a Fintech [17] application to ease the day-to-day financial transactions for both vendors and customers as the outcome.

To do a transaction the VoiceNote mobile application has implemented a monetary service API. The authors have referred many technologies to implement the VoiceNote system. After comprehensive study on available techniques, choice of selecting Data Science technique have been made by the authors, which helped to process a large number of data to get reliable results of financial transactional behaviour of the users and predictions.

Since the services provided by them are only available to limited countries and users, there is a huge advantage for the product. Discussions were made between the research team and major telecommunication companies in Sri Lanka to adapt the service. Adding to this the authors had to make sure that developing a monetary transaction managing system is not against the law of the country. Therefore, many gazettes regarding the telecommunication were referred.

In order to perform this research, the VoiceNote team used more precise data from more experienced customers and vendors and used them to de-mystify the hype. Ultimately, the result shows that QR Code and Voice transaction performs much better than other existing payment technologies.

The findings of this study will lead to a revolutionized change in the mobile payment industry and it will assist stakeholders in overcoming bottlenecks

that they are facing with the existing payment technologies.

## 6. Future work

### 6.1. Connecting the application with a POS system

Implementing a POS system within the application to manage the inventory of the vendor. By introducing this method, vendors can easily manage their inventory using QR code. It will be easy to enter, edit and remove inventory and also vendor can get the re-ordering point of the inventory. This will make vendors work easy.

### 6.2. Pay utility bills using QR code

Introducing this method users can directly pay their utility bills by scanning QR code. The relevant amount will be paid to the utility providing companies. It will be an efficient way of settling utility bills. User has to scan the QR code to pay the bill.

### 6.3. Retina scan for biometric authentication

A retinal scan is a biometric technique that uses the unique patterns on a person's retina blood vessels [8]. The application's security can be improved by using the retina scan. Logging of the application, monetary transaction confirmation, bill payments can be secured by using the retina scan.

### 6.4. Queued customers

The vendor application uses a server request to verify the data received by the customer slowing the transaction.

By introducing this method vendors are hoping to get the user information while the customer is waiting in the line. Then when the customer is ready to make the transaction the data will be already be there in a local data base making the whole transaction local and efficient.

## 7. References

- [1] Adams, D., Nelson, R. and Todd, P. (1992). Perceived Usefulness, Ease of Use, and Usage of Information Technology: A Replication. *MIS Quarterly*, 16(2), p.227.
- [2] Tomi Dahlberg, Niina Mallat, and Anssi Öörni, "Trust Enhanced Technology Acceptance Model - Consumer Acceptance Of Mobile Payment SOLUTIONS: Tentative evidence.," 2008.
- [3] Anon, (2018). [online] Available at: <http://www.nxp.com/applications/secure-transactions/mobiletransactions.html> [Accessed 12 Mar. 2018].
- [4] Doan, Ngoc, (2014). Consumer adoption in Mobile Wallet. The Turku University of Applied Sciences.
- [5] Husson, Thomas, (2015). The Future of Mobile Wallets lies beyond Payments. U.S.A.: Forrester Research Inc. ISACA, 2011. Mobile Payments: Risk, Security and Assurance Issues. U.S.A.: ISACA Emerging Technology White Paper.
- [6] Hoofnagle, C., Urban, J. and Li, S. (2012). Mobile Payments: Consumer Benefits & New Privacy Concerns. BCLT Research Paper.
- [7] Aunov, Caroline and Vickery, Graham, (2006). Online Payment systems for E-Commerce. Organization for Economic Co-operation and development (OECD).
- [8] En.wikipedia.org. (2018). Retinal scan. [online] Available at: [https://en.wikipedia.org/wiki/Retinal\\_scan](https://en.wikipedia.org/wiki/Retinal_scan) [Accessed 7 Aug. 2018].

## Passion Fruit disease detection using image processing

S.B.D.H. Dharmasiri  
 Department of Industrial Management  
 Faculty of Science  
 University of Kelaniya  
 Sri Lanka  
 sbdeshan@gmail.com

S. Jayalal  
 Department of Industrial Management  
 Faculty of Science  
 University of Kelaniya  
 Sri Lanka  
 shantha@kln.ac.lk

### Abstract

Fruit diseases are a major problem in economic and production losses in the agricultural industry worldwide. In this paper, an image processing approach is proposed for identifying passion fruit diseases. According to the Sri Lankan context, treatment details are taken by the farmers from the field officers. However, it can take a few days. So, this proposed system can be used to identify passion fruit diseases quickly and automatically. This proposed approach is composed of the following main steps; Image Acquisition, Image Preprocessing, Image Segmentation, Feature Extraction, Dataset Preparation, Training & Testing. Healthy and two types of passion fruit diseases, namely passion fruit scab and woodiness images were used for this approach. This approach was tested according to passion fruit disease type and its' stages, such as mild, moderate and severe. K-Means clustering was used for segmentation. Images were clustered according to  $k$  values, such as 2, 4, 6 and 8. Before the segmentation, images were converted to RGB,  $L^*a^*b$ , HSV and Grey colour models, because of find out the most suitable colour model for this approach. Local Binary Pattern was used for feature extraction and Support Vector Machine was used for creating the model. Seventy percent (70%) of each dataset was used to train the SVM and the other thirty percent (30%) was used to test the model. According to this approach, passion fruit diseases can be identified in the average accuracy of 79% and its' stage can be identified in average accuracy 66%.

**Keywords:** K-Means Clustering, Local Binary Pattern, Support Vector Machine,  $L^*a^*b$

### 1. Introduction

Passion fruit (*Passiflora edulis*) is a very famous fruit in modern society because of its nutrients and health benefits, such as prevention of cancer, controlling blood pressure and preventing hyperlipidemia. It has a huge demand in both local and international markets. Passion

fruit is cultivated in countries like Australia, New Zealand, Kenya, South Africa, South America, Hawaii, India and Sri Lanka. Currently, South America has become the largest passion fruit producer in the world [1]. The annual passion fruit production is around 0.8 million MT in the world.

Nowadays many people tend to cultivate passion fruit in Sri Lanka, because of favourable climate conditions, temperature and attractive market demands. There are two varieties of passion fruits in Sri Lanka. One is yellowish and another one is purplish. However, the yellowish passion fruit variety is widely cultivated in Sri Lanka, which is sourer than the purplish [2]. Kalutara, Gampaha, Ratnapura and Kurunegala are the districts, which are popular in cultivating passion fruit. The average annual passion fruit production is 500MT in Sri Lanka. In 2017, total passion fruit cultivation extent was 470ha in Sri Lanka. Figure 1 shows the total passion fruit production in each previous year, according to census and statistics data.

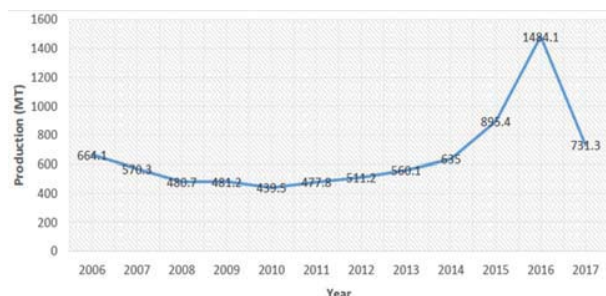


Figure 1. Passion fruit production in previous years.

Figure 1 shows how the production of passion fruit in Sri Lanka since 2006. The production of passion fruit was steadily grown between 2006 and 2014. An average 500MT can be observed during this period. In 2014, there was a 635MT production and it increased sharply until 2016 in which the production of passion fruit recorded the peak of 1484MT. After 2016, production declines suddenly. In 2017, passion fruit production was 731MT from 470.27ha extent. When considering the ratio of production and hectare in 2016 and 2017, it recorded

2.92 (MT per ha) in 2016 and 1.56(MT per ha) in 2017 respectively. The decline of this ratio is 1.36(MT per ha). When considering only the production, it was around 50.7% decline in 2017.

This production loss was around Rs.90 million in 2017 [3]. Fruit Research and Development Institute (FRDI) mentions "Pests, diseases and bad climate mainly caused the decline of the production in 2017". After spreading diseases, farmers have to use fungicides, remove infected parts and burn them with a systematic way. So, initial disease identification is very important to prevent spreading diseases. However, this prevention cannot be successful due to the lack of agricultural field officers.

Nowadays technology plays a vital role in all the fields but till today traditional methodologies are used for agriculture in Sri Lanka. However, the large scale of agricultural countries uses technologies like MRI, X-ray imaging for detecting the quality of the fruits. But these technologies are costly for farmers to afford. In Addition, to occupy large space, to have the knowledge of using and analyzing the results, have made it difficult for the users.

However, today, passion fruit disease identification is done manually by experienced people. But due to so many environmental changes and lack of resources for getting information, the prediction is becoming tougher. So, the main purpose of this research is to develop the classifier model using image processing, which will be able to identify passion fruit diseases accurately.

## 2. Literature review

Recently, many people have done researches for detecting fruit and vegetable diseases using image processing and deep learning. According to the research paper [4], authors had used image processing technology to identify the pomegranate diseases. Image preprocessing was the first step of the methodology. Image resizing was done under the image preprocessing. Because digital camera had been used to capture the images in this study, the size of those images was very large and takes more time to process. So, all images were resized to 300 x 300 PX. Morphology, colour and CCV features were used for feature extraction. K-means clustering technique was used for partitioning the training dataset according to their features. After the clustering, SVM was used for classification to identify the image as infected or non-infected. An intent search technique was provided to find the user intention. The best result was got using morphology feature extraction. Experimental evaluation of this approach was effective and 82% accurate to identify pomegranate disease.

In paper [5], the authors presented the image processing based approach for fruit disease detection.

First, read input image and transformed it from RGB to L\*a\*b colour space because, the colour information in the L\*a\*b colour space is stored in only two channels. Input images were partitioned into four segments using K-means cluster in this research because, the empirical observations, it was found that using 3 or 4 clusters yield good segmentation results. GCH, LBP, CCV and CLBP were used for feature extraction. More accurate results could be taken using CLBP feature extraction technique. K-means clustering was used for segmentation. Those segmented images were extracted to label each pixel in the image. SVM algorithm was used for training and classification of fruit disease. Authors used apple as a test case and evaluated the classification model for three types of apple diseases, which were apple rot, apple blotch and apple scab. The accuracy of this approach was achieved by up to 93%.

In paper [6], the author had used SVM classification for identifying and classifying the grape leaf diseases. Grape leaf images were taken using a digital camera and those were used to both training and testing the system. Collected images included the leaves infected by Powdery Mildew and Downy Mildew. Removing background noise and resizing to 300\*300 PX to improve the image quality were done under the image preprocessing. Gaussian filtering had been used to remove noise in the image. K-means clustering was used for segmenting an image into three groups. Features were extracted based on both colour and texture for taking accurate disease information. Finally, the classification model was used to detect the leaf disease. LSVM was used in this research for the classification of leaf diseases. This system could detect and classify the examined disease successfully. The accuracy of this system was 88.89%.

In paper [7], authors had used image processing technology for identifying the leaf diseases. First authors selected the plants, which were affected by the disease and then took the snapshot of the diseased leaf. Contrast enhancement and converting RGB to HIS was done under the image preprocessing step. K-means clustering algorithm was used to cluster the object based on the feature of leaf into k number of groups. SVM algorithm had been used in this system for classification purpose. SVM is a statistical learning-based solver. Finally, when entered a diseased leaf image to a system, the system was able to detect the leaf disease successfully.

In paper [8], authors had presented the image processing based system to identify pomegranate fruit diseases. This fruit is mainly affected by Bacterial Blight, Anthracnose and Alternaria. After capturing the disease images, image resizing, filtering, segmentation, morphological features were used to preprocess the images. Image segmentation is the process of dividing the image into multiple parts. Colour-based segmentation

was used in this research, such as clustering, YCbCr, RGB, L\*a\*b and HSV. However, the best performance in terms of segmentation error was achieved by the HSV and YCbCr. Morphology, texture and colour features were extracted for classification purpose. HIS colour model and colour histogram techniques had been used to colour feature extraction. Under the morphology feature extraction, boundary extraction was used to identify the region and shape. The eroded images were subtracted from the original image to extract a shape from healthy fruit image. Gabor filter was used to texture feature extraction. After the training and testing of images, diseased and non-diseased fruits were classified using Minimum Distance Classifier (MDC)

### 3. Methodology

There are six phases in this methodology. Those are Image Acquisition, Image Preprocessing, Image Segmentation, Feature Extraction, Dataset Preparation, Training and Testing.

#### 3.1. Image acquisition

In this phase, the sample images are collected, which are required to train the classifier algorithm and build the classifier model. Yellowish passion fruit variety was selected to take sample images because, the yellowish variety is widely cultivated in Sri Lanka. Healthy and diseased passion fruit images were taken by using 13MP mobile phone digital camera and used for both training and testing the classifier algorithm. Images were taken in different angles, under the different environmental and lighting conditions. The standard JPG format was used to store these images. In this study, images were collected from farms in different regions like Monaragala, Gampaha and Kaluthara because, according to the statistic and census data, high passion fruit production was showed in these districts in 2015. Scab disease and woodiness virus were selected for training and testing processes because, passion fruits infected by scab disease and woodiness virus, had been included in collected images. Images are shown by Figure 2, Figure 3, and Figure 4, which are the categories of passion fruits that are collected under the data acquisition



**Figure 2.**  
Healthy passion.



**Figure 3.**  
Scab disease.



**Figure 4.**  
Woodiness virus.

#### 3.2. Image preprocessing

After the image acquisition, image processing was done for improving the image quality. All original passion fruit images were stored in one folder, called “Images\_Original”. Those images were named as img\_x. “x” can take any value of numbers. Only horizontal images were rotated by 90 degrees and resized by 200x300 pixels. Vertical images were resized by 200x300 pixels and when the width and height of the image are kept same and those images were resized to 250x250 pixels. When the image size is too large, the processing task takes more time [4]. After that, one of the noise reduction methods was used to remove the noises from images and increase the sharpness of images. Later, all preprocessed images were saved in a folder, called “Images\_Preprocessed”.

#### 3.3. Image segmentation

The third phase of the methodology is image segmentation. As the first step, all preprocessed images were converted into L\*a\*b, HSV, Grey colour models and kept one in the original way (RGB) because, the identifying suitable colour model for preprocessing is one of the outcomes of this research. After converting them by using colour models, those images were renamed as mentioned in Table 1.

**Table 1. Image naming according to colour model.**

Colour model	After using colour models
RGB	img_x_0
L*a*b	img_x_1
HSV	img_x_2
Grey	img_x_3

After that, the image was converted to float32 format matrix. This matrix values were clustered using the K-means clustering algorithm. Four values: 2, 4, 6 and 8; were taken as the k value for this clustering mechanism. Clustered matrix values were converted to an image. According to the k values, clustered images were renamed as the table 2 and each image was saved in a specific folder. An L\*a\*b coloured image is mentioned in this following Table 2 as the example.

**Table 2. Image naming according to clustering.**

K value	After clustering
2	img_x_1_2
4	img_x_1_4
6	img_x_1_6
8	img_x_1_8

Figure 5 shows the sample images, which were received after the image segmentation.

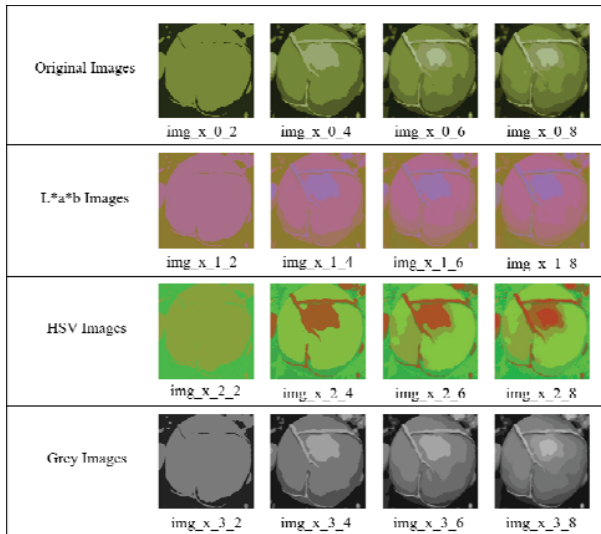


Figure 5. Image segmentation result images.

### 3.4. Feature extraction

The fourth phase of the methodology is feature extraction. Local Binary Pattern (LBP) mechanism was used for feature extraction. Segmented images were taken from folders one by one and each image was converted to grey-scale. Then a neighbourhood was selected for each pixel in grey-scale image surrounding the center. An LBP value was calculated for this center pixel and stored the output 2D array with the same width and height as the taken image.

Then calculating the LBP for center pixel was started from top right neighbourhood pixel clockwise. The results of binary values were stored in an 8-bit array, which was converted to decimal later. These converted decimal values were stored in the output LBP 2D array. A histogram for each image was computed using the output 2D array. For this implementation, the 3x3 pixel combination was used. So 256 ( $2^8$ ) number of feature histogram values were received at the end. Those histogram values were used for creating the datasets. Feature extracted values of RGB, L\*a\*b, HSV and Grey images were taken into comma-separated value (CSV) file. So at the end, four CSV files were created.

### 3.5. Dataset preparation

The fifth phase of the methodology is datasets preparation. Labels could be added to CSV files, which were created using feature extraction. However, three image sets were created to do experiments. Preparation of those image sets is discussed here. Field expertise

support was taken for the categorization of images and each image was selected randomly from the categorized sets of image.

#### 3.5.1. Considering disease name only

There are three categories in this image sets. Those are non-disease, scab disease and woodiness virus. A number of images were taken into each category is mentioned in the below tables. Table 3 shows the labels, which were assigned for each category. Table 4 shows the count of images, which were included in each dataset.

Table 3. Labelling according to disease name.

Non-disease	Scab disease	Woodiness virus	Mild	Moderate	Severe	Label
1	0	0	0	0	0	100000
0	1	0	0	0	0	010000
0	0	1	0	0	0	001000

Table 4. Three classes of datasets image amounts.

Disease name	Non-disease	Scab disease	Woodiness Virus
No. of images (1 <sup>st</sup> dataset)	13	34	40
No. of images (2 <sup>nd</sup> dataset)	26	67	80
No. of images (3 <sup>rd</sup> dataset)	39	98	120

#### 3.5.2. Considering disease names and its stages

Stages of the diseases are considered here. There are three types of stages, such as mild (I), moderate (II) and severe (III). According to those stages, there are seven categories in this image sets. Those are non-disease, mild-scab, moderate-scab, severe-scab, mild-woodiness, moderate-woodiness and severe-woodiness. Table 5 shows the labels, which were assigned for each category. Table 6 shows the count of images, which were included in each dataset.

**Table 5. Labelling according to diseases' stage.**

Non-disease	Scab disease	Woodiness virus	Mild	Moderate	Severe	Label
1	0	0	0	0	0	100000
0	1	0	1	0	0	010100
0	1	0	0	1	0	010010
0	1	0	0	0	1	010001
0	0	1	1	0	0	001100
0	0	1	0	1	0	001010
0	0	1	0	0	1	001001

**Table 6. Seven classes of datasets image counts.**

Disease	Non-disease	Scab Disease			Woodiness Virus		
		I	II	III	I	II	III
1 <sup>st</sup> set	13	12	12	10	13	15	12
2 <sup>nd</sup> set	26	24	23	20	26	30	24
3 <sup>rd</sup> set	39	36	33	29	39	45	36

These image sets were preprocessed, segmented and feature extracted separately. Four CSV files were received for every image set. Original image name and label included into one CSV file, which was called a “label file”. This file was used for giving labels for each row in feature extracted CSV file. Name column in feature extracted and label files were matched. When similar names were found, a particular label in label file was given to label column in the feature extracted file. After that, this dataset can be used for training and testing the model.

A sample feature CSV file is shown in Figure 6. Image name, feature histogram values and labels are included in this dataset.

Image Name	V_1	V_2	.....	V_255	V_256	Label
img_1_3_2	882	155	.....	421	51975	100000
img_10_0_4	1629	278	.....	750	42711	010000
img_2_1_2	1938	274	.....	886	37969	001000
img_50_2_6	2128	291	.....	989	33197	010100
img_99_3_8	988	175	.....	482	50535	010010
img_14_2_2	1698	262	.....	702	32531	001100
img_55_3_4	995	167	.....	900	33173	001010

Features histogram values

**Figure 6. Sample dataset.**

### 3.6. Training and testing

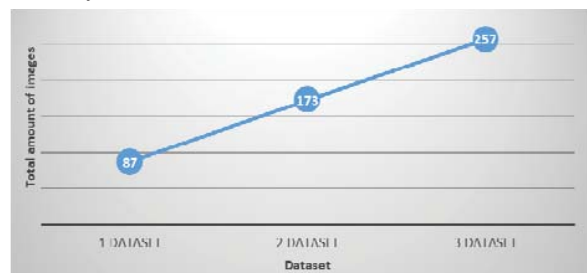
The last phase of the methodology is training and testing the model. Support Vector Machine (SVM) algorithm was used for classification as mentioned in the methodology. A set of mathematical functions are used in SVM algorithms, which are defined as kernels. The function of the kernel is to take data as input and transform those data into the required form. SVM algorithms are used with different types of kernels. Gaussian, sigmoid, hyperbolic tangent and polynomial kernel are some of them. Among those kernels, the polynomial kernel is popular in image processing. So, the polynomial kernel was used for the SVM algorithm as the kernel in this study. Each CSV file was used for training the SVM separately. Seventy percent (70%) of the data from each dataset was used for training purpose. Every training time, CSV file rows were shuffled for increasing the accuracy. When considering the testing the model, the remaining 30% of data from each dataset were used. In here, the model predicted the label for each image and those were evaluated using actual labels. The accuracy of these models can be defined as,

$$Accuracy (\%) = \frac{\text{Total number of images correctly classified}}{\text{Total number of images used for testing}} * 100$$

### 4. Experimental results

While labelling the features CSV file, three labels were used for identifying passion fruit disease according to its name. These CSV files are called as “3 classes dataset”. Another way is using seven labels for features CSV file to identify passion fruit diseases according to its stage. This CSV files are called as “7 classes dataset”. Every training and testing time, rows of CSV files were shuffled randomly for increasing the accuracy of the model. Each CSV file was trained and tested in five times and accuracy was taken. Average of those accuracies was taken as the accuracy of each model. Figure 7 shows the total number of images, which were used to create each dataset.

Using this image dataset, four types of CSV files were created. Such as Grey, HSV, L\*a\*b and RGB, which are the colour models that were used to evaluate in this study.



**Figure 7. Total number of images in each dataset.**



## 4.1. Results according to 3 classes datasets

### 4.1.1. First dataset results

Three hundred forty-eight image records are included in these CSV files and labelled by 3 labels. Each CSV file was used five times for getting accuracy and by using those accuracies; the average accuracy of the model was taken. Those models accuracies are discussed in Table 7.

**Table 7. Accuracy of models using 3 classes of first dataset.**

	1 <sup>st</sup> time	2 <sup>nd</sup> time	3 <sup>rd</sup> time	4 <sup>th</sup> time	5 <sup>th</sup> time	Average Accuracy
Grey	76.9	75.0	68.3	71.2	69.2	72%
HSV	67.3	75.9	69.2	72.1	69.2	70%
L*a*b	78.9	73.1	78.9	81.7	81.7	79%
RGB	70.2	69.2	75.0	63.5	71.2	70%

### 4.1.2. Second dataset results

Six hundred ninety-two image records are included in these CSV files and labelled by 3 labels. Each CSV file was used five times for getting accuracy and by using those accuracies, the average accuracy of the model was taken. Those models accuracies are discussed in Table 8.

**Table 8. Accuracy of models using 3 classes of second dataset.**

	1 <sup>st</sup> time	2 <sup>nd</sup> time	3 <sup>rd</sup> time	4 <sup>th</sup> time	5 <sup>th</sup> time	Average Accuracy
Grey	63.9	68.3	69.2	74.0	67.3	68%
HSV	65.4	59.1	52.4	65.4	60.9	60%
L*a*b	72.1	73.6	75.5	69.2	69.7	72%
RGB	71.2	67.3	67.3	67.3	73.6	69%

### 4.1.3. Third dataset results

One thousand twenty-eight image records are included in these CSV files and labelled by 3 labels. Each CSV file was used five times for getting accuracy and by using those accuracies, the average accuracy of the model was taken. Those models accuracies are discussed in Table 9.

**Table 9. Accuracy of models using 3 classes of third dataset.**

	1 <sup>st</sup> time	2 <sup>nd</sup> time	3 <sup>rd</sup> time	4 <sup>th</sup> time	5 <sup>th</sup> time	Average Accuracy
Grey	58.8	59.4	65.9	61.7	60.7	61%
HSV	56.1	58.4	56.5	54.9	57.4	56%
L*a*b	68.5	64.3	61.4	65.9	67.2	65%
RGB	58.8	64.3	60.4	58.1	59.4	60%

## 4.2. Results according to 7 classes datasets

### 4.2.1. First dataset results

Three hundred forty-eight image records are included in these CSV files and labelled by 7 labels. Each CSV file was used five times for getting accuracy and by using those accuracies, the average accuracy of the model was taken. Those models accuracies are discussed in Table 10.

**Table 10. Accuracy of models using 7 classes of first dataset.**

	1 <sup>st</sup> time	2 <sup>nd</sup> time	3 <sup>rd</sup> time	4 <sup>th</sup> time	5 <sup>th</sup> time	Average Accuracy
Grey	64.4	63.5	57.7	64.4	59.6	62%
HSV	63.5	57.7	55.8	52.9	56.7	57%
L*a*b	65.4	66.4	68.3	70.2	62.5	66%
RGB	63.5	63.5	59.6	70.2	61.5	64%

### 4.2.2. Second dataset results

Six hundred ninety-two image records are included in these CSV files and labelled by 7 labels. Each CSV file was used 5 times for getting accuracy and by using those accuracies, the average accuracy of the model was taken. Those models accuracies are discussed in Table 11.

**Table 11. Accuracy of models using 7 classes of second dataset.**

	1 <sup>st</sup> time	2 <sup>nd</sup> time	3 <sup>rd</sup> time	4 <sup>th</sup> time	5 <sup>th</sup> time	Average Accuracy
Grey	53.4	54.8	49.5	51.4	56.7	53%
HSV	45.7	48.6	51.9	49.0	52.9	50%
L*a*b	60.6	61.1	62.0	58.7	60.6	61%
RGB	61.5	59.1	58.2	58.7	52.9	58%

### 4.2.3. Third dataset results

One thousand twenty-eight image records are included in these CSV files and labelled by 7 labels. Each CSV file was used 5 times for getting accuracy and by using those accuracies, average accuracy of model was taken. Those models accuracies are discussed in Table 12.

**Table 12. Accuracy of models using 7 classes of third dataset**

	1 <sup>st</sup> time	2 <sup>nd</sup> time	3 <sup>rd</sup> time	4 <sup>th</sup> time	5 <sup>th</sup> time	Average Accuracy
Grey	47.1	49.7	53.3	53.9	49.7	51%
HSV	47.4	40.3	41.2	43.8	47.1	44%
L*a*b	56.8	54.9	63.6	58.8	62.3	59%
RGB	47.4	44.2	45.8	51.9	47.4	47%

### 4.3. Accuracy comparison

#### 4.3.1. According to 3 classes datasets

Figure 8 shows the accuracy comparison according to 3 classes datasets. L\*a\*b gives the highest accuracies and HSV gives the lowest accuracies among the preprocessing colour models. Grey and RGB take a similar place in Figure 8 when considered as overall. According to Figure 8, the accuracy of each model decreases within the number of images in each dataset.

#### 4.3.2. According to 7 classes datasets

Figure 9 shows the accuracy comparison according to 7 classes datasets. L\*a\*b gives the highest accuracies and HSV gives the lowest accuracies among the preprocessing colour models. When considering the grey colour model, it takes the second place of the accuracy comparison chart. The third place of accuracy comparison chart is taken by the dataset, which images were not used any preprocessed colour model. That is the RGB colour model. According to figure 9, the accuracy of each model decreases, when increasing the number of images in each dataset.

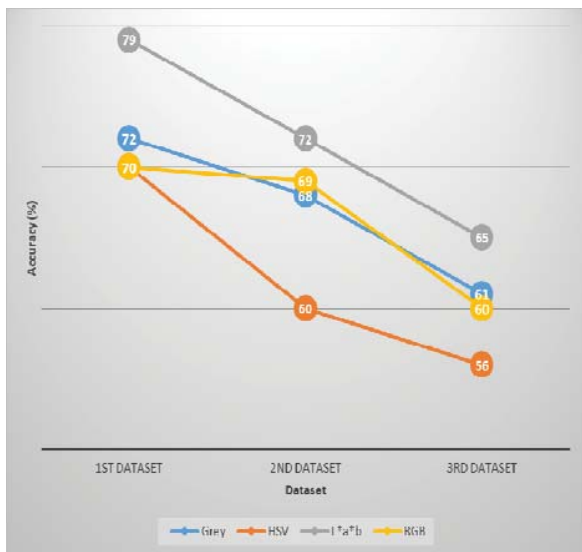


Figure 8. Accuracy of models using 3 classes datasets.

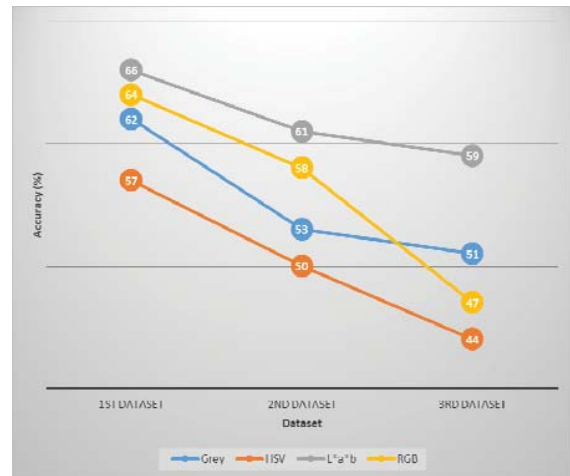


Figure 9. Accuracy of models using 7 classes datasets.

## 5. Conclusions

In this study, the support vector machine algorithm was used for creating the models, which were built according to passion fruit diseases and its stages. The passion fruit diseases can be identified in the average accuracy of 79% and its stage can be identified in the average accuracy of 66%.

When considered the 3 classes datasets, the highest average accuracy of the model was received by the first dataset, which included non-disease (13), scab disease (34) and woodiness virus (40). A number of images in each category has been mentioned in parenthesis. There are 87 images which were included in this dataset. The accuracy was 79%, which was received using L\*a\*b colour model.

When considered the 7 classes datasets, the highest average accuracy of the model was received by the first dataset, which also included 87 images. Those images were categorized as non-disease (13), mild-scab (12), moderate-scab (12), severe-scab (10), mild-woodiness (13), moderate-woodiness (15) and severe-woodiness (12). The number of images in each category has been mentioned in parenthesis. The model had 66% accuracy and it was received by the L\*a\*b colour model.

## 6. References

- [1] Sema, A. & Maiti, S. C. *Status & Prospects of Passion Fruit Industry in Northeast India*. Information bulletin, Central Institute of Horticulture, Medziphema, Nagaland, 2009
- [2] Kalubowila, I., David, S. M. G. D., & Rathnayaka, R. M. M. N. K. *Improvement of Yellow Passion Fruit through*

*Recurrent Selection*. Sri Lanka, Annals of Sri Lanka Department of Agriculture, 2017

- [3] Central Bank of Sri Lanka. *Annual Report 2017*, Colombo 01: Central Bank of Sri Lanka,
- [4] Bhangе, M. & Hingoliwala, H. A. *Pomegranate Disease Detection Using Image Processing*. India, Elsevier B.V, 2015
- [5] Dubey, S. R. & Jalal. A. S. *Adapted Approach for Fruit Disease Identification using Images*. India, International Journal of Computer Vision and Image Processing, 2012
- [6] Padol, P. B. *SVM Classifier Based Grape Leaf Disease Detection*. India, Conference on Advances in Signal Processing (CASP), 2016
- [7] Sujatha, R., Kumar, Y. S., Akhil, G, U. *Leaf disease detection using image processing*. Vellore, Journal of Chemical and Pharmaceutical Sciences, 2017
- [8] Khot, S. T., Supriya, P., Gitanjali, M., & Vidya, L. *Pomegranate Disease Detection Using Image Processing Techniques*. Pune, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2016

## Sinhala handwritten postal address recognition for postal sorting

*M.F.A. Ifhaam*

*Department of Industrial Management,  
Faculty of Science  
University of Kelaniya, Sri Lanka  
ahmedifhaam1@gmail.com*

*S. Jayalal*

*Department of Industrial Management,  
Faculty of Science  
University of Kelaniya, Sri Lanka  
shantha@kln.ac.lk*

### Abstract

*Sri Lankan post office mail sorting process is done manually, even today. Though employees are well experienced, it takes considerable time and employees need to work overtime in places like Central Mail Exchange (CME). With major issues like unclear handwriting, having trouble to recognize some uncommon or ambiguous names, and carrying these duties twice a day create a negative impact on the efficiency of the postal delivery system. In the prevailing system, forward mails and delivery mails are the two categories of separating mails at the sorting centers. Delivery mails are the posts which can be delivered to its destination directly. Forward mails are the ones which need to be sent to an appropriate post office that can deliver the particular post to its destination. Majority of Sri Lankans use Sinhala language for their day to day activities. The primary objective of the research is to identify the automatic way of forwarding the letter to the next post office from the current post office. Proposed system is focused on the recognition of Sinhala handwriting using Optical Character Recognition (OCR) and image processing technologies. Data collected under different criteria were used for training and testing the solution. Genetic Algorithm (GA) was used to generate more optimized results faster with higher accuracy. Given addresses are written in the default format. This format can be extended to more formats as improvements in future. The algorithm shows accuracy over 92% for addresses which are recognized with 3 misrecognized characters. This algorithm can be used on practice scenario as the AI Recognition has more than 79 % of accuracy.*

**Keywords:** *Image Processing, Machine Learning, Postal Address Sorting, Sinhala OCR, GA*

### 1. Introduction

Sri Lankan Postal Department is responsible for delivering posts in Sri Lanka. In order to deliver the posts, a sorting process has to be carried out. This sorting happens in all post offices in Sri Lanka including the ones located inside trains. Posts are sorted for each post offices

under two categories, namely ‘Delivery Posts’ and ‘Forward Posts’. Delivery Posts are posts which will be distributed by the postmen of that post office while the forward once are sent to another post office from the receiving post office.

During the sorting process, only the city name and the line above that; mostly the village name or the sub level of the town will be used. When a situation arises where the address written on the post does not contain the postal city name, address will be recognized with the use of a hierarchically organized locations, provided by postal department as a reference book.

This process is normally carried out twice a day. Thousands of posts are sorted each time, main post offices may have hundreds of sub post offices under their administration. Therefore, it is much valuable to introduce a ZIP CODE for the sorting process as the manual system is more time consuming. Yet, it is evident from the past records that implementing a zip code system is practically difficult in Sri Lanka as citizens are not that aware about the Zip code procedure. These details were gathered during the interaction with sorting officers of Matale post office.

Use of Character recognition and Artificial Intelligent based solution for the problem was only addressed by Karunanayaka, et al. [6]. Further, this study was mainly developed only by addressing city names of the country. However, the process cannot be completed only with the postal city name, because a deliverable address cannot be recognized with only postal city names.

According to the study “A few researches have been done on Sinhala handwritten characters. Almost all those researches are focused on identifying regular, well-defined Sinhala handwritten character recognition. Therefore, there has been no research done in cursive, unconstrained handwritten characters at all.”

### 2. Literature review

Mail sorting has been a resource consuming task for decades now. In most instances English-speaking countries have replaced the sorting process with AI based systems. Beside the postal departments, private sector also moving towards these automation processes. According to

Srihari and Kuebert, [8] The first large scale deployment of OCR (Optical Character Recognition) Equipment in the United States Postal Service (USPS) occurred in early 1980's as a part of the automation program. But still it has not been implemented in Sri Lanka. Complexity of Sinhala language is a major hurdle on this implementation. Though quiet a number of researches have been explored in this area (Sinhala character recognition) still a practically applicable solution has not been found. Only a single paper related to the domain (Sri Lankan postal address sorting) have been published, that paper also covers only postal city names, which only covers a part of the sorting process.

Since most of the research have been done on structured and/or constrained handwritten characters for Sinhala, purpose of this research is to find an applicable solution for the problem of mail sorting. Additionally, this research was aimed at extending the models for handwritten Sinhala word recognition for unconstrained and unstructured mixed script.

Silva, et al. [2] discuss the current state of Sinhala handwriting recognition of modern script. Further the study categorizes data collection methods, as structured and unstructured, based on the data collection method. It clearly defines that 'Data is retrieved in structured handwriting, if a pre-formatted form is filled to collect data'.

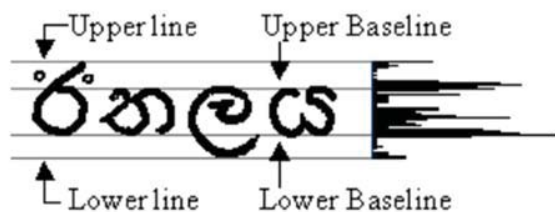


Figure 1. Sample Sinhala word with lines and horizontal projection profile.

### 3. Methodology

The process of recognizing valid Sri Lankan postal address from a handwritten address involves considerable number of steps. Each step has been elaborated as a subtopic under this heading. Overall process of the methodology has been explained by Figure 2.

As discussed in literature review most of the steps till character recognition, are based on literature. This paper can be considered as an advancement of the research published by Karunanayaka, et al. [6]. Contribution of this paper is increasing the accuracy of Sinhala handwritten postal address, using heuristics.

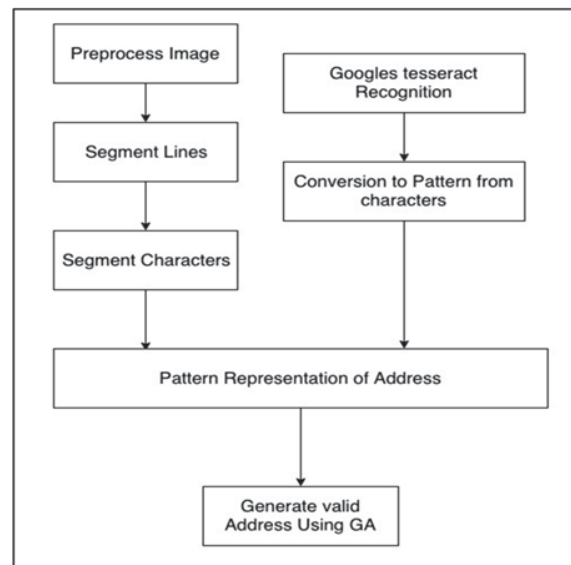


Figure 2. Methodology flow chart.

#### 3.1. Pre-processing image

Images had to be preprocessed because of the varying lighting conditions and noise created by the camera, even though camera model and angle was unalterable (fixed), Skew correction, Contrast and brightness equations were carried out to normalize the image in order to increase the segmentation efficiency.

#### 3.2. Noise removal and binarization

Even though images have been captured in full color, color details are not useful for the implementation. Hence grayscale images were used. This was pointed out in the literature review as well. Removing unwanted data, resulted in considerable increment in the performance. With the data collected via literature review and practical applications, it was decided that salt and pepper noise was the usual noise type found in the image. Mean filtering algorithms were used to remove the noise. Images were binarized using adaptive threshold method since global valued thresholding methods were not convincing in practical situation, because of the different lighting conditions in different areas of the images. Using adaptive thresholding, threshold for small regions can be calculated rather than a static threshold for complete image. This gives us a better result.

### 3.3. Horizontal projection profile for Line Segmentation

After thresholding the image, Horizontal projection of pixel values used to separate lines which contains text or to define border lines of texts found on the image. Horizontal Projection profile is basically accumulating the pixel values to understand which row of the image has more background and which has more foreground. A row of pixels that has zero-pixel value, can be used as a separation point. However, in most of the cases this is where, there is a value that is lesser than a certain value not zero value.

### 3.4. VPP and water reservoir concept for Character Segmentation

Vertical Projection Profile is same as HPP, the only difference is that columns are accumulated in VPP rather than rows in HPP. However, the issue with this method is that touching characters and overlapping characters were segmented as a single character in most of the cases. Water reservoir method was used at those situations for separation. As Karunanayake, et al. [6] explain initially character width comparison was used to check whether any segmented characters need water reservoir application.

### 3.5. Categorization of unsegmented character

Using the water reservoir method characters can be categorized to three groups. Later for each of these categories different methods can be used to separate the touching characters. This methodology for the domain has been implemented successfully by Karunanayaka, et al. [6] as explained, characters are segmented based on the categories. Each category has specific segmentation process according to the reference.

### 3.6. Recognition

#### 1. KANN Based recognition

Recognition of characters was done using the Kohonen Artificial Neural Network, this network was built using 32 x 32 input neurons and one output neuron because literature points out that this design is more effective.

Rather than recognizing the character itself, author of the research which has been taken as a base, has followed an impressive methodology, that is recognizing character groups, these groups are custom defined according to the Table 1. The system recognizes a pattern using the group of these characters. Then the system should retrieve city names from the database which is applicable for the detected pattern.

For example, the system would detect “AQTMX” for the city name written as



Major part of the methodology described up to now, has been done at least in less accurate manner. According to Karunanayaka, et al. [6] the recognition process has shown significant drawback in recognizing all similar characters like



and



Author has suggested Hybrid recognition process combining

Hidden Markov Models and Kohonen Artificial neural network for future enhancements.

Though this system was effective and successful up to a considerable level, author has not mentioned why it has not been implemented for complete address recognition. However, implementing such a methodology for bigger names is difficult and recognizing validity with just a database of city names is challenging as well.

**Table 1. Categorized characters for making the pattern.**

Symbol	Character	Symbol	Character
A	අ	Q	ඉ
B	ආ	R	ඊ
C	ඇ	S	උ
D	ඈ	s	ඌ
E	ඉ	T	ඍ
F	ඊ, උ	t	ඎ
f	උ, ඌ	U	ඏ
G	උ, ඌ	V	ඐ
H	ඌ, ඍ	W	එ
I	ඍ	w	ඒ
i	ඎ	X	ඓ
J	ඏ	x	ඔ
K	ඐ	Y	ඕ
L	එ	Z	ඖ
l	ඒ	1	඗
M	ඓ	2	඘
m	ඔ	3	඙
N	ඕ	4	ක
O	ඖ	5	ඛ
P	඗	9	ඛ, ඞ, ඟ

#### 2. Using Google's tesseract library for Sinhala

Google's open source library tesseract is an impressive tool for character recognition and now it supports Sinhala as well. Though this library is not great at recognizing handwritten characters this library performs better for printed characters. Using a mapped dictionary transformation, output recognized by the library was

converted in to the respective pattern according to Table 1. Then, the next steps were same as for the 1st method.

#### 4. Validating the address

Regardless of the option selected for recognition, output will be inaccurate. Recognizing an address with higher probability of being correct was the goal. If the address can be validated with an accuracy above 90 %, sorting process can be automated easily.

Actual postal hierarchy was defined in a data structure Figure 4 which represented exactly the same geographical hierarchy of the postal addresses. Because of the difficulty in data collection it was implemented for Matale and below, then test for the algorithm had been carried out.

Second row contains the names of the villages under Matale postal distribution zone. (Not all the villages are mentioned here). Third row villages are the ones under the distribution zone of Ukuwela sub post office.

මාතලේ										
දිස්ත්‍රික්ක	විලාසිය	ප්‍රධානවිලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය
විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය	විලාසිය

Figure 4. Example postal hierarchy.

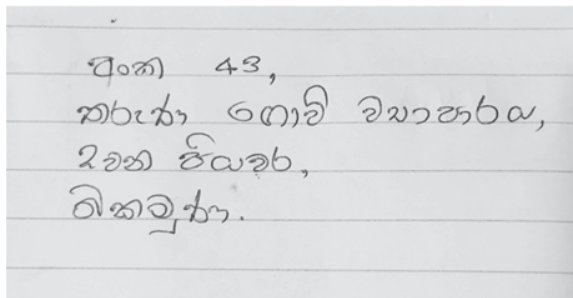


Figure 5. Sample handwritten address.

Both formats were successful but the 'Format 2' is more reliable as it will validate the address in three layers. Since each line would be assumed as a single node in the hierarchy, names separated by a comma, will not be recognized with the current implementation.

Table 2. Supported formats

Format 1	Format 2
Xx xxxxxx,xxxxxxx ආරගඟවෙල උකුවෙල	Xx xxxxxx,xxxxxxx ආරගඟවෙල උකුවෙල මාතලේ

For the address written in 'Format 1', character recognition produced the following pattern

Table 3. Recognized pattern

Pattern
Xx xxxxxx,xxxxxxx RVHHM9W T9Q9W9

That pattern denotes following address which is incorrect because of inaccurate recognition.

Table 4. Validated address

Address denoted by the pattern
Xx xxxxxx,xxxxxxx ආරගඟවෙල මාතලේ

For the incorrectly recognized pattern, algorithm proposed by this paper recognized the correct address, a modified Genetic Algorithm (GA) has been used to achieve the results.

From the misrecognized pattern GA would produce optimum results faster than normal algorithms. This is more suitable for the problem. Since character shuffling and permutations have too many combinations, traditional looping methods will not return the results as fast as GA. Since GA works with custom fitness function, address validation was implemented inside the fitness function. Hence this method was able to produce much better results than other solutions.

Fitness function was designed in a way that would assess the following constraints.

1. Characters replaced from each group make a valid city/ sub city / village / street name combined together.
2. Recognized sub city/ village / street name falls under city name or;
3. These single elements combined together make the meaningful valid address.

#### 5. Problem design

Each village, town and city were denoted as Node. This node contains 4 attributes

1. Node Value:  
This is the actual Unicode (UTF-8) representation of entity name of the city this Node represents.
2. Node Pattern:  
This is the pattern string which is the translation of node value using the category table in Table 1.

3. Children  
Children attribute is a list that contains all the nodes that come under the current node in the postal hierarchy.
4. Parent  
Parent is the node which is directly above the current node in the postal hierarchy.

## 6. Preparation

As the first step a knowledge dictionary was built for the validation algorithm that is the hierarchy of the postal system. Each entity was defined as a single node and using the above described method, data structure that is required for the GA to operate, can be prepared.

While building the structure, each of those node's pattern was added to a single list for later checks which is 'all\_nodes\_pattern' list. An index with patterns as the key of the index and node as the value, increase the efficiency of node retrieval, which was required later in the process.

## 7. Genetic algorithm representation of the problem

Problem should be represented by a chromosome at initial state. A combination of pattern char would be called a node and a collection of nodes would be called a chromosome as shown in Figure 6.

R	V	H	H	M	9	W
T	9	Q	9	W	9	

Figure 6. Genetic representation of problem

The small box around English alphabets denotes the Gene while the complete outer box denotes the chromosome. And each row has been defined as a component for easier interpretation throughout this document.

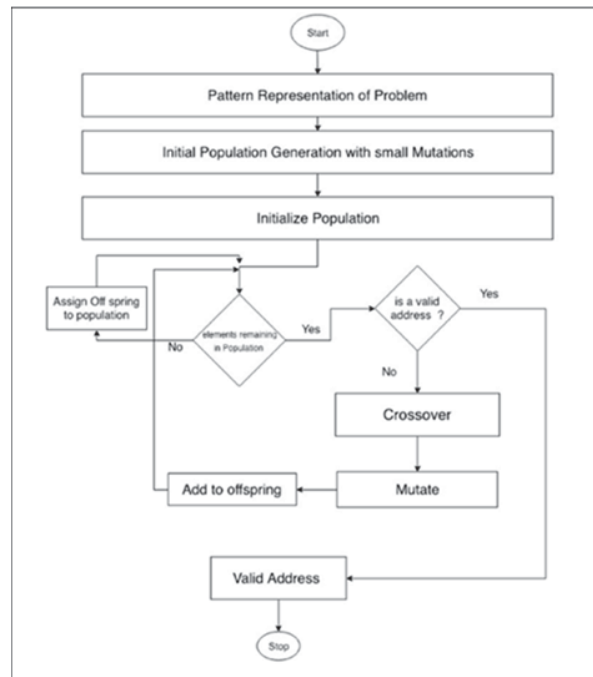


Figure 7. GA based address validation

### 7.1. Fitness function

Fitness function should check for constraints that satisfy the output. The fitness function defined here should satisfy;

1. How much each component (node\_pattern\_string) in chromosome is fitting with a node in all\_nodes\_pattern list.
2. Then check whether the component above another component is matching a child node of a component below that.
3. Then program should be stopped when satisfying result reached by the algorithm

Additionally, there is a fitness calculation for a node with a city pattern. This calculates how many letters out of the maximum number of characters match the node\_pattern\_string of current nodes. This is calculated in percentage.

### 7.2. Crossover

With crossover implementation, first check was whether there is a node that satisfies the fitness calculation, given that it is a valid node in the hierarchy it would be replaced to new cross over chromes node. For the nodes that does not satisfy, a random node will be selected. From that place it would be sliced it into two halves. After that halves of both chromosomes will be joined interchangeably.



### 7.3. Mutation

Mutation was defined from node level. Each node mutates itself. To meet the requirements of the problem, the mutation function of nodes contain two clearly defined processes.

1. Change

During this process looping through each and every gene of the nodes in a chromosome and checks a random probability, if the probability is more than .02 then function picks a random letter from pattern notations and replace it with current gene.

2. Add new

A new element to the chromosome will be added with the probability of 0.01. To overcome the issue of complete unrecognition of some letters during this recognition process, this approach was taken.

Node mutation function will be used to mutate all the components in a chromosome. So that, small changes are made to the chromosome in a positive manner.

### 7.4. Evolution

Main process of GA is evolution, this is the process that makes the problem into a solution using the above-mentioned technics and procedures. Evolution stops when it finds a valid address pattern from an invalid pattern. Termination point has been defined as 100% fitness for complete chromosome, at which state chromosome would be representing a valid address.

### 8. Data collection

Data collection has been carried out in social media since captures of images were satisfying and it is easy to find more people in there. Testing was carried out from the collection of data from 50 different people, from distant places with a variety of occupations, two addresses each. From the collection, addresses were filtered to match the testing criteria.

Line 1 : RVHHM9W  
Line 2 : T9Q9W9  
Line 3

Figure 8. Sample Recognized Pattern.

AW9XN9N  
T9N9W9

Figure 9. Validated Pattern.

### 9. Testing

This testing procedure is carried out to test the model. Valid addresses were always easy to recognize. If the model can fix the misrecognized characters by Neural network and produce an existing address that would be a valid address. According to that the results are shown in Table 5.

Table 5. Accuracy of GA.

Number of Address	Number of misrecognized characters	Number of valid addresses produced by the system	Accuracy %
10	1	10	100%
10	2	10	100%
10	3	8	80 %
10	4	4	40 %
10	5	3	30%

During the process of validating the result Table 5, an analysis on the number of generations algorithm took to reach the solution was produced by plotting the data into the following graph Figure 10, were produced to visualize the results.

This representation has no correlation with number of misrecognized letters since GA is mostly a dynamic algorithm. Indeed, it was the expectation at the beginning. It is evident that this algorithm does not consume much time or resources for the range of problem tested with this algorithm. Here the graph has reached around 30 as a maximum number of generations for a valid address. To execute 30 generation, it takes less than a minute, which is above average and considerably better than other algorithms.

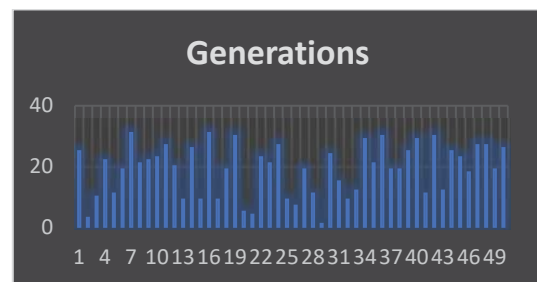


Figure 10. Max Generations Reached for Solution.

Testing process was carried out in a trial and error method. Steps involved in the testing process are:

1. Initially an address as shown in Figure 11 would be loaded.

2. A pattern recognized for the image would be Figure 12.
3. Representation of recognized pattern is  
නාවන්ද  
මානලේ
4. Validated pattern is shown in Figure 13.
5. This pattern shows exactly the same representation of digitalized image shown in Figure 11.

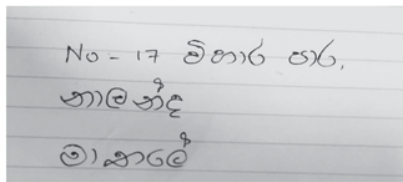


Figure 11. Initial Image.



Figure 12. Recognized Pattern.

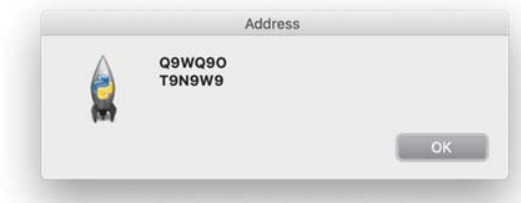


Figure 13. Validated Pattern

## 10. Conclusion

Results produced by artificial intelligence is solely based on the amount of data, quality of data as well as variety of data. Since there was significant adaptation from Karunanayaka, et al. [6], limitations of that research will be applicable for this paper as well. More than that validation procedure also added new constraints to the format of the address that can be recognized by the system. For example, restricting the address format, cities and villages to be separated by new lines between. Then the constraint of time added to the operation because of the GA's dynamic nature, a specific time cannot be defined for a 100% valid address which in case may become a problem with complete integration of postal hierarchy into the system.

However, within the limited constrains system performed very well and it was capable of extending to the larger context. But as shown in Table 5 when the number of misrecognized characters are increasing, the accuracy of the valid address is reducing so up to some extend the quality of recognition process is also impacts the address recognition.

As a conclusion this GA based address validation for handwritten address in Sri Lankan postal services is a great way of improving the accuracy of outputs produced by AI systems. Though with some limitation accuracy has been improved to 100% accuracy from 80-90% accuracy without spending a considerable time.

## 11. Future work

AI engines accuracy impacts the final output of the system. Hence an improvement to Recognition would improve the accuracy of the address validation significantly.

Also, this research has been carried out on limited number of characters in Sinhala alphabet and with a smaller number of training data. Increasing the amount of training data or increasing the alphabet supported by the algorithm may result in a proper solution for the problem.

This research is in a way increasing the accuracy of AI recognition with heuristics. If this algorithm can be applied to other domains such as voice recognition, face detection or any other recognition that would extend the capabilities and reliability of AI.

Also, since this research has been done for the fixed format addresses, additional formats can be supported with a small effort. For example, street name and Postal Box numbers can be supported.

## 12. References

- [1] Wijesooriya, V., 'Sinhala Handwriting Character Recognition and Speech Synthesis System' Faculty of Information Technology University of Moratuwa, September 2016.
- [2] Silva, C. M., Jayasundere, N. D. and Kariyawasam, C., 'State of Handwriting Recognition of modern Sinhala Script', *IJIRT - International Journal of Innovative Research in Technology*, vol. 1 no. 9, pp. 97-102, 2014.
- [3] Hewavitharana, S., Fernando, H. C., Kodikara N. D., 'Off-line Sinhala Handwriting Recognition using Hidden Markov Models' *Proceedings of the Third Indian Conference on Computer Vision, Graphics & Image Processing*, Ahmadabad, India, December 16-18, 2002
- [4] Jayasekara, B. and Udawatta, L., 'Non-Cursive Sinhala Handwritten Script Recognition: A Genetic Algorithm Based Alphabet Training Approach', in *Proceedings of the International Conference on Information and Automation*, December 15-18, 2005, Colombo, Sri Lanka pp. 292-297, 2005

[5] Dharmapala K. A. K. N. D., Wijesooriya W. P. M. V., Chandrasekara C. P. and Rathnapriya U. K. A. U., 'Sinhala Handwriting Recognition Mechanism Using Zone Based Feature Extraction' *ITRU Research Symposium*, 2015, Faculty of Information Technology, University of Moratuwa, Sri Lanka, pp. 10–15, 2015.

[6] Karunanayaka, M., Kodikara, N. and Wimalaratne, G., 'Off Line Sinhala Handwriting Recognition with an Application for Postal City Name Recognition', *Proceedings of 6th International Information Technology Conference (IITC) "From Research to Reality"* 29 Nov - 01 Dec 2004 BMICH, Colombo, Sri Lanka, 2004

[7] Kumara, T. N. and Ragel, R., 'A systematic feature selection process for a Sinhala character recognition system', 2016. *Proceedings of IEEE International Conference on Information and Automation for Sustainability: Interoperable Sustainable Smart Systems for Next Generation, ICIAfS*, 2016.

[8] Srihari, S. N. and Kuebert, E. J., 'Integration of hand-written address interpretation technology into the United States Postal Service Remote Computer Reader system', *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, 2, pp. 892–896, 1997.

## Devise best practices for mobile application development: Addressing the challenges in designing user interfaces

W. P. G. M. Weerasinghe  
Department of Industrial Management  
Faculty of Science  
University of Kelaniya, Sri Lanka  
madhumaligaya@gmail.com

D. N. Wickramaarachchi  
Department of Industrial Management  
Faculty of Science  
University of Kelaniya, Sri Lanka  
dilani@kln.ac.lk

### Abstract

*Over the past few decades, mobile device penetration has showed a significant growth. To meet consumer demands, it is essential to design mobile applications with enhanced usability and user experience. Catering to specific user requirements and designing user friendly mobile applications while overcoming limitations of mobile devices are major challenges in mobile application development. Even though there are sufficient information about UI components, there are only limited standards for mobile UI design patterns. Most of these designs are based on the desktop paradigms such as WIMP (Windows, Icons, Menus, Pointer) interface style. But mobile interface paradigm is based around widgets, touch, physical motion, and keyboards. There are limited studies available analyzing these mobile UI design patterns. Furthermore, inappropriate UI design induces usability problems to end-users. The final result is a negative user experience. This paper discusses the development of best practices which can lead to improvements in this situation. Therefore, a questionnaire was distributed among smart phone users to identify the usability issues related to smart phone applications. Through the questionnaire and the interviews with industry experts the usability issues related to UI and content, connection, navigation, user experience etc. in users' perspective had been identified. Through careful analysis of the results, the best practices that should be adopted while developing mobile applications with considerable usability were determined. In supporting this study, proposed best practices were evaluated by industry experts which indicates a higher validity of research findings.*

**Keywords:** Best practices, Mobile, Usability, User interface

### 1. Introduction

With the dawn of digitalization, mobile devices have become an essential part of our day to day lives. We rely on them for a vast range of digital services that make our

lives easy. Therefore, there has been tremendous growth in the use of mobile devices over the last few years. In contemporary time, the smartphone has become an essential device for us, since it significantly supports our daily tasks with its attributed functionalities. Hence, a smartphone is no longer just a medium of communication. This growth has motivated the development of millions of software applications for mobile devices. Companies are competing with each other to develop applications that fulfill the demands of profit maximization and capitalizing on the best potential market opportunities [1]. However, even though there is a growth in app development, there is a significant percentage of app abandonment as well [2].

Most studies in Human Computer Interaction (HCI) take interest in developing new design methodologies, experiencing with new hardware devices, prototyping new software systems and in exploring new paradigms for interaction. In order to develop a successful mobile application, user interfaces must be attractive, efficient and easy to operate. In other words, design of the user interfaces must be a user-centered design. The most common challenge is catering to specific user requirements and to design user friendly mobile applications while overcoming the limitations of mobile devices [3]. Providing users with powerful computing services and resources through small interfaces are also major challenges of mobile application development. Presenting information efficiently and effectively is even more challenging with the limited screen size. On the other hand, navigating through presented information is also a major issue. Mobile devices are operating with limited battery power. Therefore, when designing mobile applications, developers need to consider the battery consumption as well as connectivity issues. With the higher usage of mobile applications, the expectations and requirements of users become more and more complex. Hence, designers and developers have to cater to higher level of requirements. Therefore, it is very crucial for interface designers to make efforts to develop interactive mobile applications which meets usability factors such as easy to learn, effective to use, and provide an enjoyable user experience.

Hence, identifying existing usability challenges of mobile applications in users' perspective and defining list of best practices which addressed those identified usability challenges to enhance the usability of mobile applications are the main objectives of this study.

## 2. Literature review

The literature review was conducted in two parts. The first part includes key findings related to usability of mobile applications and guidelines to develop user interfaces. The second part includes psychological theories and HCI theories which support for this study.

### 2.1. Usability of mobile applications

Nilsson has presented a collection of user interface design patterns for mobile applications which gives a fairly comprehensive overview of issues when designing user interfaces for mobile applications [4]. Those issues were divided into three main areas i.e. utilizing screen space, interaction mechanism and design at large. The main focus of the patterns collection is limited to professional application in general and forms-based UIs in particular. The patterns collection must be further developed, especially catering to future challenges and new types of UIs.

A study was done to identify the effects of key sizes. It was found that if sound is added to buttons then they can be reduced in size from 5mm<sup>2</sup> to 2.5mm<sup>2</sup> (from 16x16 to 8x8 pixels) without much loss in quantitative performance [5]. The smallest key size that would not degrade performance or user satisfaction were estimated while it was shown that touch screen input keys compete with other information for limited screen space [6]. They found that spacing had no measurable effects and Entry times were longer and errors were higher for smaller key sizes, but no significant differences were found between key sizes of 20mm<sup>2</sup> and 25mm<sup>2</sup>. The key size of 20mm<sup>2</sup> was found to be sufficiently large for land-on key entry. For the experiment only twenty students participated from one university. This is a highly biased experiment group and they only used numeric keypad to do the experiment They also only focused on button size to overcome the limited screen size of mobile devices. However, further research needs to be done to identify the solutions for other aspects such as presentation styles, patterns etc. to overcome the lack of screen size.

There are HCI challenges in designing applications in both hardware and software for mobile devices. Limited input facilities and limited output facilities were identified as hardware challenges and hierarchical menus, designs for mobility, navigating and browsing, images and icons were identified as software challenges [7]. Some of the key issues related to software and hardware challenges

were addressed based on previous researches. They have also investigated principles for designing mobile interfaces. Visualizing information on mobile devices is another aspect which requires attention [8].

Another challenge is to design mobile devices in a more useful and affordable way targeting to a heterogeneous set of users. Facility to downscale the images from a regular size in a personal computer to the appropriate size in a mobile device is also an important research direction that should not be neglected.

By analyzing the classes of errors that people make with systems, it is possible to develop principles of system design in minimizing both the occurrence of error and the effects [9]. Another study addressed the development of design rules that can lead to improvements. But they mainly focused on error classes. There are studies which have given guidelines for developing software applications [10, 11, 12]. But those guidelines specially focused on desktop applications. Characteristics and limitations of current mobile device interfaces, especially compared to the desktop environment were analyzed and a set of guidelines presented for the design of handheld mobile device interfaces [13]. These guidelines are based on traditional guidelines for desktop user interfaces and published research with mobile devices and applications. They mentioned that these guidelines could be useful to practitioners who develop mobile applications, and to HCI researchers working with mobile interface design and usability. But this is not a complete set of guidelines for mobile UI design and further they have not drawn certain emphasis on interesting mobile usability issues such as connectivity, filtering, memory consumption, content and visualization etc.

Further, when developing such guidelines, usability challenges in users' perspective should be considered to enhance the user experience. But most of the guidelines in the literature did not consider those challenges. Therefore, it is very crucial to develop a comprehensive set of best practices which addressed those usability challenges.

### 2.2. Theoretical knowledge

#### 2.2.1. Fitts's law

The study intended to model the act of pointing at a target on a computer monitor. It was observed that the action of pointing to or tapping a target object could be measured and predicted mathematically. Therefore, this scientific law predicts that the time required to rapidly move to a target area is a function of the ratio between the distance to the target and the width of the target [14].

$$MT = a + b \times \log_2(D / W + 1)$$

- MT - amount of time required to complete the movement

- a, b - empirically determined regression coefficients (constants that are measured experimentally)
- D - distance from the starting point to the target object
- W - width of the target object

### 2.2.2. Importance of Fitts's law

Fitts's law is a binary logarithm. This means that the predicted results of the usability of an object and size of the object have non-linear relationship. The most prominent statement derived from Fitts's Law is that the larger a target, the faster it is to acquire. But larger targets may destroy the balance of the interface. Even if there are plenty of space to spare, constantly enlarging the target area does not make it more usable since the predicted usability of the size of a button progresses in a non-linear fashion.

The screen of the computer is limited vertically and horizontally. Similarly, the top and bottom are limited vertically. The sides and corners of a screen play an intriguing role in Fitts' law because they provide a boundary that wouldn't exist in the real world. This provide easy access due to screen limitations rather than a target in the middle of the screen. Therefore, clicking the close window button is a very easy target. But to reduce the risk, close button does not include the pixel closest to the edge. Also, this is partly as why windows start menu is in the corner of the screen and Apple's application menu always set at the top of the screen.

### 2.2.3. Miller's law

Number of objects an average human can hold in working memory is  $7 \pm 2$ . This has been referred to as Miller's law [15]. A human can remember things easily in working memory if they recall the items as groups. As an example: While recalling a mobile phone number such as 0715422889, we might break this into 071 54 22 889. Thus, instead of remembering 10 separate digits that is beyond the "seven plus-or-minus two" memory span, we remember four groups of numbers. This phenomenon is called as 'chunking'.

## 3. Methodology

A qualitative approach was employed in this research, since primarily it is expected to gather non-numeric, qualitative data from both Sri Lankan mobile application developers, UI/UX engineers, smart phone users and from the literature review. Hence, the research will be based on qualitative method. Figure 1 illustrates the high-level overview of the methodology. A detailed literature survey was carried out to determine overview of mobile application usability and identified the issues in mobile application development. Theoretical knowledge was also

acquired to formulate best practices. The expert interviews were conducted to acquire the current best practices in mobile application development and the usability issues of mobile applications from developers' perspective. Interviews were carried out in person i.e. face-to-face and also via phone.

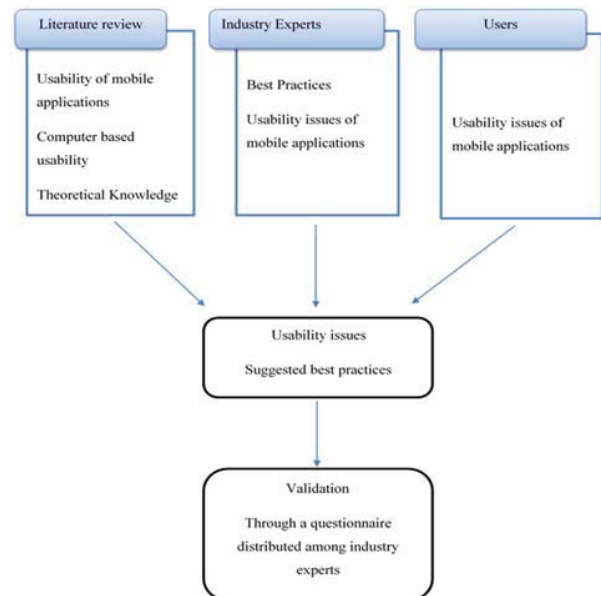


Figure 1. Research design.

A questionnaire was used to gather data from smart phone application users primarily, to identify the most frequent mobile application usability issues from user's perspective. It included multiple choice questions, short answer questions as well as open ended questions. The questionnaire was distributed via email and social media platforms such as Facebook targeting smart phone users.

### 3.1. Data analysis

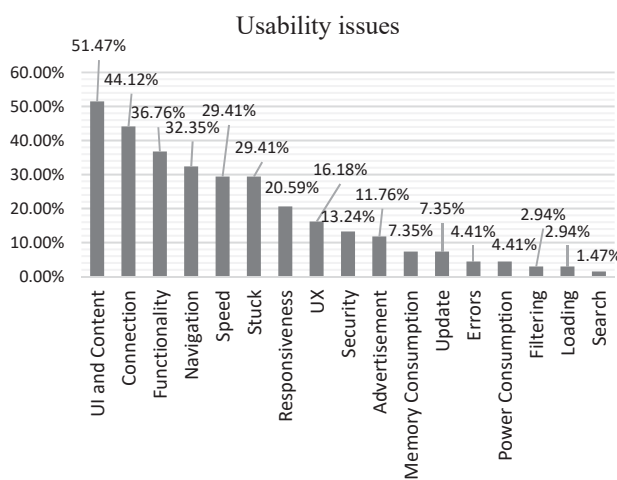
Data analysis of the questionnaire which was distributed among smart phone users, has been discussed here. Sixty eight responses were received for the questionnaire. Most of the respondents were undergraduates (63%) followed by students (18%), employees (10%), postgraduates (7%) and other (2%). Age spans were between 20-30 years. More than half of the respondents (52.9%) use 5-10 mobile applications for their day to day activities which proves that the respondents have adequate experience in using smart phone applications.

Multiple choice answer options were analyzed using quantitative methods which involved pie-charts, bar-charts and percentages. The questions related to usability issues were open-ended questions. Therefore, the

responses were in textual format. In total 221 issues were identified by users. Some of them were duplicates. After removing erroneous responses, 215 issues were considered for further review. All 215 issues were divided into 17 main categories as follows;

**Table 1. Categorization of usability issues.**

Issue	No of respondents	Percentage
Advertisement	8	11.76%
Connection	30	44.12%
Errors	3	4.41%
Filtering	2	2.94%
Functionality	25	36.76%
Loading	2	2.94%
Memory Consumption	5	7.35%
Navigation	22	32.35%
Power Consumption	3	4.41%
Responsiveness	14	20.59%
Search	1	1.47%
Security	9	13.24%
Speed	20	29.41%
Stuck	20	29.41%
UI and Content	35	51.47%
Update	5	7.35%
UX	11	16.18%



**Figure 2. Categorization of usability issues which were identified by respondents.**

Figure 2 shows that more than half of total respondents had identified issues related to User Interface and content. 44.12% of respondents have found connection issues. According to their responses, most of issues were UI & content issues, connection issues, functionality issues, navigation issues etc.

Under each category there were several usability issues. Total of 35 respondents mentioned about UI and content issues. Under UI and content issues most of the

respondents mentioned that the user interfaces of mobile applications are not user-friendly and they often touch other UI components such as buttons unintentionally. Content is difficult to read because it is not clear and there are time consuming texts to read which will lead to negative user experience ultimately.

Under the connection issues ‘not being able to work without internet connection’ was the answer of highest percentage. That means even without internet connection, the users expect an application to work. Hence, at least some functionality should be there without connection.

Most of users complained about the features of the applications. Some applications do not have the features demanded by the end users. Even though some features are unique to the application it does not add value to end users which can negatively affects the user experience. In the questionnaire, users have given examples from ‘imo’ messaging app. They mentioned that visibility of typing and deleting text to other party is an unwanted feature. If a feature cannot function well with user’s phone model, it will definitely affect the usability of the application.

Out of 32% of respondents who had identified navigation issues, 68% of them had mentioned that the navigation is difficult and some mobile applications do not have the facility to navigate anywhere from every page. As an example, since there is no home button in interfaces, they have to close all the pages to navigate back. Another issue is confusing or hidden navigations. So that users cannot identify the navigation patterns and shortcuts in the first use. This implies that navigation should be visible and users must have an idea about the navigation.

The respondents also mentioned that notifications should be understandable, timely and should not distract the user. Uber eats application does not ask ratings for the foods as soon as the delivery was ended. The rating notification will pop up after 20-30 minutes later targeting the time that user ended having the meal. Also, there should be an option to receive/ not receive notifications and filter notifications. Also, just providing user friendly notifications is not enough; but it must direct to the place where the news is after clicking the notification.

Users also do not expect traditional mobile applications. They need more user friendly, interactive applications which provide real world experience. For instance, on the Android platform, almost two-thirds of all apps are free to download [16]. To be compensated for their work, many app developers incorporate an advertisement library (also known as an ad library) in their apps. These advertisements distract the users and it negatively affects the user experience. Some advertisements collect analytics information about the users of the app. Therefore, security and privacy of the users need to be concerned.

## 3.2. Formulation of best practices

To formulate the best practices, most frequent usability issues were based on the survey were considered. The usability issues which have more than 15% frequent were considered. The key research findings and some of the theories (HCI and Psychology) have been used to formulate the best practices. Two main theories used for best practices were Fitts's law and Miller's law. Furthermore, information which were gathered through interviews, were also considered when developing best practices.

Most of the identified usability issues from the survey were covered under the proposed best practices. But especially, some of the connection issued were not covered under those proposed best practices.

As an example; Take too much time to upload/download media files. This is heavily depending on the connection type (Wi-Fi, 3G, 4G etc.), connection speed and processing power of the smart phone. Therefore, it is difficult to propose a best practice to minimize the time taken for upload and download. Developers must consider the mobile context when developing network based applications. Another issue is excess use of mobile data. This also highly depends on ISP data charges. Some of the best practices are already industry best practices. But developers sometimes do not practice those guidelines.

Following are the suggested best practices and some of the them were recommended by the industry experts.

### 3.2.1. Place rarely used/ highly risk UI components away from widely used targets to minimize human errors.

Studies have found that 49% of users hold their smart phones using one hand and most smartphones are used in portrait orientation, with the most popular size (6 to 6.5") showing 94% portrait usage [17]. Therefore, average thumb cannot reach every portion of the screen.

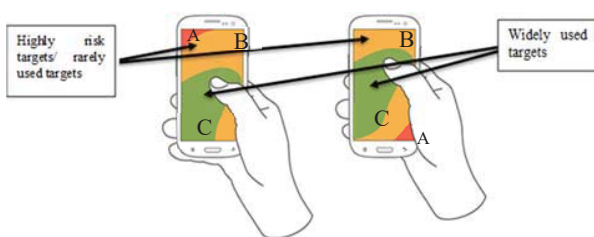


Figure 3. Easily reachable area for right hand user.

In Figure 3, 'A' area is the most difficult area to reach. 'C' area represents the easily reachable area and reaching a target in 'B' area is somewhat difficult. According to the Fitts's law, since additional muscular tension is required to reach the yellow area it will reduce the unintentional reach to that area. So, if we place highly risk targets or

rarely used targets in that area it will reduce the human errors.

### 3.2.2. Implement gesture for highly risk actions

According to Fitts's law, gestures need external muscular tension. Therefore, if the action is more critical for the application, implementing a gesture for that critical action will reduce performing the action accidentally. e.g. Swipe left to delete. But those actions/gestures are hidden; out of sight. Hence you should make them aware at the first use. Otherwise user may not be aware of those gesture based functionalities.

### 3.2.3. Avoid using hamburger menu, use tab navigation

Hamburger menus decrease the time spent on your application. (Facebook and Zeebox case studies) "out of sight-out of mind" navigation is not visible to users.

But tab navigation improves the user experience. For tab navigation use only maximum 5 items. (Miller's law) So that any individual can easily memorize the tab navigation. On the other hand, the tab navigation may affect to clear content, you can cleverly make it disappear when not using. As an example, if the home screen is a scrolling feed, the tab bar can be hidden when users flick it up to unfurl new content, and revealed if they start pulling down trying to get back to the top. In an interface like a map where full screen need to be utilized for the map, the tab bar can be hidden when a user taps, drags or when zooming.

If the navigation is too complex since tab navigation can only contain limited number of menu items, rarely used items can be added to the hamburger menu. So that tab navigation will become clearer to user.

### 3.2.4. Use radial menu

In desktop applications cursor automatically stop at the edges of the screen. Hence, commonly accessed UI elements are put on the edges. So, users can easily click on the target. Therefore, desktop menu bar is located at the top of the screen.

But when it comes to smart phones, users cannot touch top of the screen without extra force. So, easily touchable area lies with the boundary of green area and yellow area. (Figure 3) This boundary has approximately, a curved line. Therefore, use of radial menus will increase the usability.

### 3.2.5. Change placements of UI components according to the user

As mentioned before, 46% of smart phone users hold their phones using one hand. Among them 67% use right thumb to operate the phone and 33% of users use left hand



[18]. Therefore, it will be difficult to design one interface to meet individual user requirements.

As a solution, change the user interface according to the user preference. When installing the application, give option to select their preference for using left hand or right hand. Install the application depend on the user's respond. e.g. If user prefers right hand, radial menu should be located in right

### 3.2.6. Utilize every enable pixel to enlarge the tappable area

Fitts's law indicates that larger targets can be acquired faster. Even though there are restrictions to increase the button size, you can utilize every available pixel to enlarge the tappable area.

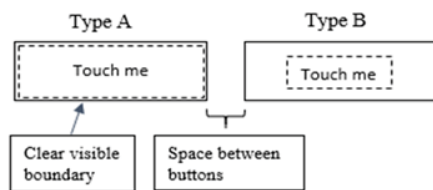


Figure 4. Utilizing pixels in buttons.

The dotted line represents the boundary of the tappable area. According to opposite of Fitts's law; a user needs more time to click smaller targets. Therefore, it should be better if you use Type A buttons as figure 4 shows for highly risk targets.

### 3.2.7. Reduce the complexity of user interfaces

Average human can only keep 5 to 9 chunks in his working memory (Miller's law). Using maximum 9 items per screen will ensure that users can easily memorize the interface. Specially this is critical for the novice users since it need practice to encode the interface into long term memory.

### 3.2.8. Protect the consistency of the interfaces

Every user interface should give the same look and feel for users. The consistent placement of the buttons on the pages will reduce the cognitive load. Similarly, if we are trying to find something, but it is in a different place or looks different from usual, we might miss it even though it is in plain view because experience tunes us to look for expected features in expected locations

### 3.2.9. Categorize items (5-9 items per group)

Human can easily recall the things which are memorized as groups, but not as individual items. (Miller's law) If you group the items your brain can

remember exactly where things are. Specially it is better to use 5 items per group because of the space limitations of smart phones and that will ensure every individual is capable of memorizing the items. e.g. Five sub menu items per main menu item.

### 3.2.10. Enable your application to detect OS level font size/ font style changes.

CSS style sheets have the power to disable the change font size option of a web browser and specify a fixed font size. About 95% of the time, this fixed size font is too small, reducing readability significantly for most people over the age of 40 [19].

Mobile applications avoid the use of fixed font size overriding the capability to change font size in OS level. It is best to respect the user's preferences and allow them resize text as needed. Further, when the OS level font size or font style change is detected, make the same changes in the application if user allows it.

### 3.2.11. Use maximum 6 words per headline

Readers do not only scan the content. They tend to scan headlines also. Researchers have found that readers absorb only first three words and last three words of a headline [20]. Therefore, the optimum number of words for the headlines should be 6 words.

### 3.2.12. Align most important contents to the left side

Researchers found that web users spend 69% of their time viewing the left half of a page and 30% viewing the right half. If you consider most common human reading patterns such as F, E and Z patterns, it will also prove that human tend to read left side than right side [19].

Therefore, if you put most important content to the left, it will ensure that the user will not miss the important facts.

### 3.2.13. Align advertisements to the right

Human storage is very fragile and information can be lost with distraction. As mentioned before, humans tend to read left hand side than right hand side. Therefore, if you align unimportant contents such as advertisements in the right hand side, readers can read the content without distraction.

### 3.2.14. Align images to the right

Images always get the high attention from users. Therefore, to utilize the screen space it is better if you can align images to right side since right side texts will not get the human attention.

### 3.2.15. Avoid bulky texts

- *Portrait Mode*

Add only 55 characters per line

Users prefer to watch videos rather than reading bulky texts. In previous research researchers have found that there is better comprehension with shorter (55 characters per line) lines rather than lengthy lines.

- *Landscape Mode*

Reading longer lines speeds up the reading speed. But longer lines increase the visual angle and so that it is difficult to locate the beginning of next line [21].

The relationship between the amplitude (A) of saccades (e.g. return sweeps) and their duration can be described by the function  $2.2A+21ms$  [21], which means that the greater number of return sweeps with shorter lines will add more to the time than longer lines. Therefore, humans prefer to read double column computer screens rather than single column.

But in mobile screens, the screen width is always limited. So, you cannot have lengthy lines as you wish. Because lengthy lines will decrease the font size resulting decrease in usability. But in landscape mode you can utilize all the screen width for text. Those recommendations can be mainly used in applications where textual display is necessary such as e-book reading app.

### 3.2.16. Use colors wisely

According to color psychology different colors have different effect on humans. When designers are selecting colors, these affects should be considered. e.g. Select blue color for banking applications will have more positive effect than using red color for banking app, likewise green color considered eco-friendly or environment friendly, healthy color. So, it is more suitable for health applications.

### 3.2.17. Give offline support

It is best recommended not to develop applications solely depending on the connectivity. It is good to implement some offline features as well. If sudden drop of connectivity is detected, the app should be able to continue the process when connection established again avoiding restart of same process. e.g. If you are posting something, it will be stored until network gets connected and you will be notified as well.

### 3.2.18. Give a user guide at first use

Novel users may take time to get used to the application. So, it is better if you provide user guidelines at the first use of installation. Further, if there is a new

feature which got updated the developer must make users aware of same.

### 3.2.19. Enhance the responsiveness of your application

When there is a background process, proper notification need to be given to user. While downloading or uploading something, user must be able to do another process. No blocking should happen to the user. But if the background process is a critical process such as payment, then it is ok to block the user and give proper indication.

### 3.2.20. Application must talk users' language

Proper success and error messages should be given. Application's language should be appropriate to the application's domain. Notifications should be given at the proper time. There should be a filtering option if the application provides too many notifications.

### 3.2.21. Enhance the user interaction

Since users are interested in having real world experiences, gathering the full use of state-of-art technologies such as gamification, social media interaction, AR/VR, chatbot etc. is important. Also, application should be compatible for screen readers. E.g. enable users to share their status, achievements etc. of your application in social media platforms. Reward (score board, badges) users if they achieve a target.

### 3.2.22. Compatible with platform

It is better if the application is compatible with every OS and every version. This is comparatively easier when it comes to hybrid application development. But in the long run using native applications gives easy maintenance, easy to fix issues, easy to support OS updates and gives more support for accessibility. On the other hand, even though you have one code base you must optimize it for each OS which may result in high cost

## 3.3. Validation

Proposed set of best practices were evaluated by industry experts using a questionnaire. The designation of the respondents was asked as a question in-order to make sure the responses were taken from the people who are eligible to answer the questions and to check whether they have a particular understanding about mobile software development. Most of the respondents were software engineers (36.4%) followed by UI/UX engineers (27.3%), mobile software engineers (27.3%) and technical leads (9.1%). Experience of respondents spans from 1-10 years.

There were six evaluation criteria and every evaluation criteria received weighted average of more than 4 out of 5.

That implies the fact that every criterion was accepted by the evaluators. Hence, we can state that the proposed best practices are accepted by the industry experts. Eleven best practices such as use colors wisely, give offline support, application must talk users' language, enhance the user interaction etc. were recommended by the industry experts.

#### 4. Conclusion

This study was conducted to identify the usability challenges of mobile applications and mainly focused on usability issues faced by the users. The primary objective of this study is to propose set of best practices to overcome those usability challenges. Eleven best practices were recommended by the industry experts and others were formulated using theoretical knowledge and previous research findings.

Most of the identified usability issues from the survey were covered under the proposed best practices. But especially, some of the connection issues were not covered under those proposed best practices. The validity of the best practices measured through a questionnaire and every criterion in the questionnaire were accepted by the industry experts. Some of the evaluators mentioned drawbacks as well. They have mentioned that as a set of guidelines, this will be more useful and accessible if these practices are categorized relating to the various touch points such as layout, aesthetics, etc. Also, some of them have mentioned that even though developers can reduce many issues by these clever best practices, it is really difficult to convince clients to change user interfaces.

#### 5. Limitations and future work

The target population of this study was all smart phone users. But the sample size was 68 which is relatively too small. Even though population is distributed among different age groups, the age variation of the sample size was 20 to 30 years. Therefore, the sample is limited to young adults. Further, the effect of voice commands in the future UI is not considered in this study. In this study a questionnaire was used to validate the best practices and only eleven industry experts contributed for the evaluation process. In future, it is better to validate these best practices by developing an application using this comprehensive set of best practices and compare it with another version of the same application which was developed without the proposed best practices.

#### 6. References

[1] Statista - The Statistics Portal for Market Data, Market Research and Market Studies <https://www.statista.com>

[2] Mobile App Marketing, Analytics, Engagement | Localytics <https://www.localytics.com/>

[3] Zhang, D. and Adipat, B., "Challenges, methodologies, and issues in the usability testing of mobile applications," *Int. J. Hum. Comput. Interact.*, 2005.

[4] Nilsson, E. G., "Advances in Engineering Software Design patterns for user interface for mobile applications," *Adv. Eng. Softw.*, vol. 40, no. 12, pp. 1318–1328, 2009.

[5] Brewster, S., "Overcoming the Lack of Screen Space on Mobile Computers," *Pers. Ubiquitous Comput. Arch.*, vol. 6, no. 3, pp. 188–205, 2002.

[6] Colle, H., "Standing at a kiosk: Effects of key size and spacing on touch screen numeric keypad performance and user preference Standing at a kiosk: Effects of key size and spacing on touch screen numeric keypad performance and user preference", 2016.

[7] Huang, K., "Challenges in Human-Computer Interaction Design for Mobile Devices," *Proc. World Congr. Eng. Comput. Sci.*, vol. 1, p. 1, 2009.

[8] Chittaro, L., "Visualizing Information on Mobile Devices," *ACM Comput.*, vol. 39, no. 3, pp. 40–45, 2006.

[9] Norman, Donald A. "Design principles for human-computer interfaces." *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 1983.

[10] Nielsen, Jakob, and Rolf Molich. "Heuristic evaluation of user interfaces." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1990.

[11] Zeng, L., "Designing the User Interface: Strategies for Effective Human-Computer Interaction (5th Edition) by B. Shneiderman and C. Plaisant," *Int. J. Hum. Comput. Interact.*, 2009.

[12] Shneiderman, B. "Designing the user interface: Strategies for effective human-computer interaction" Boston: Addison-Wesley, Inc, 1987.

[13] Gong J. and Tarasewich, P., "Guidelines for handheld mobile device interface design," *Proc. DSI 2004 Annu. Meet.*, pp. 3751–3756, 2004.

[14] MacKenzie, I. S., "Fitts' Law as a Research and Design Tool in Human-Computer Interaction," *Human-Computer Interact.*, 1992.

[15] Miller, G. A., "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychol. Rev.*, 1956.

[16] Grace, M. C., Zhou, W., Jiang, X. and Sadeghi, A. R., "Unsafe exposure analysis of mobile in-app advertisements" *In Proceedings of the fifth ACM conference on Security and Privacy in Wireless and Mobile Networks*, pp. 101-112, 2012

[17] scientiamobile.com, Smartphone and Tablet Portrait Orientation Usage 2017 Q3 <https://www.scientiamobile.com/smartphone-tablet-portrait-orientation-usage-2017-q3/>

[18] uxatters.com, “How do users really hold mobile devices”  
<https://www.uxatters.com/mt/archives/2013/02/how-do-users-really-hold-mobile-devices.php>

[19] Selected articles on Nielsen Norman Group,  
<https://www.nngroup.com/articles>

[20] buffer.com, The Ideal Length of Everything Online, Backed by Research, <https://buffer.com/library/the-ideal-length-of-everything-online-according-to-science>

[21] Dyson., M. C., “How physical text layout affects reading from screen”, *Behaviour & Information Technology*, 23(6), pp.377, 2004.

## Software usability improvements for Generation Z oriented software application

W. P. N. H. Pathirana

*Department of Industrial Management  
University of Kelaniya, Sri Lanka  
pathirana.wpnh@gmail.com*

D. N. Wickramaarachchi

*Department of Industrial Management  
University of Kelaniya, Sri Lanka  
dilani@kln.ac.lk*

### Abstract

*User Experience (UX) which focuses on having a deep understanding of users, what they need, their abilities, and their limitations is one of the most important consideration when developing software applications today. Developing sustainable software application is the aim of any software company. To achieve this, software usability should be improved in the software as one of the main nonfunctional requirements. Sustainability of software usability is determined by user experience. Main users of software applications are people who belong to the workforce since they are working in the complex nature of the businesses and with time limitations. Generation Z is the new addition to the workforce and it will be increasing in the future. Therefore, it is important to pay extra attention to increase usability of software for Generation Z who have been exposed to digital technology from an early age. Accordingly, a survey method was used to identify specific characteristics, behaviors of Generation Z when using software applications and significance of usability attributes of Generation X, Y and Z with the sample size of 77 and 25 respectively. This study found that Generation Z has special software usability requirements than Generation X and Y. Based on findings, features that need to be incorporated with Generation Z oriented software applications were proposed. Moreover, a prototype was developed to demonstrate the features identified. The proposed model together with the prototype was validated through software engineers and UX engineers and validity of the model was assured by them. The model was recommended to use when developing software application for Generation Z.*

**Keywords:** *Software usability, User Experience, Generation Z*

### 1. Introduction

User Experience (UX) has become one of the most important aspects when developing sustainable software applications as it focuses on understanding users and their behaviors. Usability is a part of UX which is one of the main nonfunctional requirement which need to be considered when developing software applications.

Therefore, deep understanding of the main users of the software will be helpful to develop sustainable software.

At present people tend to use different types of software applications in their day to day life than they did earlier. Therefore, improving software usability and user experience is a must when considering users who use software frequently. Generation Z (Gen Z) consists of people who were born with technology and they are the first generation who are widely exposed to digital technologies. In coming years, this generation will play a significant role in society, being the most dominant part as clients as well as service providers. Therefore, understanding their behaviors, characteristics and expectations are very much important to develop software applications to satisfy them by improving software usability and user experience.

### 2. Literature review

In literature, different types of studies have been done to identify Gen Z in different contexts. Most of the studies have revealed that Gen Z has a straight way of communicating with preference to flexibility, transparency, self-reliance and personal freedom [1]. Furthermore, Gen Z prefers watching videos instead of reading manuals and they are highly multi-tasking [2]. Gen Z has a casual, individual and straight way of collaborating and social networking is a vital part of their lives. In the study it was found that Gen Z tend to be more entrepreneurial, trustworthy, tolerant and less motivated by money than Gen Y [3, 1]. Due to the digital native aspect, Gen Z tends to accept online sources without any critical evaluation [4] and also Gen Z sees the technology as an instrument for them [5]. Gen Z was born into a culture of active leisure, where productivity in every moment and every activity is expected [6]. According to

Schlossberg, Gen Z has higher expectations, no brand loyalty and cares more about their experiences [7]. Literature describes that Gen Z people prefer different experiences rather than doing the same thing for a couple of times and it will be directly applicable to their expectations towards software applications. Further it has been shown that smart technologies have a significant influence for expectations of consumers in Gen Z. Moreover, this group of consumers expect various new devices and electronic processes to be widely available, thus offering consumers more autonomy and faster transactions [8].

In general, usability need to be incorporated with any software and user experience also needs to be considered when developing software applications. Today, usability and user experience have become some of the most popular factors when developing software applications, because people are in search of software applications which can be easily used. Research on “Importance of usability testing and UX testing” [9] indicates that, usability testing can be very economical if planned properly, yet greatly effective and beneficial. It, also helps to give solutions for most of the problems that user may face. Software usability can be defined in different ways and different usability components can be identified from the literature. Some of them are Effectiveness [10, ISO 9241-11], Efficiency [11, ISO 9241-11], Satisfaction [11,1993; ISO 9241-11], Learnability [10, ISO 9126], Flexibility [10], Attitude [10], Understandability [ISO 9126], Operability [ISO 9126], Attractiveness [ISO 9126], Memorability [11] etc. Moreover, usability evaluation is important because it plays a vital role in connecting user-oriented and system-oriented evaluation approaches [12]. So, different usability evaluation techniques are there in the literature such as, heuristics, cognitive walkthroughs, claim analysis and concept based analysis for surface and structural misfits [13].

Software usability and the user experience can be varied from person to person. Therefore, generation gaps can be taken as one of the main causes which affect software usability and user experience. Limited number of researches are available on software usability for Generation Z [14,15]. Usability against different generations was discussed in those research papers. However, it designates that the availability of identifying specific usability requirements for Gen Z to improve their user experience while using software applications. Therefore, it is required to have a proper methodology to develop software which suits Gen Z by incorporating software features which can address their specific usability requirements.

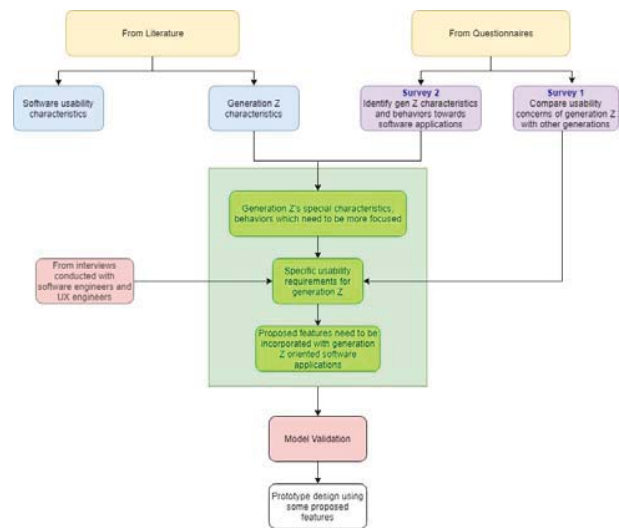


Figure 1. Methodology.

### 3. Methodology

Figure 1 shows the high-level architecture of the research design. As the first phase, usability components and Generation Z characteristics were identified from literature. Then two surveys were conducted. One of them was to identify most considerable usability factors for Gen Z than other generations (survey 1). The sample size was 25. The other survey (survey 2) was conducted to identify Gen Z’s unique characteristics and behaviors. The sample size for the survey 2 was 77. Based on the information gathered, specific characteristics and behaviors of Gen Z which need be more focused were identified. Concerning them, specific usability requirements of Gen Z, were identified with the help of industry experts. To address all these identified requirements, features should be incorporated with software applications were proposed and a prototype was developed. Finally, the proposed model along with the prototype was validated through software engineers and UX engineers.

### 4. Data collection

The data for this research was collected from different generations (X, Y and Z) which include people from general public, Software Engineers and UX Engineers through questionnaires and interviews. Initially, 25 people from Generation Z and 25 people from other generations (generation Y, X) were selected. The first questionnaire was distributed among them. This questionnaire was focused on identifying Gen Z’s most concerned factors regarding software usability than other generations. It was based on the usability components (usability attributes) and Nielsen’s 10 usability heuristics. Likert scale

questions were used to get feedback from the users as the ratings were from 1 to 5 and 1 indicated that user totally disagrees with the statement and 5 indicated that user totally agrees with the statement mentioned in the question.

The objective of the second questionnaire was to identify Gen Z's specific characteristics, behaviors when using software applications. Questionnaire was created based on the identified information regarding generation Z's characteristics, behaviors from literature and consisted of Likert scale questions, open ended questions as well as close ended questions. This questionnaire was used to emphasize the validity of the information gathered from literature.

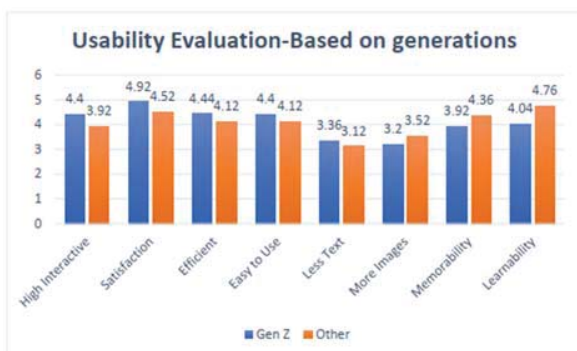
## 5. Data analysis

### 5.1. Analysis of usability factors for Gen Z and others

In the first survey, participants rated the usability components and heuristics as they expected from any software application. According to the weighted mean, the scored values are shown in the Table 1. All these scores are based on a rating score of five.

**Table 1. Usability factors for generations.**

Factor	Gen Z	Others	Difference
High Interactivity	4.4	3.92	0.48
Satisfaction	4.92	4.52	0.4
Efficient	4.44	4.12	0.32
Easy to Use	4.4	4.12	0.28
Less Text	3.36	3.12	0.24
More Images	3.2	3.52	-0.32
Memorability	3.92	4.36	-0.44
Learnability	4.04	4.76	-0.72

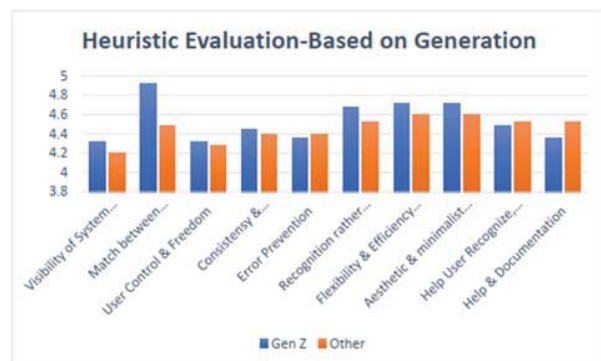


**Figure 2. Usability factors analysis for generations (graph representation).**

The analysis indicates that high interactivity, satisfaction, efficient, ease of use and less text are factors which Gen Z focuses more than, Gen Y and X.

**Table 2. Heuristics evaluation based on generations.**

Factor	Gen Z	Other
Visibility of system status	4.32	4.2
Match between system & the real world	4.92	4.48
User control & freedom	4.32	4.28
Consistency & standards	4.44	4.4
Error prevention	4.36	4.4
Recognition rather recall	4.68	4.52
Flexibility and efficiency of use	4.72	4.6
Aesthetic and minimalist design	4.72	4.6
Help user recognize, diagnose & recover from errors	4.48	4.52
Help & documentation	4.36	4.52



**Figure 3. Heuristics evaluation based on generations (graph representation).**

Table 2 shows the information on prioritized Nielsen's 10 heuristics based on the consideration of generation Z and consideration of other generations. According to the results, four factors were selected as the most important heuristics for generation Z.

### 5.2. Analysis of Generation Z characteristics

Based on the literature, different characteristics and behaviors of Gen Z were identified. Second survey was conducted for further clarification of the information collected from the literature regarding Gen Z. The second survey emphasize that characteristics and behaviors identified through literature were same as the survey findings. This survey could obtain some quantitative base for the research.

Seventy-seven Gen Z participants were surveyed and they were categorized into sub categories based on their birth year. Table 3 illustrates the key findings from the survey. The left column describes the factors that included

in the questionnaire and right column describes the related answers for the questions from Gen Z.

**Table 3. Key findings for the survey 2.**

Factors Considered	Findings
Birth year	1996 to 2001- 80.5% 2000 to 2007- 19.5%
Occupation	Undergraduate – 75.3% Student – 20.8% Employee – 3.9%
Technical Knowledge	Very good – 88.3% Good – 10.4% Average – 1.3%
Software usage in day to day life	Less than 5 software applications – 13% 5 to 10 software applications – 51.9% 10 to 15 software applications – 20.8% More than 15 software applications – 14.3%
Most common purpose for using software	Educational and entertainment – 79.2%
Most popular software applications	Instagram, Facebook, WhatsApp, YouTube
Device dependency	98.7% are very upset without their familiar device with them and when favorite software application is not supporting it

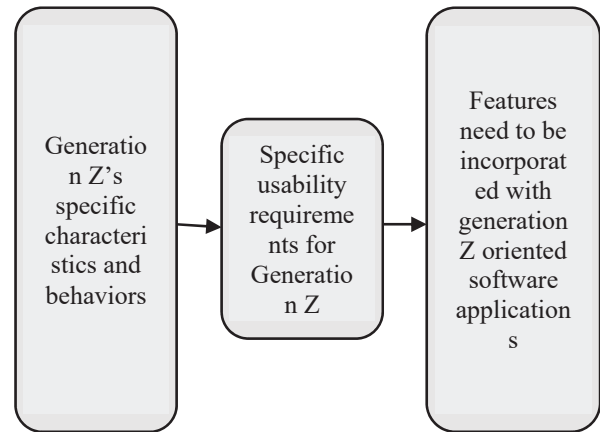
The results were from the sample of 77 Generation Z people belonging to different occupations. Based on the statistics, Gen Z can be considered as a technology sophisticated generation, and they are highly dependent on their familiar devices such as mobile phones, lap tops, tabs and etc. Moreover, their software usage was based on educational and entertainment purposes.

Furthermore, Gen Z expectations from software applications such as voice recognitions, AR/VR based solutions, helping AI bots and etc. were collected and considered when defining features for them.

## 6. Results

Considering all information gathered from literature and surveys, Gen Z's characteristics and behaviors which need to be more focused were identified. Based on them, specific usability requirements for Generation Z were identified and the software features that need to be incorporated into software applications were proposed, to address the identified specific usability requirements of them.

Proposed model is shown in the Figure 4 as the high-level diagram and Figure 5 illustrates the low-level diagram of the proposed model.



**Figure 4. High-level architecture of the model.**

When it comes to identifying Generation Z's specific characteristics and behaviors, all the identified factors were emphasized from the survey as well. Identified and emphasized Generation Z's specific characteristics are shown in the Table 4.

**Table 4. Generation Z's specific characteristics.**

From Literature	From questionnaire
Generation Z is technology sophisticated	According to the questionnaire generation Z mentioned that 88.1% have very good technical knowledge and others have rated as good and average. They have not mentioned that they have poor technical knowledge
Culturally varied	Selected sample contains people from different cultures, different religions and different ethnicities
Social networking is important and social circles not restricted by geography	Generation Z's most favorite applications were Instagram, Facebook, WhatsApp and YouTube. All are related to social networking. Therefore, we can say that they are interested in social networking
Handle short, up to date, real time information with pictures Prefer videos instead of reading text Short attention span	They have mentioned this type of things as their expectations from the software applications
Impatient, rebellious and expecting instant result	They have mentioned that they need speed and the efficiency from the software application



Experience and pleasure is more important Surprisingly social and practical	Satisfaction was chosen by generation Z as their most considerable software usability factor. They are more focus on their experience and the pleasure. They also expect real world experience from a software application	provides an option to have team collaboration, it will be a motivational factor for them to use the software application
Device dependent	They will really upset if their familiar device is no longer with them and their favorite applications are not supporting for it	Should be developed using latest technologies Is very technology savvy. Therefore, they always look for new trends, new technologies instead of using same technologies in long period of time. Therefore, software should be upgraded to the latest technologies available in the market

After analyzing and identifying most considerable factors for generation Z, specific usability requirements for generation Z, which support their characteristics could be identified. Research could identify the requirements that should be there in a software application and how it should be incorporated with software to satisfy and improve the user experience of generation Z.

Proposed specific usability requirements for generation Z according to most considerable factors of generation Z when developing software applications are as follows.

- Should be able to use user preferred languages
- Should be a team collaborative application
- Should be developed using latest technologies
- Should be integrated with social media
- Should include less text and multimedia (aesthetic and minimalist design)
- Should attractive, efficient, flexible, high interactive
- Should be able to improve satisfaction
- Should be able to provide real world experience
- Should compatible with different devices – especially mobile
- Should be able to achieve goals throughout the system working process

Identified requirements are further described in Table 5.

**Table 5. Identified specific software usability requirements for Generation Z.**

Requirement	Description	
Should be able to use user preferred languages	Culturally varied and they belong to different ethnicities, states, traditions etc. Therefore, their use of language, use of terms can be different from one person to another person. Due to that people may feel uncomfortable with the unfamiliar terms and languages. It directly affects to the user experience. Therefore, delivering software applications with language selection option will help to improve user experience of them	Should be able to provide real world experience They are very practical people who look for systems which feel like real world experience. Therefore, software need to provide real world experience for them to satisfy them. Then they will stay with the software applications
Should be a team collaborative application	Social circles are not restricted by the geography. It indicates that they are really interested in building social circles among each other. If the software application	Should compatible with different devices – especially mobile Their devices are important for them and they are highly dependent on it. When developing software applications compatibility with different devices need to be considered. If the software is not compatible with most of the devices, it will not provide good user experience
		Should be able to improve satisfaction Experience and pleasure is more important. They do things to get good experience and to be happy. If the first experience of a software use is not good for them they will not use that software application again. Therefore, software need to be developed to satisfy user and give good experience for the user. Ability to improve the user's satisfaction should be there in a Generation Z oriented software applications
		Should be able to achieve goals They are goal oriented and always try to achieve goals and develop themselves. Implementing goal oriented applications

throughout the system working process | will help to give them a challenging experience.

Then the software features that need to be incorporated with software applications were proposed based on the identified special usability requirements. Table 6 illustrates the software features needed to be incorporated with Generation Z oriented software applications.

**Table 6. Software features for generation Z oriented software applications.**

Feature	Description
Option to support multi languages	Software can be shifted to a language that user preferred. It is not feasible to develop the software for every language. But generation Z oriented software applications need to develop with the language selection option. Then users can select their familiar language and their experience will increase.
Availability of making forums, feedback, chat features	By doing this user can talk about the application's problems, good things in the forum. By developing social circles, idea sharing can be done. User experience will increase with that as well as application can be further improved using user feedback, ideas, etc.
Integrate gamification	Gamification is building a game like environment in software applications which integrating game elements like badges, points, leader boards and achievements to the software. As an example, according to user performances software can give badges, points to the users to motivate them. If it is an entertainment application, points can be given according to the number of shares made by the user etc. Likewise, different mechanics can be used to different types of software applications
Regular updates	Software development team should focus on software upgrades. Also new technologies need to be incorporated with the development.
Include graphics, animations, audios, video	Try to get the user attraction by adding multimedia instead of using text to pass the whole message.
Option to interface customization and voice recognitions	Generation Z is much more technology sophisticated and they expect ease of use, flexibility, personal freedom etc. Therefore, customization options should be there. Also, to improve the personal freedom add generation could be done without any disturbance.
Include AI bots,	To give real world experience and high interactivity, this type of features need to

simulations, 3D objects	be incorporated with generation Z oriented software applications.
Support for major platforms	Application need to be developed which supports for major platforms available. Also, responsiveness also need to be considered.
Add goals with the user profile analysis	As a main suggestion, goal establishment can be done because generation Z is more focusing on achieving goals. If the software application establishes goals to achieve user tries to achieve them. It will be a very good experience for them. As an example, for health applications can analyses user details and give a summary saying that "you have successfully burnt 110 calories per day and your target is 200 calories. You need to burn extra 90 calories within today" likewise. This type of hints, information can be included to the software application by analyzing the user.

## 7. Model validation

Once the model was created, a prototype was developed to demonstrate the software features identified and it was validated along with the model through a questionnaire.

The questionnaire was distributed among Software Engineers and UX Engineers for the evaluation of the model and the prototype. Twenty-five industry experts responded to the questionnaire and among them 60% were software engineers and 40% were UX engineers. All of them had accepted that user experience can be different from one generation to another based on their characteristics. Therefore, they agreed with the research objective, and the importance of having proper model to add software features for Generation Z oriented software applications.

Based on their comments and reviews overall acceptability and the feasibility of the model was evaluated. From the sample, 60% of engineers highly recommended this model to identify necessary software features for Generation Z based on their characteristics and their specific requirements. They also had mentioned that this model can be used to develop sustainable software for Gen Z to meet their expectations with 36% of them indicating that this model is usable to address specific usability requirements of the target audience. However, they had also mentioned that it is not always applicable for each and every software application. A small group (4%) of respondents mentioned that while this model is acceptable, when it comes to implementation some issues may need to be addressed.

Overall, their feedback indicated higher validity of this model and indicates that this model can be used to address Generation Z's specific usability requirements by

integrating software features with the generation Z oriented software applications.

## 8. Discussion and conclusion

When considering the software development industry, Generation Z will become the main target users for the industry in the near future. Therefore, Generation Z users need to be targeted and developers need to identify their characteristics related to use of software applications. Therefore, this research focused on increasing software usability, user experience for Generation Z.

The model was developed to identify Gen Z's specific characteristics, behaviors, specific usability requirements. Based on the identified characteristics, software features were incorporated with software applications to satisfy Generation Z when they use software. This model can be used to improve usability and the user experience of Gen Z by incorporating software features proposed in the model.

Proposed model is based on qualitative information from literature and quantitative information from surveys. It has made a concrete base for further research. It also has been validated by Software Engineers and User Experience Engineers. Therefore, the accuracy of the model has been established.

However, the sample is limited to a specific group and thus needs replication in other countries as well. More information could have been gathered if the sample was taken from different countries without it being limited to one country.

As future work AI bot can be implemented which can be integrated with any software applications. Furthermore, some of these information can be used for marketing related research on Generation Z.

## 9. References

- [1] Singh, A.P., Dangmei, J., "Understanding the generation Z: the future workforce", *South-Asian Journal of Multidisciplinary Studies (SAJMS)* ISSN: 2349-7858: SJIF: 2.246: Volume 3 Issue 3, 2008.
- [2] Salleh, M. S. M., Mahabob, N. N., and Baharudin, N. S., "Overview of Generation Z Behavioural Characteristics and Its Effect Towards Hostel Facility", *International Journal of Real Estate Studies*, Volume 11 Number, 2017.
- [3] Dan Schawbel. *Gen Z Employees: The 5 Attributes You Need to Know*, 2014.
- [4] Seemiller, C., & Grace, M. *Generation Z Goes to College*: John Wiley & Sons, 2016.
- [5] Van den Bergh, J., & Behrer, M. "How cool brands stay hot: Branding to Generations Y and Z", London: Kogan Page Publishers, 2016.
- [6] Merriman, M., "What if the next big disruptor isn't a what but a who? Gen Z is connected, informed and ready for business", Retrieved from [www.ey.com](http://www.ey.com), 2018
- [7] Schlossberg, M., "Teen Generation Z is being called 'millennials on steroids,' and that could be terrifying for retailers", Retrieved 9, 2016, from [Business Insider UK](http://BusinessInsider.com)
- [8] Priporas, C. V., Stylos, N., and Fotiadis, A. K., "Generation Z consumers' expectations of interactions in smart retailing: A future agenda", *Computers in Human Behavior*, 77 (2017) 374-381, 2016.
- [9] Niranjanamurthy. M., Nagaraj, A., Gattu, H., Shetty, P.K., "Research Study on Importance of Usability Testing/ User Experience (UX) Testing", *International Journal of Computer Science and Mobile Computing*, I JCSM C, Vol. 3, Issue. 10 October 2014, pg. 78-85.
- [10] Shackel, B. "Usability – context, framework, definition, design and evaluation", in Shackel, B. and Richardson, S. (Eds), *Human Factors for Informatics Usability*, Cambridge University Press, Cambridge, 1991, pp. 21-37
- [11] Nielson, J., *Usability Engineering*, Academic Press, Cambridge, MA, 1993
- [12] Joo, S., Lee, J. Y., "Measuring the usability of academic digital libraries", *Instrument development and validation, The electronic Library*, Vol. 29 ISS: 4, 2005, pp. 522-537
- [13] Blandford, A., Keith, S., Connell, I., and Edwards, H., "Analytical usability evaluation for digital libraries: a case study", in *Proceedings of the 2004 Joint ACM/IEEE Conference on Digital Libraries*, 2004, pp. 27-36.
- [14] Djamasi, S., Siegel, M., and Tullis, T. "Generation Y, Web Design, and Eye Tracking," *International Journal of Human Computer Studies*, 2010.
- [15] Paz, F., Pow-Sang, A., "Systematic Mapping Review of Usability Evaluation Methods for Software Development Process", *International Journal of Software Engineering and Its Applications*, Vol. 10, No. 1, 2016, pp. 165-178.

## Green Cloud Computing: A review on adoption of Green-Computing attributes and vendor specific implementations

J. M. T. I. Jayalath, E. J. A. P. C. Chathumali  
*Faculty of Graduate Studies and Research*  
*Sri Lanka Institute of Information Technology*  
*Sri Lanka*  
 thilini.j@slit.lk

K. R .M. Kothalawala, N. Kuruwitaarachchi  
*Faculty of Graduate Studies and Research*  
*Sri Lanka Institute of Information Technology*  
*Sri Lanka.*

### Abstract

*With cloud computing emerging as a trending topic, it has been a major point of discussion for the last few years. In regards to technological advancements, the associated shortcomings like environmental footprint caused by them also become an affair of high significance. Cloud computing itself is a much greener alternative to individual data centers with lesser number of servers being used and cloud data centers being far more efficient than those of traditional thereby reducing the carbon impact. Nonetheless, it cannot be neglected the fact that the data centers utilized by the cloud vendors are still a major source of carbon emissions due to the dirty energy usage. Therefore, the discussion of the paper is based on how green the foremost cloud providers are and the implementations of green IT attributes in the cloud infrastructure.*

**Keywords:** *Cloud Computing, Green Cloud Computing, SaaS, PaaS, IaaS, Energy Efficiency, Power Management, Virtualization, Data Centre, Resource Management, Virtual Machine(VM)*

### 1. Introduction

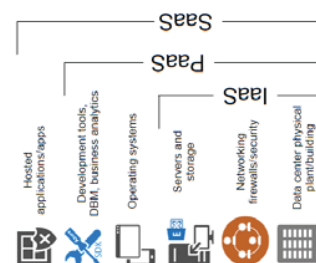
Millions of mobile and internet users share photos and documents through some applications which are supportive for the critical operations of personal or business work. For this kind of operations, it is very important to have a platform which provides accelerated access to flexible and low cost IT resources. In order to provide this platform, the novel service model called cloud computing has been introduced to the world. According to Taroub Ahmed Mustafa, cloud computing accommodates an easy way to access an extensive set of application services, databases, storage and servers over the internet. Cloud services platform provides the elasticity and rapid access for flexible cost by only paying for what the user uses. Data and services are enduring in common, scalable resource pools which are dynamically expandable, generally virtualized in a cloud computing architecture. And those data and services are available for

any verified device through the internet [1]. Google Cloud Platform, Amazon Web Services, IBM Cloud, Oracle Cloud, Microsoft Azure, Adobe, VMware, Rackspace and Red Hat are some of distinguished cloud service providers and those providers charge their customers for consuming the services according to the usage.

Another way of explaining cloud computing is, explaining the development models of cloud computing which can be described as a computing methodology that involves resource, instead of arranging the personal machine to administrate the request and the demanded applications of the user. Followings are development models of cloud:

- **Public cloud:** This will be available for any organization by allowing users to deploy and build services in the cloud environment.
- **Private cloud:** Generally, this cloud is available, maintained, controlled and deployed for a particular organization.
- **Community cloud:** Allows the users of number of specific organizations with shared concerns.
- **Hybrid cloud:** This can be a grouping of any other types of clouds (private, public or community), which have the capability over the interfaces to move and allow data or applications between more than one cloud.

Cloud computing is composed of three basic types of services as shown in the Figure 1.



- **Infrastructure as a service (IaaS):** In this case the vendor offers access for resources like networking, storage and servers for the users by providing their own applications and platform within the infrastructure. So as a replacement for buying hardware, users can pay for IaaS on demand and there is no any single possibility even for a single point of failure. This saves money and time. This service provided by the vendors' like Microsoft Azure, Amazon Web Services (AWS), Cisco Metapod, and Rackspace.
- **Platform as a Service (PaaS):** Here the third-party provider distributes software tools and hardware for the customers over the Internet. This third-party provider has hosted the software and hardware on his own infrastructure, so that user can proceed with the complete development and deployment environment in the cloud. PaaS includes business intelligence (BI) services, infrastructure networking, development tools, database management systems, middleware and more. PaaS provided by Windows Azure, AWS Elastic Beanstalk and Google App Engine.
- **Software as a Service (SaaS):** This is a main service type of cloud computing where hosts applications, and makes them available to customers over the Internet by a third-party provider. Since the user or the organization does not need to install and run applications on their personal data centers or in their computers. This reduces the expenditure of purchasing hardware, software licensing, maintenance cost, cost for installation and support. Dropbox, Google Apps and Cisco WebEx are some real world examples for SaaS.

Nowadays larger organizations like Google and Facebook, have to cater millions of services to users at the same time which requires number of data centers. Unlike traditional datacenters, with the cloud these larger organizations require only fewer servers which are more power efficient and reduce the carbon influence on the organizations datacenters. Environmental protection agency has declared that “datacenters now account for 1.5% of all electricity consumption in the U.S. and by 2020, carbon emissions will have quadrupled to 680 million tons per year”. Even though the cloud is a greener solution compared to the traditional data centers, with the number of users and the user requirements rapidly increasing, the cloud becomes a major cause of carbon emissions.

In the modern world the accumulating progress of computational demand and data storage, cloud computing has become a comprehensive area for researches. Because of these enhancements in the technology, the subsequent step should be the establishment of environmentally-safe, energy efficient and cost effective solutions. IT resources

consume a lot of energy and power. Because of that, the global climate changes such as increase of CO<sub>2</sub> emission and energy shortages can be occurred. For this reason, there is an essential requirement of “Green Cloud Computing” which is capable of producing explications that make cost effective and energy efficient IT resources. The concept of the Green Computing was started in 1987, with the awareness of sustainable development. In 1992, one consumer Energy Star plan is launched by the U.S. Environmental Protection Agency (EPA) [2].

Using the computers and further technologies in very effectual way, will increase the consumption of resources, energy efficient peripherals and decrease the electronic waste. This conception can be called as the Green Computing. Green computing will increase the resource efficiency and performance enhancements. Yashwant et al. discussed about the two aspects of Green computing as follows [20]:

- **Software:** Develop methods in a way to increase the efficiency of program, energy and storage.
- **Hardware:** Necessity of technologies to reduce the consumption of energy and economically efficient.

## 2. Applications of Green Computing

In order to increase the resource efficiency and performance there are seven application areas introduced for green IT which can be applied in Green cloud computing as follows;

- Proper management of power
- Energy efficient resources
- Design of data centers
- Virtualization of servers
- Environment sustainability design
- Eco-labeling for IT products
- Design of recycling methods

### A. Proper management of power

Since green cloud computing encourages the technology to be environmentally-safe, energy efficient and cost effective solutions, it is very important to maintain proper power management in the cloud. If all requested services are catered to users by the cloud, there can be an opportunity of services that those might not be used for the full extent. This can result in additional amount of carbon is being produced to the environment and it can be very dangerous in terms of air pollution. So that, it is essential to check whether the resources are utilized or not.

In 2015, Rakshith and Sreenivas conducted a research on power efficiency of green cloud computing and they

have found that servers run at low utilization in data centers can be the key reason of energy inefficiency in data centers. Since there is a circumstance where even though usage of resources is fewer, the power consumed is more than the power spent on usage. As per their research, energy consumption can be identified as the main reason in content distribution system and these demand an accretion of networked computing resources from one or numerous service providers on datacenters across the world. In cloud computing systems and modern data center, this above mentioned expenditure is acting as a disapproving design parameter. The high carbon emission and energy cost are depending on the power and energy consumed by the computer equipment and the connected cooling system [3].

To address these problems, Rakshith and Sreenivas has proposed a new concept for increasing power performance of data centers, cloud application. These authors mainly characterize power performance of data center and cloud computing and proposed possible techniques to minimize the power requirement. There are number of technical algorithms which can be used to manage the power consumption, but Agglomerative Hierarchical Clustering algorithm is the mostly used algorithm and it is a 'bottom up' approach where each observation starts in its own cluster, and pairs of clusters are merged as it moves up the hierarchy. When the job is submitted by the user, it will be located in job queue and later it will be assigned to Virtual Machines. Prior to that, each Virtual Machine (VM) acts as a separate cluster and separate mechanics are used to calculate the energy efficiency of each VM and machines, with same energy is combined together to form a single group. VM's are categorized into three main cluster groups called as small size, medium size and large size cluster. When the power load is low, energy values will be less and if load grows then it will be moved to Medium size cluster. This algorithm helps to reduce the CO<sub>2</sub> emission into the environment and also it avoids the problem related to carbon emission into the environment. Resources will be in idle state if those are not mandatory or used at the specific time. Because of that power utilization in the ideal method, the CO<sub>2</sub> production into the environment is decreasing [3].

### B. Energy efficient resources

With the increased usage of cloud computing, the use of huge data centers and high performing computers are expanding continuously. As a result, energy consumption by the data centers, energy dissipation in the environment, the amount of CO<sub>2</sub> emissions and the heat generated by processor chips are also rising gradually. For an example one data center occupies 50,000 square feet and requires 5-megawatt (mw) electricity which covers the power of 5000 households for one year. To avoid this problem, Green Cloud Computing has introduced. The aim of this is to develop high end computing systems like data centers and clouds to meet not only user's quality of

service requirements, but also minimize the utilization of electric power. Apart from that it has defined to manage energy effectively and efficiently [4]. Most of the time cloud data centers remain underutilized. There are huge numbers of virtual machines running on top of these underutilized machines. Therefore, it needs high energy and power. In order to preserve significant volume of energy they are migrating virtual machines to other machines and hibernating such underutilized machines. As the data centers raise highly obsessed with energy efficiency, green cloud computing provides a rare opportunity to reduce data center power bills [5].

In cloud computing there are some metrics which are used to measure the energy efficiency of a data center [4]. Some of them are Power Usage Effectiveness (PUE), Thermal Design Power (TDP), Data Center Infrastructure Efficiency (DCiE), Performance per Watt, Compute Power Efficiency (CPE), Green Energy Coefficient (GEC), Energy Reuse Factor (ERF), Carbon Usage Effectiveness (CUE), Data Center Productivity (DCP) and Space, Wattage and Performance (SWaP), Water Usage Effectiveness (WUE) [4]. According to Sasikumar et-al. [5], the most common metric is Power Usage Effectiveness (PUE). To calculate the Power Usage Effectiveness, total power entering to the data center should be divided by the power used by IT equipment.

$$\text{PUE} = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

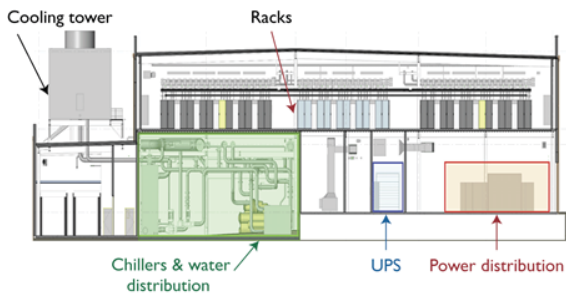
This equation can be represented as follows [5].

The power used by the support equipment is defined as overhead load and it mainly consist of power delivery, cooling systems and other infrastructure facilities like lighting.

Green cloud computing describes different techniques to minimize the power consumption. Some of them are reducing CPU power dissipation, use of renewable energy resources, use of energy efficient storages, reducing cooling requirements, use of advanced clock gating [4] etc.

### C. Design of data centers

Data centers are the heart of cloud computing and running thousands of server side workloads, data storage and large scale of data and scientific processing. Today, large data centers consume hundred Megawatts of electricity. Therefore, improving data center efficiency is a major concern in the industry world. Past researches has described about system design and analysis of power, water, and carbon usage of the Massachusetts Green High Performance Computing Center (MGHPCC), a 90,000 square feet, 15 MW. MGHPCC is a data center that uses recent advances in cooling and power distribution to improve energy efficiency.



**Figure 1. Layout of the MGHPCC data center.**

MGHPCC consumes a favorable geographical location and green design to accomplish energy efficient operation. MGHPCC is situated in Holyoke in western Massachusetts. Due to its location, MGHPCC obtains inexpensive and abundant electricity, inexpensive real estate, and proximity to fiber-optic network backbones. Western Massachusetts has a cool climate and it enables the facility to employ free cooling as shown in the Figure 2. When data centers are located in cold areas, it reduces the cooling energy consumption [1]. The power infrastructure for the data center simulates a slight scale distribution network in the electric grid. The infrastructure contains substations, feeders, transformers, and switchboards that feed power to the computing and cooling infrastructure. When developing a green data center cooling infrastructure, data center monitoring, power infrastructure and physical layout are the major areas that should be considered.

#### D. Virtualization of servers

Creating virtual resources such as desktop, operating system, file, server, storage or network are called as the virtualization. There are seven main types of virtualization as; Server virtualization, Application virtualization, Desktop virtualization, Hardware virtualization and Network virtualization. For green cloud computing server virtualization plays major role by masking of server resources, including the number and identity of individual physical servers, processors, and operating systems, from server users. In order to separate one single physical server into several remote virtual environments, administrator can use a software application. The main three approaches for the server virtualization are paravirtual machine model, virtual machine model and virtualization at the operating system layer. In 2012 Sasikumar and Yuvapriya has discussed about operating the servers in a Cloud data center. In order to minimize power consumption within the data center VM scheduling algorithm can be used. Power-aware and thermal aware scheduling are the main two types of green scheduling systems for supercomputers. In order to minimize the overall data center temperature, jobs can be scheduled and that pattern is called as the thermal-aware scheduling. It helps to decrease the energy needed to control the data

center cooling systems. Power-aware scheduling arranged to decrease the server's total power by scheduling the jobs to nodes. Operating the servers is the most valuable function within a Cloud data center. VM schedules algorithm that minimizes power consumption within the data center [13].

#### E. Environment sustainability design

The major concern of Environment sustainability design (ESD) is on reducing or completely eradicating undesirable environmental impacts in the long term with the aid of attentive design. In a study conducted by Microsoft, Accenture and WSP in 2010, it has been identified that the reduced environmental footprint of cloud computing is mainly influenced by several key factors such as Dynamic Provisioning, Multi-Tenancy, Server Utilization and Data Center Efficiency [21]. Dynamic provisioning refers to reducing the over-allocation of infrastructure for which the datacenters maintain servers considering current demand thus resulting in lesser consumption of energy. The application instances are shared between client organizations using the multi-tenancy approach which assists in leveling the peak demand as well as reducing the general energy consumption and related carbon emissions. Server utilization indicates operating servers at higher utilization rates with the aid of virtualization techniques thereby decreasing the number of active servers. Though high server utilizations imply more power consumption, highly utilized servers are capable of accommodating workload with same power usage. Moreover, data center efficiency, expressed by power usage effectiveness (PUE), fundamentally affects the energy usage in cloud computing. A higher PUE is achieved through data center infrastructure designs including modular container design, air or water evaporation based cooling, or advanced power management through power supply optimization.

#### F. Eco - Labeling

Eco-labeling refers to the mechanism of accreditation given to a product or a service by an unbiased third party which are considered to have less effects on the environment and therefore environmentally desirable. Eco-labeling comes under Type 1 of the three key voluntary environmental labeling types recognized by the International Organization for Standardization (ISO).

### 3. Vendor involvement

With the necessity for going green getting stronger than ever, many leading companies have started taking interest in environmental friendliness. Thus, in regards to green cloud computing, the foremost cloud providers such as Amazon Web Services (AWS), Microsoft, Google and IBM has taken some major steps concerning adoption of a greener cloud at the infrastructure level.

### A. AWS

Microsoft Amazon Web Services (AWS), the cloud platform provided by Amazon, is one of the leading cloud vendors offering services such as Infrastructure as a service (IaaS), Platform as a service (PaaS) and packaged Software as a service (SaaS). In regards to building a green cloud platform, AWS has taken certain prominent actions in the recent years, promising leadership in supporting clean energy while also being committed to solely rely on renewable power along with building solar and wind farms.

In a report released by environmental activist group Greenpeace in 2017, it has been discussed on how green the Amazon's cloud is along with relevant statistical data [22]. Prioritizing the environment, AWS has joined major tech companies in filing an amicus brief encouraging the Clean Power Plan of the US Environmental Protection Agency which focuses on restraining power plant emissions. Moreover, AWS is also a part of the non-profit organizations concentrating on the environment such as American Council on Renewable Energy. AWS has stated that in 2016, 45% of its energy was from renewable sources and it is expected to raise up to 50% by the end of 2017.

Nevertheless, Greenpeace states that the AWS US East region facilities located in Virginia are only about 3% clean with about a third of the power supply being associated with coal, nuclear and natural gas [22]. However, the AWS regions such as US West (Oregon) and EU (Frankfurt) are much greener with 85% of power being generated by renewables for data centers in Oregon. As per the statistics of Greenpeace, energy demands of AWS are met having 30% of coal, 26% of nuclear generation and 24% of natural gas while only 17% percent of the energy requirements are covered with clean energy. Greenpeace also specifies that Amazon is the least transparent company in regards to disclosing its energy footprint.

Nonetheless, in January 2018, AWS has been able to achieve 50% renewable energy usage [23]. Furthermore, the company has initiated 10 renewable energy projects. The projects include solar farms in Virginia such as Eastern Shore, 80 megawatt solar farm in Accomack County, and five 20 megawatt solar farms Buckingham, New Kent, Sappony, Scott and Southampton in Buckingham, New Kent, Sussex, Powhatan and Southampton counties respectively. Additionally, the projects also consist of wind farms, for example, Amazon wind farm Indiana - Flower Ridge, which is a 150 megawatt wind farm in Benton region, Amazon wind farm North Carolina - Desert Wind which is a 208 megawatt wind farm in Perquimans and Pasquotank regions and lastly Amazon wind farm Ohio - Timber Road, which is a 100 megawatt wind farm in Paulding region. The aforementioned projects are anticipated to

produce over 2 million megawatts of energy per year which should be sufficient to power a city size of Atlanta, Georgia. With these advancements, AWS states that the company is always striving toward cleaner and renewable energy sources [23].

### B. Microsoft

Microsoft is one of largest ICT Company that has committed efforts to becoming green. Their data centers and its operations have reduced carbon radiations by 9.5 metric tons. Apart from that 44 percent of Microsoft data centers come from renewable resources like new wind, solar energy and hydro. In other words, Microsoft purchased 500 Megawatts from these resources to maintain their data centers. Also, they have a mechanism to recycle consumer e-waste and nearly 10 million of them have recycled by today [6]. Recently Microsoft launched a project called 'Project Natick' to test underwater data centers by using submarine technology. Ultimate goal of this project is to take the advantage of renewable marine energy and also to cool the equipment. Microsoft eventually hopes to develop under water data centers with two decades' life span, deploy over five-year periods where computer hardware is replaced at the end of each deployment.

### C. Google

According to the statistics of Greenpeace report on the greenest IT companies of 2017, Google is the leading one [22]. Major reason for this is the sustainable data centers that are maintained by the Google Cloud. To maintain the sustainability, they are using the concepts like renewable energy, efficient data centers and circular waste management [8]. Google has maintained carbon neutral operations in last few decades and according to them, it will continue in the future as well. Google is maintaining 14 data centers all over the world. Those data centers are well designed to maximize the efficiency and minimize the environment impact. To increase the efficiency, they are using technologies like smart temperature, smart lighting controls, advanced cooling techniques and a redesign of the distribution of power to reduce energy loss. Altogether on average Google data centers use 50% less energy than others. Also, the G Suite products (Gmail, Calendar, Docs, and Drive) which are managed by Google, have reported reductions in IT energy use and carbon emissions of 65% to 85%. According to Google when the user chooses Google Cloud Platform to run computers, store data and develop applications, the digital footprint is offset with clean energy, which reduces the impact on the environment [9].

### D. IBM

IBM is another cloud vendor which has focused on "Green" for over 35 years. It has used green technology to build the company's largest data center. To do that, they have used a variety of green technologies like high-



density computing systems that use server and storage virtualization, energy efficient power and cooling systems [12]. Also, they have followed a green infrastructure which includes data center assessment and services, consolidation and Optimization Services, Tivoli energy management solutions, Tivoli Maximo asset management solutions, asset management services, Tivoli storage management and information management solutions. Apart from those, IBM has introduced sustainable solutions like strategic carbon management, carbon intelligence and sustainable logistics design and asset management [10].

#### 4. Discussion

The focus of the paper is to discuss the green IT attributes which are adopted in cloud computing domain in order to achieve green cloud computing and the efforts of the cloud vendors in becoming green. It has been discussed how several green IT aspects can be applied in cloud computing in order to achieve green cloud computing.

As far as cloud vendors are concerned, all prominent cloud providers like AWS, Microsoft, Google and IBM have made significant efforts in becoming environmental friendly. While vendors like AWS, Microsoft, Google have all committed to entirely rely on renewable power, they have also started building data centers powered by renewable sources. The Table 1 and 2 demonstrates the scores assigned by the environmental activist group Greenpeace in 2017 on clean energy usage for the major cloud vendors [22].

**Table 1: Greenpeace report statistics related to major vendors.**

	Final Grade	Energy Transparency	Renewable Energy Commitment	Energy Efficiency	Renewable Procurement	Advocacy
AWS	C	F	D	C	C	B
Google	A	B	A	A	A	A
IBM	C	C	B	C	C	F
Microsoft	B	B	B	C	B	B

**Table 2: Grades given by the Greenpeace in 2017.**

	Clean Energy Index	Natural Gas	Coal	Nuclear
AWS	17%	24%	30%	26%
Google	56%	14%	15%	10%
IBM	29%	29%	27%	15%
Microsoft	32%	23%	31%	10%

Greenpeace has evaluated the performances of each vendor based on sources such as directly offered information, public information by the vendors and reporting bodies while also taking media coverage and published reports into account. The final grades for each vendor have been assigned as per 20% weight each for Transparency, Renewable Energy Commitment & Siting Policy and Advocacy, 10% weight for Energy Efficiency & GHG Mitigation and assigning Renewable Energy Procurement, including current energy mix with a weight of 30% [22].

#### 5. Conclusion

Cloud computing is a topic which has been able to transform the organizational perspective of IT resources with it bringing out various benefits. The advantages of cloud include it being much more energy efficient than having individual data centers. However, cloud computing can still lead to certain negative environmental impacts. Hence, the paper has discussed about the adoption of green IT attributes to the cloud along with how green the cloud is based on certain criteria, taking major cloud service providers such AWS, Google, Microsoft and IBM into account. Consequently, while the major cloud providers have several undesirable effects on the environment to a certain extent, they are advancing towards becoming green aiming green cloud computing.

#### 6. References

- [1] T. A. Mustafa Sa'ed, "Toward Green and Mobile Cloud Computing," in *Proceedings of the IEEE Seventh International Conference on Intelligent Computing and Information Systems*, 2015.
- [2] Y. S. Patel, N. Mehrotra and S. Sonar, "Green Cloud Computing: A Review on Green IT Areas for Cloud Computing Environment," in *Proceedings of the 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management*, 2015.
- [3] K. Rakshith and T. H. Sreenivas, "Green Cloud Computing – Power Efficiency," *International Journal of Computer Science and Information Technologies*, vol. 6, no. 6, p. 4, 2015.

- [4] A. Jain, M. Mishra, S. K. Peddoju and N. Jain, "Energy Efficient Computing- Green Cloud Computing," 2013.
- [5] Y. Ponnusamy and S. Sasikumar, "Application of Green Cloud Computing for Efficient Resource Energy Management in Data Centres," *International Journal of Computer Science and Information Technologies*, vol. 3, no. 5, p. 7, 2012.
- [6] "PROSERVEIT," [Online]. Available: <http://www.proserveit.com/environmental-impact-green-cloud-computing>. [Accessed 05 09 2018].
- [7] "Data Center Knowledge," [Online]. Available: <https://www.datacenterknowledge.com/archives/2016/05/20/microsoft-expands-green-data-center-ambitions>. [Accessed 06 09 2018].
- [8] "Crystalloids Innovations," [Online]. Available: <http://www.crystalloids.in/blogs/4-reasons-why-the-google-cloud-is-a-sustainable-option-for-your-big-data-storage>. [Accessed 05 09 2018].
- [9] "Google Cloud," [Online]. Available: <https://cloud.google.com/environment/>. [Accessed 04 09 2018].
- [10] "IBM Green Computing – Let's Build a Smarter Planet," IBM Corporation, 2009.
- [11] "IBM," [Online]. Available: [https://www.ibm.com/ibm/green/data\\_center.html](https://www.ibm.com/ibm/green/data_center.html). [Accessed 10 09 2018].
- [12] "Network World from IDG," [Online]. Available: <https://www.networkworld.com/article/2291792/data-center/ibm-uses-green-technology-to-build-the-company-s-largest-data-center.html>. [Accessed 02 09 2018].
- [13] Y. Ponnusamy and S. Sasikumar, "Application of Green Cloud Computing," *International Journal of Computer Science and Information Technologies for Efficient Resource Energy Management in Data Centres*, vol. 3, no. 5, p. 7, 2012.
- [14] Rubyga and R. S. Dr. Ponsy, "A Survey of Computing Strategies for Green Cloud," in *Proceedings of the Second International Conference on Science Technology Engineering and Management*, 2016.
- [15] G. Jagadeeswara Rao and G. B. Stalin, "Energy Analysis of Task Scheduling Algorithms in Green Cloud," in *Proceedings of the International Conference on Innovative Mechanisms for Industry Applications*, 2017.
- [16] G. P and C. Dr. Robin, "A Comparative-Study of Load-Cloud Balancing Algorithms in Cloud Environments," in *Proceedings of the International Conference on Energy, Communication, Data Analytics and Soft Computing*, 2017.
- [17] E. Y. Goyal, M. S. Dr. Arya and E. S. Nagpal, "Energy Efficient Hybrid Policy in Green Cloud Computing," in *Proceedings of the International Conference on Green Computing and Internet of Things*, 2015.
- [18] F. Shakeel and S. Sharma, "Green Cloud Computing: A review on Efficiency of Data Centres and Virtualization of Servers," in *Proceedings of the International Conference on Computing, Communication and Automation*, 2017.
- [19] V. Londe and M. Math, "Green Computing based Cost Optimization in Cloud Computing," in *Proceedings of the International Conference on Energy, Communication, Data Analytics and Soft Computing*, 2017.
- [20] Yashwant Singh Patel, Neetesh Mehrotra and Swapnil Sonar, "Green Cloud Computing: A Review on Green IT Areas for Cloud Computing Environment," in *Proceedings of the 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management*, 2015.
- [21] "NEWS RELEASE," [Online]. Available: <https://newsroom.accenture.com/industries/systems-integration-technology/microsoft-accenture-and-wsp-environment-energy-study-shows-significant-energy-and-carbon-emissions-reduction-potential-from-cloud-computing.htm>. [Accessed 26 09 2018].
- [22] "Clicking Clean:Who is winning the race to build a green internet," [grenpeace.org](http://grenpeace.org), 2017.
- [23] "AWS & Sustainability," [Online]. Available: <https://cloud.google.com/environment/>. [Accessed 27 09 2018].

## Linguistics analytics in data warehouses using fuzzy techniques

P. P. G. D. Asanka

*Department of Computer Science and  
Engineering  
University of Moratuwa  
Moratuwa, Sri Lanka  
dineshasanka@gmail.com*

A. S. Perera

*Department of Computer Science and  
Engineering  
University of Moratuwa  
Moratuwa, Sri Lanka  
shehan@cse.mrt.ac.lk*

### Abstract

*A data warehouse is used intensively in many industry domains to gain competitive advantage over its competitors. In modern data warehouses, linguistic analytics is an important aspect, so that it has the ability to take more precious decisions. In most of the data warehouse implementations, it is designed for crisp analysis. Crisp analysis has its own limitations and boundaries with the major assumptions that every situation belongs to one state and denial to other states. Hence, crisp data warehouse does not allow to carry out linguistic analytics. When a fuzzy data warehouse is implemented, because of the fuzzy nature of the data warehouse, linguistic analytics can be done to a certain extent. In this research, non-functional requirements such as performance and configuration are also covered so that this method can be implemented in the real world.*

**Keywords:** *Data Warehousing, Fuzzy Theory, Fuzzy Membership Function, Linguistics*

### 1. Introduction

The data warehouse has become an important strategic element in the information system of the fiercely competitive market [1-3]. Data warehouse is now being used across various domains and sectors such as Agriculture [4-12], Customer Relation Management (CRM) [14], Banking [15-16] and Healthcare etc. Due to the various issues such as cost and the technical feasibility, the industry has taken some time to adapt data warehouse to the industry. Therefore, the data warehouse is not considered an emerging technology although industry now in the process of adapting to it in full scale.

Over the years, the analytics of data warehousing is limited to crisp value analysis. However, there are limited attempts made to introduce veracity aspects to many sectors by using the fuzzy theory. Fuzzy logic is proposed to mitigate uncertainty in many domains such as agriculture [17], medicine [18-19], power systems [20], production [21], sports [22], transportation [23] etc.

In the field of classical data warehousing, crisp values are used for analytical purposes. For example, when there is a need to analyze some measures (assume sales revenue) with the age of the customer; customer age can be configured as nominal values such as Young, Middle and High. Depending on the domain and the situation, data ranges for Young, Middle and High will be different. When the nominal labels are used, it is obvious that all the analysis is not precise. To introduce veracity which is a challenging factor in data warehouse, fuzzy logic [24] can be used. For example, thirty years of age can be considered as 0.3, 0.7 weightages for MEDIUM and YOUNG respectively whereas in case of crisp set analyze, thirty years of age will be categorized as MEDIUM and only MEDIUM. By doing this, YOUNG or OLD contribution of the age is ignored or considered to be zero weightage to YOUNG and OLD. Due to this configuration, only limited linguistics analytics can be done. For example, with the crisp data warehouse implementation, possible linguistic analytics are NOT OLD Age, OLD OR MEDIUM Age etc.

In the case of classical data warehousing, linguistic variables cannot be used to much extent. Only possibilities are negative and conjunctions. Since variables are crisp sets in classical data warehousing, the analysis is limited. However, in the case of fuzzy techniques, linguistic variables play major roles when doing further analysis.

In this research paper, the current research status of linguistics analytics is discussed in the State of Art – Linguistic Analysis. In the Linguistic Variable section concept behind the linguistic analysis, definitions are discussed. The methodology is discussed in the following section while configuring of fuzzy membership is discussed in the next section. Non - Functional requirement is discussed later and finally, the conclusion and future work is discussed.

### 2. Linguistics variable

Linguistics variables are used to build a relationship with a different state. For example, when defining membership function for temperature, there can be a state called Hot. Also, there can be two adjacent states such as Less Hot and Very Hot respectively. These two states can be a function of the hot state which can be more accurately defined using linguistic fuzzy logic.

Less, High, Low, Much, More and Less are termed as atomic terms.

These atomic terms may be grouped into three main classes.

- Primary terms: labels of fuzzy sets in  $X$  associated with their corresponding meaning, as high and low in the preceding example.
- Hedges: act as a linguistic modifier of the primary terms such as very, much, more or less.
- Connectives and Negation: the connectives and/or and the negation do not act as linguistic modifiers. In the following, it is shown by a series of examples using typical hedges how computations can be done. Concentration (CON) and Dilation (DIL) are the main two operations that were used in this research [25].
- If  $T = \text{very } L$ , then  $T(x) = \text{CON\_}L(x)$  thus very acts as an intensifier. For example, considering the linguistic variable temperature, we may have temperature = very high with  $p = 2$ .
- If  $T = \text{plus } L$  then  $T(x) = \text{CON\_}L(x)$  with  $p = 1.5$ .
- If  $T = \text{minus } L$ . then  $T(x) = \text{DIL\_}L(x)$  with  $r = 0.75$ .

This modifies plus and minus which can be used to define some hedges whose meaning differs marginally from others. For instance, highly could be defined as highly = plus very or possibly as highly = [minus very very].

- If  $T = \text{more or less}$  then  $T(x) = \text{DIL\_}L(x)$  with  $r = 0.5$

FUZZY is another fuzzy operation which is not used in this research.

### 3. State of art - linguistic analytics

As discussed in the introduction section, there are a lot of domains which have discussed the usage of the data warehouse. Since most of the research papers are published in recent years, it can be concluded safely that data warehousing is still a popular and used technology in the industry today.

There are a few which have tried to introduce linguistics analytics for the data warehouse. By using Natural Language Processing and Text Analytics, there was an attempt to performance text analytics [26]. In this research, test analytics was done on the unstructured data. XML tagging is the technique proposed in this research. However, extracting unstructured data and storing them is a challenging task. This research does not derive linguistic values from the existing dimension attributes.

Another research done by the same authors of this research [27], carried out a survey to find out what is the domain where fuzzy data warehouses can be implemented. For that research paper, several sectors were selected such as Retail Sales, Inventory, Order Management, Customer Relationship Management, Human Resource Management, Financial Services and Health care.

Another research paper [28] has proposed an extension to the popular Ralph Kimball four-step methods for dimensional modeling. In this extended and modified process, the research has proposed a step to identify a grain of the crisp and fuzzy data. In the dimension selection process, it was recognized that a fuzzy dimension is nothing but an extension of a crisp dimension. However, philosophically crisp is an extension of fuzzy modeling as most of the real-world cases fall into fuzzy dimensions. In other words, crisp is a summarized version of fuzzy. Therefore, it is scientifically not correct to extend the crisp dimensions to fuzzy dimensions. In this research, the definition of parameters for fuzzy membership is stored and can be modified in case of a need. This will make easy the configurations, so that users can change the parameters accordingly.

In another research paper titled, a data warehouse model for integrating fuzzy concepts in meta table structures [29], defined guidelines for modeling the fuzzy meta-tables and a meta-model for the fuzzy data warehouse are also outlined in this paper. In this extended and modified process, research has proposed a step to identify a grain of the crisp and fuzzy data. In the dimension selection process, it was recognized that a fuzzy dimension is nothing but an extension of a crisp dimension. However, philosophically crisp is an extension of fuzzy modeling as most of the real-world cases are falling into fuzzy dimensions.

Research paper on OLAP (Online Analysis Process) Over Uncertain and imprecise data [30] is more about representing data ambiguity, specifically imprecision and uncertainty. In this research, uncertain measures are stored as a range of possible value, together with the likelihood of each possible value. Those values are represented as a value for an uncertain measure as a probability distribution function over values from an associated base domain. This research has generalized the OLAP model to represent data ambiguity.

Another research [31] has been identified that in most cases fuzzy inferencing is defined in an ad-hoc manner entirely based on an expert's knowledge. When domains consist of common-sense knowledge in Business Intelligence (BI), expert's decisions are also consistent. Therefore, it has been argued that the use of fuzzy inferencing for modeling common sense should be extended in BI.

The research titled Measures of Uncertainty in Decision Making [32] has emphasized the use of fuzzy techniques for decision making. It has provided a solution to analyze the data using fuzzy techniques, whereas this research has placed its importance on designing the warehouse. The player interchange was with constraint rules implemented. In this research, data storage is relational while it supports fuzzy queries.

Another research [33] by the same authors of this research, has introduced a design strategy for fuzzy warehouse, where it has indicated that the entire end to end process of data warehousing stages is considered.

Furthermore, a data-driven methodology is introduced for fuzzy membership function where, in many other researches subjective mechanisms are used.

Next literature view was done on linguistic fuzzy research papers as fuzzy technique is used to implement linguistic analysis in this research paper.

A research paper by V.N. Huynh [34] a model was proposed for the parametric representation of linguistic hedges or linguistic modifiers in Zadeh's fuzzy logic. In the proposed model truth value for every linguistic truth value is identified depending on the primary terms.

In another research paper [35] three-layered data warehouse semantic model consisting of quantitative summarization, qualitative summarization and quantifier summarization is proposed. Also, these research paper SQL queries are also proposed for better retrieval of results. In this research FUZZ-TERM and FUZZ-QUANTIFIER are introduced to enable warehouse users to declare a basic set of linguistic terms. In this FUZZ-TERM is nothing but, the membership function and FUZZ-QUANTIFIER is the linguistics definition of the FUZZ-TERM. However, in this research, the physical data model is not introduced.

A research paper [36] rule-based technique was introduced for different scenarios. This research paper has proposed an optimized rule-based technique with reduced rules in linguistic analysis. However, this research has not discussed the physical implementation of these models with respect to data warehouses.

In the literature review, it was discovered that there are not many discussions have happened with respect to physical modeling on linguistics analysis of data warehouses. Also, configurability aspects and performance of the entire system are not considered although those aspects will play a key role when it comes to implementation and usability.

#### 4. Objective and methodology

The research objective is to derive a data-driven method to perform linguistic analytics on the data warehouse.

Detail research was done by the same authors of this research paper, on the implementation of fuzzy data warehouse [33] and in the implementation of fuzzy membership functions for data warehouse [37] are the core areas for this research. Linguistic fuzzy membership function defines the fuzzy data warehouse. Fuzzy data warehouse design is a requirement for the proposed linguistic analysis. Since linguistics analytics will be done on dimensions, these dimensions should be defined with the use of fuzzy membership function. The data-driven membership function is the function which was designed by analyzing the data set. There are several methods of determines the fuzzy membership function using a data-driven technique such as Neural Network [38], Box Plot [39] etc. Also, the same authors have published a research paper on determining fuzzy membership function for fuzzy data warehouse which has discussed how to derive membership function using

data-driven, subjective, hybrid and derived methods. Using one or two fuzzy membership function(s) linguistics membership function will be defined. Then values will be generated to the relevant attributes' values of dimensions in order to improve performance.

To validate this technique, plantation attendance data for more than 10 years, is used. This data set is a real-world data, not sample data so that the practical implementation of this technique is evaluated.

#### 5. Linguistic membership function definition

Linguistics Membership functions are defined by considering the basis of existing membership function which can be defined by using data-driven, subjective, hybrid or derived mechanisms.

As described in the Linguistics Variable section, there are fuzzy operations such as Concentration (CON), Dilation (DIL), are used. For more accuracy, the fuzzy state is used rather than using the entire fuzzy membership function as shown in the below equations.

$$\begin{aligned} \text{Linguistic State} &= \text{CON}(\text{Base State}, n) \quad n > 1 \\ \text{Linguistic State} &= \text{DIL}(\text{Base State}, n) \quad 0 < n < 1 \end{aligned} \quad \text{EQ. (1)}$$

For example, High Medium of Age can be defined by applying the Concentration operation for the Medium State of Age with  $n = 1.15$ . Similarly, Low Medium of Age can be defined by using the Dilation operation for Medium states with  $n = 0.65$ . This can be represented by the following equation.

$$\begin{aligned} \text{High Medium} &= \text{CON}(\text{Medium}, 1.15) \\ \text{Low Medium} &= \text{DIL}(\text{Medium}, 0.65) \end{aligned} \quad \text{EQ. (2)}$$

Apart from concentration and dilation, negation is also used as a linguistic analysis.

$$\text{Linguistic State (Negate)} = 1 - \text{Base State} \quad \text{EQ. (3)}$$

Also, in this research combination of states are used to define another linguistic state. Combination state can be base state or linguistic state. In this research, as a measure to reduce the complexity and to improve the performance, the limitation was introduced where only two base states can be used to define the linguistic state.

Following are defined linguistics functions.

$$\begin{aligned} \text{Low OR High} &= (\text{Low}) \text{ OR } (\text{High}) \\ \text{Low AND Very Low} &= (\text{Low}) \text{ AND } (\text{Vey Low}) \end{aligned} \quad \text{EQ. (4)}$$

In the first example, Low and High are two base functions whereas in the second example, Low is a base function but Very Low is previously defined linguistics function.

## 6. Physical model and implementation

Since this is used in data warehouses, it is essential to have configurability, so that parameters can be modified depending on the requirement. Table schema to store Linguistic Configuration values is shown in Figure 1.

FuzzyMembershipFunctionLingusiticConfiguration	
ID	
Function	
State	
BaseFunction	
BaseState1	
BaseState2	
LinguisticType	
Parameter1	
isActive	
ValidStartDate	
ValidEndDate	
IsCurrent	

Figure 1. Table definition for linguistic fuzzy membership function.

In the above configuration table, the Function is used to relate to the primary function of the fuzzy function. For example, Age, Experience will be the primary function. The state will define the linguistic state such as Low Medium, Not Low OR Not High etc. Base Function is to provide the flexibility where multiple functions can be linked. Two bases states (Base State1 and Base State2) are included, so that multiple states can be linked. For Example, Not Low OR Not High state both Low and High states are linked. Linguistics Type and its parameters are defined in Linguistic Type and parameter columns. Linguistic Types are standard operations in Fuzzy sets which were discussed before.

In the above design, only two base membership functions are possible as two base membership functions are the most common scenario. Sample configuration of Linguistics Fuzzy Membership Function is shown in Table 1.

Table 1. Sample data set for linguistic fuzzy membership function.

State	Base State1	Base State2	Linguistic Type
Low Medium	Medium		DIL
High Medium	Medium		CON
Low AND Very Low	Low	Very Low	NOT
Low OR High	Low	High	
Not High	High		NOT
Not Medium	Medium		NOT

Table1 displays the sample configuration for Function1 in which base function is a different membership function.

By using this configuration, the contribution for each value is calculated and stored in the table shown in Figure 2.

FuzzyContribution	
Function	
Value	
Description	
Contribution	

Figure 2. Table definition for fuzzy contribution

This is done in order to improve the performance of the system. If the weightage of contribution is calculated during the run-time, analytics will be slower. Considering the fact that, data warehouse queries are executed on a large volume of data, this approach will suit from the performance viewpoint.

Figure 3 shows the relationship of linguistic fuzzy membership functions to the other fuzzy membership functions.

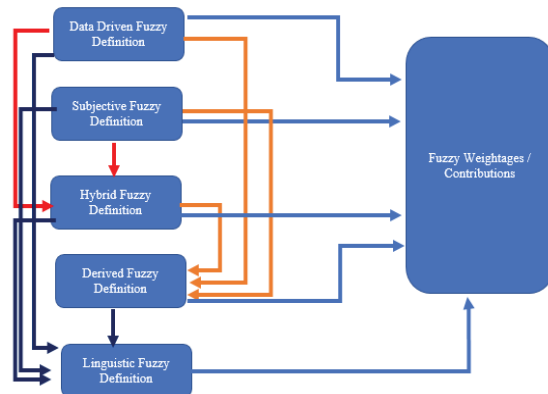


Figure 3. Table definition for fuzzy contribution.

As shown in Figure 3, linguistics fuzzy membership functions can be derived from using Data-driven, Subjective, Hybrid, Derived membership functions.

Table 2 shows a sample set of data for linguistic calculation weightage. This shows weightages for the value 2 and 3 for a given function Experience.

A subset of the above data is pivoted in the following table for better presentation. Summation should be 1 for the base states. However, that rule does not apply for the linguistic states as shown in Table 3.

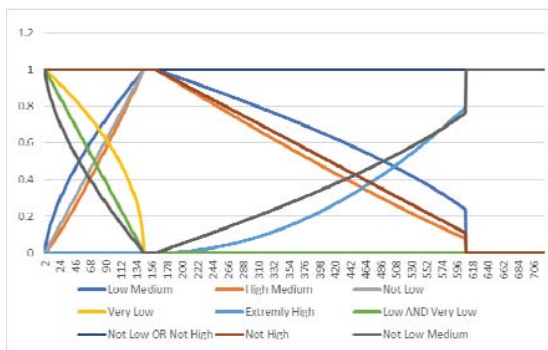
**Table 2. Sample data set for linguistic fuzzy experience membership function.**

Value	State	Contribution
3	High Medium	0.0028
3	Low AND Very Low	0.994
3	Low Medium	0.036
3	Not High	1
3	Not Low	0.006
3	Not Low Medium	0.964
3	Not Low OR Not High	1
3	Very Low	0.997
4	High Medium	0.0068
4	Low AND Very Low	0.987
4	Low Medium	0.0594

**Table 3. Sample data set for linguistic fuzzy membership function.**

	High Medium	Low	Low AND Very Low	Low Medium	Not High
2		1	1		1
3	0.0028	0.994	0.994	0.036	1
4	0.0068	0.987	0.987	0.0594	1

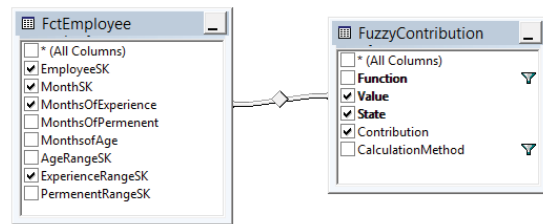
Unlike data-driven fuzzy states, the profile of linguistics states does not follow a smooth curve as shown in Figure 4.



**Figure 4. A sample of linguistic states profile.**

## 7. Evaluation

These linguistics states are joined with fact tables for the analysis. As shown in Figure 5, FCTEmployee fact table is joined with the fuzzy contribution table with a filter on the Function, in the example shown it is the experience of the employees.



**Figure 5. Sample for schema for data warehouse.**

From the above table schema, different types of linguistics analytics can be done on the fact tables which is shown in Figure 6.

State /Year	2000	2001	2002	2003	2004	2005	2006	2007	2008
Extremely High	651	719	775	832	849	847	839	828	812
High	651	719	775	832	849	847	839	828	812
High Medium	1067	1272	1364	1495	1526	1529	1689	1755	1773
Low	308	450	507	623	650	682	849	938	960
Low AND Very Low	308	450	507	623	650	682	849	938	960
Low Medium	1067	1272	1364	1495	1526	1529	1689	1755	1773
Medium	1067	1272	1364	1495	1526	1529	1689	1755	1773
Not High	1068	1276	1388	1505	1532	1558	1704	1785	1786
Not Low	1068	1276	1388	1505	1532	1558	1704	1785	1786
Not Low Medium	1068	1276	1388	1505	1532	1558	1704	1785	1786
Not Low OR Not High	1068	1276	1388	1505	1532	1558	1704	1785	1786
Very Low	308	450	507	623	650	682	849	938	960
Grand Total	9699	11708	12715	14038	14354	14559	16108	16875	16967

**Figure 6. Analysis with linguistic analytics on data warehouse.**

As shown in Figure 6, linguistics analytics is done for the estate sector employees' age between 2000 and 2007. By analyzing this linguistic data set, there is a tendency that VERY LOW age group employees are increasing which will be a negative social impact.

## 8. Non-functional requirements

In this implementation, there are non-functional requirements were achieved in order to have a scalable implementation of fuzzy data warehouse if not this technique will not be able to implement in the real world.

Typical type 2 slowly changing dimensions [40] is used in the data warehouse design so that historical aspect of fuzzy linguistic analytics. In order to cater to historical aspect, in the configuration tables, ValidStartDate, ValidEndDate, IsCurrent columns are used to configure the historical aspects of configurations as there can be situations where configurations are changing over the time.

Performance is a key factor in any system. In the case of data warehousing, there are special effort needs to be done since it is dealing with a large volume of data which is in many tera bytes in nature. More consideration should be done for a fuzzy data warehouse with respect to performance. Typically, indexes are the key element when it comes to database and data warehouse performance. Index strategy is different in data warehouse [41] as the data warehouse has very less writes. Since Index improves retrieve queries and has a

negative impact to write queries, many numbers of indexes can be included in data warehouses. Also, column store indexes are heavily used in data warehouses as data warehouse scan through the entire fact table most of the time [42]. In the case of fuzzy data warehousing, row-based indexes and column store indexes can be utilized. Addition to those indexes, there are other key approaches which can be taken to improve the fuzzy data warehouse so that user experience can be improved.

## 9. Conclusion and future work

In this research, a different approach was proposed for the linguistic analysis of a data warehouse using the fuzzy technique. This research has utilized existing fuzzy membership functions to derive linguistic functions. Since it is used in the same table structure, users have the option of the same table for the analysis.

This research has considered the non-functional requirements such as performance, configurability and the slowly changing dimension type 2. By considering the non-functional requirement, it was confirmed that this can be used in real-world scenarios.

This research indicates that data can be analyzed in linguistics on top of the fuzzy data warehouse. In order to facilitate real-world implementation, non-functional requirements also achieved.

Currently, linguistics parameters are pre-defined but the data-driven technique can be introduced to identify those parameters.

## 10. References

- [1] Chaudhuri, S., Dayal, U., "An Overview of Data Warehousing and OLAP Technology", *SIGMOD Record*, vol 26, no 1, 1997, pp. 65-74.
- [2] Mumick, L. S., Quass, D., Mumick, B. S., "Maintenance of Data Cubes and Summary Tables in a Warehouse", in *Proceedings of the ACM SIGMOD International conference Management of data*, June 1997, pp. 100-111.
- [3] Wisdom, J., "Research Problems in Data Warehousing, Proceedings at International Conference", *Information and Knowledge Management*, 1995.
- [4] Chai, K. L., Costello, T.A., Wells, B. R., Norman, R.J., "Expert System for Fertilizer Management of Rice", *App- Eng-agric.*, November 1994, pp. 849-855.
- [5] Coble, H.D., *Proceedings of the annual meeting of the Northeastern Weed Science Society*, 1994, 48, pp. 155-159.
- [6] Jørgensen, L. N., Noe, E., Langvad, A. M., Jensen, J. E., Ørum, J.E., Rydahl, P., *EPPO Bulletin*, 2007, 37(2), 374-377.
- [7] Detlefsen, N. K. and Jensen A. L., *Agricultural Systems*, 2004, pp. 55-72.
- [8] Bazile, D., Le Page, L., Dembélé, S., Abrami, G., "Perspectives of modelling the farmer's seed system for in conservation of sorghum varieties in Mali", 5th Conference of the European Federation for Information Technology in Agriculture, Food and Environment and 3rd World Congress on Computers in Agriculture and Natural Resources., 2005, Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal.
- [9] Mosseddaq, F., Dnidane, S., Lahlou, M., "Integrated Wheat N Nutrition Management in Morocco: A Decision Support Model", *Progress of Information Technology in Agriculture Proceedings of the 4th International Symposium on Intelligent Information Technology in Agriculture (ISIITA)*. 2007.
- [10] Riberir, R.P., Borges, J.G., "MetaForest- a web-based decision system to support forest management involving multiple-ownership", *EFITA/WCCA*, 2005, Vila Real, Portugal.
- [11] Ibragimov, T. Z., Ibragimova, I. T., Sanin, S. S., "Decision support systems for cereal crop disease control", *EFITA/WCCA*, 2005, Vila Real, Portugal.
- [12] Sodtke, R., "Multifunctional of Landscapes - Analysis, Evaluation and Decision Support", *International Conference, Univ. Giessen*, 2005, 193.
- [13] Suarez de Cepeda, M., Recio, B., Rubio, F., "Decision Support System for Farm Mechanization", *5th Conference of the European Federation for Information Technology in Agriculture, Food and Environment and 3rd World Congress on Computers in Agriculture and Natural Resources*, 2005, Vila Real, Portugal.
- [14] Cooper, B.L., Watson, H.J., Wixom, B.H., Goodhue, D.L., "Data Warehousing supports Corporate strategy at First American Corporation", *MIS Quarterly*, Minneapolis, vol.24, Issue 4, 2000, pp. 547-567.
- [15] Chavda, M. B., Solanki, A., "Innovative banking products: Win-Win situation for Customers and Banks", *IRACST – International Journal of Commerce, Business and Management (IJCBM)*, ISSN: 2319-2828 Vol. 3, No.6, December 2014.
- [16] Gadda, K. R., Dey, S., "Business Intelligence for Public Sector Banks in India: A Case study- Design, Development and Deployment", *Journal of Finance, Accounting and Management*, 5(2), 2014, pp. 37-58.
- [17] Ashraf, A., Akram, M., Sarwar, M., "Type-II Fuzzy Decision Support System for Fertilizer", *Punjab University College of Information Technology*, University of the Punjab, April 8, 2014.
- [18] Sapna, S., Pravin, K., "Integration of Fuzzy Clustering Technique with Big Data for Disease Diagnosis", *International Journal of Electrical, Computing Engineering and Communication (IJECC)* Vol. 1, Issue. 3, June 2015 ISSN (Online): 2394-8310.
- [19] Dong, W., Huang, Z., Duan, H., "A genetic fuzzy system for unstable angina risk assessment", *Cardiology Department of Chinese PLA General Hospital*, Beijing, China, 2014.
- [20] Wang, Q., Wang, Y., Zhang, H., Sun, T., "Load Forecasting Research of Power System Based on Fuzzy Sets Algorithm", *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol.9, No.6, 2016, pp.283-292, ISSN: 2005-4254.
- [21] Bhargava, A. K., Singh, S. R., Bansal, D., "A Fuzzy Goal Programming Model for Bakery Production", *International Journal of Computer & Mathematical Sciences, IJCMS*, ISSN 2347 – 8527 Volume 3, Issue 4, June 2014.



- [22] Zeng, W., Li, J., "Fuzzy Logic and Its Application in Football Team Ranking", *The Scientific World Journal, Hindawi Publishing Cooperation*, Volume 2014, Article ID 291650, 6 pages.
- [23] Wang, Y. C., Chien. C. J., Wang, C. H., "A Fuzzy-Neural Adaptive Iterative Learning Control for Freeway Traffic Flow Systems", in *Proceedings of the International Multi Conference of Engineers and Computer Scientists*, 2016 Vol I, IMECS 2016, March 16 - 18, 2016, Hong Kong.
- [24] Zadeh, L. A., "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes", *IEEE Transaction Systems, Man and Cybernetics*, Vol 3, No 1, 1973.
- [25] Kumar, S., Biswas, R., Roy, A. R., "Some Operations on Intuitionistic Fuzzy Sets", 2000, *Elsevier*.
- [26] Prasad, K .S. N., Ramakrishna, S., "Text Analytics to Data Warehousing", *International Journal on Computer Science and Engineering*, Vol. 02, No. 06, 2010, 2201-2207.
- [27] Dinesh Asanka, P. P. G., Perera, A. S., "Feasibility of Fuzzy Data Warehouse", *International Journal of Research in Computer Applications and Robotics*, ISSN 2320-7345, Vol. 5 Issue 11, November 2017.
- [28] Sapir, L., Shmilovic, A., "A Method for Design of a Fuzzy Data Warehouse", Department of Information Systems Engineering, Ben-Gurion University, Israel, 2008.
- [29] Medina, J. M., Vila, M. A., Cubero, J. C., Pons, O., "Towards the Implementation of Fuzzy Relational Database Model", *Fuzzy Sets and Systems*, ELSEVIER, 273-289., 1994.
- [30] Burdick, D., Deshpande, P. M., Jayaram, T. S., "OLAP Over Uncertain and Imprecise Data", 31st VLDB Conference, Trondheim, Norway, 2005.
- [31] Faisal D., Shahzad K., A Data Warehouse Model for Integrating Fuzzy Concepts in Meta Table Structures, Information Systems Research Group, Department of Informatics, University of Fribourg, Switzerland.
- [32] O. Sedlak, M. Cileg, KisT, Z. Ciric., Measures of Uncertainty in Decision Making, International Seminar on Operational Research on Operational Research, 2013.
- [33] Dinesh Asanka, P. P. G., Perera, A. S., "Design Strategy for Fuzzy Data Warehouses", *2nd International Conference on Innovative Research in Science, Technology & Management*, National University Singapore, 29-30 September 2018.
- [34] Huynh, V. N., Ho, T. B., Nakamori, Y., "A Parametric Representation of linguistic Hedges in Zadeh's Fuzzy Logic", *International Journal of Approximate Reasoning*, 2002, Elsevier.
- [35] Feng, L., Dillion D. T., "Using Fuzzy Linguistic Representations to Provide Explanatory Semantics for Data Warehouses", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 15, No 1, January/February 2003.
- [36] Gegov, A., Babushka, R., Verbruggen, H. B., "Linguistic Analysis of Interactions in MIMO Fuzzy Systems", *14th Triennial World Congress*, Beijing, China, 1999.
- [37] Dinesh Asanka, P. P. G., Perera, A. S., "Defining Fuzzy Membership Functions for Fuzzy Data Warehouses", *4th I2CT IEEE Conference*, SDMIT Ujire, Mangalore, India, October 2018.
- [38] Wang, S., "Generating Fuzzy Membership Functions: A Monotonic Neural Network Model", *Fuzzy Sets and Systems*, Vol. 61, Issue 1, 1994 January.
- [39] Dinesh Asanka, P. P. G., Perera, A. S., "Defining Fuzzy Membership Function Using Box Plot", ISSN 2320-7345, Vol. 5 Issue 11, November 2017.
- [40] Dinesh Asanka, P. P. G., "Loading Historical Data into a SQL Server Data Warehouse", <https://www.mssqltips.com/sqlservertip/5505/loading-historical-data-into-a-sql-server-data-warehouse/> (Online), Accessed: 2018-11-22.
- [41] Aouiche K., Darmont J., "Index and Materialized View Selection in Data Warehouses", 2006.
- [42] Amro, A. A., Abutabh M., "Case Study: Comparison between Traditional Indexing and Column store Indexes in SQL Server 2012", *Conference: International Conference on Information & Intelligent Systems*, Volume: N&N Global Technology for Computer and Information Service & IEEE Computational Intelligence Society, March 2013.

## An application of transfer learning techniques in identifying herbal plants in Sri Lanka

Y. R. Azeez  
Department of Industrial Management  
Faculty of Science  
University of Kelaniya  
Dalugama  
rameela20@gmail.com

C. Rajapakse  
Department of Industrial Management  
Faculty of Science  
University of Kelaniya  
Dalugama  
chathura@kln.ac.lk

### Abstract

Sri Lanka has a considerable collection of plant species that have been utilized for generations as medicinal treatments. Knowledge regarding herbal plants is restricted mainly among practitioners in traditional medicine. Available systems studied; had no proper methodology to search information regarding herbal plants, which can be identified through analyzing an image of an herbal plant given. Systematic literature review was done based on herbal plants in Sri Lanka, transfer learning and plant image recognition and two open ended interviews were conducted with traditional medicine practitioners. As main objective of the study, reorganization of Information was done building a technique to enhance capability of identifying herbal plants based on deep convolutional neural networks and image processing techniques which would ultimately assist more locals with identification. Five herbal plant types were chosen to analyze further in detail and the images of the plants were acquired from web and also images photographed via 13MP camera creating a data set validated through traditional medical practitioners. Images were preprocessed and retrained on Inception-v3, Resnet, MobileNet and Inception Resenet V2 based on transfer learning. Algorithm was fine-tuned using image processing techniques for preprocessing and prototype was tested 5 times reaching highest average accuracy of 95.5% on Resnet for the identification of 5 different plant types. Conclusively, this study enhanced the capability of searching herbal plants by reorganizing the information.

**Keywords:** *deep convolutional neural networks, transfer learning, Inception-v3*

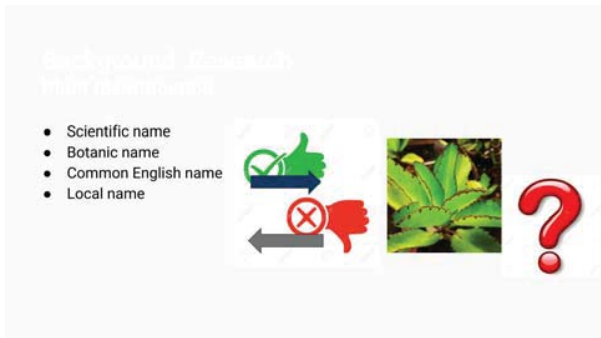
### 1. Introduction

The study has considered the identification of herbal plants in Sri Lanka, where transfer learning was used. The term of medicinal plants include various types of plants used in herbalism and some of these plants have medicinal activities. Medicinal plants are the “backbone” of traditional medicine, which means more than 3.3 billion people in the less developed countries utilize medicinal plants on a regular basis [1].

The knowledge regarding identifying herbal plants is mostly restricted to traditional medicinal practitioners. People have become more urbanized and spend busy livelihoods. So to conserve this precious knowledge a need was identified to introduce a proper methodology to identify herbal plants that have great value in a developing nation like Sri Lanka.

There were websites such as [instituteofayurveda.org](http://instituteofayurveda.org), [doshaguru.com](http://doshaguru.com) and etc. where anyone can find out details about it by searching the name of the plant. But, the need is to find out what the plant is. Since the information is unknown to the people who are not familiar with the medicinal plants, first identification should be addressed. So the problem is the way information is organized to be delivered is not in the needed order to help the natural process of identifying an unknown plant when someone see and want to find out what it is.

Currently, there is not a proper application that assist the people to identify plants during their day-to-day lives. This study focuses to contribute by introducing a technique that will aid plant image recognition. The model transfer learning was used for the development of the model.



**Figure 1 : Problem of the information organization.**

Transfer learning is a new approach used in machine learning that can be used to solve different, but related problems via transferring knowledge from source domain to the target domain.

In this research transfer learning techniques are proposed to use in building the deep learning model, assisting the Plant Image Identification of 5 different plant types.

## 2. Literature review

### 2.1 Herbal plants in Sri Lanka

In 2004, the Sri Lanka Conservation and Sustainable use of Medicinal Plants Project was completed and the report stated the following: It was evident from discussions during field visits to Rajawaka that generally the younger generation had been distanced from the traditional systems of medicine and healthcare. It did come as a surprise that they had not been familiar with common medicinal plants. On reflection, this could be ascribed to the steady erosion of awareness, knowledge and dependence on the traditional systems. So, especially younger generation has a knowledge gap regarding herbal plants.

As published on Pharmaceutical Journal of Sri Lanka 2015, a descriptive cross-sectional study was conducted during June and July 2013 at Geli-Oya Muruthagahamula division (rural) and Bahirawakanda (urban) Gramaniladhari areas in Kandy District, Central Province, Sri Lanka. Volunteers (n=400) between 18 -65 years of age were selected by a random sampling method. An interviewer administered questionnaire was used to collect data from voluntary participants via a face to face interview. Among the 400 participants (200 from each division) 82% had the opinion that self-medication is enough to cure common ailments.

However, the usage of Traditional Medicine (TM) is comparatively high in rural area than the urban. This may be because of the herbs which were used for TMs are

found easily in the nearby forest around the rural area and readily available in the courtyard of houses. This enables easy access of TMs.

So, there seemed a clear knowledge gap regarding herbal plants when younger generation and urban population considered.

### 2.2. Plant image identification

When in Plant Image recognition, researchers have used mainly leaf recognition and they identified, tried to address the challenges. And the information about related work as follows.

Several methods to identify plants have been proposed by several researchers. Commonly, the methods did not capture color information. Because, color was not recognized as an important aspect to the identification. In this research, shape and vein, color, and texture features were incorporated to classify a leaf. In this case, a neural network called Probabilistic Neural network (PNN) was used as a classifier. The experimental result shows that the method for classification gives average accuracy of 93.75% when it was tested on Flavia dataset that contains 32 kinds of plant leaves. It means that the method gives better performance compared to the original work [7].

In a research done in 2012, first mobile app for identifying plant species using automatic visual recognition was introduced. The system – called Leafsnap – identifies tree species from photographs of their leaves. Key to this system were computer vision components for discarding non-leaf images, segmenting the leaf from an untextured background, extracting features representing the curvature of the leaf's contour over multiple scales, and identifying the species from a dataset of the 184 trees in the Northeastern United States. Their system obtained state-of-the-art performance on the real-world images from the new Leafsnap Dataset – the largest of its kind. Throughout the paper, they document many of the practical steps needed to produce a computer vision system, which at the time had nearly a million users.

Although, the dataset and algorithm were available to analyze. And also, the dataset was specific to Northeastern United States plants [8].

Therefore, taking images of the whole herbal plant to be analyzed and has not been explored but the leaf image is mostly explored. Technique using deep learning and computer vision techniques needed to be explored for the images of whole plant.

### 2.3. Transfer learning

The researchers tend to explore the most effective methods that can be used for deep learning approaches. Mostly in Convolutional Neural Networks (CNN) a large

set of training data is needed and where resources such as GPUs and longer training time is required. Transfer learning is an approach that helps to solve such scenarios [9].

Image retraining on tensorflow uses transfer learning concepts introduced in the paper published in 2013 regarding DeCAF. They evaluated whether features extracted from the activation of a deep convolutional network trained in a fully supervised fashion on a large, fixed set of object recognition tasks could be repurposed to novel generic tasks. Their generic tasks may differ significantly from the originally trained tasks and there may be insufficient labeled or unlabeled data to conventionally train or adapt a deep architecture to the new tasks. They investigated and visualized the semantic clustering of deep convolutional features with respect to a variety of such tasks, including scene recognition, domain adaptation, and fine-grained recognition challenges. They compared the efficacy of relying on various network levels to define a fixed feature, and report novel results that significantly outperform the state-of-the-art on several important vision challenges. They released DeCAF, an open-source implementation of these deep convolutional activation features, along with all associated network parameters to enable vision researchers to be able to conduct experimentation with deep representations across a range of visual concept learning paradigms. And tensor flow had based this for image retraining [6].

In India, a Survey on Ayurvedic Medicine Classification using TensorFlow has conducted during 2017. So, the researchers have trained GoogleNet network with the dataset they created. After training, the Network was ready to recognize the medicinal plants for which it was trained. So they would be giving image to the network as an input and the name of the most similar medicinal plant will come in the output was the thought which they had. First, GoogleNet is an inception module given by Google which is a 22 layer CNN (Convolutional Neural Network) [10].

### 3. Proposed architecture

Transfer learning is a machine learning method which utilizes a pre-trained neural network. For example, the image recognition model called Inception-v3 consists of two parts:

- Feature extraction part with a convolutional neural network.
- Classification part with fully-connected and softmax layers.

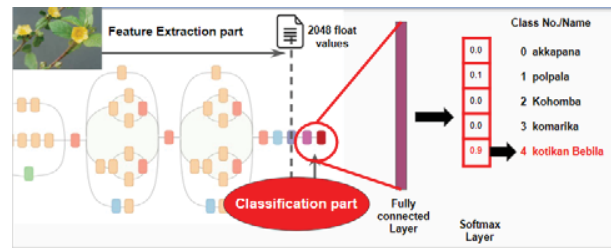


Figure 2 : Retraining on Inception-v3.

The pre-trained Inception-v3 model achieves state-of-the-art accuracy for recognizing general objects with 1000 classes.

By using the method of transfer learning, the parameters of the previous layer were unchanged, but only the last layer was trained. The last layer was a softmax classifier. This classifier was 1000 output nodes in the original network (ImageNet has 1000 classes), so, there was a requirement to remove the last network layer, and then retrained the last layer. The reason for retraining the last layer was to work on the new object classes. As this study has chosen 5 herbal plant types, there were 5 classes. And number of nodes were 5 [3, 4].

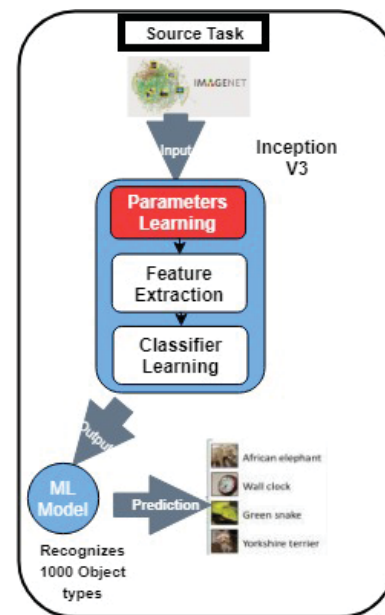


Figure 3 : Source task on Inception-v3.

By looking at the source task, consider the scenario of retraining images on Inception-v3 model.

Inception-v3 is a convolutional neural network that is trained on more than a million images from the ImageNet

database. So, the input images were the images taken from ImageNet database.

Since the inception v3 required 299X299 image input size images were resized to fit the requirement.

### 3.1 Parameters learning

Ability to automatically learn hierarchical feature representations was one important aspect of this deep learning model. Features computed by the earlier layers were general and could be reused in various problem domains, while features computed by the last layer were specific and depend on the chosen dataset and task.

For an example, inception v3 first layers had learnt to detect colour changes, shape detection, edge detection kind of tasks and these parameters were more general for a lot of images.

And, in the last layers it had learnt how to detect a human, a chair... etc. which were very specific.

So, in this phase of parameter learning weights and biases for more than 5 million of parameters were finalized to identify the needed feature representations from the input images fed to inception v3; which required a more labeled training data and a lot of computing power (hundreds of GPU-hours or more) for the ImageNet Large Visual Recognition Challenge using the data from 2012 (ILSVRC-2012-CLS). And also there were deep learning models such as Resnet, MobileNet and Inception Resnet v2 available on tensorflow which were also trained on ILSVRC-2012-CLS dataset.

These learned parameters were frozen in the inception v3 pretrained model during the source task of training for general object recognition model which could be applied for the target task of training Medicinal Plant identification model [3, 4].

So, only feature extraction to bottlenecks and Classifier Learning happened during the Inception v3 model retraining. So, only classifier part was trained from the Inception v3 model during target task.

### 3.2 Image preprocessing and preparation for input

For training to work well, gathering at least a hundred photos of each kind of plant type we wanted to recognize was recommended. And also needed to make sure that the photos were a good representation of what the application would actually encounter. For example, taking all the photos indoors against a blank wall for the training and when users try to recognize objects outdoors, they do not see good results using the trained model.

Collected image data from web and photographed images using smartphone (OPPO A3s) 13 MP camera were validated from an Ayurveda doctor and image files were copied into the respective folder named according to the relevant plant type.

Images were organized into root folder plant\_images. And 5 directories were created to represent the 5 plant types, i.e. akkapana, Kohomba, Komarika, kotikan\_bebila, and polpala. And image type conversion eg: PNG to JPG and background removal was done.

It is advisable to always capture images with distinct colour background and without other plants except the plant considered.

### 3.3 Feature extraction

These images were given as input into the feature extraction part of Inception-v3 which converted the image data into feature vectors consisting of 2048 float values for each image. And these features were extracted depending on the feature representations learnt via parameter learning phase in source task. A feature vector represented the features of the image in an abstract manner. It can better classify images based on these vector values rather than raw image data. These feature vectors are called bottlenecks.

### 3.4 Classifier learning

By default, the script to retrain model ran 4,000 training steps. Each step chose ten images at random from the training set, found their bottlenecks from the cache, and fed them into the final layer to get predictions. Those predictions were then compared against the actual labels to update the final layer weights through the back-propagation process. As the process continued the users could see the reported accuracy improve, and after all the steps were done, a final test accuracy evaluation was run on a set of images kept separate from the training and

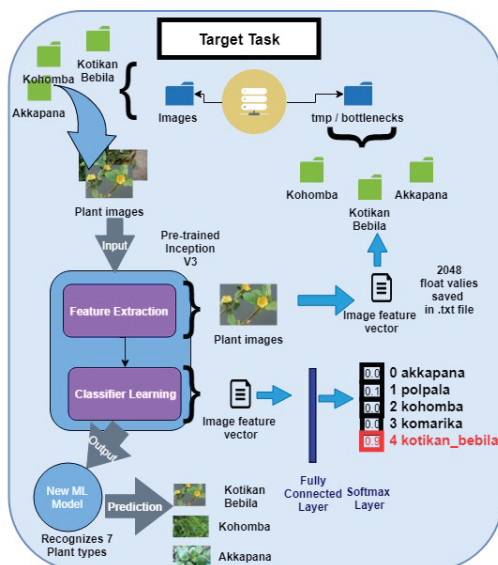


Figure 4 : Target task on Inception-v3.

validation pictures. This test evaluation was the best estimate of how the trained model would perform on the classification task.

The usual split was to put 80% of the images into the main training set, keep 10% aside to run as validation frequently during training, and then had a final 10% that were used less often as a testing set to predict the real-world performance of the classifier. So, that provided validation percentage as 10% and testing percentage as 10%.

### 3.5 Prediction

The new model contained both the TF-Hub module inlined into it, and the new classification layer. When a test image was given to the model to be predicted top 5 predictions for the medicinal plant types were given.

## 4. Results

Different models were retrained to evaluate how the testing accuracy could be improved. Testing percentage was given as 10%, since the percentage split was depicted as 10%; out of the total image set 10% of images were randomly selected as test set.

$$Accuracy (\%) = \frac{\text{Total number of images correctly classified}}{\text{Total number of images used for testing}} * 100$$

Following depicts the deep learning models used for retraining herbal plant images.

**Table 1 : Deep learning models used for retraining on TensorFlow.**

Deep learning model	Tf-Hub Module used
Inception-v3	<a href="https://tfhub.dev/google/imagenet/inception_v3/feature_vector/1">https://tfhub.dev/google/imagenet/inception_v3/feature_vector/1</a>
Resnet	<a href="https://tfhub.dev/google/imagenet/resnet_v2_50/feature_vector/1">https://tfhub.dev/google/imagenet/resnet_v2_50/feature_vector/1</a>
MobileNet	<a href="https://tfhub.dev/google/imagenet/mobilenet_v2_100_224/feature_vector/2">https://tfhub.dev/google/imagenet/mobilenet_v2_100_224/feature_vector/2</a>
Inception Resnet v2	<a href="https://tfhub.dev/google/imagenet/inception_resnet_v2/feature_vector/1">https://tfhub.dev/google/imagenet/inception_resnet_v2/feature_vector/1</a>

Three datasets were used for retraining. Number of images used according to plant type (Scientific name and Sinhalese names used) are as following [5].

**Table 2: Number of plant images used in each dataset.**

Plant type	Dataset 1	Dataset 2	Dataset 3
<i>Kalanchoe pinnata</i> (Akkapana)	68	68	68
<i>Aerva lanata</i> (Polpala)	63	63	63
<i>Azadirachta indica</i> (Kohomba)	81	81	81
<i>Aloe vera</i> (Komarika)	77	77	77
<i>Sida acuta</i> (Kesara bebila)	0	0	97
<i>Sida rhombifolia</i> (Kotikan bebila)	80	80	80
<i>Sida cordifolia</i> (Sulu bu bebila)	0	86	86
<b>Total</b>	369	455	552

Dataset 1 – Only contain images of 5 distinct plant types (5 classes)

Dataset 2 – Contains images of 4 distinct plant types and 2 similar plant types (6 classes)

Dataset 3 – Contains images of 4 distinct plant types and 3 similar plant types (7 classes)

These 3 datasets were trained 5 times each on the 4 existing models on tensorflow. And average accuracies of the 5 runs were calculated for each model.

**Table 3 : Average test accuracies of the models retrained.**

Deep learning model	Dataset 1	Dataset 2	Dataset 3
Inception-v3	88.6	86.18	71.4
Resnet	95.5	78.6	71.4
MobileNet	86.4	81	64.3
Inception Resnet v2	81.8	83.3	73.2

For the models retrained for 5 distinct plant types highest accuracy was given by Resnet which is 95.5%.

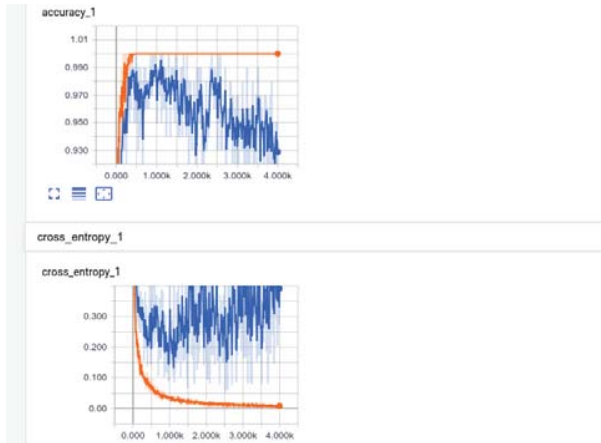
For the models retrained for 4 distinct plant types and 2 similar plant types highest accuracy was given by Inception-v3 which is 86.18%.

Finally, for the models retrained for 4 distinct plant types and 3 similar plant types highest accuracy was given by Inception Resnet v2 which is 73.2%.

When the number of similar plant types considered were increased; highest accuracy of the best model had been decreased. This is due to over fitting. To increase accuracy number of images should be increased and

should acquire images considering variations as much as possible.

Tensorboard was used to visualize how training and validation accuracy varies over the training period. Training accuracy is shown by orange colour while validation accuracy is shown by blue.



**Figure 5 : Visualizing training and validation accuracy of best model trained on Resnet for dataset 1.**

## 5. Conclusion and future works

In this study the technique is proposed based on transfer learning. Existing deep learning models are re-trained to specialize the models for identifying 5-7 medicinal plant types in Sri Lanka. Parameter Learning, Feature Extraction and Classifier Learning happens in order and parameters learnt in source task is transferred to target task of medicinal plant identification. Different experiments and testing scenarios were carried out to fine tune a model with better performance. Mainly following can be summarized as the important points that should be considered regarding the proposed technique.

1. Image data should be collected for as many variations as possible.
2. Training a model from scratch requires a lot of infrastructure when compared with transfer learning.
3. Deep learning models retrained consist of mainly 2 parts i.e. Feature extraction part and classification part. Here only classification part is trained.
4. During feature extraction bottlenecks are cached. And they will be reused during another iteration of training.
5. Model becomes over fitting when there is not enough data available to understand the features better.

6. For similar plant types identification model should be improved.
7. Distortions can be applied to improve the accuracy of the model.

This study shows that a general object recognition model trained for millions of parameters can be retrained and repurposed to recognize specific type of target task such as medicinal plant identification.

There are few suggestions for the further development.

1. Model's accuracy can be improved using more training data avoiding over fitting.
2. Distortions should be applied and model should be fine-tuned. eg: random crop, random brightness, flip left, flip right
3. Prototype need to be tested in a mobile based environment
4. Cloud based technique should be improved.
5. Multiple models can be trained parallelly and an integrated model can be considered for better accuracy [2].

## 6. References

- [1] Davidson-Hunt I, Ecological ethno botany: stumbling toward new practices and paradigms, *MASA Journal*, vol.16, 2000, pp. 1–13.
- [2] Herve Goeau, Pierre Bonnet, Alexis Joly, Plant identification based on noisy web data: the amazing performance of deep learning, *CLEF 2017 - in Proceedings of the Conference and Labs of the Evaluation Forum*, Sep 2017, Dublin, Ireland, pp. 1-13.
- [3] How to Retrain an Image Classifier for New Categories, [https://www.tensorflow.org/hub/tutorials/image\\_retraining](https://www.tensorflow.org/hub/tutorials/image_retraining)
- [4] Image classification Transfer Learning with inception v3, <https://codelabs.developers.google.com/codelabs/cpb102-txf-learning/index.html>
- [5] Jayaweera, MEDICINAL PLANTS, Indigenous and Exotic Used in Ceylon, 1993.
- [6] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, Trevor Darrell, DeCAF: a deep convolutional activation feature for generic visual recognition, in *Proceedings of the 31st International Conference on International Conference on Machine Learning*, June 21-26, 2014, Beijing, China.
- [7] Kadir, Abdul & Nugroho, Lukito & Susanto, Adhi & Santosa, Paulus. (2011). Leaf Classification Using Shape, Color, and Texture Features. *International Journal of Computer Trends and Technology*, vol. 1, pp. 225-230.

[8] Kumar, Neeraj & Belhumeur, Peter & Biswas, Arijit & W. Jacobs, David & John Kress, W & C. Lopez, Ida & V. B. Soares, João. (2012). Leafsnap: A Computer Vision System for Automatic Plant Species Identification, *Computer Vision - ECCV 2012*, 7573. 502-516. 10.1007/978-3-642-33709-3\_36.

[9] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2818-2826.

[10] Vishal Batvia, Drashti Patel, Dr. Avani R. Vasant "A Survey on Ayurvedic Medicine Classification using Tensor flow", *International Journal of Computer Trends and Technology (IJCTT)*, V53 (2): 68-70, November 2017. ISSN: 2231-2803, [www.ijcttjournal.org](http://www.ijcttjournal.org), Published by Seventh Sense Research Group.



## Impact of emerging motorbike swarms on transportation requirements in Sri Lanka: An agent based simulation study

E. R. Kulathilake  
Department of Industrial Management  
Faculty of Science  
University of Kelaniya, Sri Lanka  
kulathilake.roshan@gmail.com

C. Rajapakse  
Department of Industrial Management  
Faculty of Science  
University of Kelaniya, Sri Lanka  
chathura@kln.ac.lk

### Abstract

Transportation system in Sri Lanka is increasingly getting dependent on private modes rather than public modes. Among them, the motorbike has become the most popular mode of transportation among the middle income category of the society. Rush hour traffic congestion as well as the growing number of motorbike sales outlets appearing across the country, poor quality of public transport would be the main reason for this increasing trend of motorbike swarms. This research is based on the question where would this end up if this trend continues. Therefore, Agent-Based Modeling and Simulation methodology was used as the computational method to foresee the future by creating simulation models in order to study macro level emerging patterns in a system. The primary data to construct the model as well as to calibrate its parameters were collected through a questionnaire distributed among motorbike riders. The data collected, include the background of using motorbike as the primary mode of transportation, the issues in the public transportation system as well as the common riding patterns of individual motor biker. Simulation results of the constructed model were compared with the observations in the real environment to validate the model. The validated agent-based simulation model can be useful for policy makers to make predictions about the future and arrive at conclusions about the future traffic patterns in order to come up with better strategies to reduce the congestion.

**Keywords:** Traffic congestion, Agent-Based Simulation Model, Motorbike swarms

### 1. Introduction

Globalization encourages the community to optimize their resources and time in means of transportation aspect. The complexity of lifestyles of people is a major concern that drives the aforementioned need. People face difficulties in managing time for their day-to-day activities and in the meantime, they feel uncomfortable

with current transportation system. Development of infrastructure was a prominent factor continued over the past decade. One such example is emergence of possibility to travel from Colombo to Galle within two hours. However, the road traffic and necessity of having own transportation method are concentrating the focus on coming up with different travelling methods. As a result, Transportation system in Sri Lanka is increasingly getting dependent on private modes rather than public modes. As many East Asian countries such as Vietnam, Indonesia and Thailand are currently suffering from the largely unorganized traffic resulted from motorbike swarms, we see the investigation of the possible impact of this growing tendency to use motorbikes in Sri Lanka on the transportation system as a critical issue that the policy makers should be concerned of. As, most of people tend to use their own methods, the limitations such as high price of vehicles, traffic in urban areas and continuous increase of price on fuel will stand against the decision of purchasing a vehicle. On this ground, buyers have to consider all options for a beneficial and affordable decision to be made. Among them, motorbike has become the most popular mode of transportation among the middle-income category of the society, since automobile industry brings an easy and affordable solution called two wheelers to avoid such problems for everyone. Therefore, Sri Lankan two-wheeler customers do think that two-wheeler is the economical mode of travelling and a user friendly approach that counterparts [1]. Due to these factors, number of motorbikes are getting increased with reference to new registrations year by year as shown in the Figure 1.

Hence developing a simulation model for identifying growing trend and assessing reasons for this growing trend was the primary purpose of the study. Results of the validated model will be discussed in this paper.

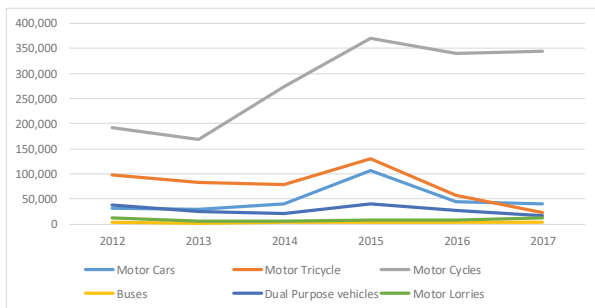


Figure 1. New Vehicle Registrations in Sri Lanka.

### 1.1. Factors affecting on motor-biking

The history of the motor bicycle begins in the second half of the 19<sup>th</sup> century. Now it has been observed that many people have started using motor bicycles to fulfill their daily transportation requirements. In Sri Lanka, motor bike has become the family vehicle too. Even though motor bike carries some risk, people use it without considering the risk it brings. Considering the factors impact on the purchasing decision of a motorbike, real median household income has been getting increase by last couple of years as shown in the Figure 2.

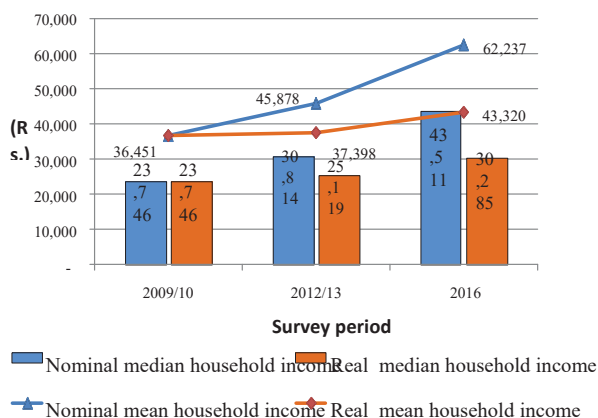


Figure 2. Household income level of Sri Lanka.

Poor quality of public transport, availability of finance and leasing facilities, availability of more brands and the promotions being offered to customers are the key factors affecting on a purchasing decision of motorbike.

### 1.2. Traffic congestion

Needless to say, contemplating how to travel is a common problem which is faced by people daily in Sri Lanka. As Sri Lankans, most of them spend a lot of time on the road when travelling. The intense traffic people face consist of all vehicles, whether public or private.

According to estimated statistics, 87.1% of vehicles on our roads fall under private transport, including cars, motorbikes and trishaws. The number of private vehicles on our roads greatly exceeds the minimum number necessary for a smooth flow of traffic. People travel on roads blocked by hundreds of cabs, vans, trishaws, cars, and motorbikes, with each vehicle switching lanes, creeping through small spaces and definitely taking a lot more space on each road than public transport would cover. This unordered traffic congestion also leads to think people about private mode of transport.

### 1.3. Computational intelligence

Generally, computational intelligence is a set of computational methodologies and approaches, which are inspired by nature to address complex real-world problems. Indeed, many real-life complex problems cannot be translated into binary language for computers to process it. Therefore, computational Intelligence provides solutions for such problems. Agent-Based Simulation method was proposed as the computational methodology to study this motor bikes growth.

### 2. Literature review

A study is conducted about “Psychological and social factors influencing motorcycle rider intentions and behavior” published in 2007 has assessed the psychological and social factors influencing motorcycle rider behavior for the Australian context. It has two studies and one of them has identified six key aspects of rider behavior considered to influence safety. Study two was survey-based and examined the psychological and social factors influencing these behaviors and other relevant psychological constructs, such as sensation seeking and aggression [2].

A study is conducted about “Motorbikes against Ho Chi Minh? Or the Consumption Icons of a Social Transformation in Vietnam” published in 2009 has addressed the question of the values and social transformations within the post-socialist society in Vietnam, symbolized by the consumerism practices of motorbikes. It also explores the process of classification and differentiation and the consumerism significance related to motorbikes. Vietnam’s recent economic and social transformations are manifested in the streets of its capital city through millions of motorbikes and a rapidly growing presence of cars. Therefore, motorcycle is an unavoidable item in contemporary Vietnam. Motor cycle has become the most visible consumption good in Vietnam [3].

A study is conducted about “Factors affecting purchase decision for Indian two wheelers in Sri Lankan market” published in 2015 has observed that within the

past decade, two wheeler usage indicates a rapid growth in Sri Lankan market. Among the two wheeler population, Indian two wheeler brands represent a huge portion. In this research study researcher put effort to find out what are the factors effect on decisions of consumers on Indian two wheelers as a novel research. Main purpose of this study was to identify why Indian two wheeler brands have become more popular and which factors effect on the purchasing decision and open up the gateway to study on this area during this study. Data was gathered from two wheeler users and non-users who are having either two wheeler or willing to have a two wheeler. Further contribution of the demographic factors such as age, gender, income level, education level were considered. Furthermore, how far marketing managers with their strategies impacted on the purchasing decisions of the two wheelers is also considered. Ultimately the study provides important insides how the variables impact on purchasing decisions towards the Indian two wheelers at the Sri Lankan market [1].

A study is conducted about “Urban traffic congestion: The problem & solutions” published in 2004 has revealed that motorization within Colombo city has increased with the increase of travel demand as many people refrained from using public transport mainly due to its inefficiency and poor quality. However, the increased motorization has led to many negative implications such as congestion, environmental pollution, road accidents, etc. Thus, introduction of an efficient and high quality public transport system that is capable of attracting private vehicle users has become an urgent need. This paper examines prevailing passenger transport demand, drawbacks and viable improvements required to the existing public bus transport in Colombo city [4].

A study is conducted about “Computational Social Science: Agent-based social simulation” published in 2007 has discovered that Social simulation is the application of computer-based methods and technologies to replicate human social behavior in various environments and scenarios. Agent Based Social Simulation is one of the key areas of social simulation. The construction of computer programs that simulate aspects of social behavior can contribute to the understanding of social processes [5].

### 3. Methodology

Research methodology of the research was simulation approach. Simulation will be implemented using agent-based modeling. In this model each agent has their own

behaviors. When simulating, all agents act as a group. There, agent’s behaviors will be considered in order to address the macro level environment problems. To match the simulator with real world situations, calibration is needed. Therefore, to set the model, parameters like factors which effect on motor biking are required. Those data will be collected through interviews and questionnaires. After designing simulator, validation is needed. Therefore, developed model is validated using some real world scenarios, which have been already collected through questionnaire.

Agent is a software entity that functions continuously and autonomously in an environment in which other processes take place and other agents exist. Agents are representation of real-world individuals. There are five agents, identified in the model. They are financial institutes, purchasing decision maker, motorbike users, general public and government. From the agents, parameters were identified. Those parameters will change the mind of a decision maker agent. Those parameters are loan/leasing rates, fuel price, peer influence, tax rates, road traffic and the quality of the public transport. Final conceptual model is shown in the Figure 3.

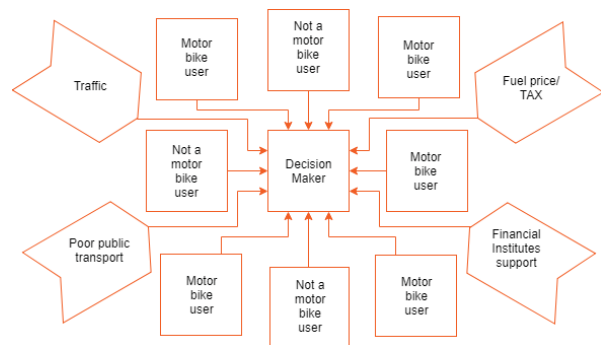


Figure 3. Final conceptual model.

### 4. Implementation

To construct the agent model, conceptual model will be drawn using analysis in order to make model design easier. To implement the identified simulator in to a model tool IDE is needed. There are several agent-based modeling IDEs. NetLogo was selected as the IDE since NetLogo is Agent-based programming language, integrated modeling environment, free and open source, and fully programmable. Language is Logo dialect extended to support agents. The NetLogo environment enables exploration of emergent phenomena.

After identifying the agents, those agents have to be placed in a network. That network is known as social network. It is a social structure made up of a set of social actors such as individuals or organizations. Social network can be represented with help of cellular automata model. Cellular automata model is a mathematical model can be used to investigate complex patterns of natural system. Final agent-based simulation model will be shown in Figure 4.

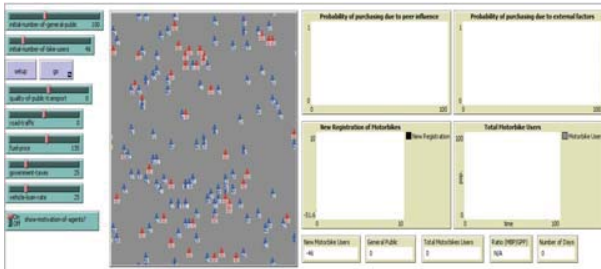


Figure 4. Agent-based simulation model.

At the left hand side there are some input controls for the model. General public is represented in blue and motorbike users are represented in red. General public is the people whose age range from 20-50 years and they belong to the potential cluster for purchasing a motorbike. Cluster was derived from the collected data. Depending on the external environmental changes they will decide whether to buy a motorbike or to use another method as their mode of transport. For everyone there is a motivation factor to purchase a motorbike. It is represented by a number 1-100. Every motorbike user has already purchased a motorbike. Therefore, their motivation is 100. General public was given a motivation value 1-100 randomly. According to the given inputs those motivation values will get changed. When it reached to 100, that individual purchase a motorbike. As an example if quality of public transport will get poor day by day, if I am the decision maker my motivation to buy a motorbike will be increased day by day. Model will be simulating for 3650 ticks (10 years). For every year population will grow by 0.76% according to the Sri Lankan population growth. At the initial point ratio between motorbike users and general public is 0.46.

## 5. Testing and validation

Real word data was used for testing and the validation of the developed model. According to the statistics provided by the Ministry of Transport & Civil Aviation, Sri Lanka, the motorbike usage growing trend was decreased highly in 2013. Considering the budget 2013, import duties of all vehicle imported was increased by 10-20 percent and Motor vehicle revenue license fee was

also increased. Due to these factors, there is decrease in growing trend as shown in the figure 1.1. Government tax parameter will be increased to see whether same results are giving or not. When the tax rates are high, motivation to buy a motorbike will be low as shown in the Figure 5.

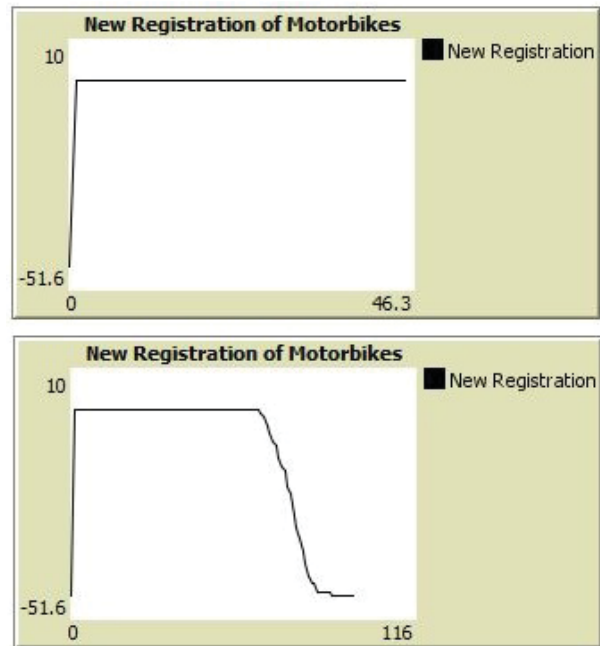


Figure 5. Model results due to parameter changes.

Model was tested for different scenarios. They are, changing the quality of public transport, changing road traffic condition, changing price of fuel (1L of petrol), changing government taxes and changing loan/finance rates.

Apart from the model findings, following probabilities and purchasing power percentages are found from the collected data. Purchasing power percentages are shown in the Table 1.

Table 1. Population with purchasing power.

Age group	Male	Purchasing power	Female	Purchasing power
20-29	1424234	64%	1515277	8%
30-39	1372396	11%	1574468	9%
40-49	1351926	3%	1440872	2%

From the collected data, 20-50 age groups have the motivation and willingness to buy a motorbike. Using conditional probability,

$P(A)$  – Probability of being an individual who can have the purchasing power

P(B) – Probability being an individual to their respective age group

$$P(A|B) = P(A \wedge B) / P(B)$$

Ex: for age group 20-29 Male

$$P(A \wedge B) = 1424234 * 0.64 / 8679679 = 0.105$$

$$P(B) = 1424234 / 8679679 = 0.164$$

$$P(A|B) = 0.105 / 0.164 = 0.64$$

Calculated conditional probabilities are shown in the Table 2.

**Table 2. Population with probabilities.**

Age group	Male	Probability of purchasing	Female	Probability of purchasing
20-29	1424234	0.640	1515277	0.080
30-39	1372396	0.107	1574468	0.089
40-49	1351926	0.029	1440872	0.002

Therefore, age 20-29 male age group has the highest probability to purchase a motorbike due to environmental changes as discussed.

## 6. Conclusion

In this study, simulator was developed using Netlogo. Agents and factors were identified through the questionnaire, which was distributed among motorbike riders. Then those agents and their behavior were modelled in the agent-based model. The simulator has been calibrated and validated with real data through responses given by motorbike users. Test result has shown that the simulator accurately models motorbike population growth in a variety of conditions. Conditions will be handled by different agents. Finally, this model would be useful for policy makers when there are actions need to be taken at different conditions. Then, this agent-based simulation model can be applied to analyze patterns. When applied the simulator to analyze pattern regarding motorbike population, it proves to be a useful tool to identify solutions for traffic congestion resulted by motorbike swarms. Policymakers can see the visualized output of different scenarios using the simulator.

## 7. References

- [1] Mendis, A. S. M. G. and Weerasiri, R. A. S., “Factors affecting purchase decision for Indian two wheelers in Sri Lankan market”, *Kelaniya Journal of Management*, Vol. 4 No. 2, July-December 2015, pp. 11-15.
- [2] Watson, B., Tunnicliff, D., and White, K., “Psychological and social factors influencing motorcycle rider intentions and behavior”, *ATSB research and analysis report road safety research grant report*, August 2007, pp. 10-21.
- [3] Freire, A. D., “Motorbikes against Ho Chi Minh? Or the Consumption Icons of a Social Transformation in Vietnam”, *The Copenhagen Journal of Asian Studies*, January 2009, pp. 02-17.
- [4] Kumarag, A. S., “Urban traffic congestion: The problem & solutions”, *Economic Review-Sri Lanka*, January 2004, pp. 02-07.
- [5] Gilbert, N., “Computational Social Science: Agent-based social simulation”, *The University of Surrey*, May 2007, pp. 01-11.



# **Systems Engineering**





## GPS guided auto sensing system for motor vehicles

D. C. R Ranasinghe  
Samson Rajarata Tiles (pvt) Ltd.  
Sri Lanka  
c2rishere@gmail.com

W. K. I. L. Wanniarachchi  
Department of Physics,  
University of Sri  
Jayewardenepura,  
Sri Lanka  
iwanni@sjp.ac.lk

U. A. D. N. Anuradha  
Department of Industrial  
Management  
Faculty of Science  
University of Kelaniya, Sri Lanka  
nilupulieanuradha@yahoo.com

### Abstract

*Driver errors are the most common cause of traffic accidents. Mobile phones, in-car entertainment systems, traffic volume increases, road systems becoming complicated contribute towards such driver errors. This paper introduces developing a GPS guided auto pilot system for vehicles. This system will be a driverless vehicle system. The user has to give the location of the destination to the system and then the system will automatically navigate to the given destination. These systems are currently used in aircraft, submarines and ships but not used in ground vehicles. Use of such systems in the open street is more complex than use of such systems in air or marine systems. The possible route to the destination must be selected by the vehicle after the destination coordinates are given by the user. Then the vehicle navigates through the open streets without colliding with other moving or non-moving objects. GPS sensor takes real time coordinates of the vehicle and decide the direction to be moved with respect to the given destination coordinates and pass control signals to the motor controller. While navigating, the vehicle keeps appropriate safe distance and speed with the vehicles in front of it. If the lane is not clear, the vehicle applies breaks to avoid collisions. Sonar sensors are used to detect the object in the road as they are more convenient in the outdoor applications. With further developments, this system will be able to assist drivers who drive long trips and play a vital role in minimizing road accidents.*

**Keywords:** Global Positioning System, auto pilot system, vehicle navigation

## 1. Introduction

### 1.1. Background

An autonomous robot is a self-piloted vehicle that does not require an operator to navigate and accomplish its tasks. There are two categories in autonomous navigating named as indoor navigating and outdoor

navigating. Indoor navigation can be seen as line follower robots, RFID based navigation, Wireless network-based navigation etc. Outdoor navigation is widely based on GPS. In the outdoor navigation category, auto-piloted vehicles are a recently developed subset of robotics and can come in three general forms; air, ground and submarine. Auto-pilot technology is widely used in both air and submarine forms except in ground form because in ground form there are more parameters that have to be considered when auto-piloting rather than the other forms. The term “Auto-pilot” is not a term widely used when referring to ground vehicles. Implementing of auto-pilot system for ground vehicles is more complex, since roads have more traffic density than the air and the marine systems. But introduction of these systems will be helpful for many motorists such as older people who are not fit enough to drive long distances or have some physical handicap. These systems take decisions after obtaining information from the environment, and do not take unjustified risks. Therefore, number of potential accidents will be lowered when such systems are used.

Global Positioning System is a widely-used tool in navigation, cartography and land surveying. United States Department of Defense initially developed the GPS technology for their military activities. But they had applied a selective error in those systems to avoid having more precise data for the hostile forces. In May 01, 2000, US president Bill Clinton ordered the US defense to stop jamming signals and to turn off selective error function that prevented other users to have pinpoint accuracy [1]. There are twenty-four satellites placed in the orbit to transmit data for the GPS receivers [2].

To calculate the location of a receiver it must receive signals from at least three satellites for a 2D fix and at least four satellites for a 3D fix [3].

There are several advanced studies that have been done in this field in both laboratory and commercial scale. Most of the leading automakers such as Mercedes, Audi, and BMW invest on this field. Among them Tesla introduced an auto-pilot system called advanced driver assistance system for their vehicles in October 2015 [4].

Yamaha Motors developed a robotic chauffeur named Motobot which is able to ride a motor bike without any aid from the human being. Apart from developing an auto-pilot system, they have created a robot who can ride the bicycle on the race track with the performance similar to a professional racer. That robot also can detect objects, detect curves and curbs and navigate to the given destination [5]. Department of the Transactions of the Chinese Society of Agricultural Engineering has developed an autonomous navigating system for linear move sprinkler machine based on GPS. That rover was also capable of navigating to the given GPS location in a linear way, autonomously. It was equipped with a microcontroller, navigation controller, speed steering controller, GPS sensor and a digital compass sensor [6].

## 1.2. Objectives

The purpose of this study was to develop an algorithm to identifying the location given by the user and to analyze the behavior of the moving and non-moving objects on its way for steering the vehicle through the streets. The vehicle must be able to identify the angle which it is heading according to the destination. The rover should be able to go forward if its' lane is clear. Otherwise it must stop or maintain appropriate speed and distance with the front vehicle. It should not be collide with the any of the objects on its track.

## 2. Problem statement

Developing an auto-pilot system for ground vehicles will address several problems in the open streets. Number of road accidents occur because of the risky decisions taken by the drivers. Sometimes they cannot take correct actions or reactions within a limited time period. Lack of vision, slow reaction speed are main causes for those accidents. The proposed system acts as an assistant to make those decisions. The system takes decisions based on the sensors mounted to the vehicle and the given algorithm. So, this system helps to take more precise and timely decisions.

There are several factors affecting the precision and accuracy of GPS data. They are,

- Number of visible satellites
- Satellite geometry
- Satellite clock errors
- Receiver errors

Because of these errors, it cannot be solely dependent on GPS data to navigate a vehicle. Therefore, some other sensors must be added to evaluate the surrounding of the vehicle. Sonar sensor is the best match to "know" the surrounding. Then with the aid of sonar sensors vehicle can navigate to the given destination without any

collisions. That means sonar sensors aid the main objective, which is self-navigating to the destination.

## 3. System modeling

### 3.1. Overview

This study consists of several parts as listed below,

- Designing and fabricating the rover's chassis.
- Selecting the appropriate components.
- Building the Collision avoiding algorithm.
- Building the GPS navigating algorithm.
- Programming the rover

### 3.2. Methods and techniques

#### 3.2.1. Designing and fabricating the rover's chassis

To demonstrate the developed algorithm, the rover must act similar to an actual vehicle. Most of the RC cars in the open market fulfill that requirement. It has single motor to power the rear wheels with a differential system. Steering is done using another motor. It consists front and rear spring suspension system which assist smooth mobility on the streets. The study used the RC car and disassembled to fix sensors and controllers.

The motors must be operated according to the signals given by the microcontroller. Microcontroller unit analyses the data from the GPS sensors and the sonar sensors according to the developed algorithm and gives command to the motors.

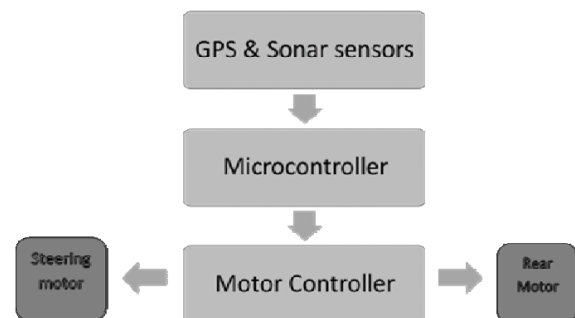
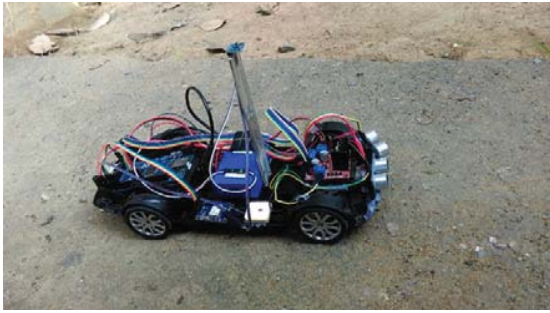


Figure 1. Block diagram of the rover.

#### 3.2.2. Selecting the appropriate components.

In this rover, several electronic components have been used.

- Arduino Mega Microcontroller
- L239D Motor controller
- GPS sensors
- Sonar sensors
- HMC 5883L Compass Sensor



**Figure 1. Assembled Rover.**

a) Arduino Mega Microcontroller

Arduino Mega 2560 is used as the microcontroller in this system. It gets input signals from the GPS sensor and the Sonar sensors. Then it processes the signal according to the given algorithm and supply output signals to the motor controller to drive the motors. Microcontroller is programmed using Arduino IDE.

b) L293D Motor controller

L293D is a typical Motor driver or Motor Driver IC which allows DC motor to drive on either direction. L293D is a 16-pin IC which can control a set of two DC motors simultaneously in any direction [7].

c) GPS sensor

Technical data sheet describing the cost effective, high-performance u-blox 6 based NEO-6 series of GPS modules that brings the high performance of the u-blox 6 positioning engine to the miniature NEO form factor. These receivers combine a high level of integration capability with flexible connectivity options in a small package. This makes them perfectly suited for mass-market end products with strict size and cost requirement.

d) Sonar sensors

The HC-SR04 ultrasonic sensor uses sonar to determine distance to an object like bats or dolphins do. It offers excellent non-contact range detection with high accuracy and stable readings in an easy-to-use package. From 2cm to 400 cm or 1" to 13 feet. Its' operation is not affected by sunlight or black material like Sharp rangefinders are (although acoustically soft materials like cloth can be difficult to detect). It comes complete with ultrasonic transmitter and receiver module [8].

e) Digital compass

The HMC5883L sensor is a 3-axis digital magnetometer IC designed for low-field magnetic sensing. Communication with the HMC5883L is simple and all done through an I2C interface. Need to connect power, ground and only two cables to Arduino board (SDA, SCL).

The HMC5883L board can be powered up by 5V or 3.3V pins of Arduino board.

### 3.2.3. Building the collision avoidance algorithm

This algorithm depends on data received by the sonar sensors. There are three sonar sensors mounted in the rover. One on the front and another two on the left and the right hand sides.

Depending on the data being received by the sonar sensors, algorithm divides into two parts.

a) Braking

b) Keeping the appropriate distance with the front vehicle

a) Braking

Braking of the vehicle is done when the vehicle has not enough space to move in front. For an example, vehicle in the front is stopped and the next lane is not clear to overtake, then vehicle should have to brake and stop itself. For applying the brakes, vehicle should determine its braking distance with respect to the current velocity. In this demonstration system, braking is done by cutting out the current supply to the motors by PWM signals.

b) Keeping the appropriate distance from the front vehicle

If there is another car is moving in front of the rover, it should maintain its distance with respect to the front car. A constant value is given to the rover which is known as the safe distance. If the distance between two cars is equal to that value, rover decelerate to a safe speed in order to maintain the distance.

### 3.2.4. Building the GPS navigating algorithm

a) Autonomous GPS navigation is done in several parts.

b) Identifying the current location

c) Identifying the destination

d) Selecting the possible route

e) Identifying the current headed direction with respect to the destination direction.

f) Steering the rover

The rover can identify its current location using the satellite data from the GPS sensor. It takes its latitude and longitude to determine the location. When the user enters the destination, it takes the latitude and longitude values of the destination. The rover determines the direction which it should move and emit control signals to the motor controller. While navigating, the rover obeys the algorithm for collision avoidance.

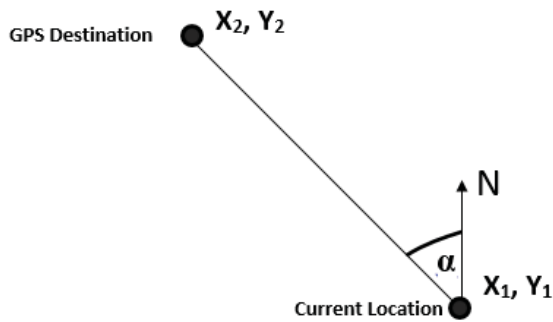


Figure 2. Turn angle between the destination and the north direction.

$$\tan \alpha = \frac{(Y1 - Y2)}{(X1 - X2)}$$

$$\alpha = \tan^{-1} \frac{(Y1 - Y2)}{(X1 - X2)}$$

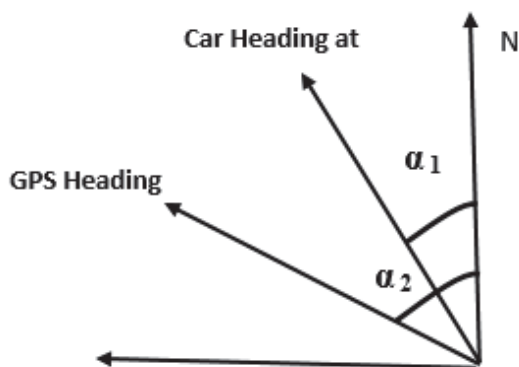


Figure 3. Angle difference between the destination and the rover direction with reference to the north direction.

- If  $\alpha_1 > \alpha_2$  Vehicle turns right
- If  $\alpha_1 < \alpha_2$  Vehicle turns left

The north of planet and magnetic north is not the same point. Magnetic north is some distances off the North Pole. Therefore, some value called “declination angle” must be added to get actual north when we use compass. While testing, the rover declination angle was calculated using the data from an online reference map.

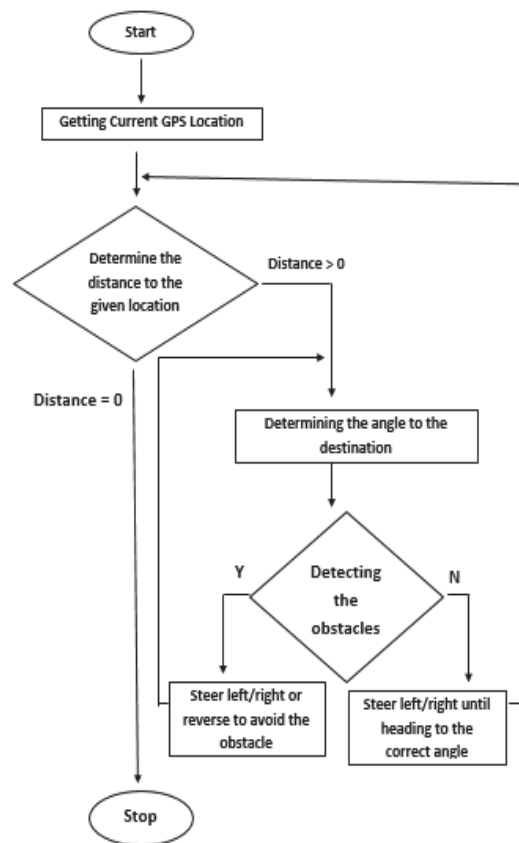


Figure 5. Flow chart of the algorithm.

#### 4. Speed analysis of the developed system

Table 1. Test data.

Action	Distance of the Path (m)	Percentage of Successes trials at 0.5ms <sup>-1</sup> Speed (%)	Percentage of Successes trials at 0.25ms <sup>-1</sup> Speed (%)
1	50	55	70
2	50	70	80
3	50	60	70
4	50	60	70
5	50	65	75
6	50	70	80

#### Action Description

- 1 - Autonomous GPS navigation done in several parts
- 2 - Identifying the current location
- 3 - Identifying the destination
- 4 - Selecting the possible route
- 5 - Identifying the current headed direction with respect to the destination direction
- 6 - Steering the rover

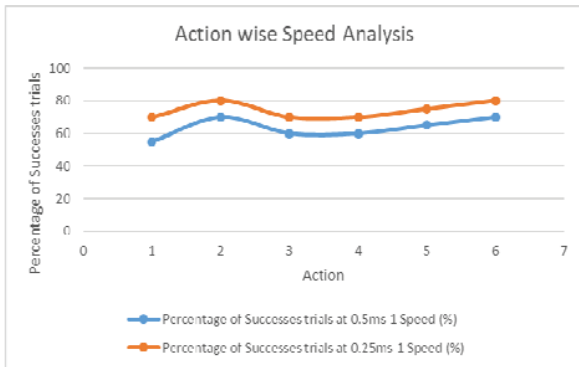


Figure 4. Action wise speed analysis.

Following hypotheses are considered to check the significance of different actions

#### Hypothesis

**H<sub>0</sub>** – There is no significant difference in the speed at 0.5ms<sup>-1</sup> and developed system based on action.

**H<sub>1</sub>** – There is a significant difference in the speed at 0.5ms<sup>-1</sup> and developed system based on action.

Chi-squared test was applied to check the above test hypothesis and results are shown in Table 2.

Table 2. Chi-Square Goodness-of-Fit Test for Observed Counts in Variable: Success of trials at 0.5ms<sup>-1</sup>

Action	Observed	Test Proportion	Expected	Contribution to Chi-Square
1	56	0.166667	64	1
2	72	0.166667	64	1
3	63	0.166667	64	0.01563
4	60	0.166667	64	0.25
5	65	0.166667	64	0.01563
6	68	0.166667	64	0.25

Results of chi-squared test indicate that there is no significant difference in success of trial according to the given 6 actions at 0.05 significant level (P-value = 0.772)

#### Hypothesis

**H<sub>0</sub>** – There is no significant difference in the speed at 0.5ms<sup>-1</sup> and developed system based on action.

**H<sub>1</sub>** – There is a significant difference in the speed at 0.25ms<sup>-1</sup> and developed system based on action.

Chi-squared test was applied to check the above test hypothesis and results are shown in Table 3.

Table 3. Chi-Square Goodness-of-Fit Test for observed counts in variable: Success of trials at 0.25ms<sup>-1</sup>.

Action	Observed	Test Proportion	Expected	Contribution to Chi-Square
1	70	0.166667	74.1667	0.234082
2	80	0.166667	74.1667	0.458801
3	70	0.166667	74.1667	0.234082
4	70	0.166667	74.1667	0.234082
5	75	0.166667	74.1667	0.009363

Results of chi-squared test indicate that there is no significant difference in success of trial according to the given 6 actions at 5% significance level (p-value = 0.898)

#### 4.1. Collision analysis

Active collision avoidance system is tested against both moving and non-moving objects. For non-moving objects, the number of collisions, when the rover moves at 0.5 ms<sup>-1</sup> is 10%, and for moving objects, the number of collisions when the rover moves at 0.5 ms<sup>-1</sup> is recorded as 40%. Even if the rover moves at 1 ms<sup>-1</sup>, the proportion of non-moving objects remains the same, but in the case of moving objects, the collision ratio increases up to 60%.

#### 5. Result and discussion

In general, the rover could move through a given path without collisions with the objects in its path. It could stop when the lane is not clear and could maintain safe distance with the vehicle in front of it. Steering was done smoothly.

When considering the GPS navigation, it was not a success due to various reasons. The current location of the rover was not precisely accurate and had an error of about 25m air distance. For that level of inaccuracy of the data as mentioned in the problem statement, there could be several reasons this might affect the results. Another major error occurs when the on-board CMOS battery drains. The battery is there for retaining data and nonvolatile storage for configuration. Normally the battery is a rechargeable one. It takes more than two hours to recharge at the initial starting point. The battery status affects also the data accuracy. To have a better accuracy, the battery must be at 5v.

For obstacle avoidance and collision avoidance, sonar sensors are the best match rather than IR sensors, Sharp IR sensors or Laser range finders. As this system is for the outdoor applications, IR based sensors are not suitable because it has considerable amount of noise occurrence with sunlight. Laser range finders are accurate, but cannot be used to detect objects with a small width such as poles. But the sonar sensors have no interaction or noise generation due to sunlight, colour of the material or the width of the object. It has little amount of noise effect due to the echo signals from the other sound generating object. But relatively to the other sensors the error due to noise is negligible.

The compass module was initially mounted on the rover chassis. But when the rover is moving, magnetic field which was generated from the motors affected the accuracy of the compass. Therefore, the compass was mounted high from the motors, to reduce the magnetic field effect.

## 6. Conclusion and future work

Driverless vehicle is a task which almost all the vehicle manufacturing companies are trying to develop. They use most advanced technologies in their research studies. Most important in this kind of system is the accuracy as the system directly deals with human lives. So, when the accuracy is higher, the probability to make accidents become minimal. That's why the pinpoint accuracy of data is needed in such a system.

This system can be applied to the motor vehicle as an auto piloting system such as in aircraft. One problem the drivers face most of the time is getting bored while driving for long distances such as in highways. By implementing this system to the vehicles, drivers can take a rest whenever they want without stopping the vehicle. Handicapped people also can easily travel wherever they want because this system assists the driver while driving. As speed controlling, braking is done automatically after analyzing the surrounding, chances for accidents to occur will be decreased.

In the proposed system, collision avoidance algorithm is solely relying on the sonar sensors. But accuracy of collision avoidance can be increased by combining the vision-based sensor, such as real time image processing techniques with the sonar sensors. Then the vehicle will receive better understanding about the surrounding.

In this study, destination which the user needs to go is given by the latitude and the longitude values. This is not user-friendly. Hence, this research can be extended to input the destination directly by name of the place or selecting the destination in a map.

## 7. References

- [1] Cameron. L, *Geocaching Handbook: The Guide for Family Friendly, High-Tech Treasure Hunting. 2nd ed.* Montana: Morris Book Publishing (2011).
- [2] Anon., *National Coordination Office.* (2016) [Online] Available: <http://www.gps.gov> [Accessed 07 October 2016].
- [3] Weston. E, *2-D vs. 3-D Fixes* (2015) [Online] Available: <http://www.gpsnuts.com/mygps/gps/technical/ed.htm> [Accessed 07 October 2016].
- [4] *Autopilot*, Retrieved from [www.tesla.com](http://www.tesla.com): (2016, 1229). Available: <https://www.tesla.com/presskit/autopilot> [Accessed: 2 Feb. 2017].
- [5] *Future Tech.* (2015, 11 4). Retrieved from Pauley: <http://www.pauley.co.uk/techreview/future-tech/self-driving-car-or-a-robotic-chauffeur-autopilot>.
- [6] Liu. K, Wu. P, Zhu. D, Chen. J, *Design and experiment of autonomous navigation system for solar powered linear move sprinkler machine based on GPS.* (2016). Retrieved from [www.researchgate.net](http://www.researchgate.net): [https://www.researchgate.net/publication/306229755\\_Design\\_and\\_experiment\\_of\\_autonomous\\_navigation\\_system\\_for\\_solar\\_powered\\_linear\\_move\\_sprinkler\\_machine\\_based\\_on\\_GPS](https://www.researchgate.net/publication/306229755_Design_and_experiment_of_autonomous_navigation_system_for_solar_powered_linear_move_sprinkler_machine_based_on_GPS)
- [7] Ron Robotics, *L293D Motor Driver IC* (2013). [Online] Available: <http://www.rakeshmondal.info/L293D-Motor-Driver> [Accessed 12 October 2016].
- [8] Technologies, *Product User's Manual – HC-SR04 Ultrasonic Sensor*, Cytron Technologies: (2013)

## Industry 4.0 maturity assessment of the Banking Sector of Sri Lanka

O. K. K. Bandara  
*Department of Industrial  
 Management  
 Faculty of Science  
 University of Kelaniya  
 Sri Lanka*  
 oshadhikavindra@gmail.com

V. K. Tharaka  
*Department of Industrial  
 Management  
 Faculty of Science  
 University of Kelaniya  
 Sri Lanka*

A. P. R. Wickramarachchi  
*Department of Industrial  
 Management  
 Faculty of Science  
 University of Kelaniya  
 Sri Lanka*

### Abstract

*Technological advancements have created massive changes in the way of performing businesses in this era. Whilst in creating the transition of physical world to a virtual world, industry 4.0 plays a significant role. Thus, the development of the concept of industry 4.0 revolutionize the way of conducting businesses. This concept was first limited to manufacturing sector thus with the evolvement of customer behaviour, service sector also applied these concepts to offer a better customer satisfaction. As modern customer expectations have risen with technology, Sri Lankan banking sector focussed to deliver their services strategy by enabling advanced technologies. Hence there is an essential need to define, formulate a set of guidelines in order to assess the progress of the current state of Sri Lankan banking sector in their journey of adapting industry 4.0. So, this scrutiny assesses the maturity of Sri Lankan banking sector by application of an industry 4.0 maturity model developed by the authors. The results of this study indicates that Sri Lankan banking sector is in the third maturity level of the model "Defined" as the overall maturity is 3.668.*

**Keywords:** *Industry 4.0, Sri Lankan banking sector, technological advancements*

### 1. Introduction

Since the beginning of the industrialisation, technological dives have created massive changes in the way of performing the businesses, hence in retrospect, these deviations can be defined as industrial revolutions. Industrial revolutions first started as industry 1.0, where the concepts of steam power, water power and mechanization was practised. The second was the concepts of mass production, assembly line and electricity. Then it evolves as third industrial revolution (Industry 3.0), the introduction of computerization and automation though the fourth is the move towards digitization [9]. Industry 4.0 (I4) is the concept of automation and data exchange in the manufacturing technologies, which enables the use of Internet of

Things (IoT), Cyber-Physical Systems (CPS), big data analytics, cloud computing and cognitive computing in order to move towards a smart world [12].

The dispersion of the concept of industry 4.0 to the service sector was begun recently thus it evolves in the financial sector as well. As modern customer expectations have risen with the technology improvements, they value relationship centric experience that hinges on trust and personalization while receiving unlimited accessibility, convenience and speed of service. This will challenge banks to deliver their services strategy in order to survive in the competition. Therefore, banking sector has been integrated the concepts of industry 4.0 into their processes to offer better customer satisfaction.

Accordingly, there is an essential need to assess the performance of the banking sector in the journey of industry 4.0, with the aim of continuous improvement. Thus, this study assesses industry 4.0 maturity of Sri Lankan banking sector by applying a model developed by the authors. The summary of the developed model was given in the next section of this paper. Section 3 of the paper provides a brief description of the methodology that the researchers have followed and the section 4 provides an analysis of the collected data. The final section of this paper provides a discussion of the findings of the study followed by a conclusion and future research perspectives.

### 2. Maturity model

The applied maturity model consists of seven dimensions which are further assessed by its respective explanatory factors. The model was customised to measure the maturity of the banking sector. The maturity assessment of the model is done based on the assumption of each dimension is equally important. Table 1 provides an overview of seven dimensions of the maturity model.

**Table 1: Maturity Dimensions.**

Dimension	Explanation
Products and services	Customised products, Digitalised products, Data-driven services, Digital marketing approaches
Technology and Resources	Use of advance technologies (PSD II, mobile wallet, mobile Apps, cryptocurrencies, biometric technologies, behavioural analytics through AI and etc.), use of mobile devices, cloud technologies
Strategy and organisation	Implementation I4 roadmap, Available resources for realization, Adaption of business models, organisation culture for innovation, Alignment of I4 with organisational goal
Operations	Decentralization of processes, Interdisciplinary, Interdepartmental collaboration, Data driven excellence, New digital business models, Equipment readiness for I4
Customers	Personalised services, Utilization of customer data, Digitalization of services, Customer's Digital media competence, customised Apps.
Governance	Labour regulations for I4, Technology standards, Real time Decision making, Data governance (Data collection, Usage, Data analytics and use of Big Data tools, and Data-driven services), Information Security
Employees	Knowledge, Expertise and Competences of employees to new technologies, User training, ensure service accuracy

The applied model consists of five maturity levels such as initial, managed, defined, established and optimised.

(I) **Initial Level:** The processes are poorly controlled and process management is reactive. The organization mainly focusing on the fundamental operations that helps to run the business. There is no proper organizational and technological tools and infrastructure to enhance the smooth run of operations.

(II) **Managed Level:** The process is moderately planned and implemented. The transformation has been started and process management is in considerable level thus limited tools for organization and enabling technologies. Represent a partial maturity in managing the technological infrastructure development. The vision of Industry 4.0 and the idea for a I4 roadmap exists thus it is not fully implemented.

(III) **Defined Level:** The processes are defined and considerable practices for planning and management procedures are employed. Limited management of the processes due to constraints on technical and resource needs. Practise for integration and interoperability concepts of the information exchange. Data driven services are taken to consideration. The vision of Industry 4.0 and the idea for a I4 roadmap exists thus, it is not fully implemented.

(IV) **Established Level:** The vision of Industry 4.0 and the idea for a I4 roadmap is implemented. Key activities, value added operations are well-defined and employed technologies to increase efficiency of the processes respect to best practices in the industry. The processes and operations are considered to be follow standardization. The concepts of integration and interoperability to improve information interchange is highly practised

(V) **Optimised Level:** The organisational processes are digitally oriented thus based on a solid technology infrastructure targeting on high potential growth organization. High level of integration of the concepts and principles of I4 thus personalised, data driven service is provided. Collaboration with experts to offer better services. Information, service accuracy is guaranteed. Data security is a major concern. Automated, user friendly, customised processes.

### 3. Methodology

The initial step of this study is to comprehensively analyse the literature according to the areas of Industry 4.0, Industry 4.0 applications in the banking sector and maturity models for assessing industry 4.0 maturity levels. Total of 26 articles were analysed based on their relevancy and used them as a basement for developing the proposed maturity assessment model. Accordingly, this study employed a primary data collection through a questionnaire survey to assess the maturity of Sri Lankan banking sector by using the developed model. Judgemental sampling was used to select the sample of 10 banks out of the population of 26 licensed commercial banks in Sri Lanka. This sample includes both private and state owned banks in Sri Lanka. Thus, data was collected through interviewing the employees of the head offices of each bank which were located in Colombo, Sri Lanka. The data collection process involves both quantitative and qualitative approaches where the analysis was performed based on both of the results. Semi-structured questionnaire was used as the research instrument. The questionnaire was prepared based on the maturity dimensions and its explanatory criteria of the applied model and comprised of both open and closed ended questions. Closed-ended questions employed a Likert scale reaching from 1- "Very Poor" to 5- "Very Good". Thereby, the collected data was analysed using a mathematical equation, in order to gain



the maturity of each dimension as a weighted average of all explanatory factors in respected dimension. The weighted average equation is shown below.

$$M_D = \frac{\sum_{i=1}^n M_{E_i} \times W_{E_i}}{\sum_{i=1}^n W_{E_i}}$$

**M** – Rating factor  
**D** – Maturity Dimension  
**E** – Explanatory Factor  
**W** – Weighted Score  
**n** – # of Explanatory factors

The overall maturity of the banking sector is calculated using the average maturity of all seven dimensions.

#### 4. Analysis and results

The maturity variation of the sample banks respected to dimension 1, 2 and 3 are displayed in the following figures. Figure 1 displays the maturity variation of first dimension: Product and Services respected to selected sample of 10 banks in Sri Lanka. The numbers 1 to 5 displayed in the radar chart (figure 1) are the explanatory factors considered to measure the maturity of the dimension 1 and A to J are selected sample of 10 banks.

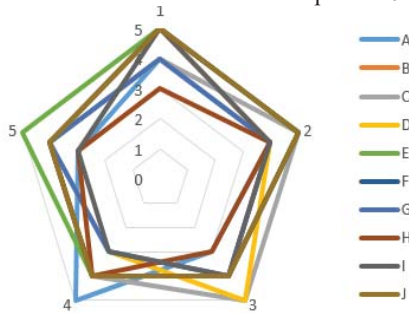


Figure 1: Product and services.

Figure 2 displays the maturity variation of the second dimension: Technology and Resources respected to selected sample of 10 banks in Sri Lanka. The numbers 1 to 6 displayed in the radar chart (figure 2) are the explanatory factors considered to measure the maturity of the dimension 2

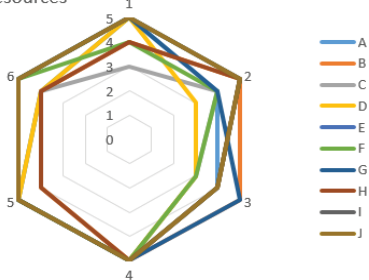


Figure 2: Technology and resources.

Figure 3 displays the maturity variation of the third dimension: Strategy and Organisation respected to selected sample of 10 banks in Sri Lanka.

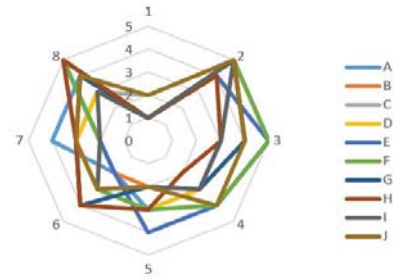


Figure 3: Strategy and organisation.

The numbers 1 to 8 displayed in figure 3 are the explanatory factors used to measure the maturity of third dimension: Strategy and organisation. As the figure 3 shows some of the explanatory factors in dimension 3 are performing low when considering the banking sector of Sri Lanka due to factors like less focussed on using industry 4.0 road map etc. These maturity variations are graphed in radar charts based on the data collected from the sample.

The overall maturity calculations related to each dimension is displayed in table 2.

Table 2: Overall maturity.

Bank	Dimension							Average
	1	2	3	4	5	6	7	
A	2.947	3.481	4.760	4.231	4.560	3.524	3.111	3.802000
B	3.095	3.444	4.417	4.200	4.185	3.048	2.684	3.581857
C	3.045	3.750	4.143	3.769	3.964	4.000	3.045	3.673714
D	2.950	3.440	4.320	4.240	3.920	3.455	3.000	3.617857
E	2.957	3.517	4.667	4.280	4.444	3.455	3.100	3.774286
F	2.864	3.692	4.607	4.214	4.333	3.609	2.857	3.739429
G	3.000	3.552	4.308	4.154	4.250	3.333	3.000	3.656714
H	3.059	3.500	4.280	4.185	4.000	3.238	2.950	3.601714
I	3.211	3.517	4.091	3.955	3.714	3.789	2.857	3.590571
J	2.864	3.517	3.926	4.120	4.345	3.632	3.143	3.649571
Average	2.9992	3.541	4.3519	4.1348	4.1715	3.5083	2.9747	3.668771

As the table 2 shows the overall maturity of the Sri Lankan banking sector is calculated as 3.668.

#### 5. Discussion

The maturity levels of each dimension was calculated in this study using the mathematical equation where the findings were displayed in table 1. Thus, the calculated results reveal that under the dimension 1: products and services, the selected banks in the sample is performing in the maturity range between 2.8 to 3.2 where B, D, G, H and I were performing better than A, D, E, F, J. But, the results are quite same where the variation is 0.4. The figure 1 better visualises the responses of the sample regarding the dimension 1. Hence, the results reveal that all the banks are considering the advance technology based service offering in order to provide a better customer satisfaction based on the assessment of products and services.

When considering the outcome of the maturity levels of the dimension 2: technology and resources with reference to table 2, Sri Lankan banking sector is performing better where the results variates between 3.4 to 3.8 with a variation of 0.4. The figure 2 better visualises the responses of the sample regarding the dimension 2. As per the results, out of the sample of ten, A, B and D are showing lower result than others where their technology and resource aspects should be better improved. PSD II, mobile apps, biometric identifications, internet of things and use of Big data can be identified as use of technologies and concepts of industry 4.0 when considering Sri Lankan banking sector.

Considering the results of the maturity dimension 3: strategy and organisation as displayed in table 2, Sri Lankan banking sector is performing better as the result varies between 3.9 to 4.7 with a variation of 0.9. As per the results, out of the sample of ten, G, H, I and J are lagging behind where A, B, C, D, E and F are performing at similar level. Thus, it reveals that more of the organisations strategy and vision is now aligned with the industry 4.0 vision. Whilst the selected sample is aware about the industry 4.0 road map but still not practise to use it, neither the documentation of activities.

According to the results displays in table 2 in dimension 4: Operations, selected sample is performing good. The maturity values are varying in between 3.7 to 4.2 where the variation is 0.5. Out of the sample of ten, A, B, D and E are performing at a similar maturity level where sample C and I are resulting low performance. Hence, as the high performance organisations revealed that their processes were decentralised where the employees are empowered to take their decisions thus the power and the authority were given. Thereby, the decision making was done based on the local market environment. Although most of the organisations practised interdisciplinary processes in order to facilitate decision making via cross functional or expertise teams thus technology supports in collaborative decision making.

When considering the results of the dimension 5: customers, the maturity value varies in between 3.7 to 4.5 where the variation is 0.8. Thus, out of the sample of ten, B, C and I were resulted low performance and the others were performed at the same level with small deviations. As stated by the high performed organisations under this dimension, they practised using the concept of personalised service offering to the customers where the all processes were more customer centric. Although apart from their branches, they provide personalized services through other channels like telephones, emails, video conferencing and smart phone apps.

According to the results of the dimension 6: Governance, the performance of the selected banks varies in between 3.0 to 4.0 where there is a huge variation of 1.0 compared to other dimensions. Thus, out of the sample of ten, A, C, E, I and J are performing

similar ranges where their performances are quite similar. The facts like maintaining minimum technology standards in the business models were becoming a qualifying factor for all banks today with the advancement of technology. Even though when comparing private banks vs state banks in 5 to 10 years ago, private banks performance was really good compared to state banks regarding the technology aspect. Whilst today, both are similarly performing in almost all the business processes. Although in mobile banking, internet and online banking the data governance are practising within the bank, when data collection from consumers for usage and data analytical purposes thus data security is a major concern for all of the banks.

According to the results of the dimension 7: Employees, the results vary in between 2.6 to 3.1 with a variation of 0.5. As the results indicated, out of the sample of ten, B, F, H and I show low performance comparing to others. Hence, as most of the respondents of the of the sample mentioned that their employees are aware about the use of advanced technologies whilst the traditional and older generation employees are reluctant to use new technologies even though the competency and expertise of the employees are relatively low regarding the introduction of new technologies and concepts. Hence, it is important to provide better training about these concerns in order to offer a better service without errors as well as to reduce the restriction arise form employees.



Figure 4: Overall maturity variation.

The overall maturity of the Sri Lankan banking sector under each dimension varies as displayed in figure 4. This was graphed by taking the average values of each bank within each dimension. Hence, it is visible the performance of the Sri Lankan banking sector regarding the first dimension “product and services” is 2.99, where it is a low value. Although dimension 7: employees also performing at a low value, 2.97 compared to other dimensions. Thereby, the reason for this result can be due to low focus on these dimensions when implementing industry 4.0 technologies in their processes. Thus, the results indicate that personalised products and service offerings of the banking is comparatively low and practising service oriented

architecture and digitisation of products and services are low as well. Although, there should be a more focus on the importance of the knowledge of the employees and training given on industry 4.0 activities. The overall maturity level of the Sri Lankan banking sector is 3.668. Thus, it reveals that Sri Lankan banking sector is in the third maturity level “Defined” of the industry 4.0 maturity assessment model. Though results can be interpreted that the processes of the bank are defined and considerable practices for planning and management procedures are employed within the organisation. Limited management of the processes due to constraints on technical and resource needs. Although practise of integration and interoperability concepts of the information exchange within the banking processes are visible. Data driven services were taken to consideration. The vision of Industry 4.0 and the idea for Industry 4.0 roadmap exists thus it is not fully implemented or practised within the banking sector.

Hence, it is important to enhance industry 4.0 performance to the next maturity level of the maturity assessment model to comply with world class performance as well as for continuous improvement. The senior management of the banking sector could gain an understanding of the current progress of their journey to digitization through the results of this study. Thus, the identified weak points and low focus areas must be considered. Therefore, it is important to implement and practise industry 4.0 roadmap for planning and controlling industry 4.0 activities within the banking sector while giving equal importance for all seven-dimension’s performance improvement to provide a satisfied customer service.

## 6. Conclusion

The main purpose and the theoretical contribution of this study is to develop and validate an industry 4.0 maturity assessment model that best fit for the banking sector where it was applied and tested using the banking sector of Sri Lanka.

Hence, this study reveals the importance of enhancing industry 4.0 performance of the banking sector to the next maturity level of the maturity assessment model to comply with world class performance as well as for continuous improvement. Whilst senior management of the banking sector could gain an understanding of the current progress of their journey to digitisation through the results of this study. As the modern customer expectations have evolved with the advancement of technology, they demand more relationship experience that hinges with trust and personalization. Hence, in order to attract, leads as well as to retain customers it is important to understand current performance of industry 4.0 in order to focus on continuous improvement to offer better customer satisfaction.

This study only focusses to assess industry 4.0 maturity of the Sri Lankan banking sector via selected sample of 10 banks with the time and resource constraints. Hence, this can be further extended to measure the industry 4.0 maturity of the other financial institutions available in Sri Lanka in order for a better generalisation of the model.

The overall maturity of the banking sector was calculated in this study, based on the assumption of all seven dimensions are similarly important. Thus, this assessment can be further optimise by studying the importance of each dimension separately when calculating the overall maturity as a future research perspective.

## 7. References

- [1.] Alaeddin Koska, N. G. (2017). Measuring the Maturity of a Factory for Industry 4.0. *International Journal of Academic Research in Business and Social Sciences*, 7(7).
- [2.] Anirudh Agrawal, D. S. (n.d.). Incorporating Industry 4.0 in Corporate Strategy.
- [3.] Anna De Carolis, M. M. (2017). A Maturity Model for Assessing the Digital Readiness of Manufacturing Companies. *IFIP International Federation for Information Processing* (pp. 12-20). Springer.
- [4.] Bradley, A. (2014). *The connected enterprise amaturity model*. USA: Rockwell Automation.
- [5.] Christian Leyh, T. S. (n.d.). SIMMI 4.0 – A Maturity Model for Classifying the Enterprise-wide IT and Software Landscape Focusing on Industry 4.0. *Federated Conference on Computer Science ans Information Systems* (pp. 1297-1302). FedCSIS.
- [6.] Cook, S. (2017). *Selfie banking: is it a reality?* Biometric Technology Today.
- [7.] Diogo Proença, J. B. (2016). Maturity Models for Information Systems - A State of the Art. *Conference on Computer Science* (pp. 1042-1049). Science Direct.
- [8.] Dong-Young Kim, G. G. (2010). E-government maturity model using capability maturity model integration. *Journal of Systems and Information Technology*, 12(3), 230-244.
- [9.] Deloitte (2015). *Industry 4.0 : Challenges and solutions for digital transformation and use of exponential technologies*.
- [10.] Ebru Gökalp, U. Ş. (2017). Development of an Assessment Model for Industry 4.0: Industry 4.0-MM.
- [11.] Ernst Pessl, S. R. (2017). Roadmap Industry 4.0 Implementation Guideline for Enterprises. *International Journal of Science, Technology and Society*, 193-202.
- [12.] Hercko, J., & Hnat, J. (2015). Industry 4.0- new era of manufacturing. *InvEnt 2015*. Demänovská dolina.
- [13.] Hyman, G. (n.d.). *Are You a Banker or a Visionary?* Fintech Focus.
- [14.] (2015). *Industry 4.0 : Challenges and solutions for digital transformation and use of exponential technologies*. Deloitte.
- [15.] Jacapo Brunelli, V. L. (2017). *Five lessons from the Frontlines of Industry 4.0*. Boston Consultant Group.
- [16.] Karl Lichtblau, R. B. (2015). *INDUSTRIE 4.0 READINESS*. Aachen, Cologne: Foundation for mechanical engineering, plant engineering, and information technology.

- [17.] Kermer-Meyer, A. (2017). Industry 4.0 Maturity Assessment. *HANNOVER MESSE*. ResearchGate.
- [18.] KPMG. (2017/2018). *Technology challenges for Dutch banks in the digital era*. Banking Systems Survey KPMG.
- [19.] Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research Issues. *Journal of Industrial Information Integration*, 6, 1-10.
- [20.] Onur Agca, J. G. (2017). *An Industry 4 readiness*. Coventry: International Institute for Product and Service Innovation, University of Warwick.
- [21.] Rajan, S., & Kuder, D. (2017). *How Advances in Machine Intelligence Will Benefit Banks*. FinXtech.
- [22.] Roblek, V., Mesko, M., & Krapez, A. (2016). A Complex View of Industry 4.0. *SAGE Open*.
- [23.] Saurabh Vaidya, P. A. (2018). Industry 4.0 - A Glimpse. *2nd International Conference on Materials Manufacturing and Engineering* (pp. 233-238). Elsevier BV.
- [24.] Schumacher, A., Erol, S., & Sihm, W. (2016). A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises. *Changeable, Agile, Reconfigurable & Virtual Production Conference* (pp. 161-166). Elsevier B.V.

## Integrating smart transportation system for a proposed Smart City: A mapping study

H. K. S. K. Hettikankanama  
Department of Computing & Information  
Systems  
Sabaragamuwa University of Sri Lanka  
Belihuloya, Sri Lanka  
keshalasantuni@gmail.com

S. Vasanthapriyan  
Department of Computing and Information  
Systems  
Sabaragamuwa University of Sri Lanka  
Belihuloya, Sri Lanka  
priyan@appsc.sab.ac.lk

### Abstract

*Smart transportation is playing a vital role as an important building block in smart city, providing solutions to many issues that relate to traffic on the road. This influences safety and quality of living (QoL) the main goals of smart city development. In such case, investigation of different aspects like different functionalities of smart transportation, purposes, research methods and technologies are obligatory. This study presents a literature review on existing researches on smart transportation in smart city to identify the area as well as the future research needs and gaps to be fulfilled. Here it summarizes different perspectives of smart city and smart transportation considered in studies, types of the researches performed, reported problems which are addressed in studies, purposes of deploying smart transportation, reported benefits and problems and technologies which are compatible in this research area.*

**Keywords:** Smart transportation, Smart city, mapping study

### 1. Introduction

Within the most recent 50 years, city measurements have been expanding to an ever increasing extent, everywhere throughout the world. By 2050, 70 % of the populace will live in urban areas [1]. Urban areas can be identified as both spots of opportunities and spots of maladies. Opportunities, since urban areas are places where individuals live and meet, where organizations are settled and schools and colleges are generally situated [2]. Maladies that spread due to city traffic jams, contamination and waste generation have gone more terrible than elsewhere and furthermore the average cost for living is high [3].

It has become a much challenge to public administration and municipalities to create a practical and sustainable urban development considering the need of

both making openings for work and protecting nature, providing individuals in city with the best living conditions [4, 5].

Smart City is viewed as a triumphant urban technique which utilizes technology to uplift the personal satisfaction or QoL in urban area, both enhancing the environmental quality and conveying better service to citizens [6].

Sri Lankans spend a lot of time on roads for travelling. The stringent traffic comes across in slowing the pace of every vehicle, despite public or private [7]. Traffic congestion, air pollution and loss of much needed man hours occur due to daily unmanageable entrance of more than 500,000 vehicles [8]. Some government statistics show that the total vehicle population has increased from 3.39 million in 2008 to 6.33 million in 2015 in Sri Lanka and also states that 65% of the road space is used by 38% of the passengers [9]. Thus, utilization of these is highly needed.

The congestion increases as people shift from public transport towards private transport. Thus, a solution for the high traffic is to improve public transport where buses can carry 40-50 passengers for one bus. It will help in reducing the number of vehicles on the road, thus making usage of road space more efficient [10, 11]. The fact that level of service provided by the public transportation sector is low is the widely accepted opinion in the present Sri Lankan society. To curb this massive, unproductive cost, it is vital to take measures to modernize and improve public and private transport [12].

The ultimate goal of this research is to understand and find out the key challenges and unfilled gaps in smart transportation in Sri Lanka. By identifying those key factors based on previous literature, the objective is to suggest solutions to overcome these barriers and fill the existing gaps between smart transportation and general transportation system in Sri Lanka. This paper provides a literature survey which provides studies of smart transportation in smart city. Purpose of this is to find the gaps and research areas where high concern is needed

and to understand methods techniques used in performing studies in same area.

The rest of this paper flows as following. Section 2 presents the background of smart city, smart transportation and also related research. Section 3 provides a view about the research method applied to perform the mapping study. Results of this mapping study are presented then in Section 4. Section 5 presents discussion and conclusion and present possibilities and directions for future work.

## 2. Background

### 2.1. Internet of Things

Internet of Things or mostly referred to as IoT is an progressive paradigm that visualize not so distant future, where daily assets will contain digital communication transceivers, microcontrollers and appropriate protocols enabling them to communicate with each other and users. [13]. IoT provide range of devices like home devices, vehicles, supervising cameras, sensors and many which give service to public administrations, companies and citizens. This mostly searches application in some different domains, such as automating home, industry automation, medical aids, elderly assistance, intelligence energy management and smart grids, healthcare, traffic management, automotive, and many others [14].

### 2.2. Smart City

It is not practical to provide a common unique definition for “Smart City,” but it is accepted that ultimate objective of this is to provide better use of public resources, by accelerating quality of services given to people, by lowering costs of administration. This aim can be achieved by the deploying an urban IoT [15].

Different smart city aspects can be identified as waste management, structural health of buildings, air quality, traffic congestion, noise monitoring, smart parking, smart lighting and city energy consumption [16].

### 2.3. Smart Transportation

Smart transportation is enabled by number of technologies like IoT, GPS, Wireless Technologies and Sensing Technologies. Applications of smart transportation are commonly wide scoped and complicated as those are working with heterogeneous and dynamic devices and data. Smart Parking systems, intelligent public transportation systems, traffic management systems, smart taxi applications, smart traffic light systems, disaster precaution systems are common functionalities of smart transportation [17].

## 2.4. Related work

In this part it represents a secondary study. A study that is based on analyzing existing research papers (referred to as primary studies). Mapping studies are intended to provide an overview of a topic area and identify whether there are sub-topics where more primary studies are needed. Before acquiring the secondary study presented here, a tertiary study which searched for secondary studies examining the state of the art in smart transportation initiatives in smart city considered [18]. As there is a research gap in Sri Lanka on smart transportation for smart cities the search was done on all smart transportation and smart city related studies regardless of the country. In this, the search string shown in Table 1 is used in three metadata fields (title, abstract and keywords). The search string was applied in the following electronic databases: IEEE Xplore, ACM Digital Library, Springer Link, Emerald Insight, Science Direct and Research Gate.

**Table 1. Search terms of smart transportation for smart city.**

Areas	Search Terms
Smart City	“Smart City”, “IoT and Smart City”, “Digital City”
Smart Transportation	“Smart Transportation”, “Intelligent Transportation Systems”, “Smart Traffic Monitoring”
Review	“systematic literature review”, “systematic review”, “systematic mapping”, “mapping study”, “systematic literature mapping”

Search string: (“smart city”, “IoT for smart city” OR “digital city”) AND (“Smart Transportation” OR “smart traffic monitoring”) AND (“systematic literature review” OR “systematic review” OR “systematic mapping” OR “mapping study” OR “systematic literature mapping”)

As in some databases there were lack and none results on this search string smart transportation and smart city were searched separately. For the tertiary study that looks for secondary studies in Smart Transportation, the search strings shown in Table 2 are used and for secondary studies in Smart Cities, search strings shown in Table 3 are considered. The same six electronic databases were searched returning 388 results. From them 37 was selected by applying the selection criteria and considering abstract, keywords and title. From the left 11 was selected by considering full text. From snowballing and direct searching from researchers and research groups finally 13 studies were selected which are highly related.

**Table 2. Search terms of smart transportation.**

Areas	Search Terms
Smart Transportation	“Smart Transportation”, “Intelligent Transportation Systems”, “Smart Traffic Monitoring”
Review	“systematic literature review”, “systematic review”, “systematic mapping”, “mapping study”, “systematic literature mapping”

Search string: (“smart transportation”, “intelligent transport systems” OR “smart traffic monitoring”) AND (“systematic literature review” OR “systematic review” OR “systematic mapping” OR “mapping study” OR “systematic literature mapping”)

**Table 3. Search terms for smart city.**

Areas	Search Terms
Smart City	“Smart City”, “IoT and Smart City”, “Digital City”
Review	“systematic literature review”, “systematic review”, “systematic mapping”, “mapping study”, “systematic literature mapping”

Search string: (“smart city”, “IoT for smart city” OR “digital city”) AND (“systematic literature review” OR “systematic review” OR “systematic mapping” OR “mapping study” OR “systematic literature mapping”)

### 3. Research method

The research method for this mapping study is defined based on the guidelines given by Kitchenham and Charters [19]. In selection at first current status on studies on smart transportation in smart city is considered. Then secondly the studies are taken through a selection process to retrieve the relevant. Here following selection criteria was used in: (i) defining terms and search strings (ii) selection of sources (iii) defining the inclusion and exclusion criteria (shown in Table 4) and (iv) defining the way of data storing. Thirdly in this selection stage, publications until June 2018 were considered and a total of 388 publications were returned as result as 189 from IEEE Xplore, 73 from ACM Digital Library, 20 from Springer Link, 87 from Science Direct, 13 from Emerald Insight and 6 from Research Gate were found. To extract most relevant, sub selection process was performed on these selected publications.

In the first stage, duplications were removed, retaining 115 publications (reduction of 70%). In the second stage, selection criteria (inclusion and exclusion criteria) on the title, abstract and keywords is applied, resulting in 37 papers (reduction of 68%). In the third stage, considering full text the selection criteria were applied resulting 11 studies (reduction of approximately 70%). As 4th stage, snowballing was performed resulting in 7 papers and by applying selection criteria over title,

abstract and keywords, 3 papers remained (reduction of 57%) and on them selection criteria were applied considering the full text and only one paper selected (overall 86% reduction from snowballing).

**Table 4. Inclusion and exclusion criteria.**

Inclusion criteria	
IC1	The study discusses about smart transportation in smart cities
Exclusion criteria	
EC1	The study which does not contains an abstract
EC2	The study which is only published as an abstract
EC3	Study which is not written using English Language
EC4	The study which is an older version or previous version of an study that already considered
EC5	The study which is not a primary study, such as workshops, editorials, tutorials and summaries of keynotes

Finally, from 12 papers which came up till the 5th stage, direct search was done on publications authored by the researchers and research groups involved in these studies which resulted 4 papers and by applying all selection criteria 1 most appropriate paper was selected. At the end of this searching process totally 13 papers were retrieved as the final result including, 11 from the sources, 1 from snowballing and 1 from direct searching researchers and research groups. Summary of the selections in selection stages are shown in Table 5 clearly emphasizing the progressive elimination of studies in the selection process. Table 6 presents bibliographic references of selected studies. Unique identifier (#id) is given to each study in this table and those identifiers are used to represent those studies in the rest of this paper.

#### 3.1. Limitations of this mapping study

This section presents about the limitations of method.

1. As major steps of this mapping was performed by one author a subjectivity could happen when selecting and extracting data.
2. As only few specific databases were selected to perform searches which may lead to loss some important studies which are in other databases. But as snowballing and direct search was performed using references of selected studies this limitation was addressed up to some extent.
3. When performing direct search on researchers and research groups that information was taken only

from the selected studies there may be a chance of missing some important studies due to that.

- When search is performed on some databases as a filter option the year range was given (2010-2018). This will create a huge chance of missing some related and well performed studies which were done before the lower year margin.

#### 4. Results and discussion

In this section, it presents the results retrieved for the research questions. A table is used here with the id of the paper, its bibliographic reference, and including the facets of the classification schema aforementioned. This form was used to extract the answers for each research question.

As Figure 1 shows, here the selected studies are from 2014 onwards that is very recent. As this topic is subjected to many researches this time it is much better to choose most recent studies as the relevant technologies also developing rapidly. The selected studies were published in tow main vehicles: Journals and Conferences. Journals have been the main forum for presenting Smart Transportation in Smart Cities, encompassing 61.5% (8 studies in 13). Conferences are the publication forum of 38.5% (5 out of 13).

Table 7 shows the distribution of the studies according to the research focus from the Smart City perspective. Most of the papers address a specific aspect of smart city (11 out of 13 – 84.6%) for providing smart transportation facilities. With respect to Smart Transportation, as Table 8 shows, the great majority of the studies discussed aspects related to the Smart Transportation process as a whole or focusing on one of its activities.

Table 9 presents the distribution according to the research types. Nine studies out of all considered 13 (69.2%) studies propose some solution for Smart Transportation in Smart City. In addition to presenting a solution proposal, most of them also discuss some sort of evaluation: 6 studies also report an Evaluation Research (46.2%), while 4 studies report a Validation Research (30.8%). The 6 studies with an Evaluation Research discuss a practical implementation and its consequences. Three studies (23.07%) present only a solution proposal.

Table 10 shows the distribution over the years considering the problems reported related to smart transportation in smart cities. Table 11 presents the distribution over the years considering the purposes of having Smart Transportation. It illustrates that “Prevent unneeded driving time and fuel consumptions” (9 studies – 69.2%) and “Reduce vehicle waiting time on traffic jam” (5 studies – 38.5%) have the largest focus.

Table 12 shows the types of transportation on which the studies mainly focus on. Public transportation is discussed in 4 studies (30.8%). When considering public transportation there were 6 main purposes separately as bus trip monitoring , bus fee online payment, bus arrival time estimation, accident reporting and safety, alcohol detection in driver, Seat availability. On the other hand, 9 studies out of 13 (69.2%) have discussed private transportation. The technologies used to implement Smart Transportation in Smart Cities are shown in Table 13.

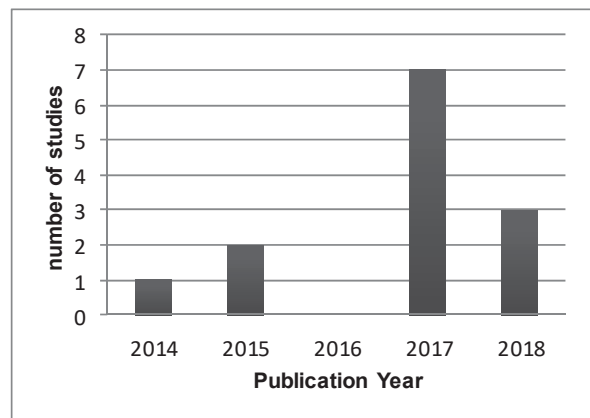


Figure 1. Distribution of studies over the years.

Table 14 shows how the main aforementioned benefits are perceived in the selected studies.

- Time Saving:** Smart Transportation system helps to estimate traffics and then drivers can choose alternative paths and also some systems automatically suggest shortest paths thus it reduce the time spent on roads unwantedly. And also smart public transportation systems help to view bus trips thus people do not have to spend time on a queue.
- Cost Reduction:** These systems will reduce the driving time thus fuel consumption is less it will lead to save money.
- Eco Friendly:** As the driving time and waiting time on traffic are reduced CO2 emission by burning fuel will reduce.

The most highlighted and discussed benefit is “Time Saving” (76.9%).

Table 15 shows how the main problems aforementioned are perceived in the selected studies.

- Security Breaches and Attack Vulnerability:** These systems contain sensitive information of public. Mostly all stored information is recorded in cloud architecture thus it has to be more secure to prevent from breaches if not public data will be compromised.



B. Digital Divide: Here it talks about public acceptability towards the system.

C. Limited Resources when Comparing to Increasing Number of Transportation Assets

**Table 5. Results of the selection stages.**

Stage	Criteria	Analyzed Content	Initial no of studies	Final no. of studies	Reduction (%)
1 <sup>st</sup>	Remove duplicates	Title, Abstract, Keywords	388	115	70.4
2 <sup>nd</sup>	IC1, EC1, EC2, EC5	Title, Abstract, Keywords	115	37	67.8
3 <sup>rd</sup>	IC1, EC4 and EC3	Full text	37	11	70.3
4 <sup>th</sup> (a)	Snowballing, EC5	Title, Abstract, Keywords	7(added by snowballing)	3(added by snowballing)	57.1
4 <sup>th</sup> (b)	Snowballing, IC1	Full text	3(added by snowballing)	1(added by snowballing)	66.7
5 <sup>th</sup>	Research group, EC4	Full text	6(added by research groups)	1(added by research groups)	83.3
Final result			388 (sources) + 7 (snowballing) + 6 (research groups) = 401	11 (sources) + 1 (snowballing) + 1(research groups) = 13	96.75

**Table 6. Selected studies.**

ID	Bibliographic reference
#1	Zanella, A., Bui, N., Castellani, A., Vangelista, L. And Zorzi, M., 'Internet Of Things For Smart Cities - IEEE Journals & Magazine' (Ieeexplore.Ieee.Org, 2014)
#2	Matthew.N.O. Sadiku, Adebawale. E. Shadare, & Sarhan M. Musa. Smart Transportation: A Primer. (2017, March).
#3	Murad, D. F., Abbas, B. S., Trisetjarso, A., Suparta, W., & Kang, C. Development of smart public transportation system in Jakarta city based on integrated IoT platform - IEEE Conference Publication. (2018, April 30).
#4	Pham, T. N., M. T., Nguyen, D. B., Dow, C., & Deng, D. A Cloud-Based Smart-Parking System Based on Internet-of-Things Technologies - IEEE Journals & Magazine. (2015, September 09)
#5	Gawade, P., & Meeankshi, A., Prof. IOT Based Smart Public Transport System. (2017, July).
#6	Sherley,J., & Somasundareswari,D. INTERNET OF THINGS BASED SMART TRANSPORTATION SYSTEMS. (2015, October).
#7	Talukder, M. Z., Towqir, S. S., Remon, A. R., & Zaman, H. U. An IoT based automated traffic control system with real-time update capability - IEEE Conference Publication. (2017, December 14).
#8	Latif, S., Afzaal, H., & Zafar, N. A. Intelligent traffic monitoring and guidance system for smart city - IEEE Conference Publication. (2018, April 26).
#9	Vidyasagan, S., Devi, S., Varma, A., A., & Charan, H. 'A low cost IoT based crowd management system for public transport - IEEE Conference Publication. (2017, November 24).
#10	Jiménez, A., García-Díaz, V., & Anzola, J. Design of a System for Vehicle Traffic Estimation for Applications on IoT. In <i>Proceedings of the 4th Multidisciplinary International Social Networks Conference (MISNC '17)</i> . ACM, New York, NY, USA, Article 15, 6 pages. (2017, July 17).
#11	Murad, D. F., Meyliana., Hidayanto, A. N., & Prabowo, H. IoT for Development of Smart Public Transportation System: A Systematic Literature Review. <i>International Journal of Pure and Applied Mathematics</i> . (2018).
#12	Wu, H., & Horng, G. Establishing an Intelligent Transportation System With a Network Security Mechanism in an Internet of Vehicle Environment - IEEE Journals & Magazine. (2017, September 14).
#13	Kim, T., Ramos, C., & Mohammed, S. Smart City and IoT Elsevier B.V. <i>Future Generation Computer Systems</i> . (2017, July).

**Table 7. Research focus from the Smart City perspective along the years.**

Research Focus	2014	2015	2016	2017	2018	Total (%)
City Aspect Converted to Smart	#1	#4 #6		#2 #5 #7 #9 #10 #12	#3 #8	11(84.6)
Smart Mobility Objectives		#4 #6		#2 #5 #7 #9 #10 #12	#3 #8	10 (76.9)
Techniques Used	#1			#2 #5 #7 #9 #10 #12	#8	8 (61.5)
General				#13	#11	2 (15.4)

**Table 8. Research focus from the Smart Transportation perspective along the years.**

Research Focus	2014	2015	2016	2017	2018	Total (%)
Smart Traffic Control	#1	#6		#2 #13 #7 #10 #12	#8 #11	9 (69.2)
Smart Parking		#4 #6				2 (15.4)
Dynamic Traffic Light Sequence	#1			#12		2 (15.4)
Vehicle Notification Systems				#5 #9 #12	#3 #11	5 (38.5)
Automatic road Enforcement	#1			#10		2 (15.4)
Smart Ticketing				#5 #9	#11	3 (23.0)
Smart Policing.				#5		1 (07.7)

**Table 9. Distribution over research type**

Research Type	2014	2015	2016	2017	2018	Total (%)
Solution Proposal	#1	#4 #6		#5 #7 #9 #10 #12	#8	9 (69.2)
Validation Research				#2 #13	#3 #11	4(30.8)
Evaluation Research		#4		#5 #7 #9 #10	#8	6 (46.2)

**Table 10. Distribution of related problems.**

Problem	2014	2015	2016	2017	2018	Total (%)
Time and money spent uselessly on roads because of traffic	#1	#4 #6		#2 #5 #9 #13 #7 #10	#3 #8 #11	12 (92.3)
Digital divide				#5	#3	2 (15.4)
Security aspects				#12	#3	2 (15.4)
Accuracy				#5 #7 #12	#3	4 (30.8)
Limited resources in automating		#4 #9				2 (15.4)

**Table 11. Distribution of purposes.**

Purpose	2014	2015	2016	2017	2018	Total (%)
Prevent unneeded driving time and fuel consumptions	#1	#4 #6		#5 #7 #10 #12	#3 #8	9 (69.2)
Validate proposed researches				#2 #13	#11	3 (23.1)
Reduce vehicle waiting time on traffic jam	#1	#6		#7 #10	#8	5 (38.5)
Answer uncomfortable-ness of public bus transportation				#5 #9	#3	3 (23.1)
Automatic shortest path suggestions		#4 #6		#10	#8	4 (30.8)

**Table 12. Distribution types of transportation.**

Type	2014	2015	2016	2017	2018	Total (%)
Public Transportation Four papers out of 13 (30.8%)	bus trip monitoring			#9	#3 #11	3 (23.0)
	bus fee online payment			#5		1 (7.7)
	bus arrival time estimation			#5	#3	2 (15.4)
	accident reporting and safety			#5		1 (7.7)
	alcohol detection in driver				#5	
Private Transportation Nine papers out of 13 (69.2%)	Seat availability			#5 #9		2 (15.4)
	shortest path suggestion	#1	#4 #6	#10	#8	5 (38.5)
	traffic jam reporting	#1	#6	#2 #7 #10 #13 #12	#8	8 (61.6)
smart parking			#4 #6			2 (15.4)

**Table 13. Distribution of technologies used.**

Technology	2014	2015	2016	2017	2018	Total (%)
RFID		#4				1 (7.7)
Sensors		#6		#2 #5 #7	#3 #8	6 (46.2)
Arduino				#7 #9 #10		3 (23.1)
GSM/GPS				#2 #5		2 (15.4)
WSN		#4				1 (7.7)
Mobile Software					#3	1 (7.7)
Study does not address this issue	#1			#13 #12	#11	4 (30.8)

**Table 14. Distribution of identified benefits.**

Benefits	2014	2015	2016	2017	2018	Total (%)
A	#1	#4 #6		#2 #5 #7 #10	#3 #8 #11	10 (76.9)
B	#1	#4 #6		#5		4 (30.8)
C	#1	#6				2 (15.4)

**Table 15. Distribution of identified problems.**

Technology	2014	2015	2016	2017	2018	Total (%)
A		#4		#2	#8	3 (23.1)
B					#3	1 (7.7)
C		#4		#7#12 #13	#3 #8 #11	7 (53.8)

## 5. Conclusion

This study presents a systematic mapping on factors influencing Smart transportation system in smart city. Nine research questions were defined and addressed investigating the following facets: (i) distribution of the selected studies over the years; (ii) research focus from the Smart City perspective; (iii) research focus from the Smart Transportation perspective; (iv) research type; (v) reported problems; (vi) purposes to employ Smart Transportation in Smart Cities; (vii) types of Transportation typically considered in converting to smart; (viii) technologies used in implementing smart transportation in smart city; (ix) main conclusions (benefits and problems) reported on the implementation of smart transportation in smart cities.

In this context, we highlight the following conclusions: (i) Smart transportation in smart cities is a mostly considered topic; (ii) the major problem in society to have this kind of system is to save time and cost that uselessly spent for transportation; (iii) prevent unneeded driving time and fuel consumptions is the main purpose of having smart transportation system in smart cities; (iv) there is a great concern with public transportation in order to convert in to smart transportation and also private transportation in smart parking, traffic monitoring and automatic path suggestion; (v) advanced technologies that used to provide smart transportation in smart cities are sensors, Arduino, GSM/GPS, WSN, Mobile software. The main contribution of this work is

on making evident some aspects associated to smart transportation in smart cities, in order to drive future research in this area. Most of the studies consider about providing smart parking, shortest path suggestion systems, and smart traffic light systems. As a future work many of them point out the need of smart public transportation system which can overcome the issues related to security breaches, limited resources, and low public awareness on new technologies.

## 6. Acknowledgement

Authors of this study thank Sabaragamuwa University of Sri Lanka for facilitating the research.

## 7. References

- [1] Dameri, R.P., "Comparing smart and digital city: initiatives and strategies in Amsterdam and Genoa. Are they digital and/or smart?", In: Dameri, R.P., Rosenthal-Sabroux, C. (eds.) *Smart City. How to Create Public and Economic Value with High Technology in Urban Space*, Springer, Heidelberg, 2014, pp. 45–88.
- [2] Krajzewicz, D., Erdmann, J., Behrisch, M. and Bieker, L., "Recent Development and Applications of SUMO - Simulation of Urban Mobility", in *International Journal on Advances in Systems and Measurements*, December 2012, pp. 128-138.
- [3] Wegner, A., Piorkowski, M., Raya, M., Hellbruck, H., Fischer, S. and Hubaux, J.P., TraCI: "An Interface for Coupling Road Traffic and Network Simulators", in *Proceedings of the 11th Comm. and Networking Simulation Symp. (CNS 08)*, April 2008.

[4] Pan, J., Khan, M., Popa, I., Zeitouni, K., and Borcea, C., "Proactive vehicle re-routing strategies for congestion avoidance", in *Proceedings of the 2012 IEEE 8th International Conference on Distributed Computing in Sensor Systems*, Zhejiang, China, May 2012, pp. 16-18.

[5] Hannan, M. A., Mustapha, A. M., Hussain, A., and Basri, H., "Intelligent Bus Monitoring and Management System", in *Proceedings of the World Congress on Engineering and Computer Science 2012 Vol II WCECSf 2012*, San Francisco, USA, October 2012, pp. 24-26.

[6] Hall, P., "Creative cities and economic development", *Urban Stud.* 37, 2000, 633–649.

[7] Hiranand, V., "Sri Lanka's public transport system – How do we travel?", *Pulse*, 2016.

[8] Sundaytimes.lk. (2016). "Sri Lanka to introduce advanced public transportation system" *The Sunday Times Sri Lanka*.

[9] Samarakoon G.S., Weerasinghe W.M.J.P., Kosgoda K.M.D.L. and Vimarshani H.K.M., "Improving public transport as a solution for traffic congestion: Attitudes first", *Island.lk*, 2017

[10] Nafeel, N., "Towards a better public transport service", *Daily News*, 2016.

[11] Huang, Y., Wang, L., Hou, Y., Zhang, W., and Zhang, Y., "A prototype IOT based wireless sensor network for traffic information monitoring", in *International journal of pavement research and technology*, March 2018, pp. 146-152.

[12] Aishwarya, S. R., Rai, A., Prasanth, M. A., and Savitha, S. C., "An IoT Based Accident Prevention & Tracking System for Night Drivers", 2015 pp. 3493–3499.

[13] Atzori, L., Iera, A., and Morabito, G., "The internet of things: A survey," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 54, no. 15, October 2010, pp. 2787–2805.

[14] Bellavista, P., Cardone, G., Corradi, A., and Foschini, L., "Convergence of MANET and WSN in IoT urban scenarios," in *IEEE Sensors Journal*, vol. 13, no. 10, October 2013, pp. 3558–3567.

[15] Cuff, D., Hansen, M., and Kang, J., "Urban sensing: Out of the woods," in *Communication of the ACM*, vol. 51, no. 3, March 2008, pp. 24–33.

[16] Zanella, A., Bui, N., Castellani, A., Vangelista, L. And Zorzi, M., "Internet of Things for Smart Cities" in *IEEE Journals & Magazine*, 2014.

[17] Mahmood, Z., "Smart cities development and governance framework", in *Springer International Publishing AG, part of Springer Nature*, 2018.

[18] Landis, R., and Koch, G.G., "The measurement of observer agreement for categorical data", *Biometrics*, 33, 1977, pp. 159-174

[19] Kitchenham, B., Charters, S., "Guidelines for performing Systematic Literature Reviews in Software Engineering",

School of Computer Science and Mathematics Keele University and Department of Computer Science University of Durham, UK, 2007

## Modelling & simulation of the hydration process of cement for rapid concrete constructions

S. M. D. T. Sandaruwan  
 Department of Chemical & Process  
 Engineering,  
 Faculty of Engineering,  
 University of Peradeniya, Sri Lanka  
 smdtsandaruwan@eng.pdn.ac.lk

A. Pallegedera  
 Faculty of Engineering,  
 University of Peradeniya, Sri Lanka  
 achalap@pdn.ac.lk

### Abstract

Hydration is series of exothermic reactions which occurs between cement and water. It is difficult to predict the exact behaviour of hydration process. It was modelled using Affinity Hydration model and was simulated using finite element approach and which was then validated with realistic parameters. Time dependent simulation has been carried out for various geometries to non-isothermal conditions inside the model. Modelling and simulation was performed in both polystyrene insulation and wooden insulation for different environmental conditions. Temperature, degree of hydration and rate of hydration was obtained using simulation and was verified with data from the realistic experimental data from literature.

**Keywords:** Modelling & simulation, cement hydration Affinity Hydration model

### 1. Introduction

Concrete is a widely use building material in the world for centuries of years. It consists with OPC (Ordinary Portland Cement), Water, Coarse and fine aggregates. Aggregates are not reactive materials while OPC and water reacts with each other. These reactions are called cement hydration reactions (CHR). OPC contains Tri Calcium Aluminate (C3A), Tri Calcium Silicate (C2S), Di Calcium Silicate (C2S) and Tetra Calcium Aluminoferrite (C4AF) as coarse and active elements.

Hydration reactions are exothermic. Therefore, temperature rise of concrete can be experienced while hydration take place. It will influence properties of concrete such as cracks and creeping.

Actual chemical-physical composition of concrete is difficult to determine. Therefore, predictions of hydration using kinetics are not possible and various empirical models were developed for the very purpose. Affinity model, Kinetic model and reactive porous model are frequently used relationships for concrete hydration calculations [1, 2].

The input parameters of Affinity model are relatively less than that of other relative models.

Therefore, Affinity hydration model is a common method used for modelling Concrete Hydration [1]. Temperature of concrete have rapid and non-linear variation during Hydration time and afterwards. Therefore, non-isothermal conditions must be used in analysing cement hydration.

The affinity hydration equation is quite complex and non-linear in nature. Therefore, drawing an analytical solution is highly unlikely due to the intricate nature of the problem. Therefore, a numerical method for simulating is proposed, which is reliable in this type of scenarios and cases to be characterized and analysed. COMSOL Multi-physics is widely accepted by researchers to be used for Simulations in Finite Element Method (FEM) approach. Hydration of Cement was analysed for different conditions using FEM COMSOL Multi-physics software.

### 2. Theory

#### 2.1. Hydration

Hydration was calculated by Affinity Hydration model. Rate of Hydration is called as the Chemical Affinity of the concrete. It can be used to analyse all steps in concrete Hydration.

$$\frac{d(DoH)}{dt} = \dot{A}_r(DoH)$$

DoH means degree of Hydration. It can be obtained using following equation

$$DoH = \frac{m_{hydrated}}{m_{total}}$$

The equation to Affinity at 25°C ( $\dot{A}_{25}$ ) was proposed by [3] and it was developed by [2] as,

$$\dot{A}_{25} = \beta_1 \left( \frac{\beta_2}{DoH_{\infty} + DoH} \right) (DoH_{\infty} - DoH) \exp \left( -\eta \frac{DoH}{DoH_{\infty}} \right)$$

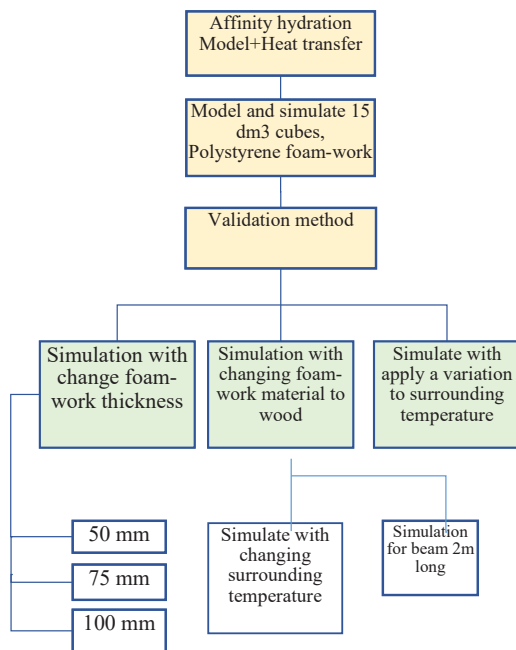


Figure 1. Methodology.

$B_1$  and  $B_2$  are depended upon cement properties,  $DoH$ ,  $DoH_\infty$  and  $\eta$  represents degree of hydration, ultimate hydration degree and micro diffusion of free water through formed hydration products correspondingly.

To calculate non-isothermal affinity of cement, Arrhenius equation can be used.

$$A_T = A_{27} \exp \left[ \frac{E_a}{R} \left( \frac{1}{298.15} - \frac{1}{T} \right) \right]$$

$E_a$  is given for activation energy of the Hydration.  $R$  is the universal gas constant

## 2.2. Heat transfer

The following general form of heat equation is used for modelling heat-transfer in COMSOL Multi-physics. Heat transfer of convection, conduction and radiation are represented in different usual terms defined in physics.

$$V\rho C_p \frac{\partial T}{\partial t} + \nabla \cdot (k \nabla T) + \nabla \cdot (k_{eff} \nabla T) \times V = QV$$

## 3. Method

Simulations were performed to  $14 \text{ dm}^3$  cube and  $2 \text{ m} \times 24 \text{ cm} \times 24 \text{ cm}$  beam which were surrounded by Polystyrene and Wooden Foam-work for different foam-work thickness. The methodology and approach illustrated as schematic block modules in Figure 1.

According to Figure 1, the method was validated initially. Then the proposed method applied to various geometrical and environmental conditions as illustrated in Figure 1.

Value of specific heat capacity of the concrete was used as  $870 \text{ J/kg.K}$  and value of thermal conductivity of concrete was used as  $1.8 \text{ W/m.K}$ , [1]. Potential hydration energy of concrete is  $518.37 \text{ J/g}$  [4].  $1 \text{ m}^3$  of concrete was designed to create  $420 \text{ kg}$  of binder inside it. Binder contains about 45 % of non-reactive materials. Therefore, effective material inside the  $1 \text{ m}^3$  of concrete is around  $230 \text{ kg}$ .

The parameters in affinity hydration model were used as  $B1 = 0.0007 \text{ 1/s}$ ,  $B2 = 6.0e-5$ ,  $\eta = 6.1$ ,  $DoH_\infty = 0.85$  and  $E_a = 38.3 \text{ kJ/mol}$ . [1]

Density of polystyrene is at  $28 \text{ kg/m}^3$  [5]. Specific heat capacity is about  $1300 \text{ J/kg.}^\circ\text{C}$  and thermal conductivity is about  $0.033 \text{ W/m.K}$  [6]. Density of wood is about  $765 \text{ kg/m}^3$  heat capacity is about  $1760 \text{ J/kg.}^\circ\text{C}$  and Thermal conductivity is about  $0.082 \text{ W/m.K}$  [6]

In COMSOL Multi-physics software, Affinity hydration model was applied as a Domain ODE.

## 4. Results and discussion

### 4.1. Validation

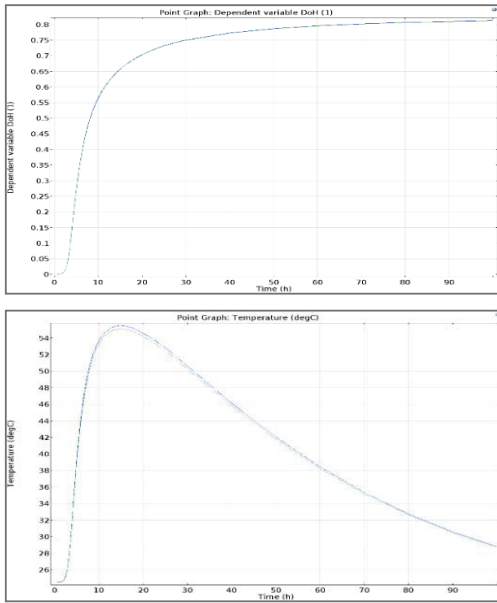
Initially, simulations were performed to  $14 \text{ dm}^3$  cube surrounded with  $100 \text{ mm}$  thickness polystyrene foam-work. The obtained Temperature variation and Hydration Curve of cement within first  $100 \text{ hours}$  are shown in Figure 2. Maximum temperature of concrete cube was obtained as  $54.8 \text{ }^\circ\text{C}$  at  $15.1 \text{ hours}$ . Those results were similar to obtained results in [1].

Temperature and Hydration of concrete cube at different times was shown in Figure 2.

Figure 3, illustrated hydration at different stages of an inside plane of concrete cube insulated with  $100 \text{ mm}$  polystyrene foam-work. Top figure shows hydration after  $1 \text{ hour}$ . Hydrated product percentage of cube is around  $0.0062\%$ . Second figure shows hydration after  $3 \text{ hours}$ . Around  $4\%$  of cement mass got hydrated according to that  $26-27\%$  of cement got hydrated after  $5 \text{ hours}$  to  $100 \text{ hours}$ . It was illustrated in third figure. Eighty one percent of cement got hydrated after  $100 \text{ hours}$  which shown in Figure 2.

Initially, Hydration in all areas of concrete cube is same. But, due to heat generation during hydration temperature gets increased. Heat in middle of concrete cube is more difficult to transfer outside, than heat in outer side of concrete cube Therefore, cement inside concrete cube quickly hydrated.

Rate of hydration is reasonable for hydration. In Affinity hydration model, affinity represents rate of hydration.



**Figure 2. Temperature variation and hydration within 100 hours in validation model.**

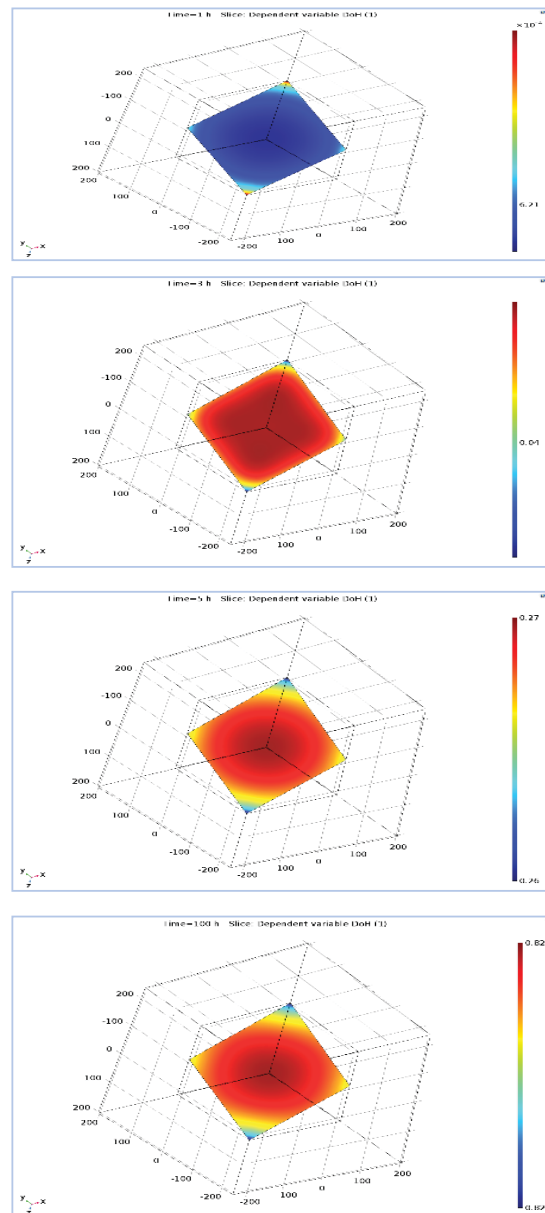
Initially Affinity at all areas of cube is same. But due to heat generation of hydration, temperature increase. Middle of cube shows highest temperature at a time. According to Arrhenius equation, the affinity at middle of concrete cube is highest.

However, after certain time considerable cement in middle of concrete cube got hydrated. Then Affinity at the middle becomes low. This variation is illustrated in Figure 4.

Figure 4 illustrated affinity of concrete cube at different stages. Figure at left side of top shows Affinity after 1 hour. It is around  $4 \times 10^{-6}$  (1/s). Figure at right side of top shows Affinity after 5 hours. According to it, Affinity is between  $3.15 \times 10^{-6}$  (1/s) to  $3.24 \times 10^{-6}$  (1/s). Figure at bottom left side shows Affinity after 10 hours. In this duration, pattern of Affinity, changed. Before 10 hours elapsing, Affinity at middle was the highest. But after 10 hours Affinity at middle is shown to be the lowest.

Affinity is around  $7.62 \times 10^{-6}$  (1/s). Bottom right side figure illustrated affinity after 15 hours. Value of affinity is between  $3.36 \times 10^{-6}$  and  $3.39 \times 10^{-6}$ .

Affinity variation during 100 hours is illustrated in Figure 5. It begins at zero degrees and rapidly goes high.



**Figure 3. Hydration at different stages of concrete cube.**

Maximum value of Affinity is  $3.6 \times 10^{-5}$  (1/s). Then it lower rapidly at first, after it become lower slowly.

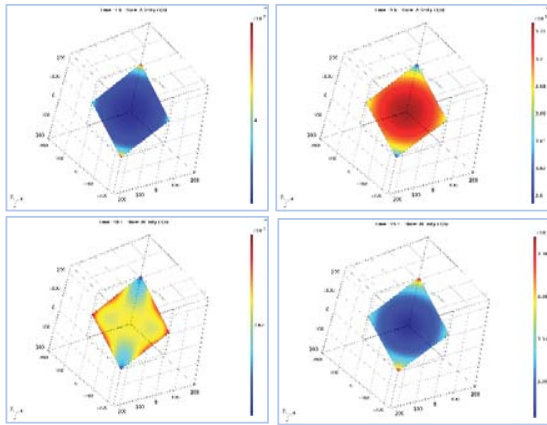


Figure 4. Affinity at different stages.

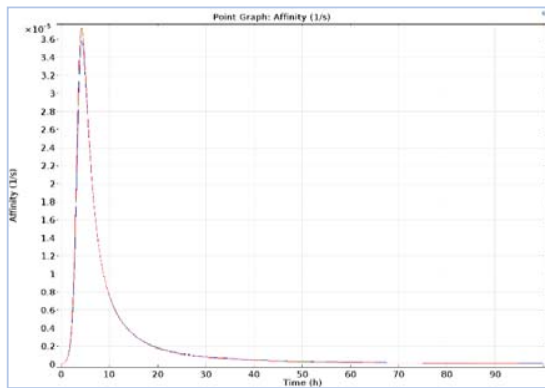


Figure 5. Affinity variation in 100 hours.

#### 4.2. Changing form work thickness

For this purpose, 50 mm, 75 mm and 100 mm used as foam-work thickness. Highest temperature, respective time and hydration are illustrated in Table 1.

Table 1. Maximum temperature, respective hydration and time.

Thickness	Hydration	Temperature	Time (h)
50 mm	0.60	325.33 K	12
75 mm	0.65	327.33 K	14.6
100 mm	0.66	328.66 K	15.1

From results obtained, maximum foam-work thickness gave highest maximum temperature and highest respective hydration. Increase of form work thickness resists transferring heat to outside. Therefore, middle temperature becomes high. The temperature variation for 100 hours related to above thickness illustrated in Figure 6.

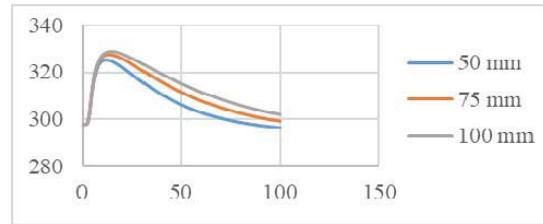


Figure 6. Temperature variation for different thickness, Polystyrene foam-work.

Thermal stress leads concrete to cracks with time. Standard deviation of temperature gives idea about temperature variations inside concrete cube. Low temperature variation means low thermal stress inside concrete cube.

Standard deviation of Degree of Hydration gives a variation of Degree of Hydration at a certain time. Low standard deviation of DoH means uniformity of Hydration is high.

Table 2 illustrated standard deviation of Temperature and respective Degree of Hydration at maximum temperature. Standard deviation of obtained results were calculated considering the results of centre of the cube, results at middle of an edge and results of an edge was used for respective calculations.

Table 2. Standard deviation of temperature and DoH at maximum temperature.

Thickness	Standard Deviation	
	DoH	Temperature
50 mm	0.0089	1.75
75 mm	0.0072	1.52
100 mm	0.0063	1.35

Form work at 50 mm thickness gave high Standard deviation of Temperature and Degree of Hydration at maximum temperature. It means variation of Temperature and Degree of Hydration is comparatively high in concrete cube insulated with 50 mm form work.

#### 4.3. Changing form work material

Although polystyrene is a good insulation material, wood form work is mostly used material in foam-work material in local applications.

Comparing to polystyrene, wood is high in density, and a high thermal conductive material. Heat capacity of wood is also higher than that of polystyrene.

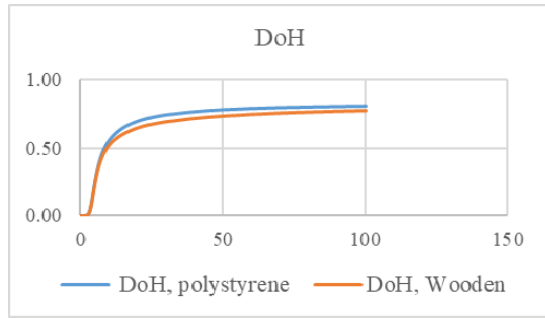
Higher density increases mass of structure. Therefore, high density, high heat capacitive material can store more thermal energy than low density, low heat capacitive material. That means wood can store more thermal heat than Polystyrene.

Higher heat conductivity also leads to higher heat flux. Therefore, wooden foam-work let high thermal energy to transfer outside from concrete cube.



The simulations were executed for concrete cube insulated with wooden foam-work; 50 mm, 75 mm and 100 mm are used as foam-work thicknesses.

Compared to polystyrene form-work, wooden foam-work gives low values for maximum temperature and respective degree of hydration. This is due to above mentioned reasons and characteristics. Due to high heat flux, temperature inside the cube is low.



**Figure 7. DoH of polystyrene foam-work and wooden foam-work.**

Figure 7 illustrated Degree of hydration of concrete increments from the centre of cube insulated with 100 mm thickness foam-work made up with polystyrene and wooden material.

According to Figure 7 concrete cube is insulated with polystyrene foam-work observed to be hydrated quickly. According to Arrhenius equation high temperature leads to high affinity of concrete. The impact of temperature for hydration is considerable and hence experiential.

#### 4.4. Change surrounding temperature

Initially, Surrounding temperature was applied as 20 °C. After that, it was changed to 40 °C. Polystyrene and Wooden foam-works were used. Form work thickness was used as 100 mm.

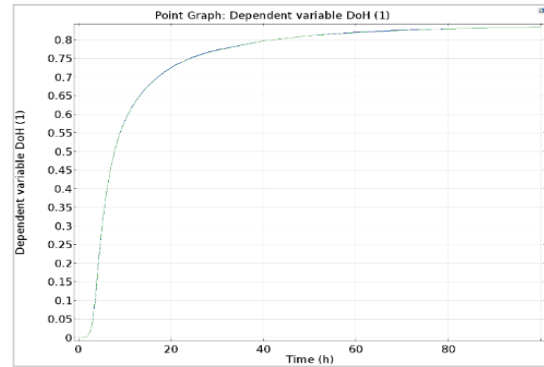
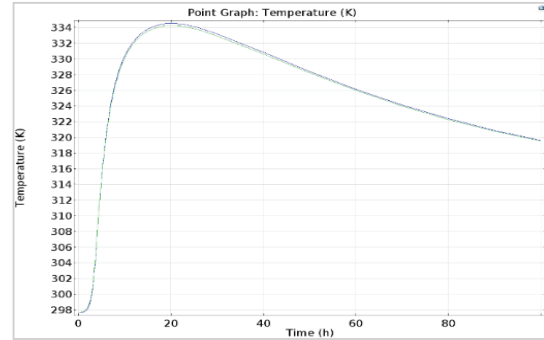
The hydration is rapid compared to 20°C because Temperature difference of outward heat flux is small when surrounding temperature is at 40°C.

**Table 3: Maximum temperature, respective DoH and time for wooden foam-work**

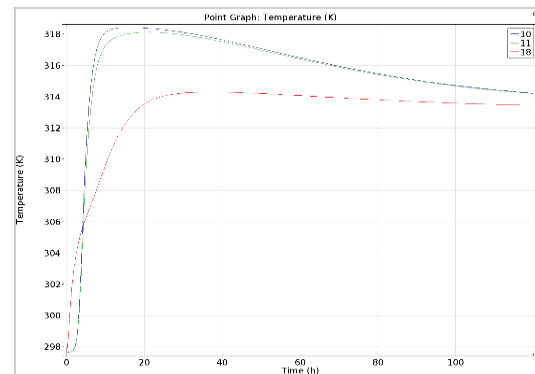
Thickness	Hydration	Temperature	Time (h)
50 mm	0.46	313.31 K	8.2
75 mm	0.50	316.2 K	9
100 mm	0.51	316.62 K	9.4

Hydration curve and Temperature profile within 100 hours of 14 dm<sup>3</sup> concrete cube insulated with 100 mm thickness polystyrene foam-work, when surrounding temperature is around 40 °C was

illustrated in Figure 8. The top graph of Figure 8, illustrates the temperature variation and bottom graph illustrates the hydration. According to Figure 8, maximum temperature reaches 61 °C in polystyrene insulated model when surrounding temperature is at 40°C. After 100 hours' time Degree of Hydration reaches to 0.83.



**Figure 8. Hydration curve and Temperature profile of 100 hours when surrounding temperature is 40 °C, Polystyrene foam-work.**



**Figure 9. Degree of Hydration during 100 hours, wooden foam-work for surrounding temperature at 40 °C.**

According to Figure 9, maximum temperature was obtained is 45 °C where, 81% of cement was hydrated in 100 hours and is given according to Figure 10.

The inside temperature is high when surrounding temperature is high on all two models. Hydration is quite faster when surrounding temperature is high. Degree of hydration and temperature profiles of wooden foam-work insulation that was in at  $40^{\circ}\text{C}$  surrounding temperature; is illustrated in Figure 9.

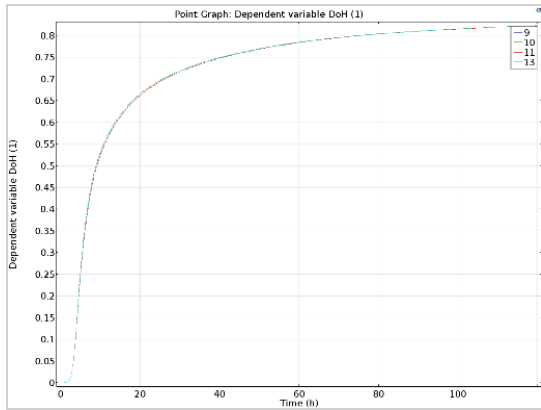


Figure 10: Temperature variation during 100 hours, wooden foam-work surrounding temperature at  $40^{\circ}\text{C}$ .

#### 4.5. Variation of surrounding temperature

Environmental temperature is not constant and is difficult to predict. Environment temperature rises during the daytime and becomes low at nighttime. To illustrate the effect of this temperature change, outside temperature of foam-work was applied to model as sinusoidal variation.

Initial variation considered as  $20^{\circ}\text{C}$ . Sinusoidal variation which has magnitude of  $5^{\circ}\text{C}$  was considered. Both positive and negative Sinusoidal variation were considered. The period of recurrence was considered as 1 day.

Figure 11 illustrated the temperature variation of middle of the concrete cube which is insulated with wood form-work. Red colour line shows the temperature of the centre of cube when there is a positive sinusoidal variation. Blue colour line shows the temperature variation of centre when the outside temperature remain constant at  $20^{\circ}\text{C}$ . Ash colour line shows temperature variation of centre when there is negative sinusoidal variation present.

Initially, outside temperature is kept at  $20^{\circ}\text{C}$ . In positive sinusoidal variation temperature was high. After 6 hours, outside temperature has been reached to  $25^{\circ}\text{C}$ . Then it became significantly low. After 12 hours it became  $20^{\circ}\text{C}$ . After 18 hours it reached its global minimum. Value of temperature at that state is  $15^{\circ}\text{C}$ .

However, in Figure 11, temperature of the centre is kept higher than that of the outside temperature variation condition until 1<sup>st</sup> 20 hours of time taken.

Increment of temperature increases affinity of the model that had positive sinusoidal surrounding

temperature variation. Therefore, heat generation became higher than other models. Due to that reason temperature became comparatively higher. Therefore, Affinity became excessively higher. This happened until considerable amount of cement was hydrated.

However, after one day, there is insufficient amount of cement to increase Affinity, higher. Therefore after one-day, temperature of concrete cube varied according to the surrounding temperature conditions. Magnitude of temperature variation in concrete cube is observed to be lower than the magnitude of temperature variation in surrounding.

#### 4.6. Simulation of 2-meter long beam

For, modeling and simulation 2m long beam was considered. Cross sectional area of beam was  $240\text{ mm} \times 240\text{ mm}$  and was illustrated in the Hydration of a beam Figure 13.

Figure 13 illustrates, temperature variation during 100 hours. Blue line shows temperature variation of the centre and yellow line shows temperature variation of an edge.

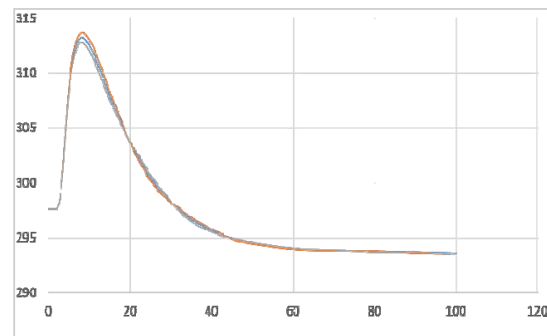


Figure 11: Variation of inside temperature for 100 hours when different temperature variation.

Maximum temperature obtained is given as  $43^{\circ}\text{C}$ . The temperature variation and Degree of hydration is similar to the cube, but Distribution of temperature and Degree of hydration has differences.

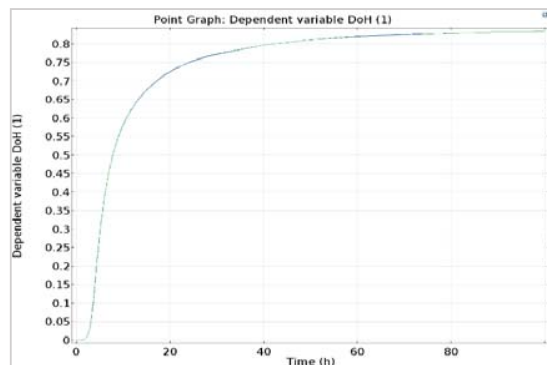


Figure 12: Hydration of the 2 m beam.

Third figure illustrated hydration at 15 hours. 0.67 of cement hydrates at that time. Figure 12 describes hydration at 100 hours. Between the values of DoH as given from 0.82 to 0.83, concrete was hydrated only after 100 hours.

Figure 15 illustrated affinity at different stages of beam. In there also first affinity at middle was high but after considerable amount of cement got hydrated affinity at middle becomes lowest. In Figure 15, figure at left side top, illustrated the affinity at 1 hour. Affinity is between  $3.95 \times 10^{-7}$  to  $4 \times 10^{-7}$  at 1 hour, Figure at right side top illustrated the affinity at 5 hours. Affinity is between  $3.23 \times 10^{-5}$  to  $3.28 \times 10^{-5}$ . Similarly bottom figures illustrated affinity at 10 hours, 15 hours, 20 hours, 25 hours, 30 hours, and 100 hours respectively. Corresponding affinity values distribute within  $7.7 \times 10^{-5} - 7.82 \times 10^{-5}$ ,  $3.4 \times 10^{-5} - 3.48 \times 10^{-5}$ ,  $8.37 \times 10^{-7} - 8.64 \times 10^{-7}$ ,  $4.5 \times 10^{-7} - 4.73 \times 10^{-7}$ ,  $2.73 \times 10^{-7} - 2.95 \times 10^{-7}$ ,  $6.27 \times 10^{-8} - 7.48 \times 10^{-8}$ .

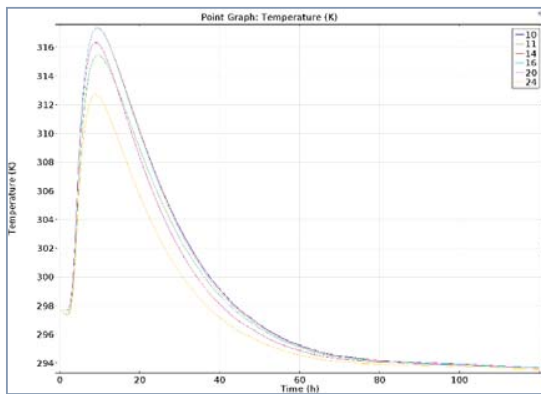


Figure 13: Degree of hydration and Temperature variation of 2 m long beam.

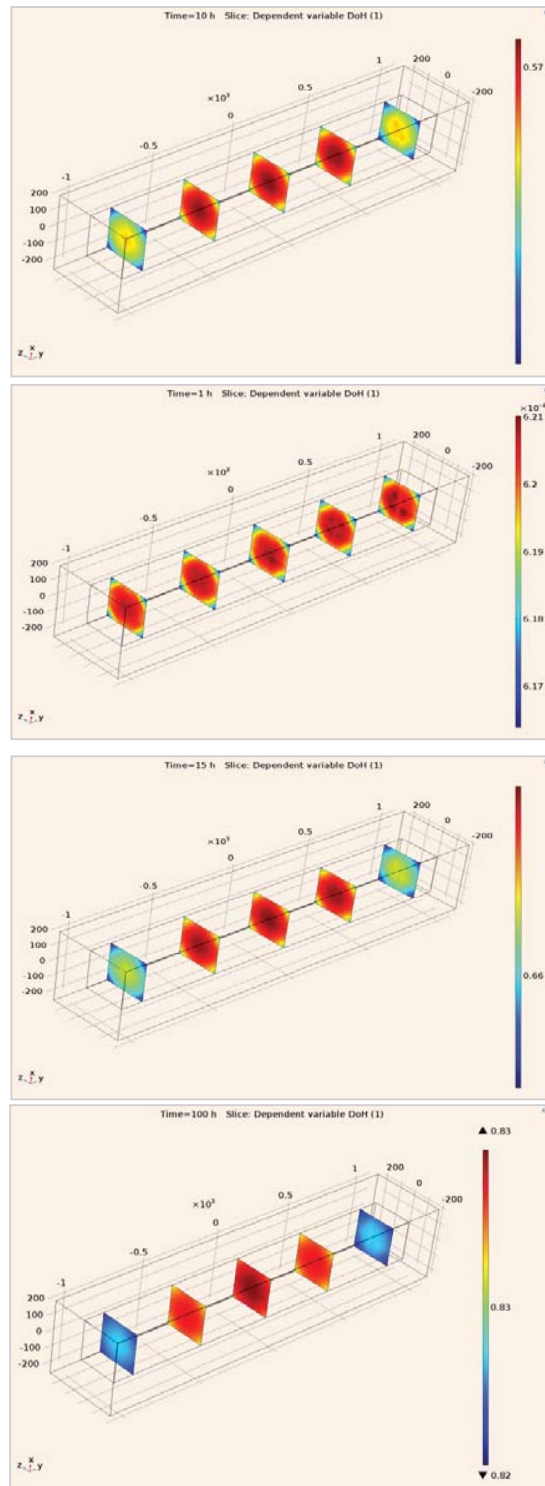


Figure 14: Hydration of a 2 m long beam at different stages.

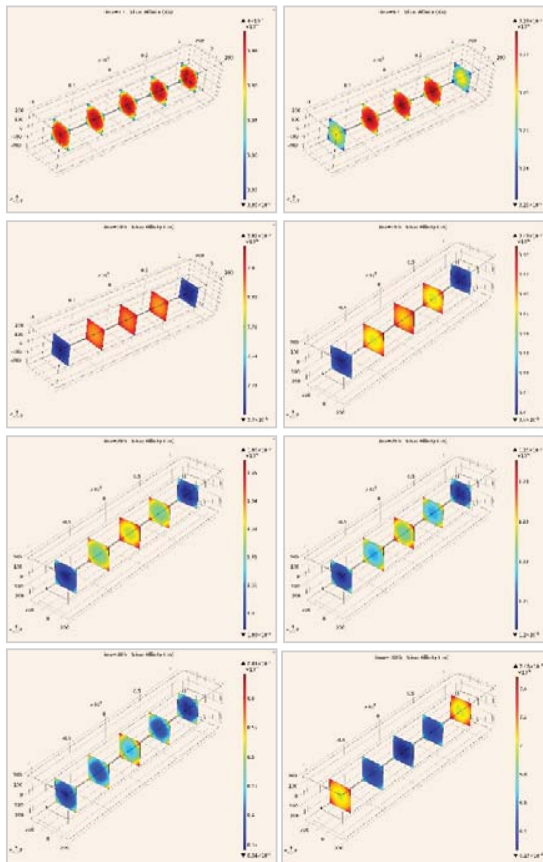


Figure 15: Affinity at different stages of 2 m long beam.

## 5. Conclusion

Concrete cubes of  $14 \text{ dm}^3$  with Polystyrene foam-work gave rapid hydration than concrete cube with wooden foam-foam work. Increment of thickness of foam-work gave quick and uniform hydration. Uniformity of temperature was maintained by foam-work thickness and material.

Surrounding temperature is affects hydration. When surrounding temperature is  $20 \text{ }^\circ\text{C}$ , 80% of cement got hydrated in 100 hours.

Variations of temperature lead to the high temperature inside concrete cube.

Therefore, it can identify an influence of Foam-work thickness, Foam-work Material, Variation of Temperature Surrounding Temperature effect to Concrete Hydration using affinity hydration model and simulated using COMSOL® Multiphysics Software.

Therefore, this method can be used for prediction of cracks in concrete [7]. The method proposed in [8] can be used to increase accuracy and reliability of the proposed method.

## 6. References

- [1] Ricardo, W., Da Silva, L., Smilauer, V. and Stemberk, P., 2015. "Upscaling semi-adiabatic measurements for simulating temperature evolution of mass concrete structures". *Materials and Structures*, 48(4), 1, April 2015, pp.1031-1041
- [2] Gawin, D., Pesavento, F. and Schrefler, B.A., "Hygro-thermo-chemo-mechanical modeling of concrete at early ages and beyond. Part I: hydration and hygro-thermal phenomena" *International Journal for Numerical Methods in Engineering*, 67(3), 16, July 2006, pp. 299-331.
- [3] Cervera M, Oliver J, Prato T (1999) "Thermo-chemo-mechanical model for concrete. I: hydration and aging." *J Eng Mech ASCE* 125(9), September 1999, pp.1018–1027.
- [4] Brazilian Association of Technical Standards: NBR 5736: Pozzolanic Portland Cement-Specification (1991). Rio de Janeiro, Brazil
- [5] *Insulation materials and their thermal properties.* (2018, 05 08). Retrieved from <http://www.greenspec.co.uk/building-design/insulation-materials-thermal-properties/>
- [6] Khan, M.A., Idriss Ali, K.M. and Rahman, M.S. " Swelling and thermal conductivity of wood and wood-plastic composite." *Polymer—Plastics Technology and Engineering*, 36(2), 1, March 1997, pp.179-187.
- [7] Li, C.Q. and Yang, S.T., "Prediction of concrete crack width under combined reinforcement corrosion and applied load". *Journal of engineering mechanics*, 137(11), 24, June 2011, pp.722-731.
- [8] da Silva, W.R.L. and Šmilauer, V., 2015. "Fuzzy affinity hydration model". *Journal of Intelligent & Fuzzy Systems*, 28(1), 1 January 2015, pp. 127-139.

## A simulation model to analyse Sri Lankan Megacity logistics behaviour: Megapolis logistics of Sri Lanka

U.G.M.D. Dias  
*Department of Industrial  
Management  
University of Kelaniya, Sri  
Lanka*  
mdheemadias@gmail.com

V. K. Tharaka  
*Department of Industrial  
Management  
University of Kelaniya, Sri  
Lanka*  
kasuniv@kln.ac.lk

J. N. Nanayakkara  
*Department of Industrial  
Management  
University of Kelaniya, Sri  
Lanka*  
julian@kln.ac.lk

### Abstract

*Colombo, the commercial capital of Sri Lanka is gradually being transformed into a Megacity with increasing urbanization and industrialization and investment in infrastructure development. Having experienced the consequences of unplanned and non-systematic nature of freight transportation in Colombo city, authorities have come up with a megacity development strategic plan, namely, The Western Region Megapolis Master Plan in order to reduce drawbacks associated with inland freight transportation in urban areas. It focuses on a two tier transportation network structure where the freight is consolidated and value added at hub locations via planned road network and transported to the port for shipping. In this scenario, the transportation structure which connects the consolidation hub with the port and different spokes of the network can be considered as the elementary operational unit of the logistics framework. For analysing operational feasibility of this scenario, simulation-based modelling can be utilized in order to predict the future operational performance of the proposed logistics plan. Having identified the elementary operational unit of functionality in the proposed logistics framework, a computer-based simulation model of the elementary unit is developed using Arena software by Rockwell Automation and validated in order to predict the performance of the two-tier transportation network. This quantitative model has the capability to provide strategic decision-making ability by analysing different scenarios, in terms of traffic congestion, road network utilization and the impact of different truck type combinations for optimum utilization of truck fleet.*

**Keywords:** *Computer simulation, Hub, Inland freight, Megacity logistics, Megapolis, Transportation network.*

### 1. Introduction

Megacity logistics aims at developing urban logistics system that is effective and efficient as well as environmentally friendly [1]. The Sri Lankan initiative in developing a futuristic megacity with an effective

logistic platform has been planned via the Logistics corridor of Western Region Megapolis Planning Project. The proposed logistics corridor is aimed at tapping into the strategic opportunity to grow logistics industries and manage the existing sporadic logistics activities. The project will consolidate and link existing and potential logistics activities in the Western Region and the country [2].

The main freight transportation and distribution strategies related to a logistics framework are direct shipping and multi-echelon distribution. While indirect shipping vehicles are transported directly to the customers, two-echelon distribution system which is one form of multi-echelon transportation, delivers freight to an intermediate depot and from that depot to customers. Thus, the transportation network is decomposed into two levels, the 1<sup>st</sup> level connecting the depot to the intermediate depots and the 2<sup>nd</sup> connecting the intermediate depots to the customers. The objective is to minimize the total transportation cost of the vehicles involved in both levels [3]. The intermediate depots discussed by [3], shows a similar idea to the hub concept discussed by [4], in which hubs are special facilities that serve as switching, transshipment and sorting points in many-to-many distribution systems. Instead of serving each origin–destination pair directly, hub facilities concentrate flows in order to take advantage of economies of scale. The above explained theoretical framework can be observed in the proposed Megapolis megacity development initiative in Sri Lanka. According to [2], proposed logistics corridor has identified three major clusters for logistics purposes in three strategic locations in Western Province. Furthermore, it is planned to apply the concept of Logistics Park in the above-mentioned areas. Logistics park creates economies of agglomeration for freight activities thereby lowering operational costs. Furthermore, to support this logistics facility, new road network has been designed connecting hub to port for efficient access of cargo to the deck. But according to [2], the new bonded highway that has been considered for container transportation in 2030 must need a clear and thorough detailed feasibility study to identify whether it will facilitate the future container demand. Therefore, it is evident that even though a

solution has been designed, it has not been validated for the operational feasibility in this context.

A simulation of a system is the operation of a model of the system. The model can be reconfigured and experimented even with; impossible, too expensive or impractical to do in the system it represents. The operation of the model can be studied and hence, properties concerning the behaviour of the actual system or its sub-system can be inferred. In its broadest sense, simulation is a tool to evaluate the performance of a system, existing or proposed, under different configurations of interest and over long periods of real time [5]. Therefore, simulation can be used as a tool to analyse the operational feasibility of the above-mentioned bonded highway proposed by the logistics corridor master plan.

Hence, the objective of this study is to develop a simulation model for the purpose of analysing operational feasibility of the transportation network proposed by Megapolis Logistics Corridor Master Plan.

### 1.1. Research problem

The Megapolis master plan elaborates on a logistics framework consisting of hubs that are connected to the Colombo port and Bandaranaike airport by a road network. It further focuses on a bonded highway connecting these facilities for optimization of transportation. Yet the master plan itself doubts the operational feasibility of the proposed transportation framework connecting these logistics facilities.

The research problem of this study was developed based on this context. It is to develop a simulation model to mimic the logistic framework proposed by Megapolis master plan. The model is limited to the export procedure of marine logistics context based on Colombo port.

When analysing the proposed plan, an elementary unit was identified as the operational unit of logistics plan. It consists of three entities: origin, hub and port and the road network connect these entities with one another. Figure 2 demonstrates the elementary unit together with the research boundary. The corridors mentioned in the model are proposed transportation network by Megapolis transportation corridor development project. As the elementary unit is repetitively utilized throughout the logistics plan, researcher decided to model the elementary operational unit in simulation environment in order to mimic the operational behaviour of the proposed framework.

When analysing the research gap, this study fulfils an empirical research gap as there does not exist any procedure to analyse the operational feasibility of the above-mentioned proposed plan. On the other hand, this study contributes to fulfilling a theoretical research gap as well. It is because of unavailability of researches conducted in the context of megacity logistics in Sri Lanka.

The remainder of the paper is organized as follows; literature review, model design, simulation model, validation of the model and closure of the paper by

offering conclusions and an attempt to provide some perspectives on future research.

## 2. Literature review

### 2.1 Megacity logistics

The transportation system is one of the basic components of an urban area's social, economic and physical structure. One of the major challenges being faced today is to improve the quality of service in urban transportation systems in order to make them competitive and attractive for more and more passengers [6]. Megacity logistics aim to reduce the nuisances associated to freight transportation in urban areas while supporting the economies and social development of the cities. The fundamental idea is to view individual stakeholders and decisions as components of an integrated logistics system. This implies the coordination of shippers, carriers and movements as well as consolidation of loads of several customers and carriers. City logistics explicitly aims to optimize such planning issue, the integrated scheduling of operations and management of resources [7].

### 2.2 Hub based cargo transportation network

Hub or central transshipment facilities, allow the construction of a network where large numbers of direct connections can be replaced with fewer indirect connections. Hub and spoke configurations reduce and simplify network construction costs, centralize commodity handling and sorting, and allow carriers to take advantage of scale economies through consolidation of flow [10].

Flow of people, commodities, information and energy all require a complex network of inter-linkages between origin and destination [10]. The planning of operations for the supply or regulation of transportation services aim to ensure an optimal allocation and utilization of resources to achieve the economic and customer service goals by optimizing the routing of goods through the service network [11].

In freight transportation, there are two main distribution strategies: direct shipping and multi-echelon distribution. In direct shipping, vehicles starting from a depot transport their freight directly to the customers, while in multi-echelon systems the freight is delivered from the origin to the customers through intermediate depots. The growth in the volume of freight traffic, as well as the need to take into account factors such as environmental impact and traffic congestion, has led research in recent years to focus on multi-echelon distribution systems, and, in particular, two-echelon systems. In two-echelon distribution systems, freight is delivered to an intermediate depot and from this depot to customers [7].

## 2.3 Megapolis logistics

Megapolis logistics corridor master plan was initiated based on the theoretical foundation of two concepts mentioned above: megacity logistics and hub-based cargo transportation network. Megapolis project has targeted only commercial city in Sri Lanka that matched up to the status of a megacity: Colombo and devised this plan to optimized logistics functions in order to address issues arisen in the present context due to unstructured and non-systematic logistics arrangements. The proposed logistics framework can be elaborated as follows.

The concept of Logistics corridors represents a geographical area with its main function being a logistical platform providing appropriate logistics infrastructure and physical facilities and substantial existing logistical services. In addition, there are a range of related business, value-added services and social infrastructure components that are necessary elements for the logistics corridor to act as an integrated concept [2]. Proposed logistics corridor has comprised around 126 sq. km<sup>2</sup> of land of Western Region which includes several land parcels of Wattala, Ja-Ela, and Minuwangoda. The projected residence population in proposed corridor is about 680,000 which is around 8% of Western Region population. The estimated employment in the Zone would be approximately 250,000. Logistics corridor between Port & Airport has proposed to link with: National Expressway System, National Railway Line, National Highways and alternative roads which are linked with main urban centres & logistic clusters [2].

Proposed Logistics corridor has identified three major clusters for logistic purposes; Cluster 01 - Ja Ela - Ekala Logistic Cluster, Cluster 02 - Welisara - Mabola Logistic Cluster and Cluster 03 - Peliyagoda logistic Cluster. It has identified that the concepts of inland port; logistic park can be applied in above mentioned areas. [2].

Based on the above plan, it is evident that each hub-based road network acts as a two-tier logistics framework connecting origins of cargo to its destination: Colombo port.

Even though this framework has been conceptualized, it is mentioned in the [2] that the proposed model needs to be validated for its' feasibility in the operational environment of 2030.

## 2.4 Modeling and simulation

Modelling is the process of producing a model; a model is a representation of the construction and working of some system of interest. A model is similar to but simpler than the system it represents. One purpose of a model is to enable the analyst to predict the effect of changes to the system [5].

A simulation of a system is the operation of a model of the system. Simulation is used before an existing system is altered or a new system built, to reduce the

chances of failure to meet specifications, to eliminate unforeseen bottlenecks, to prevent under or over-utilization of resources, and to optimize system performance [5]. The steps involved in developing a simulation model are as follows.

- Step 1. Identify the problem.
- Step 2. Formulate the problem.
- Step 3. Collect and process data
- Step 4. Formulate and develop a model.
- Step 5. Validate the model.
- Step 6. Document model for future use.

Based on the above steps, a computer-based simulation model is designed and developed to mimic the operational behaviour of hub based two tier logistic framework of proposed Megapolis logistics master plan. For this purpose, Arena modelling and simulation software by Rockwell Automation is utilized.

When considering literature available in simulation of logistics frameworks, [8] has considered supply-chain modelling and simulation of an analysis framework for port operations. The simulation approach serves two purposes in this study: modelling a supply-chain network in quantity approach and evaluating its supply-chain performance based on proposed strategies. Based on simulation results, effects of various strategies are used to guide the way to administrate the supply chain in the different objectives.

In the supply chain simulation study for the Portuguese automotive supply chain, a subset of the automotive supply chain involved in the case study is a three-echelon supply chain, composed by one automaker, two 1st-tier suppliers, two 2nd-tier suppliers, and one outsource entity [9]. The purpose of the study is to evaluate alternative supply chain scenarios for improving supply chain resilience to a disturbance and understanding how mitigation strategies affect each supply chain entity performance.

When considering the scope of this paper, it focuses on an empirical gap as there are no studies conducted on developing simulation-based model or any other initiative in analysing operational feasibility of Megapolis logistics framework of Sri Lanka.

## 3. Methodology

The methodology adopted for development of the research is based on following steps as in Figure 1.

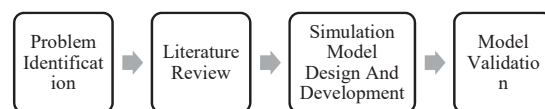


Figure 1. Research methodology.

Initial step of the research procedure is to identify the problem of interest. The seed for problem identification was observation of unstructured and non-systematic

nature of truck transportation in the present context. When further researching on the problem, it was found out that an initiative has already been made by Government of Sri Lanka via Logistics Corridor Master Plan, but yet it lacks the proof of operational feasibility. Therefore, this situation was subjected to the study for further research. A literature survey was conducted on the following areas of concern; Megacity logistics, logistics hub, two-tier transportation network and simulation-based modelling. Furthermore, documents related to Megapolis and Logistics corridor were referred in order to get a better understanding of the research context. Then the results of the literature review were related to the study context in order to build up the model. The third step of the research; simulation-based modelling was conducted using Arena software by Rockwell Automation. The model design is elaborated in (4) below. The final step of research is validation of the developed model for its' reliability and accuracy. The validation procedure is elaborated in (5) below.

#### 4. Model design

Logical model design of the simulation model is elaborated based on the research problem subjected to the study.

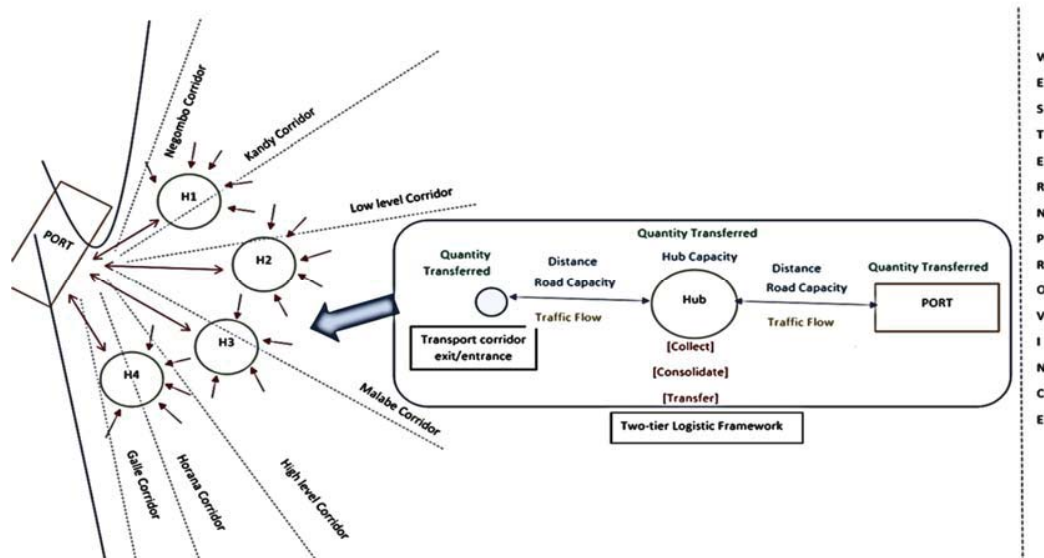


Figure 2. Elementary operational unit of Megapolis logistics framework.

#### 4.1. Activity model

The activity model was developed using a standardized format to aid in the formulation of operational unit of two-tier logistic framework. Integrated DEFINITION (or IDEF) provides a visual depiction of the logistics framework and related inputs, outputs, controls, mechanisms, and metrics. With a visual depiction of the proposed simulation model, a

comprehensive analysis of the key elements and components can be addressed through the entire logistics framework. An overview of the logical model designed based on IDEF notation is demonstrated in Figure 3 and Figure 4 and Figure 5. Figure 3 demonstrates the procedure of cargo coming from origin points up to the start of shipping documentation and authorization process. Figure 4 demonstrates from the document authorization process up to hub exit while Figure 5 depicts transportation from hub exit until the destination.

#### 4.2. Simulation model

Simulation model was developed based on the logic elaborated above in the IDEF model. This model is developed in Arena platform. The developed Arena model can be discussed in three aspects: inbound logistics process, hub process and outbound logistics process.

The inbound logistics process depicts functions from origin up to the hub. According to the proposed logistics framework, mixed traffic is expected in this part of the road network. Therefore, all vehicle categories are originated and at the hub entrance, the other vehicle types are filtered out and cargo vehicles are considered for further processing. This sub model is depicted in Figure 6.

The hub-based logistics process includes many

complex processes that do not occur in line with the cargo flow within a unit time period. Therefore, such processes (value adding functions and warehouse handling activities) are omitted from the simulation model and only the cargo generated into and out of those processes to the general cargo flow is considered. This sub model is depicted in Figure 7. The documentation procedure is expanded in Figure 8.



The outbound logistics procedure from hub to port is simulated considering different speed patterns of vehicles. The simulation model is displayed in Figure 9.

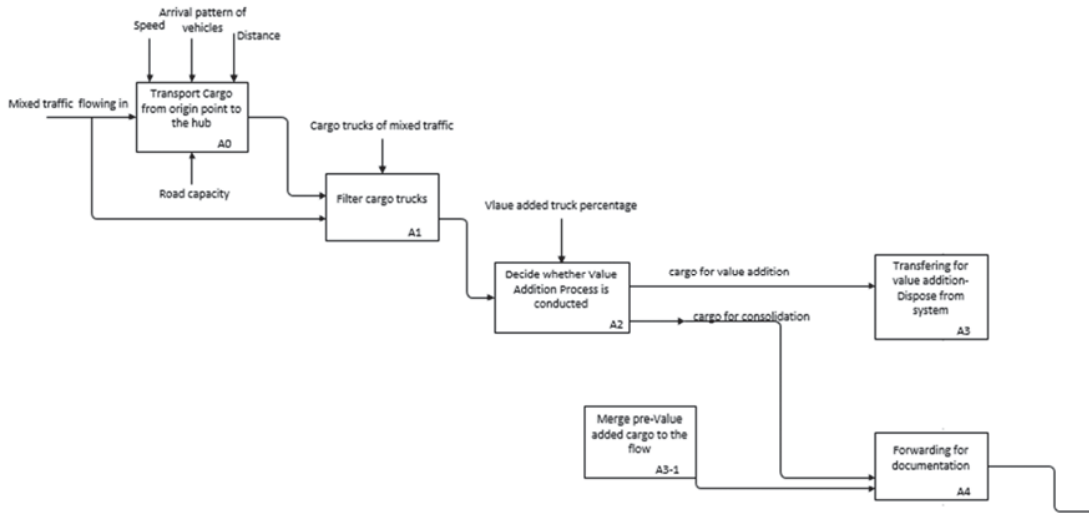


Figure 3. IDEF model from origin to the start of documentation process.

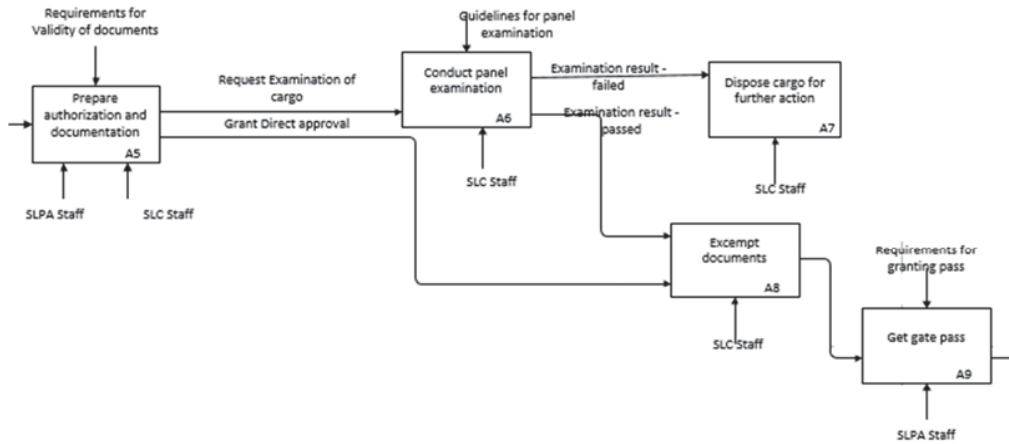


Figure 4. IDEF model from documentation process to hub exit.

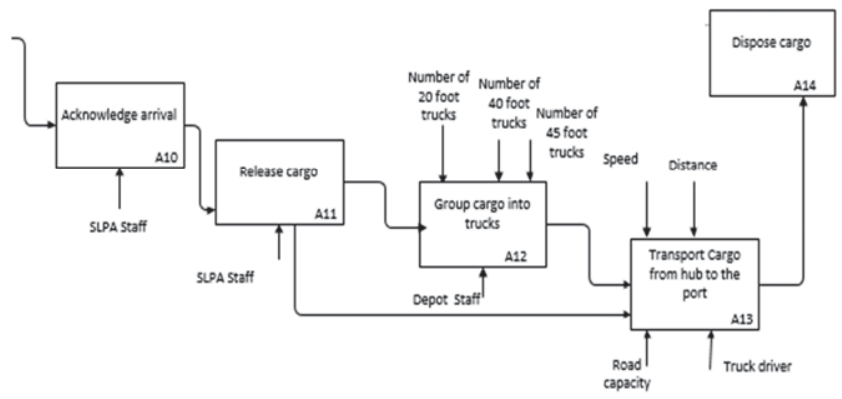


Figure 5. IDEF model from hub exit to port entrance.

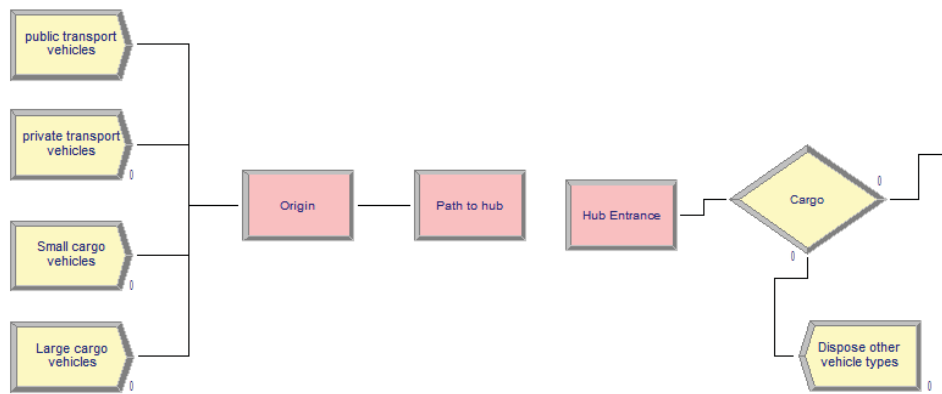


Figure 6. Arena simulation sub model from origin to hub.

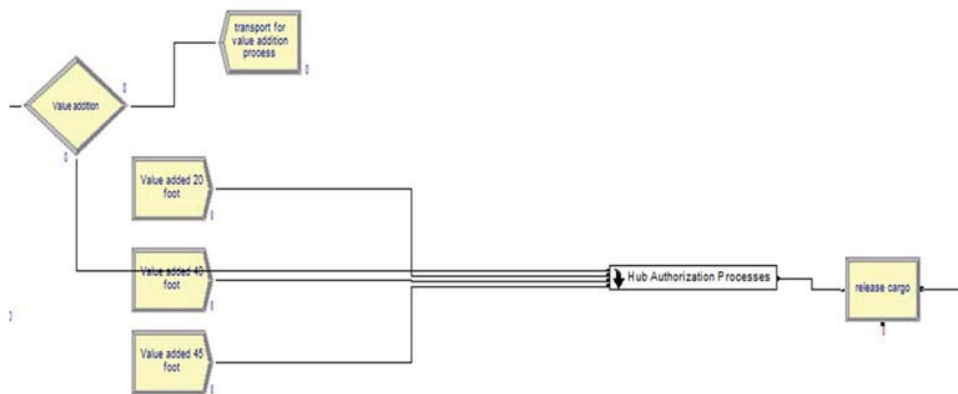


Figure 7. Arena simulation sub model of hub processes.

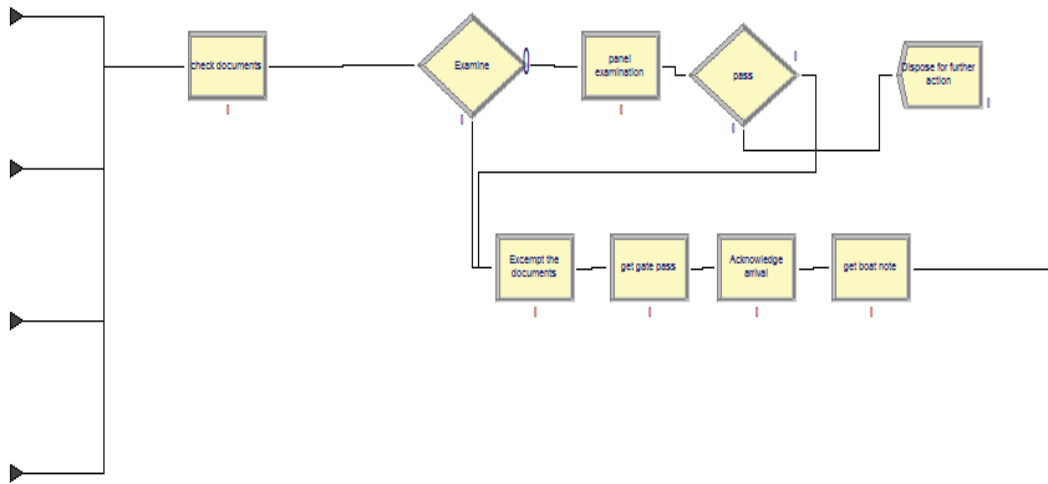


Figure 8. Arena simulation sub model of documentation process in the hub.

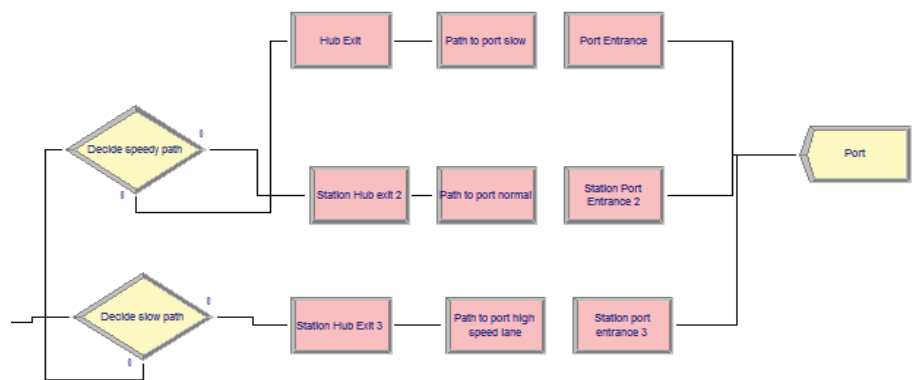


Figure 9. Arena simulation sub mode from hub to port.

## 5. Model validation

Once the model is designed using Arena, it needs to be validated to ensure that the model output will provide reliable results in the experimentation scenario.

Therefore, for the validation purpose a real-world scenario that matches with the simulated environment was selected. As the model focuses on entity flow from origin to destination with merges and distributions, a vehicle pattern that depicts similar behaviour was selected. It is the vehicle behaviour of the southern expressway of Sri Lanka. The reason behind selecting expressway instead of any other Sri Lankan road is to reduce the variability occurring due to noise variables.

Data required for validation was collected from expressway management system and they were provided as inputs to simulation model in order to analyse the

mean number of units outputted by the system and to get the transfer time of entity by the system. Then those figures were compared with the real-world data obtained.

As a result, it was found out that there exists a variability of 4.72% when considering the mean transfer time of entity and a variability of 8.1% when considering the mean output by the simulation model versus real world data. Therefore, it can be concluded that the Arena simulation model will provide results well within 90% of accuracy.

## 6. Conclusion

This paper focuses on developing a computer-based simulation model to analyse two-tier logistics transportation network proposed by Megapolis Logistics Corridor Master Plan, having identified the requirement to analyse its 'operational feasibility'. As the output of

this study, the researcher has been successful in delivering a validated simulation model. There exist many opportunities to utilize the devised model in order to analyse the operational feasibility in the above-mentioned scenario.

This simulation model has the capability to provide strategic decision-making ability in terms of traffic congestion, road network utilization and the impact of different truck type combinations for the purpose of truck fleet optimization. Therefore, for further research there exists opportunity to utilize this model as a tool and analyse the operational feasibility in varying road conditions and truck fleet combinations. Furthermore, there exists an opportunity to devise an aggregated model of the hub network and analyse its' operational feasibility.

## 7. References

- [1] Saragih, N., Bahagia, S., Suprayogi and Syabri, "City Logistics for Mega City: A Conceptual Model (Case Study: DKI Jakarta)," in *Joint International Conference on Electric Vehicular Technology and Industrial, Mechanical, Electrical and Chemical Engineering (ICEVT & IMECE)*.
- [2] "Logistic Corridor Master Plan," Ministry of Megapolis & Western Development, 2016.
- [3] Perboli, G., Tadei R., and Vigo, D., "The Two-Echelon Capacitated Vehicle Routing Problem: Models and Math-Based Heuristics," *Transportation Science*, vol. 45, no. 3, p. 364–380, 2011.
- [4] Alumur, S., and Kara, B., "Network hub location problems: The state of the art," *European Journal of Operational Research*, vol. 190, no. 1, pp. 1-21, 2008.
- [5] Maria, A., "INTRODUCTION TO MODELING AND SIMULATION," in *Winter Simulation Conference*, 1997.
- [6] Dridi, M., Mesghouni, K., and Borne, P., "Traffic control in transportation systems," *Journal of Manufacturing Technology Management*, vol. 16, no. 1, pp. 53-74, 2005.
- [7] Crainic, T., Ricciardi N., and Storchi, G., "Models for Evaluating and Planning City Logistics Systems," *Transportation Science*, vol. 43, no. 4, pp. 432-454, 2004.
- [8] Lee, Park N. and Lee, D. "A simulation study for the logistics planning of a container terminal in view of SCM," *Maritime Policy & Management: The flagship journal of international shipping and port research*, vol. 30, no. 3, pp. 243-254, 2003.
- [9] Carvalho, H., A. P. Barroso, V. H. Machado, S. Azevedo and V. Cruz-Machado, "Supply chain redesign for resilience using simulation," *Computers & Industrial Engineering*, vol. 62, p. 329–341, 2012.
- [10] O'Kelly, M. and Miller, H., "The hub network design problem: a review and synthesis," *Journal of Transport Geography*, vol. 2, no. 1, pp. 31-40, 1994.
- [11] Crainic, G., "Service network design in freight transportation," *European Journal of Operational Research*, vol. 122, pp. 272-288, 2000.

## Factors influencing Enterprise Information Systems adoption of Small and Medium Enterprises (SMEs): A case study on Sri Lankan manufacturing sector

T. R. Kaluarachchi  
*Dept of Industrial Management*  
*University of Kelaniya, Sri Lanka*  
*thisurikaluarachchi@gmail.com*

W. M. S. K. Weerabahu  
*Dept of Industrial Management*  
*University of Kelaniya, Sri Lanka*  
*samanthiw@kln.ac.lk*

L. D. J. F. Nanayakkara  
*Dept of Industrial Management*  
*University of Kelaniya, Sri Lanka*  
*julian@kln.ac.lk*

### Abstract

*Small and Medium Enterprises (SME) play a vital role in the Sri Lankan economy in terms of national output and employment as well as incubating innovative capabilities. In today's highly competitive economy, small and medium-sized enterprises lack the resources and technologies to compete with large enterprises, although having a relatively high casualty rate. In order to survive in the competitive market and exploit opportunities, many small and medium-sized enterprises adopt Information Technology (IT) related applications. However, in the current context of IT Application usage, such as Enterprise Information Systems (EIS), the SME sector in Sri Lanka is lagging compared to other countries. Therefore, the need of adopting to EIS / other Information Systems related technologies is becoming a must or an urgent need in the context of establishing a competitive SME sector. Based on empirical evidence and review of literature this study captures significant factors that influence EIS adoption by SME in Sri Lankan context. The objective of this study is to recognize the real need of EIS based applications for the SME sector while identifying and defining the effectiveness of driving and hindering factors which affect the focus of SMEs adoption towards EIS based applications. Since cost based constraints was identified as a major barrier to adopting EIS based solutions, the increase of productivity and sales profitability are stated as main driving forces by the SME owners and industry experts. The results are expected to provide a practical contribution in the area of EIS adoption in the Sri Lankan Small and Medium manufacturing SMEs for better reinforcement strategies for successful implementation.*

**Keywords:** *Enterprise Information Systems (EIS), Manufacturing, SMEs*

### 1. Introduction

In the digitalized economy, Small and Medium Enterprises need to develop competitive advantages based on adequate and intensive use of Enterprise

Information Systems (EIS), which is an essential element of success in today's market. In every developed or developing country SMEs are considered as the backbone of the economy. The SME sector in Sri Lanka accounts for more than 50 percent [19] of GDP while being the key contributor for the economic growth. SMEs have gained wide recognition as a major source of employment, income generation, poverty alleviation and regional development. The SMEs cover broad areas of economic activity such as agriculture, mining, manufacturing, construction and service sector industries.

It was also identified that SMEs were a fragile sector on economical perspective [5], probably due to the difficulties they face in financial funding and other regulatory subjects. According to government findings [5] on the SME sector, access to finance, regulatory and institutional barriers, access to markets and lack of technological support were identified as critical factors which influenced the growth of this sector.

Most significantly, technological support was identified as one of the key factors for the success of SME sector as they must target Information Technologies as a business driver [1]. However, in developing countries many enterprises in this sector do not obtain any kind of knowledge, related to the latest technologies and some enterprises do not allow these kind of innovative technologies or systems because of their budgetary constraints [3].

As [4] has mentioned, the country can benefit via export-oriented small enterprises rather than depend on foreign exchange savings. The Balance of Payment problem may be reduced via promotion of export oriented small enterprises in this current globalization context. Therefore, establishing export oriented SMEs which adopt Enterprise Information Systems / Technologies is becoming a must or an urgent need. Many IT organizations have identified this need and are developing business solutions according to their information requirements.

Most significantly, there needs to be a proper understanding of the appropriate needs to develop customized business solutions for handling manufacturing SME's specific information requirements.

Therefore, this study is focused on identifying and defining the effectiveness of driving and hindering factors which affect the focus of SMEs adoption towards EIS based applications.

## **2. Small & Medium Scale enterprises in Sri Lanka**

SMEs play an important role in the economy of both developed and developing countries. Their importance is reflected by the number of establishments, generation of employment, contributions to GDP, embarking in innovations, stimulating of other economic activities [6]. However, in Sri Lankan context there is no clear definition of Small and Medium Scale Enterprises and most studies have identified the emerging need of creating a national policy framework and separate authority for betterment and rapid growth of the SMEs.

The Government of Sri Lanka recognizes SMEs as the backbone of the economy, as it accounts for more than 75% of the total number of enterprises, provides 45% of the employment and contributes to 52% of the Gross Domestic Production (GDP) [19]. SMEs promote broad based equitable development and provide more opportunity for Women and Youth to participate in the economic development of the country. The Central Bank of Sri Lanka (CBSL) defines SMEs as enterprises that have an annual turnover of less than Rs 600 million and borrowings below Rs 200 million, while the Ministry of Industry and Commerce describes them as enterprises which employ less than 300 employees and which have an annual turnover not exceeding Rs 750 million.

Different countries use different parameters to define SMEs such as number of employees, amount of capital invested, amount of turnover, etc. At the same time different definitions are also used by different organizations within the same country to define SMEs for different purposes as stated above and this is common scenario in Sri Lankan context as well. International experience, reveals the use of a variety of working definitions depending on country specific objectives and structures.

## **3. Role of Enterprise Information Systems (EIS)**

An enterprise information system (EIS) is defined as software systems for business management, encompassing modules supporting organizational functional areas such as planning, manufacturing, sales, marketing, distribution, accounting, financial, human resources management, project management, inventory management, service and maintenance, transportation and e-business [7]. The term 'enterprise' is defined to include any business activity or enterprise engaged in

industry, agri-business and/ or services whether single proprietorship, partnership, or corporate venture.

It includes commercial software packages that enable the integration of transactions-oriented data and business processes throughout an organization (and perhaps eventually throughout the entire inter organizational supply chain). Enterprise systems include ERP (Enterprise Resource Planning) software and such related packages as advanced planning and scheduling, sales force automation, customer relationship management, and product configuration [8].

Thus EIS is software that enables the user to be able to access data from the database as required. The user can access the data easily (by asking questions such as 'what were the sales in May') without having the need to have a knowledge of programming languages. The results are also provided in a simple format which is easy to comprehend even for a layman (e.g. by way of a graph or chart).

EIS plays a major role in supporting the SME sector for today's rapid development as a developing nation. It also assists them to face competitive challenges in the business world. Enterprise information systems, such as ERP, manufacturing executive systems (MES), customer relationship management (CRM) and so on, are obtaining increasing attention, due to their ability to improve production and business performance, and raise competitive advantage for enterprises [9].

These EIS based applications and IT (Information Technology) systems normally enable companies to respond faster and more effectively to customer requirements. However, if we are to do justice to the money that a SME invests on IT, the key delivery should be a robust and dependable IT software system that fulfills the increasing customer demand and overall business requirements of manufacturing SMEs. As they can be used as a key driver of business than just being using as a data storage.

## **4. EIS adoption in manufacturing SMEs**

A viable and dynamic SME sector is essential for economic development of developing countries. Therefore, promotion and development of small and medium enterprises have been emphasized by policy makers, government, non-governmental organization and international aid agencies as a strategy to achieve high economic growth and poverty reduction [4]. Consequently, it is acknowledged that there is a burning need for adopting Technological enhancements for this sector while investigating above mentioned facts.

To remain competitive in the global market place, SMEs have to embrace Information Technology or its' related systems as a business driver, than an information filing system which is happening currently [4].

While ICT (Information and Communication Technology) can benefit SMEs in multiple ways, they have to identify their own obstacles in IT/ EIS adoption [2], considering their own business priorities and strategies. The major constraints being identified are level of initial investment, organizational culture, readiness, top management support, relative advantage, complexity, and compatibility, the level of external IT (Information Technology) support, economic and economic growth, and competitive pressure [14].

Further, it was found that SMEs differ from large companies in important ways affecting their information-seeking practices [10]. Thus, the adoption of EIS innovations in manufacturing SMEs cannot be a miniaturized version of larger manufacturing organizations despite advances in IT and the acceptance by large organizations of such technologies. Therefore, the same level of adoption is not evident among SMEs.

This also suggests that SMEs face significant and unique challenges in adopting EIS, ICT and e-commerce. This low level of adoption, particularly inhibits SMEs in developing countries. Adoption rate of ICT and e-commerce in SMEs developing countries are different from the developed countries. Furthermore, manufacturing SMEs in developing countries fall behind with adoption due to barriers, both internal and external, prominent and inherent in most cases [3].

Since significant variables that influencing internet adoption by SME in a Sri Lankan context are captured by another study [15] which includes the benefits of internet, complexity, business orientation, newwork practice adoptability and ICT costs.

Normally in Sri Lankan scenario many manufacturing SMEs are mostly centered on doing their daily business transactions or manufacturing processes while not considering latest technological innovations and enterprise applications. It is perceived that manufacturing SMEs encountering new technological enhancements and adopting them are more essential for the growth of this sector. The sector in the country has a low adoption rate compared to other countries possibly due to financial or infrastructural issues and policies [4].

Thus, there is urgent need to further improve the inherent capabilities of the manufacturing SMEs in order to enhance their role in the economic growth and poverty reduction.

## **5. Adoption benefits of EIS for SMEs:**

### **Reasons for change**

In today's highly competitive environment, small and medium-sized enterprises lack the resources to compete with large enterprises. In order to survive, many small and medium-sized enterprises adopt information technology (IT) because it can help them to exploit

opportunities and strengthen competitive capabilities. In addition to these advantages of IT, many studies have attempted to find the factors influencing IT adoption in SMEs [11]. Appropriate ICT can benefit SMEs in cost cutting by improving their internal processes, improving their product through faster communication with their customers, and better promoting and distributing their products through online presence. ICT has in fact the potential to improve the core business of SMEs at every step of the business process [17].

Adoption of internet and related technologies, facilitate firms to identify new investment opportunities in local and overseas markets for the SMEs, thereby allowing them to be successful. The real need to accelerate the adoption of internet in developing countries is needed now more than ever as a means to reduce economic disparity among nations and the responsibility lies with the respective governments to identify and to provide the necessary opportunities through a facilitation process which may involve significant national investments to uplift local knowledge base relating to the scope [15].

Many small and medium-sized companies are implementing ERP/ EIS relate technologies to increase their competitiveness. The challenge is to find the best fit between target business process and ERP system logic [12].

On the other hand, use of IT makes small production runs easier and, hence, SMEs will be able to be more competitive as they can introduce new manufacturing technologies quickly. These new technologies enable SMEs to be more flexible and be able to respond to customer needs. Generally, SMEs view investments in IS (Information Systems) /IT in the same way as they invest for their production systems – they expect them to last for a considerable time. This supports the contention that SMEs are not particularly flexible as they see themselves continuing in their chosen market. An alternative view of lack of investment in IT might be that SMEs attempt to remain flexible by holding cash rather than becoming committed to a particular system [17].

The willingness of SMEs to integrate e-business practices depends on how much it can directly improve their core business and how potential benefits outweigh the costs. By recognizing these differences and focusing their efforts on removing the constraints, governments can play an important role in encouraging SMEs to become more effective users of EIS/ ERP [14].

It was currently identified that implementing EIS related applications failed most times in large organizations. Due to uncertainty in the business environment, most ERP vendors are moving their attention towards SMEs by offering cheaper and simplified solutions that appeal to this market segment. The offerings include ERP system with compact

packages, flexible pricing policies, new implementation methodologies and more specialized functionalities [18].

Previous findings also suggest that many organizations adopting EIS have serious conflicts with their business strategies and the majority of EIS projects are often characterized by delays and cost overruns [14].

Many SMEs either do not have sufficient resources or are not willing to commit a huge fraction of their resources due to the long implementation times and high fees associated with ERP/EIS implementation. The resource scarcity, the lack of strategic planning of information systems (IS), the limited expertise in IT and also the opportunity to adopt a process-oriented view of the business are among the factors that strongly influence, either positively or negatively, ERP adoption by SMEs [13]. However, SMEs are less inclined to radical change and less aware of the organizational impact caused by the implementation of an ERP system. Similarly, SMEs' traditional focus on operations onto day-by-day management, coupled with a lack of strategic view of ICT, could be partially accountable for these findings. Indeed, SMEs could simply prefer to continue doing business as they are used to, by refusing to adopt solutions that could change their "course".

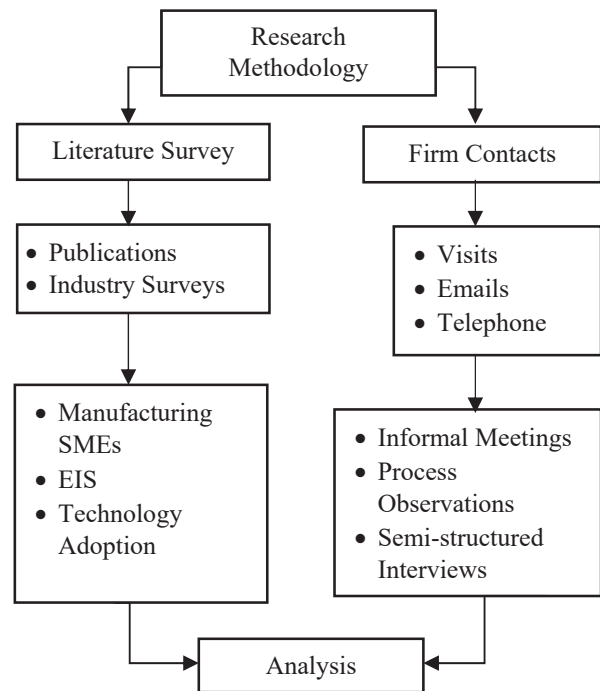
These results could imply that SMEs structural peculiarities are a real obstacle to ERP diffusion and lack of adequate strategy by vendors [16]. Moreover, the decision process regarding the adoption of ERP systems within SMEs is still more affected by exogenous reasons or "opportunity of the moment" rather than on business-related factors contrary to large companies that are more interested in managing process integration and data redundancy/inconsistency through EIS related technologies (ERP) implementation [10].

Most of them have implemented ERP to integrate the existing information system on SMEs and found ERP implementation mainly beneficial in reducing inventory, improving customer services and improving communications. Top management support, and user involvement and participation are found to be the major contributors of ERP/EIS success [18].

In the present era of globalization, it is obvious that the survival of the manufacturing SMEs will be determined and will have to be addressed extensively. Accessibility to the right kind of information at right time is the first and foremost in their quest to manufacture/supply more at competitive cost, quicker delivery times, minimum defects and using fewer resources. In order to face this challenge, the issue of ERP is the need of the hour and hence the need of EIS/ERP in manufacturing SMEs is beyond doubt [18].

## 6. Methodology

This study is based on identifying the current literature and empirical consequences that have been performed/resulted on EIS adoption in SMEs and how those different driving and hampering factors affect successful adoption of EIS in manufacturing SMEs.



**Figure 1. Scenario of the research methodology**

A mixed method of research approaches was obtained to analyze factors of the study. This consists literature review, observations and interviews conducted on SMEs practices on EIS. The literature review conducted broadly at the initial stages of this research demonstrated the existence of gaps in the knowledge of adoption factors. Since, the review has not provided general factors about EIS adoption within organizations, further data was collected based on case study to validate the literature findings as well as generating new knowledge with empirical evidences.

The population definition focused on industrial establishments stated at [19] and designated the sample from a specific industrial sector which consists of manufacturing SMEs in the Gampaha district, Western Province of Sri Lanka. The data collection process involved SMEs sample in Industrial Estate situated in Gampaha district under the Industrial Development Board (IDB) of Sri Lanka. Hence, consultations from industry experts on EIS implementations on SMEs and Board of Directors of IDB were also involved in the data



collection procedure. Visitations to ten SME manufacturers enabled the review processes and personal interviews and surveys conducted. The case study organizations mostly consisted those in chemical, petroleum, rubber and plastics and metal based products manufacturing.

However, this research study uses two main methods i.e. the secondary data and documents, and primary data which was obtained by the researcher through direct observations and interviews as shown in Figure 1.

Based on the findings from the literature review as well as feedback from case study industries, a content analysis was done and a conceptual framework was developed considering the different factors obtained from relevant studies.

The observations and interviews considered following characteristics on case study companies in manufacturing SME sector.

- Have ensured the survival in the industry  
As for their evolving or failure status, the requirement on IT related systems are not always high in SMEs. Therefore, it is assumed that only the fittest SMEs will be survived after 6 years of establishment.
- Have initial awareness on EISs or are informed about the ICT  
Most SMEs are either not aware or not willing to invest or participate in these developments.
- Have considered the Sri Lankan born companies since multinational companies and its subsidiaries will tend to follow higher standards and practices.

## 7. Results and discussion

According to the previous research publications and conducting pilot surveys and expert opinion interviews, the following driving and hindering factors were identified. A preliminary survey or pilot survey was done to clearly address on an important and relevant issues targeted by this research. However, it was used to gather information prior to conducting a survey on a larger scale. Because of the effectiveness of a pilot study in identifying flaws in a surveys and structure of interviews, appropriate amendments can then be utilized in a full study [20].

This research investigation has realized following main factors in both sides hence critical for successful adoption of EIS within SMEs. Among these factors the industry owners were given basic ranking structure on these factors in respect of the influence to successful EIS adoption of their firms.

In light of the factors affecting the better utilization and adoption of EIS solutions in Sri Lankan SMEs, a conceptual framework was developed. This identified the driving forces and hindering factors behind the successful adoption and utilization of sophisticated EIS solutions

within Sri Lankan SMEs as SMEs' development is an essential element in the growth strategy of many economies, including Sri Lanka. These factors were extracted and prioritized among other factors according to the ranking structure given to the SME owners and industry experts accordingly.

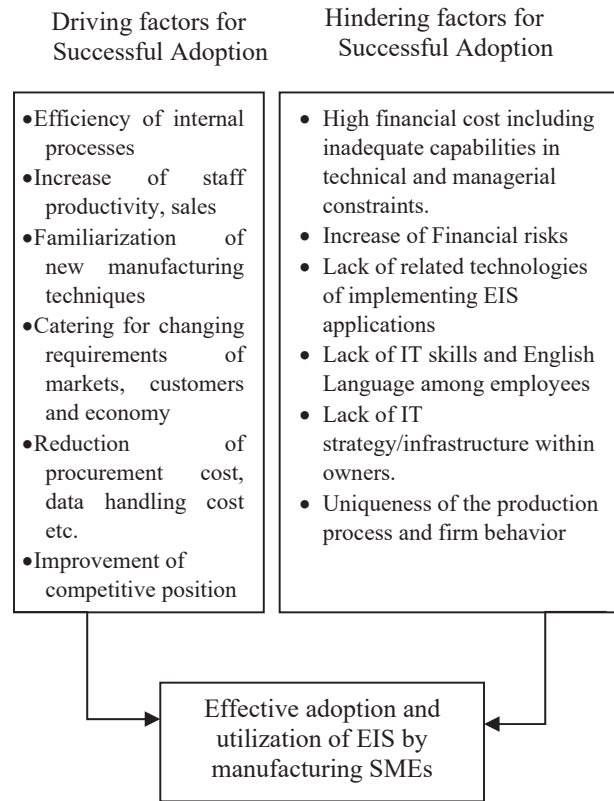


Figure 2. Conceptual framework.

Following factors (in Figure 2) were recognized the major driving and hindering forces behind the effective adoption and utilization of EIS by manufacturing SMEs most importantly. According to the analysis of driving factors, efficiency of internal processes, increase of staff productivity, sales and catering for changing requirements of markets, customers and economy has gain higher percentage compared to other driving factors. Mainly productivity improvements, to maintain efficient internal processes and increase of sales through EIS based technologies are the major driving forces or benefits they gain through EIS adoption according to the sample data in manufacturing sector. This can be explained through current context of Sri Lankan SME sector because the major concern arose from market orientation. As for the respondents, it was revealed that they strongly believe that EIS related technologies are required to achieve competitiveness of Small and Medium scale manufacturing enterprises. Therefore, it

can be concluded that there is an actual requirement of adopting EIS based applications in our context as well.

Table 1 delivers the literature findings of the driving factors which were identified mostly on empirical evidences. However, most of the organizations don't yield the true benefit of EIS solution after implementation, due to lack of continual improvement of the implementation process. As indicated by ranking sequence, SMEs always expect benefit realization through EIS adoption process. But most of the time this can have a lengthy time period which SMEs cannot afford as its' initial investment cost exceeds revenue. Therefore, it also becomes one of the hampering factors on EIS adoption in SME manufacturing sector.

**Table 1. Literature findings on driving factors.**

Driving Factor	Source
Efficiency of internal processes	[21, 22]
Increase of staff productivity, sales	[23-25]
Familiarization of new manufacturing techniques	[26, 27]
Catering for changing requirements of markets, customers and economy	[28, 32]
Reduction of procurement cost, data handling cost etc.	[29-32]
Improvement of competitive position	[26]

High cost of computer equipment and high charges on connection services, High initial set up cost and high in-house development cost and financial constraints are the main hampering factors in our sample data indicated. This factors were mainly recognized in expert interviews as well. Since EIS Adoption in Sri Lankan SME sector is far beyond compared to other countries mainly because of the immense financial constraints as focused on major constraint for successful utilization. Meanwhile there is an accepted level of demand of EIS technologies there is a gap in their investment potential and initial cost of that kind of system. Consequently, there seems a significant deviation of EIS adoption mostly because of these reasons. Table 2 represented the hindering factors identified in different manufacturing industries which had successful EIS adoption, while these were further validated by empirical evidences on the manufacturing SMEs in Sri Lanka.

**Table 2. Literature findings on hindering factors.**

Hindering Factor	Source
High financial cost including inadequate capabilities in technical and managerial constraints.	[10, 24, 25, 33]
Increase of Financial risks	[27, 29]
Lack of related technologies of implementing EIS applications	[34]
Lack of IT skills and English Language among employees	[30, 35, 37]
Lack of IT strategy/infrastructure within owners.	[34, 36]
Uniqueness of the production process and firm behavior	[38, 40]

## 8. Conclusion

The main objective of the research was to examine how SMEs are receptive for the so called business solutions or applications and what are the factors associated with this technology adoption. Therefore, the proposed conceptual framework was structured and validated through comprehensive content analysis. The purpose of this research is to identify the context and causes for relatively less adoption level of Small and Medium Enterprises towards Enterprise Information System Applications in Sri Lanka, while identifying the factors that contribute towards adoption.

However, the generalization of this framework reveals many important factors which demonstrate that SMEs do not process relatively same level of IT adoption compared to local large companies. While investigating on the business solutions provided by IT vendors in Sri Lanka, it has an urgent need to introduce more scale down version of enterprise information systems which are provided by large scale ERP vendors. SMEs can benefit either as adopters of EIS and producers of EIS related technologies. According to this study, financial resource constraints and uncertainty of financial environment were major concerns, in adopting these EIS technologies. In the current phase of digitalization era, SMEs can clearly have more benefit compared to larger companies if they come-up with their own strategies. As experts in the industry stated that Sri Lanka has an emerging speed of technology adoption compared to other countries in South Asian region. Therefore, there is a vast opportunity for maximizing their production capabilities and gain competitive advantage with this massive spread of technological innovations throughout future generations.

In initial discussions with manufacturing SME owners revealed that they have a positive attitude on adopting/using EIS technologies to their business as they believe in gaining competitive advantages through them.

Therefore, it is much needed to bring forward the willingness to adopt EIS related technologies by providing necessary guidance to SMEs and encouraging them to become benefitted by those EIS/ERP systems where some SMEs already have basic usage of IT related applications like email, MS office packages, accounting software etc. Hence, it is much beneficial for successful adaptation in other related technologies as well.

Some of these factors were initially highlighted in expert interviews and further introduced as cost factors. User resistance is another major component in EIS adopting besides these identified factors. Simultaneously, the financial constraints immensely affected successful adoption on these related technologies on SMEs because they didn't have the potential of investing in those huge projects.

In conclusion it is hoped that the developed conceptual framework and the empirical results presented here provide a useful starting point for future empirical studies that examine successful EIS related technology adoption by SMEs in different sectors in Sri Lanka other than manufacturing sector.

## 9. References

- [1] Dasanayaka S., "Critical success factors affecting the development of clusters for small and medium scale information technology firms in Sri Lanka", *International Journal of Entrepreneurship and Small Business*, vol. 17, no. 1, p. 118, 2012. Available: 10.1504/ijesb.2012.048652.
- [2] Barba-sánchez, V., Martínez-ruiz, M. P., & Jiménez-zarco, A. I. (2007). Drivers, Benefits and Challenges of ICT Adoption by Small and Medium Sized Enterprises (SMEs): A Literature Review Drivers, Benefits and Challenges of ICT Adoption by Small and Medium Sized Enterprises (SMEs): A Literature Review, (January).
- [3] Kapurubandara, M. and Lawson, R., 2006. "Barriers to Adopting ICT and e-commerce with SMEs in developing countries: an exploratory study in Sri Lanka." *University of Western Sydney, Australia*, 82, pp.2005-2016.
- [4] Vijayakumar, S., 2013. "The status of Small and Medium Enterprises and promotions for their growth in Sri Lanka." *International Journal on Global Business Management and research*, 1(2), pp.1-12.
- [5] "National Strategy for Small and Medium Enterprise Sector Development in Sri Lanka White Paper - Institute of policy studies Sri Lanka", Institute of policy studies Sri Lanka. [Online]. Available: <http://www.ips.lk/national-strategy-for-small-and-medium-enterprise-sector-development-in-sri-lanka-white-paper/>. [Accessed: 23- Jan- 2019].
- [6] Gamage, A.S., 2003. "Small and medium enterprise development in Sri Lanka: a review." *Meijo University, Faculty of Business Management, Nagoya, Japan*. [http://wwwbiz.meijo-u.ac.jp/SEBM/ronso/no3\\_4/aruna.pdf](http://wwwbiz.meijo-u.ac.jp/SEBM/ronso/no3_4/aruna.pdf).
- [7] Rashid, M.A., Hossain, L. and Patrick, J.D., 2002. The evolution of ERP systems: A historical perspective. In *Enterprise Resource Planning: Solutions and Management* (pp. 35-50). IGI Global.
- [8] Markus M., Tanis C. and van Fenema P., "Enterprise resource planning: multisite ERP implementations", *Communications of the ACM*, vol. 43, no. 4, pp. 42-46, 2000. Available: 10.1145/332051.332068.
- [9] Liang C. and Q. Li, "Enterprise information system project selection with regard to BOCR", *International Journal of Project Management*, vol. 26, no. 8, pp. 810-820, 2008. Available: 10.1016/j.ijproman.2007.11.001 [Accessed 23 January 2019].
- [10] Buonanno, G., Faverio, P., Pigni, F., Ravarini, A., Sciuto, D. and Tagliavini, M., 2005. "Factors affecting ERP system adoption: A comparative analysis between SMEs and large companies." *Journal of Enterprise Information Management*, 18(4), pp.384-426.
- [11] Shiau, W.L., Hsu, P.Y. and Wang, J.Z., 2009. "Development of measures to assess the ERP adoption of small and medium enterprises." *Journal of Enterprise Information Management*, 22(1/2), pp.99-118.
- [12] Vilpola I., Kouri I. and Vaananen-Vainio-Mattila K., "Rescuing Small and Medium-Sized Enterprises from Inefficient Information Systems--A Multi-disciplinary Method for ERP System Requirements Engineering", 2007 *40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, 2007. Available: 10.1109/hicss.2007.464
- [13] Al-Mashari M., "Enterprise resource planning (ERP) systems: a research agenda", *Industrial Management & Data Systems*, vol. 102, no. 3, pp. 165-170, 2002. Available: 10.1108/02635570210421354.
- [14] Malawige I. and L. Nanayakkara, "Drawbacks in SMEs" EIS applications: framework for development of EIS for SMEs in Sri Lanka", *International Journal of Multidisciplinary Studies*, vol. 1, no. 1, p. 63, 2014. Available: 10.4038/ijms.v1i1.40.
- [15] Suriyapperuma, H.P., Ab Yajid, M.S., Khatibi, A. and Premarathne, S.P., 2015. The Impact of Internet Adoption on SME performance in Sri Lanka: Development of a Conceptual Framework. *International Journal of Arts and Commerce*, 4(1), pp.46-58.
- [16] Ratnayake, L.S., 2012. "The Influence of complexities of ERP for effective implementation at medium sized companies in Sri Lanka."
- [17] Levy M. and Powell P., *Small Business Economics*, vol. 11, no. 2, pp. 183-196, 1998. Available: 10.1023/a:1007912714741 [Accessed 23 January 2019].
- [18] Kale, P.T., Banwait, S.S. and Laroija, S.C., 2010. Performance evaluation of ERP implementation in Indian SMEs. *Journal of Manufacturing Technology Management*, 21(6), pp.758-780.

- [19] "Annual Survey of Industries Final Report 2016", Department of Census and Statistics, Ministry of National Policies and Economic Affairs, Sri Lanka, 2016.
- [20] Hassan, Z.A., Schattner, P. and Mazza, D., 2006. "Doing a pilot study: why is it essential?" Malaysian family physician: the official journal of the Academy of Family Physicians of Malaysia, 1(2-3), p.70.
- [21] Raymond, L., Rivard, S. and Jutras, D. 2004. 'Small enterprises' predisposition to adopt an ERP'. Sixth International Conference on Enterprise Information Systems (ICEIS), Portugal, 10-14 April, pp.614-618.
- [22] Argyropoulou, M., Loannou, G. and Prastacos, G.P. 2007. "Enterprise Resource Planning implementation at Small and Medium Sized Enterprises: an initial study of the Greek market," International Journal of Integrated Supply Management, vol. 3, no.4, pp.406-425.
- [23] Bernroider, E., and Koch, S. 2001. 'ERP selection process in mid-size and large organizations'. Business Process Management Journal, 7(3), pp. 251-257.
- [24] Chong, S. and Pervan, G. 2007. "Factors Influencing the Extent of Deployment of Electronic Commerce for Small-and Medium-Sized Enterprises," Journal of Electronic Commerce in Organizations, vol. 5, no.1, pp.1- 29.
- [25] Tagliavini, M., Faverio, P., Ravarini, A., Pigni, F. and Buonanno, G. 2002. 'Exploring the use of ERP systems by SMEs'. In Callaos, N. (eds.) SCI 2002: Proceedings of the 6th World Multi-Conference on Systemics, Cybernetics and Informatics, Orlando, FL.
- [26] Levy, M. and Powell, P. 2000. 'Information systems strategy for small- medium-sized enterprises: an organizational perspective'. Journal of Strategic Information Systems, 9(1):63-84.
- [27] Raymond, L., Rivard, S. and Jutras, D. 2004. 'Small enterprises' predisposition to adopt an ERP'. Sixth International Conference on Enterprise Information Systems (ICEIS), Portugal, 10-14 April, pp.614-618.
- [28] Branzei, O. and Vertinsky, I. 2006. 'Strategic pathways to product innovation capabilities in SMEs'. Journal of Business Venturing, 21: 75-105
- [29] Koh, S.C.L. and Saad, S.M. 2006. 'Managing uncertainty in ERP-controlled manufacturing environments in SMEs'. International Journal of Production Economics, 101: 109-127.
- [30] Raymond, L. and Uwizeyemungu, S. 2007. 'A Profile of ERP adoption in manufacturing SMEs,' Journal of Enterprise Information Management, vol. 20, no.4, pp.487-502.
- [31] Koh, S.C.L. and Simpson, M. 2005. 'Change and uncertainty in SME manufacturing environments using ERP.' Journal of Manufacturing Technology Management, 16(6): 629-653.
- [32] Palaniswamy, R. and Frank, T. 2000. 'Enhancing Manufacturing Performance with ERP systems'. Information Systems Management, 17: 43-55.
- [33] Bernroider, E., and Koch, S. 2001. "ERP selection process in mid-size and large organizations," Business Process Management Journal, Vol. 7, No.3, pp.251-257.
- [34] Lertwongsatien, C., & Wongpinunwatana, N. (2003). E-commerce adoption in thailand: An empirical study of small and medium enterprises (smes). Journal of Global Information Technology Management, 6(3), 67-83. <https://doi.org/10.1080/1097198X.2003.10856356>
- [35] Ward, J. M. (2002). Understanding the successful adoption and use of IS/IT in SMEs: an explanation from Portuguese manufacturing industries, 121-152.
- [36] Thong, J.Y. and Yap, C.S., 1995. "CEO characteristics, organizational characteristics and information technology adoption in small businesses." Omega, 23(4), pp.429-442.
- [37] Apulu, I., Latham, A. and Moreton, R., 2011. "Factors affecting the effective utilisation and adoption of sophisticated ICT solutions: Case studies of SMEs in Lagos, Nigeria." Journal of Systems and Information Technology, 13(2), pp.125-143.
- [38] Van Everdingen, Y, van Hillergersberg, J. and Waarts, E. 2000. 'ERP adoption by European mid-size companies'. Communications of the ACM, 43(4): 27-31.
- [39] Motwani, J., Mirchandani, D., Mandal, M., Gunasekaran, A. 2002. 'Successful implementation of ERP projects: evidence from two case studies'. International Journal of Production Economics, special issue IT/IS in 21<sup>st</sup> century manufacturing, 75: 83-96.
- [40] Grandon, E., & Pearson, J. M. (2004). E-commerce adoption: Perceptions of managers/owners of small and medium sized firms in Chile. *Communications of the Association for Information Systems*, 13, 81-102.

## Process aligned business rule framework for immunization services in Sri Lanka

Gilmini Dantanarayana  
 Department of Computer  
 Science  
 Faculty of Science  
 University of Ruhuna  
 Sri Lanka  
 gilmini@dcs.ruh.ac.lk

Prasad Jayaweera  
 Department of Computer  
 Science  
 Faculty of Science  
 University of Sri  
 Jayawardenepura  
 Sri Lanka  
 pja@sjp.ac.lk

Anusha Walisadeera  
 Department of Computer  
 Science  
 Faculty of Science  
 University of Ruhuna  
 Sri Lanka  
 waindika@dcs.ruh.ac.lk

### Abstract

*Identification of healthcare (HC) service processes is a key requirement of successful e-Health system engineering. Ensuring compliance of the HC service process with medical knowledge and guidelines integrated from various written documents and expertise knowledge is a fundamental challenge to e-solution development. Current e-Health solutions often rely on technical level communication and there is no systematic methodology to derive these low level requirements from higher level knowledge base and guidelines. Thus, it is necessary to design a systematic approach to fill the gap between higher level service policies and expertise knowledge with lower level service processes. Therefore, the proposed framework could be considered as a bridge, facilitating e-Health solution for developers to systematically realize high level requirements on underlying technologies. Formalizing care pathway is a promising approach to meet this requirement. However, currently the formalizing care pathway is incomplete due to the complexity in HC domain. The methods also used to implement electronic clinical pathways in other countries are difficult to adopt in Sri Lankan context. This study is an initiative contributing to designing a framework of formalizing a clinical pathway with three main consecutive layers for higher level policy driven service process designing.*

**Keywords:** e-Health solutions, HC service process, service policies

### 1. Introduction

Process modelling is used to represent procedures and workflows of a business. Therefore, it is a vital modelling step in requirement engineering of Healthcare (HC) e-solution development. Nowadays, rule based process modelling is a fundamental requirement in successful e-solution design and development. However, rule aligned service process designing in HC domain is challenging due to the interoperable and dynamic

behaviour of HC services. HC service processes must undergo through a set of complex policies, laws or acts which are difficult to use as sources, to refine a set of business rules as a requirement of process modelling. A notable global standardization effort in this direction is SAIF (Service Aware Interoperable Framework) [1, 2]. Most of the research considers higher level policies related to information security perspectives with the acts such as Health Insurance Portability and Accountability Act (HIPAA) [3]. One promising approach to realize the higher level policies and guideline in healthcare is clinical pathways [4]. The clinical pathway is the practical path which conform all these guidelines and policies stated in different resources [4, 5]. When healthcare service in Sri Lanka is considered, e-Health system engineering is performed by considering only low level process requirements. According to the information provided by the Sri Lankan Healthcare service providers, implementing clinical pathways lack in considering specific standard guidelines. Specifically, information is missing in certain situations where complex decisions are required to manage. Therefore, it is necessary to perform system engineering mainly based on clinical pathways while ensuring standard guidelines. However, these efforts are still incomplete and no systematic guidelines or approaches that could facilitate healthcare policy-based and rule-based service designing and deployment are available. As a remedy, this study introduces an architectural framework to map higher-level service policies into service processes with the set of mapping guidelines. Compared to other countries, Sri Lankan e-Health solutions should consider some of the separate set of guidelines and pathways due to a number of factors. Socio-economic factors, cultural factors, and environmental factors are some of them. Therefore, the work proposed in this study, focuses on the application scenarios in Sri Lankan context while addressing its specific set of guidelines. The proposed study based on the principals of global HC standardizations efforts and existing business rule and process modelling ontologies which are introduced briefly in the next section of the paper.

The remainder of this paper is organized as follows. Succinct analysis of related research is presented in the second section. Section 3 discusses compliance/rule management of HC Service solutions and proposed framework for rule aligned process model in Healthcare. Finally, Section 4 concludes the paper and presents directions for future work.

## 2. Background; modelling ontology in HC and formalizing clinical pathways

There is an increasing demand in designing and implementing computer-based clinical pathways [4, 5]. Automated clinical pathways act as powerful tools in supporting care delivery standardization and improvement [5, 6]. A clinical pathway consists of a set of systematically developed statements to assist practitioners decisions, in order to provide optimal care with respect to a specific clinical condition [7]. Care pathways are adaptations of clinical guidelines tailored to respond to the healthcare and economic requirements defined for a local clinical context [7]. Therefore, it is required to define clinical pathways considering local context. National Program for Information Technology (NPFIT) initiated by NHS (National Health Service), UK mainly considers care pathways as core elements in achieving personalized healthcare [8]. They also highlight the necessity of electronic care pathways in practicing evidence-based medicine. Availability of electronic care pathways may result in a number of benefits to both patients and healthcare practitioners [7, 8]. Continuity of care, involved patient in planning care, receiving own care pathway are some benefits for patients. Receiving most up-to-date guidelines and best practices, monitor progress, ability to share experience to improve the process, adhere to the standards, and understand the deviations of the standard processes are major benefits to the healthcare practitioners.

There are many on-going projects in Sri Lanka to set up local set of guidelines and pathways for some of the critical diseases. However, the most common set of guidelines used in Sri Lanka is the guidelines specified by WHO (World Health Organization) [9]. Nowadays, few information systems are used in Sri Lanka. But, they are not enough to facilitate the decision making pathways. Process oriented system development is unable to provide decision making assistance when handling complex situations where multiple parties and resources are involved. The importance of automating clinical pathways, conforming to country specific guidelines is a major requirement in Sri Lanka HC sector [9], [10]. Currently, guideline manuals are developed by Sri Lankan HC authorities for some chronic diseases such as strokes, hypertension and diabetes. Diabetes is one of the most common chronic diseases in Sri Lanka. The correct pathways depend on a number of factors. Most of the decisions depend on medical history of patient, some physical factors, demographic factors, economic factors, etc.

Application of process modelling tools is more common in designing clinical pathways. Business Process Modelling Notation (BPMN) is the most emerging process modelling tool used by many recent studies [11–13]. Clinical pathways usually based on various clinical protocols and guidelines that may be retrieved from informal higher level decisions, are not supported only by process modelling techniques. The process modelling mainly assists in identifying and representing apparent processing tasks. Therefore, it is required to extend the existing process modelling methods in healthcare in order to meet the higher level decisions that cannot be presented only by process models. Some recent studies explored extensions to process models in order to handle complex situations in medical domain. A research study highlighted the requirement of mapping higher level decisions with lower level process modelling. The study uses Decision Modelling Notation (DMN) for higher level decisions and BPMN for process modelling [12]. Even though, the work provided a complete set of steps, it focuses mainly on the required sources for particular decision. It is important to validate whether the defined rule is realized in the ultimate system.

Clinical pathways automation must be done according to the global HC standards. Importantly, Health level 7 (HL7) and Service Aware Interoperable Framework (SAIF) are related global standards specifically designed for health domain [14, 15]. SAIF resulted as an extension for the HL7 and designed to perform healthcare service modelling based on working interoperable behaviour in this domain [16]. SAIF is a global HC framework that focuses on interoperable behaviour in clinical context. SAIF discusses interoperable behaviour of healthcare via four frameworks, Enterprise Conformance and Compliance Framework (ECCF), Governance Framework (GF), Information Framework (IF) and Behavioural Framework (BF). Most of the e-health systems focuses on behavioural and informational aspects of healthcare scenarios with limited study on higher level governance aspects. This research work introduces a suitable higher-level framework that assists to address less traceability among higher-level rule statements, lower level rules and processes. Therefore, it is necessary to find applicability of existing mapping methods of trading sector and then to extend them to meet the requirement alignment issues associated with different rule levels. These different rule levels include acts, statements, and constraints that govern HC service processes. They also have no clear distinction or relationship when performing requirement engineering.

It is important to analyse the tools and methods in the trading sector which are appropriate for adopting for HC sector. Business rules are guidelines and restrictions for the processes in an organization that represent terms of business policies or constraints. Business rule is defined in various ways and different methods are used to analyse and represent them in business context. According to the [17], business rules are statements of goals, policies and

constraints which are required for compliance management in business organizations. However, specificity of the business rules to HC service must be discussed in advance after carefully analysing service processes in HC. In contrast to trading sector business rules, HC business rules are complex and no complete methodology or guidance is used to represent them with respect to each service process. Inadequacy of understanding distinction among higher level rule statements and lower level rule statements leads to difficulty of deriving a complete set of rules that controls HC service process. However, few studies focused on understanding and realizing business rules in healthcare context. This research attempts to investigate the appropriateness of existing ontologies for the HC sector that are successfully used in mapping between process models and rule models in trading sector. Object Management Group (OMG) standards such as SBVR and DMN provide a comprehensive guidance to analyse and represent structural and behavioural rules in business context [18, 19]. There are also existing translations between process models such as PBMN or BPEL and business rule models such as SBVR [20]. In addition to adaptation of these existing rule modelling methodologies, the study focuses to be in line with the guidelines in global HC standard, SAIF that developed as an interoperable framework for HC services. The applicability of existing business ontologies in adopting and realizing HC service rules should be examined according to the proposed framework.

Rule aligned process designing in Information System Engineering is a proven theory in trading sector [21, 22]. The same requirement is identified in Health domain too [12]. Due to the complexity of collaborations in healthcare service, it is difficult to identify an approach for this misalignment between processes and rules.

### 3. Proposed framework

The initial requirement of proposing complete mapping guidance among existing ontologies is to analyse the generic layers to assign those existing ontologies. Therefore, this research introduces three major layers which can be modelled using relevant OMG standards [23].

The proposed model includes three layers, such as HC Governance rules, HC service rules and HC service process. Corresponding to the guidelines of each sub framework in SAIF, some of the common business modelling tools can be adapted to HC domain as depicted in Figure 1. The work is described by referring infant immunization case study in Sri Lanka.

#### 3.1. Infant immunization in Sri Lanka

National expanded program on Immunization, Sri Lanka is published by Epidemiological Unit, Ministry of Health, Nutrition and Welfare. The program has been funded by World Health Organization. The program document is a manual that provides a comprehensive and

clear guideline about immunization specifically focus on vaccine preventable infectious diseases in Sri Lanka. This study concerns a set of guidelines of standard vaccination procedure that is described in the manual.

There are series of vaccines which are given to an infant by considering many factors in parallel. The vaccine time is defined primarily with the age. However, some other factors such as weight, allergies, etc. also are considered for each individual infant. Several decisions should be performed conditionally and there should be individual monitoring systems. Maximized efficiency can be obtained by retrieving higher level process requirements from higher level guidelines indicated in manuals and other sources. Particularly, when there are many authorized actors required to collaborate with certain situation, the accurate and the most practicable pathway can be selected using this method. The infant is unwell in the allocated date or the infant is having treatment which lowers immunity or the infant has had severe reaction to any vaccine are some critical situations that are important to know in deciding vaccinations.

The following three layers of the proposed model are explained with the vaccination guidelines defined for the Sri Lankan context.

#### 3.2. HC governance rules

HC Governance layer referred to the objectives, policies and guidelines of the HC organization or collaborative parties. Mainly governance establishes rules for higher level decisions with responsible authority of each decision. SAIF\_CD specification explained the governance framework language by means of precepts, people, processes and metrics with referred to the recent research [24]. Governance layer must explicitly define responsibilities or authority of members in order to reach objectives or goals of the organization as an input to HC Service Rules. Furthermore, these higher level decisions must be extracted from the sources such as laws, acts and statements of the organization. Business Motivation Model (BMM) and Responsibility Assignment Matrix (RAM) [25] could be suitable methods to capture high level healthcare service rules and guidelines. Higher level constraints related to service development, administration and authorization could be modelled as these types of rules. In this work, we propose the utilization of RAM (Responsibility Assignment Matrix) as an appropriate approach in modelling higher level rules with respect to the activities of HC services.

As reported in Table 01, the governance rules can be exemplified as below. Rules can be categorized into three phases: Pre-operational, operational and post-operational. Operational rules rely on pre-operational rules. Post-operational rules always depend on the results of operational rules.

**Table 1: Categorization of a set of immunization rules / guidelines.**

Rule/Guideline	Rule Type
Provide details to parents on risks of vaccination and risks of not being vaccinated	Strictly-enforced rule
Ensure that valid consent is given	Strictly-enforced rule
Check the identity of the recipient	Strictly-enforced rule
Administer the vaccine, using the correct technique	Strictly-enforced rule
Immunization should be postponed if the subject is suffering from any established acute illness. Minor infections without fever or systemic upset are not contraindications.	Strictly-enforced rule
Children with minor illness may be vaccinated safely. But, if they are suffering from a major illness or high fever (102°F) they should not be vaccinated. These children should be vaccinated after they recover.	Pre-authorized override
When infants and children have missed scheduled vaccine doses, every opportunity should be taken to check vaccination status and to provide missing doses as early as possible. The next dose should be scheduled after an appropriate minimal interval.	Post-justified override
If the reaction is severe, it may be appropriate to omit further doses of the vaccine. If there has been a severe reaction to Triple Antigen (DTP), it may be necessary to use a preparation without pertussis vaccine (DT) instead of Triple Antigen (DTP). Such reactions should be reported to the medical officer of health.	Pre-authorized override

According to Table 1, the rules found in the immunization guidelines manual in Sri Lanka can be categorized under different enforcement levels. Identifying enforcement levels leads to the requirement of representing a rule in more informative way. Authorization of HC rule is important when performing the corresponding action of the rule.

The rule must be defined with a set of information. Usually, healthcare rules are enforced by a set of actors

who have different roles. In some situations, all roles are played by a single actor. However, most of the rules have a set of actors who have separate roles. Actor who is responsible, informed, accountable can be different or same for particular rule. Referred to the Responsibility Assignment Matrix (RAM), the specified enforcement rule structures are shown in Table 2.

**Table 2. Informative description of rules.**

Pre-authorized override	<condition : the reaction is severe for vaccine > <authorized action : omit further doses of the vaccine> <Override action : use a preparation without pertussis vaccine (DT) instead of Triple Antigen (DTP)> <informed actor: Medical Officer Health>
Post-justified override	<condition : the reaction is severe for vaccine > <authorized action : omit further doses of the vaccine> <Override action : use a preparation without pertussis vaccine (DT) instead of Triple Antigen (DTP)> <informed actor: Medical Officer Health>

### 3.3. HC service rules

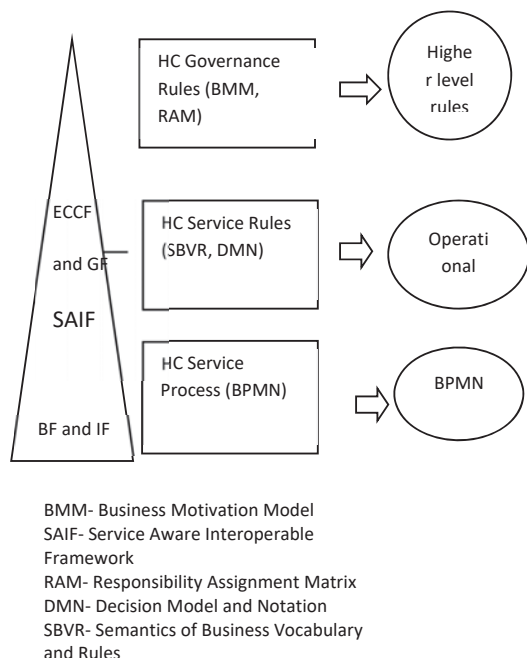
Intentional rule statements resulted from governance layer can be further decompose into operational rules conforming to structural and informational frameworks requirements of SAIF. SBVR is widely used business rule ontology of OMG that covers many aspects of business rules. Furthermore, the operational rules that can be modelled with respect to few enforcement levels provide a sufficient knowledgebase of business rules. According to SAIF, behavioural framework covers dynamic semantics of HC enterprise. It is necessary to combine these dynamic semantics with static semantics that are defined in IF sub framework in SAIF. Proposed framework attempts to model these operational rules with SBVR and DMN conforming the requirements of both IF and BF.

### 3.4. HC service process

The process of providing service care to the patient consists of four main services. Admit, detect, treat and discharge are the four main phases which act as major work flow and they must be composed with the set of defined service process rules in the particular context. Operational rules identified in the HC service rule layer could be modelled further as service process rules in



terms of association between actor, activity and information objects. These rules are necessarily constrained by relevant structures defined in the first two layers of the proposed model.



**Figure 1: Relationship between HC ontologies and Business rule ontologies.**

#### 4. Results and discussion

This paper discusses the proposal for aligned business rule framework for healthcare service solutions. The proposed framework uses different modelling ontologies for designing artefacts of each layer. The framework consists of three modelling layers: HC Governance (HG), HC Service Rules (HSR) and HC Service Process (HSP). Principals of global HC service interoperability standards (SAIF) and business modelling ontologies are considered when designing each layer of the framework. Governance layer is designed based on the principals of SAIF governance layer that refers higher level policy decisions and constraints with considering responsible authorities. This study explains how the higher level guidelines in immunization scenario are modelled using RAM and using enforcement level of BMM. Further, the work reported here introduced a contribution in an endeavour to develop a complete and sound business rule oriented service designing framework. Yet another commendable contribution is the facilitation to bi-directional traceability between these modelling layers that designers could achieve with the adoption of the proposed framework.

The higher level rule framework for HC service solutions is introduced in this research. Furthermore, it is a partial endeavour of ongoing research work to fulfil the

requirement of aligning HC service process with the HC service rules which are informally defined as higher level decisions. System developers are the target audience of the framework as it facilitates an efficient approach to capture the system requirements from higher level rule statements or constraints. The proposed framework that consists of a set of guidelines along with clinical pathway is represented in terms of rule ontology. Then a web based system is implemented based on the rule ontology. Clinical expertise is benefited since the correct clinical pathway is captured and implemented. All the actors who are assigned with different responsibilities and delegation of authorities could be well defined in a single system while accountability of each task is maintained. Decision making assistance is another major facility of formalisation of automated clinical pathways. The importance of automating clinical pathways for certain diseases in Sri Lankan context is highlighted in the study. It is also necessary to extend the work until the process requirements are completely identified and validated, corresponding to the higher level rules and guidelines for selected diseases.

#### 5. References

- [1] Zhang, J., Madduri, R., Tan, W., Deichl, K., Alexander, J. and Foster, I., "Toward semantics empowered biomedical web services," in *Proceedings of the IEEE 9th Int. Conference on Web Serv. ICWS 2011*, pp. 371–378, 2011.
- [2] Iqbal, A. M., "LNCS 6719 - An OWL-DL Ontology for the HL7 Reference Information Model," *Lncs*, vol. 6719, pp. 168–175, 2011.
- [3] Annas, G. J., "legal issues in medicine HIPAA Regulations - A New Era of Medical-Record Privacy?," *N. Engl. J. Med.*, pp. 1486–1490, 2003.
- [4] Wang, H. Q., Li, J. S., Zhang, Y. F., Suzuki, M. and Araki, K., "Creating personalised clinical pathways by semantic interoperability with electronic health records," *Artif. Intell. Med.*, vol. 58, no. 2, pp. 81–89, 2013.
- [5] Lenza, R.; Blaserb, R.; Beyerb, M.; Hegerb, O.; Biberc, C.; Baumleinc, M.; Schnabelc, M. IT support for clinical pathways—Lessons learned. *Int. J. Med. Inf.*, 2007, vol. 76, no. SUPPL. 3, pp. 397–402,
- [6] Isern D. and Moreno, A., "Computer-based execution of clinical guidelines: A review," *Int. J. Med. Inform.*, vol. 77, no. 12, pp. 787–808, 2008.
- [7] Launer, J., "Care pathways," *Postgrad. Med. J.*, vol. 84, no. 993, pp. 392–392, 2008.
- [8] Wachter, R. M., "Making IT work : harnessing the power of health IT to improve care in England. Report of the National Advisory Group on Health Information Technology in England," p. 71, 2016.
- [9] N. Expanded, "Immunization Handbook National Expanded Programme on," *Nutrition*, 2002.
- [10] N. K. Foundation, "C Linical P Ractice G Uidelines for H Emodialysis," *Group*, no. July, 2016.
- [11] Combi, C., Oliboni, B., Zardini, A. and Zerbato, F., "Seamless Design of Decision-Intensive Care Pathways," in *Proceedings of the IEEE Int. Conference on Healthc. Informatics, ICHI 2016*, pp. 35–45, 2016.
- [10] N. K. Foundation, "C Linical P Ractice G Uidelines for H Emodialysis," *Group*, no. July, 2016.

- [11] Combi, C. , Oliboni, B. , Zardini, A. and Zerbato, F., "Seamless Design of Decision-Intensive Care Pathways," in *Proceedings of the IEEE Int. Conference on Healthc. Informatics, ICHI 2016*, pp. 35–45, 2016.
- [12] Combi, C. , Oliboni, B. , Zardini, A. and Zerbato, F., "A Methodological Framework for the Integrated Design of Decision-Intensive Care Pathways - an Application to the Management of COPD Patients," *J. Healthc. Informatics Res.*, pp. 157–217, 2017.
- [13] Zerbato, F., Oliboni, B., Combi, C. , Campos, M. and Juarez, J. M., "BPMN-based representation and comparison of clinical pathways for catheter-related bloodstream infections," in *Proceedings of the IEEE Int. Conference on Healthc. Informatics, ICHI 2015*, pp. 346–355, 2015.
- [14] Zhang, Y. F., Tian, Y., Zhou, T. S., Araki, K. and Li, J. S., "Integrating HL7 RIM and ontology for unified knowledge and data representation in clinical decision support systems," *Comput. Methods Programs Biomed.*, vol. 123, pp. 94–108, 2016.
- [15] A. Review and W. Group, "HL7 Service-Aware Interoperability Framework - Canonical Definition , Release 1 HL7 Informative Document," pp. 1–63, 2011.
- [16] Landgrebe J. and Smith, B., "The HL7 Approach to Semantic Interoperability," 2011.
- [17] Number O. M. G. D. and Files, M. C. "Business Motivation Model," no. May, 2014.
- [18] Object Management Group, "Decision Model and Notation," vol. 1.0, no. September, p. 172, 2015.
- [19] Nelson M. L. and Sen, R., "Business rules management in healthcare: A lifecycle approach," *Decis. Support Syst.*, vol. 57, no. 1, pp. 387–394, 2014.
- [20] Group, O. M., "Business Motivation Model," no. May, 2015.
- [21] Rathnayake N. and Jayaweera, P., "Rule Framework for Motivation / Service Process Alignment," pp. 2–7.
- [22] Andersson, B., Johannesson, P. and Zdravkovic, J., "Aligning goals and services through goal and business modelling," *Inf. Syst. E-bus. Manag.*, vol. 7, no. 2, pp. 143–169, Mar. 2008.
- [23] Group, O. M., "Semantics of Business Vocabulary and Business Rules TM," no. May, 2017.
- [24] C. Mead *et al.*, "Service - Aware Interoperability Framework - Canonical Definition," pp. 1–61.
- [25] Melnic, A. S. and Puiu, T., "The Management of Human Resources within Projects: the Structures of the Project Team, the Responsibility Assignment Matrix," *Econ. Transdiscipl. Cogn.*, vol. 14, no. 1, pp. 476–484, 2011.

## Automated, low-cost pedestrian crossing carriage for efficient traffic control and pedestrian safety

B. A. O. S. Yaraswin  
 University of Sri Jayewardenepura  
 Gangodawila, Nugegoda.  
 Colombo, Sri Lanka  
 selaka.onila@gmail.com

T. M. K. K. Jinasena  
 University of Sri Jayewardenepura  
 Gangodawila, Nugegoda.  
 Colombo, Sri Lanka  
 kasun@dscs.sjp.ac.lk

### Abstract

*Disorganized city planning and a huge rise in use of automobiles on the road have caused massive traffic congestion in cities across the world. Pedestrian crossing designed to facilitate movement across the road network, have unfortunately become a hindrance to movements of traffic. Though the smart zebra lines had been introduced, it has not contributed much to reduce the time that holds vehicle lines under the traffic lights. Moreover, the establishment of transfer hubs and underground crossings remain silent in local context because of their cost. Even though there are plenty of pedestrian crossing mechanisms available, they are not secure and not ideal for the disabled, elderly, children, and the sick. Considerable numbers of police officers have to spend their time on traffic controlling duties though it's inefficient and wastage of human resources. A number of studies have focused on automated vehicles and robots as tools to ease problems of congestion. This paper, it focuses on the design of an automated guided carriage system for pedestrian transportation in an efficient and secure manner.*

**Keywords:** Pedestrian carriage, Zebra lines, Traffic congestion

### 1. Introduction

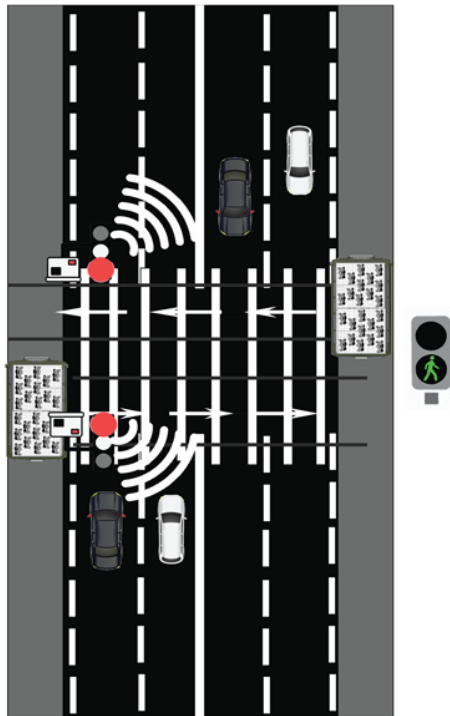
The increase of the middle class, especially in developing countries has led to an increase in the number of automobiles on the roads which have caused enormous traffic congestion [1]. Moreover, the inappropriate city planning, rapid increase in population, urbanization and limited road capacities in major cities has expanded this matter [2]. Simultaneously pedestrians become a deciding factor for traffic congestion [3]. As roads are getting wider it's time-consuming to move across a zebra line where streets are broader. That factor is consolidated because, in order to extend transportation facilities with industrialization, streets are being made broader with multiple lanes. With that kind of broad

streets, time taken to cross the zebra line may vary for elderly, children and the disabled. Thereby, the vehicles will have to halt in front of the zebra lines for a longer time. This may cause huge vehicle queues all along the streets and many zebra lines mean more vehicle holding time next to zebra lines until they get the green light. Similarly, streets that have multiple zebra lines will increase vehicle holding time. The time that is wasted can be reduced by using the proposed pedestrian carrier. These carriers can be placed at the end of the zebra line, on top of the pavement. When it reaches the maximum pedestrians which are designed to be transported, carriage starts to drive itself along the automated destination location across the road. This process can be completed within two to three seconds. Thereby, vehicle won't need to wait for a long time till all the pedestrians have crossed. It allows the vehicle to travel maximum distance without stopping. Because the proposed system is an intelligent solution it can be synchronized with other similar carriage systems which are located nearby and it clears the way for vehicles in a more efficient manner by synchronizing them and letting the traffic flow smoothly. It can reduce traffic congestion by a considerable amount [4]. These traffic congestions cause wastage of resources such as fuel and vehicle parts. This issue not only wastes thousands of man-hours but also is a devastating factor for environmental pollution.

### 2. Overview

Introducing subway tunnels to cross is not a sustainable solution since the budget allocation for such projects is very expensive and time-consuming [5].

Even overhead bridges are of the same nature. Climbing over such bridges are not easy for disabled or elderly people while it restricts usage by persons on wheelchairs.



**Figure 1. Illustration of the proposed system.**

Most of the time pedestrians are not willing to use overhead bridges and subway tunnels to across the road since they have to exert extra energy. Using such links might be time-consuming for the pedestrians as well. Because of that, pedestrians try to take the shortest path even though they have to break road traffic rules. Such actions may end up with sudden deaths, mortal wounds or minor injuries. Most of the time sudden weather changes can be expected. Under such bad weather conditions like monsoon showers, it is troublesome to use pedestrian crossing techniques mentioned earlier. Even with the evolvement of the smart city concept based on the Internet of Things, no new proposals designed to overcome the issue have been seen which are efficient, accurate and secure. Studies have reported smart cities incorporating zebra lines which are transmitting precautions like lighting up the zebra line while pedestrians are crossing [6]. But those techniques were neither secure nor efficient. Proposed carriage works as an elevator that is moving horizontally. This carriage needs to be placed at the end of the zebra line, top of the pedestrian platform. Once it reaches the destined amount of pedestrians, carriage takes off from the parked location within few minutes.

Once it crosses the first lane vehicles which are holding in front of that particular zebra line idly will be allowed to move on. Similarly, it can minimize the time

that holds the vehicles off. This carriage will facilitate disabled people and even those who travel via wheelchairs. Similarly proposed carriage is safer than walking along the zebra line with great exposure to vehicles. In case of collision with a vehicle, it might be fatal to a pedestrian. That can be minimized from the proposed system because accidental impacts which happen directly with carriage versus vehicle will be less fatal than impact with a vehicle. This can also minimize the death rate that has shown an annual increase.

### 3. Literature review

In order to ensure pedestrian's safety and reduce the traffic congestion there are some techniques that have been used previously. A few studies done regarding these aspects are described in this section.

Pedestrian behavior is dependent on gender. Pedestrian speeds are also precisely related to pedestrian age. The speeds of pedestrians are slower as they get older [7]. And also Decision-making power, time and space synchronizations have issues when it comes to females.

The importance of enforcing the recognition of crosswalk area using light emitting platform marker was articulated and hence, vehicle driver can be easily informed where the actual zebra line is on the road [8]. Similar to this solution, 3D lines to show visual illusion had been introduced. It acts as a roadblock on the road so that vehicles slow down anyway. They can be called as precaution mechanisms for pedestrian crossings.

Another study showed that using illuminator provides three times longer distance of noticing the pedestrian and provides ample time to slow down the speed of the vehicle. [9] But this is just a precaution mechanism and it doesn't ensure the safety of the pedestrians.

The usage of pedestrian bridges minimizes the risk of pedestrians and also the flow of the traffic [10].

The decision to build a footbridge is based on the width of the road, traffic volume and the numbers of pedestrians that crossed the road [10]. But building up a footbridge is always an expensive solution which is sometimes cannot be applied to and each every road because of their width.

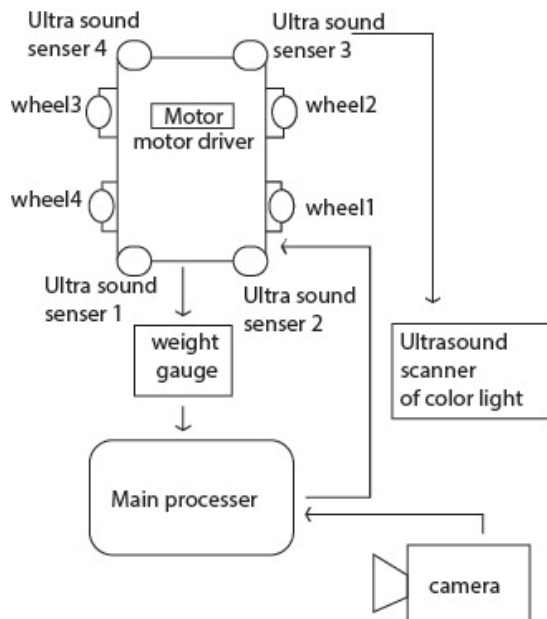


Figure 2. Block diagram of the system.

## 4. Methodology

### 4.1. Structure

The proposed carriage is to be built as a shell which can be afforded a sudden impact that can cause a vehicle accident. The shell needs to be protected by airbags for the safety of the passengers.

This carriage has a ramp which facilitates disable personnel and wheelchair travelers to get in. For the ramp and door opening mechanism, DC (direct current motors) can be used.

### 4.2. Transmission

The proposed carriage is designed to travel top of a railing. It helps to reduce the friction; thereby the transportation is comfortable to passengers similar to a train. Using iron wheels can extend the comfort as well. Further, less friction helps to reduce power consumption as well. There would be two brushless DC electric motors to power up the whole carriage. Brushless DC motors have a good dynamic response to propel electric cars as well [11]. So it also more suitable to use DC motors for this carriage as well because, the control of the motors have to be designed carefully. Because, other wheels are running on a top of a rail, and by using such motors, the speed of the carriage can be increased instantly and also it won't be generating any uncomfortable vibes and sounds. The carriage will also be allowed to start the ride only after the camera which

is placed in front of the carriage provides the signal that the path is clear.

### 4.3. Sensors

Ultrasound sensors need to be placed at the end of both sides and both ends of the carriage to detect when it is passing by the nearest traffic light and it would be signaling to the vehicles that are holding in that lane to pass through.

### 4.4. Traffic lights

Traffic lights will be turned to red when the carriage starts. They will be turned to green when the carriage is passing a particular lane. Cameras placed on top of signal detect the vehicles which are moving ignoring the signal to stop and passes a caution message. The carriage can be halt even though the signals are green for the carriage. Similarly cameras which are placed at the top of the signal lights can be also used to monitor the traffic flow and detect the rule violations like detecting vehicles which are neglecting commands that are generating through the signal lights. Such camcorders can also be used to detect high-speed movement of vehicles and for surveillance purposes as well.

### 4.5. Power source

Since the motors are powered up by the electricity, Lithium-ion batteries are used because they hold large amounts of energy, which makes them excellent choices for the carriage as well. The power source can be charged using two charging points which are placed in the source location and the destination location of the carriage. Charging can be done by also using wireless charging mechanisms like inductive charging; by that electricity can be transferred via an air gap from one magnetic coil located in the wireless charger to a second magnetic coil mounted to the carriage. When the carriage

is stopped at the destination location and when the coils are aligned, charge begins [12].

Simultaneously charging points can be power up by solar power because it can be placed on both ends of the street and easily can be exposed to the sunlight. Thereby, power source can be charged by using solar panels.

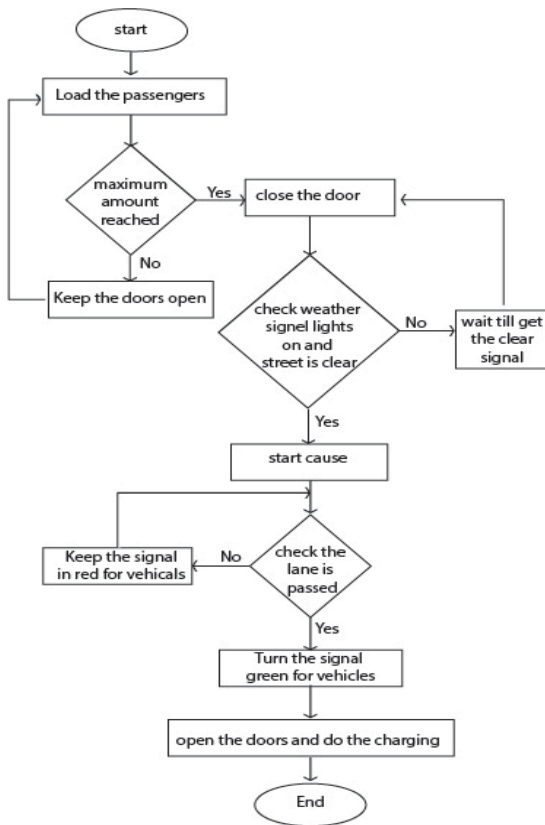


Figure 3. Flow chart of the system functionality.

The carriage can be placed on the platform and electronic door can slide off when it completes the maximum number of passengers that can be carried at one instance. Not only the crossing time of the passengers can be reduced by fifty percent but also the holding time of the vehicles can be reduced since the traveling speed of the carriage can be increased rapidly.

#### 4.6. Crowd density and size of the carriage

According to UK event guidance, two people per square meter is the low risk standing formation for human beings. But this can be further enhanced up to 2.5 person per square meter and three people per square meter in a secure manner. So if the carriage can be made with an area of ten square meters (2×10m) it can easily host 20-30 passengers without any risk. Accordingly, the

area of the carriage can be designed with respect to the width and type of the road. Simply the size of the carriage can be changed with respect to the type of the road and crowd density of that particular area.

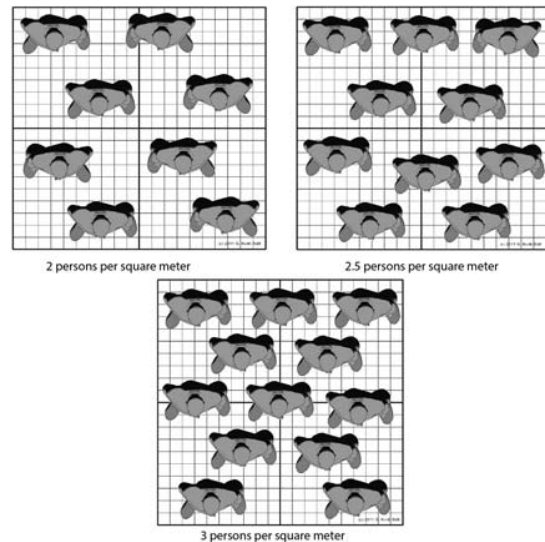


Figure 4. Crowd density variation per square feet.

#### 4.7. Efficiency

On average, a man's walking stride is 30 inches while a woman's average stride length is 26.4 inches. According to wrist-worn IMU, the authors have recently developed a method that uses the arm movements during walking to estimate walking speed. They found that the accuracy of human walking is about 5cm per second [13]. Accordingly, precise time which takes to walk ten meters can be calculated as follow.

Table 1. Average crossing time.

Gender	Human step length	Time takes for step	Length walked	Time takes
Male	30 inches	15sec	10meters	3.28 min
Female	26.4 inches	13.4sec	10 meters	3.33 min

Other than this disabled people and elderly people take more time to cross over line ten-meter zebra line. Probably it can take nearly four and a half minutes. So

willingly or unwillingly drivers have to stop their vehicles until they pass. But using this carriage this time can be minimized by nearly one minute. Therefore, it would obviously be efficient than walking over the zebra line.

Lanes of the road -3 lanes.  
Width of the road = 10 m  
Average speed of the carriage = 18 kmph = 5 m/s  
Time takes to cross 10 m = 2 s

Width of the crossing distance = 10 m  
Number of passengers per round = 40

Average speed of a vehicle = 36 kmph = 10 m/s  
Number of crossings per hour = 20 (every 4 min)

It can save 18 sec from one crossing  
Time takes to cross 10m without carriage = 20s  
Time takes to cross 10m when using carriage = 2 s

Number of extra vehicles passes through per hour  
=  $3 \times 20 \times 18 = 1080$  additional vehicles per hour.

Number of passengers cross through per hour  
=  $40 \times 20 = 800$

**Figure 5. Calculation which emphasize efficiency.**

## 5. Experimental results

At experiment level, we have developed a prototype by using Arduino based guided vehicle which follows a line sequence as the first stage. Then the carriage is being designed to run on top of a rail, which is currently ongoing. This carriage is integrated with ultrasound sensors to identify the vehicle when passing by the single lights. The designed carriage is programmed to reach the destination in a few seconds. Further on image processing mechanism is proposed to identify the obstructions which are lying in front of the carriage.

## 6. Conclusion

The proposed solution can minimize the fatal accident damages because the exposure of direct impact with a vehicle can be avoided by the carriage's safety shell. This will ensure the security of pedestrians. This carriage also helps to minimize the traffic flow by reducing the vehicle line holding time. Implementing these carriages is also cost-effective, because building an overhead bridge or underground tunnel is far more expensive than the proposed mechanism. So this can be introduced as a practical and low budget solution for both traffic congestion and fatal accidents. It ensures the

pedestrian's safety since the carriage can be protected using airbag system which is triggered by accidental impact with a vehicle. A strong frame rail can be used as a security mechanism which could withstand even a critical impact. This carriage is a comfortable transportation method for pedestrians even under any kind of weather condition because of its shell. Furthermore, proposed application can be used to collect the data sets of traffic flow and pedestrian behavior in order to use data mining and big data analytics as well. Efficiency depends on the synchronization among the nearest carriage systems. The synchronization can facilitate vehicles to move on uninterruptedly to their destination. Moreover, pedestrians can cross the road in a few seconds without any delays. Thereby, this application is a topical suggestion for pedestrian safety and efficient traffic control.

## 7. References

- [1] Fanghua , X., and Wenghong , L. F. L., "Desing of Inner Intelligent Car Parking Syatem," in *Proceedings of the International Conference on Information Management, Innovation Management and Industrial Engineering*, 2008.
- [2] Balta, M., and Ozcelik, I., "Traffic Signaling Optimization for Intelligent and Green Transportation in Samart cities.," in *Proceedings of the International Conference on Smart City and Emerging Technology (ICSCET)*, Mumbai, India, 2018, pp. 1-6.
- [3] Yong, L. and Xiuchun, G., "Research of Pedestrian Traffic Characteristics in University Campus," in *Proceedings of the Fourth International Conference on Digital Manufacturing & Automation*, Qingdao, China, 2013, pp. 1022-1024.
- [4] Dissanayake, H. and Koggalage, R., "A Cost Effective Intelligent Solution to Reduce," in *Proceedings of the Seventh International Conference on Wireless and Optical Communications Networks - (WOCN)*, Colombo, Sri Lanka, 2010, pp. 1-6.
- [5] Li, Z., Hong, J., Han, D., and Zhang, L. Z. R., "Research on Characteristics of Pedestrian Traffic and Simulation in the Underground Transfer Hub in Beijing," in *Proceedings of the Fourth International Conference on Computer Sciences and Convergence Information Technology*, Seoul, South Korea, 2009, pp. 1352 - 1357.
- [6] Jeong, Y., J., Lee, S. W. and An., D. S. K., "Smart Crossing System using IoT," in *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2017, pp. 392 - 393.
- [7] Tabish, S., and Kumar, E. M., "Research paper on study of pedestrian crossing behavior analysis at intersection," *International Journal of Latest Research in Science and Technology*, 2017, pp. 43-47.
- [8] Zhang, X., Skinner, N. P., and Bullough, M. S. R. J. D., "Evaluation of visual performance from pedestrian crosswalk lighting," in *Proceedings of the Transportation Research*

Board 91st Annual Meeting, Washington DC, United States, 2012.

[9] Jeong, Y. J., Lee, S. W., and KwangEun, D. S. A., An, "Smart Crossing System using IoT," in *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE)*, 2017.

[10] Kadzim, N. H. B., "A Study on Effectiveness of Pedestrian Bidge Utilization," 2012.

[11] Purwadi, A., Desmana, T.R., and Haroen, Y., "Dynamic response analysis of Permanent Magnet Brushless DC Motor drives for city electric car based on Matlab/Simulink," in *Proceedings of the International Conference on Rural Information & Communication Technology and Electric-*

*Vehicle Technology (rICT & ICeV-T)*, Bandung, Indonesia, 2013, pp. 1-5.

[12] Thai, V. X., Rim. C. T. and Kim, J. H., "Free space wireless charging in cars with fishbone DQ IPT coils," in *Proceedings of the IEEE 2nd Annual Southern Power Electronics Conference (SPEC)*, Auckland, New Zealand, 2016, pp. 1 - 6.

[13] Zihajehzadeh, S., and Park, E. J., "A Gaussian process regression model for walking speed estimation using a head-worn IMU," in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo, South Korea, 2017.



## Analysis of factors influencing the Virtual Learning Environment in a Sri Lankan higher studies institution

Rajaratnam Charanya  
 Department of Physical Science  
 Vavuniya Campus of the University of Jaffna  
 Vavuniya, Sri Lanka  
 charanyarj@gmail.com

Manoharan Kesavan  
 Department of Construction Technology  
 Wayamba University of Sri Lanka  
 Kuliypitiya, Sri Lanka  
 kesavan@wyb.ac.lk

### Abstract

*A project is commonly acknowledged as a successful project when the aim of the project is achieved positively. A system called Virtual Learning Environment (VLE) was designed among the students and university academic staff to encourage a positive approach in knowledge achievement and support active learning within the university. This study was carried out to analyze the factors influencing the VLE system and explore the relationship between the students and university academic staff on the system. The factors influencing VLE were identified through the literature review and the interviews which were conducted among the university academic staff and the industry experts. A paper-based questionnaire survey was carried out among the students and university academic staff in order to measure the severity of the factors influencing the VLE system. The respondents chosen for this study were the undergraduate students and university academic staff from Vavuniya Campus of the University of Jaffna, who used the above created VLE system. There were 120 responses from the students and 30 responses from the university academic staff. The students and university academic staff were requested to indicate their level of contribution on various factors in the survey questionnaire with a 5-point Likert scale and the Relative Importance Index (RII) was calculated for each factor. The severity of each factor was identified based on its RII value. The factors were ranked based on their severity and Spearman's rank correlation coefficient was calculated. It was found that there was 26.9% of positive degree of agreement between the students and university academic staff on the factors influencing VLE. This paper also explores some recommendations to improve the usage of VLE system.*

**Keywords:** *Virtual Learning Environment, Learning Management Systems, Learning Tools*

### 1. Introduction

One of the most important problems that may arise in the university is the interface between the student and the university academic staff member is not much fair. In order to maintain a good relationship in teaching and learning activities among students and university academic staff, a system called VLE (Virtual Learning Environment) can be used.

VLE is defined as an alternative of the classical learning environment. It can be used as a virtual source for learning activities. A user can log on to the system through the supporting applet interface. If the user uses the system for the first time, he/she should register by filling a registration form in the provided login interface. After the registration takes place, the profile of corresponding student or university academic staff is updated to the database. Then the users are enabled to access their personalized learning course materials from anywhere. VLE enables the structure of user login, admin panel, post students' problem, course materials, quizzes, chatting tools, calendar view, reminder notifications, comments and short message service for important posts.

Within a span of a decade, an LMS (Learning Management Systems) has moved from a support tool to the learning process. All learning management systems are not alike; they can be used in numerous ways. However, a common idea behind the virtual learning is that e-learning is organized and managed within an integrated system. Different tools are integrated in a single system which offers all necessary tools to run and manage e-learning activities. All learning activities and materials are managed and organized within the system.

### 2. Literature review

Frasen [5] studied on the critical success factors to enhance the quality of web supported learning opportunities in a blended learning environment in higher education. The factors were organized into six categories and determined how these factors are affecting the performances. In order to evaluate the effectiveness of LMS, one needs to compare student engagement,

motivation and performance in their learning environment. The review of factors and related researches indicate that these two aspects are an essential part of student web supported learning.

Sanders et al. [2] found that student attitudes toward the internet and web based courses can influence the future use of the web based instructional materials and how educationally beneficial web based resources are there for students. Instructor attitude toward web based instruction also affect student's experiences with web based courses.

In addition to web based courses consisting of an instructor-created web-site, many instructors are now using proprietary LMS as course delivery platforms. Wernet et al. [3] reported on a survey examining the satisfaction 32 levels and perceptions of 39 social work students regarding the use of WebCT in social work education course. All students responded that they found the course materials on the course website helpful.

Mafuna et al. [4] carried out the study through the questionnaire survey and it was found that students were ignorant about the new technology in their university. It was found that the availability of resources is very important in the acceptance of LMS and it should be prioritized and utilized effectively.

Swan et al. [6] focused the factors affecting the success of online learning through an investigation of relationships between student perceptions and course design factors in one of the largest learning networks in the city. It was found that three factors such as consistency in course design, interaction with course instructors and active discussion have been consistently shown to significantly influence the success of online courses. It was suggested that the reason for these findings relates to the importance of building community in online learning.

### 3. Objectives

The following objectives are developed to achieve the aim of the research.

- To identify the significant factors influencing the VLE system
- To measure the severity of the factors influencing the VLE system
- To determine the degree of agreement between the students and university academic staff on VLE system

### 4. Methodology

The VLE system was designed as a web version according to the requirements of the students and the university academic staff of Vavuniya Campus of the University of Jaffna. Knowledge areas needed for

designing this project were Software Engineering, Web Engineering, UI/UX Engineering, and Database Management System. PHP, AJAX, JavaScript, CSS and HTML were the languages used to design VLE system. Laravel 5.2.4, JQuery and Bootstrap were the frameworks used for the development of the system. MYSQL was used for the data vendors. The database and the application both were installed on a server. Passwords were saved on the database by encrypting Bcrypt method. Each and every activity can be backed up on a separate cloud outside the server.



Figure 1. User interface.



Figure 2. VLE home page.

After the completion of VLE system, the interviews and discussions were conducted among the university academic staff and the industry experts to check the performance of the system. The suggested modifications were made on the system and it was given to the university to implement for the academic usage.

The significant factors influencing VLE were identified as follows.

- System quality
- Appropriate use of technology
- Reliability
- Availability of usage
- Problem solving

- Assignment strategies
- Learning resources
- Self-directed learning
- Frequent feedback
- Academic background
- User friendliness
- SMS alert notification
- Reminder
- Discussion
- Download demands
- Flexible Learning
- Sustainability
- Time saving
- Collaborative learning
- Use of media
- Evaluation of teaching competence
- Infrastructure
- Motivation of learning activities
- Ease of assignment submission
- Learning practice

A questionnaire was designed to measure the severity of the factors influencing VLE. The questionnaires were distributed to the students and the university academic staff members who used the above created VLE system. The students and university academic staff were requested to indicate their level of contributing on the above various factors in the survey questionnaire with a 5-point Likert scale type as a measurement from 1 to 5 (very low effect to very high effect).

There were 120 responses received from the students and 30 responses received from the university academic staff members. The collected data was analyzed using MS Excel. This data analysis was used to determine the relative importance of the various factors influencing VLE. The following steps were followed in the analysis of data.

- Relative Importance Index (RII) was calculated.

$$RII = \frac{\sum W_i X_i}{\sum X_i}$$

Where:

i - Response category index

w<sub>i</sub> - Weight assigned to i<sup>th</sup> response =1, 2, 3, 4, 5 respectively

x<sub>i</sub> - Frequency of the i<sup>th</sup> response given as percentage of the total responses for each factor

- The factors were ranked in each category based on their Relative Importance Index (RII)

Kesavan et al. [7] Spearman's rank correlation is a non-parametric test. Correlation is a relationship measure among different parties or factors. This correlation is computed by the following formula.

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Where:

r<sub>s</sub><sup>1</sup> - Spearman rank correlation coefficient between two parties

d - Difference between ranks assigned to variables for each cause

n - Number of pairs of rank

This correlation coefficient is used to show the degree of agreement between the different parties. It varies between +1 and -1, where +1 implies a perfect positive relationship (agreement), while -1 results from a perfect negative relationship (disagreement). The value near to zero indicates little or no correlation. In this study, this correlation coefficient was determined to find out the degree of agreement between the students and university academic staff.

## 5. Results

The data was analyzed using Microsoft Excel. All the factors influencing VLE system were ranked based on the Relative Importance Index (RII). Ranking of factors influencing VLE system based on the responses from the students shows in Table 1.

**Table 1. Ranking of factors influencing VLE system based on the responses from the students.**

Factors	RII	Rank
Infrastructure	4.60	1
Time saving	4.56	2
Collaborative learning	4.50	3
Learning practice	4.45	4
Use of media	4.37	5
SMS alert notification	4.35	6
Reminder	4.26	7
Motivation of study	4.25	8
Reliability	4.24	9
User friendliness	4.23	10
Frequent feedback	4.20	11
Availability of usage	4.10	12
Problem solving	4.07	13
Ease of assignment	4.06	14

submission		
Academic background	3.98	15
System quality	3.89	16
Appropriate use of technology	3.77	17
Self-directed learning	3.57	18
Download demands	3.21	19
Sustainability	3.20	20
Evaluation of teaching capability	3.11	21
Flexible learning	3.09	22
Learning resources	3.07	23
Discussions	3.02	24
Assignment strategies	2.76	25

Ranking of factors influencing VLE system based on the responses from the university academic staff shows in Table 2.

**Table 2: Ranking of factors influencing VLE system based on responses from university academic staff.**

Factors	RII	Rank
Collaborative learning	3.43	1
Learning practice	3.35	2
Evaluation of teaching capability	3.26	3
Flexible learning	3.25	4
Frequent feedback	3.24	5
Availability of usage	3.22	6
Reminder	3.20	7
User friendliness	3.15	8
Infrastructure	3.12	9
Motivation of study	3.08	10
System quality	3.05	11
Use of media	3.04	12
Self-directed learning	3.03	13
Download demands	3.02	14
Academic background	3.02	15
Ease of assignment submission	3.01	16
Reliability	2.96	17
Discussions	2.91	18
Appropriate use of technology	2.80	19
Sustainability	2.78	20
Problem solving	2.66	21
Learning resources	2.48	22
Time saving	2.44	23
Assignment strategies	2.03	24
SMS alert notification	2.02	25

The Spearman's rank correlation coefficient is applied to measure the degree of agreement or disagreement associated with the importance rankings. The following table shows the ranking of factors and the Spearman's rank correlation coefficient was derived.

**Table 3: Summary of ranking factors influencing VLE.**

Factors	Students	Lecturers	d <sup>2</sup>
	Rank	Rank	
System quality	16	11	25
Appropriate use of technology	17	19	4
Reliability	9	17	64
Availability of usage	12	6	36
Problem solving	13	21	64
Assignment strategies	25	24	1
Learning resources	23	22	1
Self-directed learning	18	13	25
Frequent feedback	11	5	36
Academic background	15	15	0
User friendliness	10	8	4
SMS alert notification	6	25	361
Reminder	7	7	0
Discussions	24	18	36
Download demands	19	14	25
Flexible learning	22	4	324
Sustainability	20	20	0
Time saving	2	23	441
Collaborative learning	3	1	4
Use of media	5	12	49
Evaluation of teaching capability	21	3	324
Infrastructure	1	9	64
Motivation of study	8	10	4
Ease of assignment submission	14	16	4
Learning practice	4	2	4

$\Sigma d^2$

1900

Using the Spearman's rank correlation coefficient, the degree of agreement or disagreement between the students and university academic staff was determined.

Here,  $\Sigma d^2 = 1900$  and  $n = 25$   
 $r_s I = 1 - [6 \times 1900] / [25 (25^2 - 1)] = 0.269$

It was found that there was a positive relationship between the students and university academic staff on the

VLE system. The results present that 26.9% of degree of agreement is there between them.

## 6. Conclusion

The students stated that infrastructure, time saving, collaborative learning, learning practice and use of media are the most significant factors influencing the VLE system. In the students' point of view, assignment strategies, discussions, learning resources and flexible learning were the less influencing factors. Collaborative learning, learning practice, evaluation of teaching capability, flexible learning and frequent feedback were identified as the most critical factors influencing the Virtual Learning Environment system for the university academic staff. SMS alert notification, assignment strategies, time saving, learning resources, problem solving, sustainability and appropriate use of technology were found as very less importance for the university academic staff. It was observed that there was a positive relationship between university academic staffs and students. It was 26.9% of positive degree of agreement found between them.

## 7. Recommendation

This paper explores some recommendations to improve the usage of VLE system in order to improve the VLE system.

So, the system has to be developed in order to attract the students to put more involvement in assignments, discussions and learning resources. The system can be modified to send a sticker, when the students undergo assignments and discussions. Then the students will do the assignments and the discussions as a competition by collecting the high number of stickers. This will lead the students to involve in assignments and discussions with high attention level. It is suggested to add a game facility which has a limited time with a set of levels and each level should be unlocked, when the students undergo assignment and discussion activities. This system can be linked with the library facilities to add more learning resources.

In addition, it is highly recommended to arrange any training facilities or workshops for the students and the university academic staff separately, by considering their academic activities within the university. It is important to give the basic training to use this type of VLE systems effectively.

## 8. Acknowledgement

The authors express gratitude to the undergraduate students and the academic staff members of Vavuniya

Campus of the University of Jaffna for providing valuable support during this study.

## 9. References

- [1] Farman Ali Khan, Sabine Graf, Edgar R. Weippl, and A Min Tjoa, "An Approach for Identifying Affective States through Behavioral Patterns in Web-based Learning Management Systems", in *Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services*, iiWAS'09, 2009, pp. 431-435.
- [2] Diana W. Sanders and Alison I. Morrison-Shetlar, "Student Attitudes toward Web-Enhanced Instruction in an Introductory Biology Course", *Journal of Research on Computing in Education*, 33:3, 2001, pp. 251-262.
- [3] Wernet S. P., Olliges R. H. and Delicath T. A, "Postcourse evaluations of WebCT (Web Course Tools) classes by social work students", *Research on Social Work Practice*, 2000;10(4), pp. 487-504.
- [4] M.Mafuna and N.Wadesango, "Students' Acceptance and Experiences of the New Learning Management System (LMS) -Wiseup", *Anthropologist*, 14(4), 2012, pp.311-318.
- [5] Frasen, J.W., "Narrative Inquiry: a Research Method and Process Enabling Life-long Learning Influencing Workplace Communities and Cultures", in *Proceedings of the 37th Annual SCUTREA Conference*, 3-5 July 2007, Queen's University Belfast, Northern Ireland, 2007
- [6] Karen Swan, Peter Shea, Eric Fredericksen, Alexandra Pickett, William Pelz and Greg Maher, "Building Knowledge Building Communities: Consistency, Contact and Communication in the Virtual Classroom", *Journal of Educational Computing Research*, Volume: 23 issue: 4, 2000, pp. 359-383.
- [7] Manoharan Kesavan, Navanesan Gobidan and Pujitha Dissanayake, "Analysis of Factors Contributing Civil Engineering Construction Project Delays in Sri Lankan Building Construction Industries", *Journal of Industrial Engineering Research*, 1(7) Special 2015, pp. 5-11.
- [8] Steel C., "Reconciling university teacher beliefs to create learning designs for LMS environments", *Australasian Journal of Educational Technology*, Vol. 25, 2009, pp. 399-420.

## Predictive Maintenance 4.0 as next evolution step in industrial maintenance development

P. Poór  
*Department of Industrial  
 Engineering*  
*University of West Bohemia*  
*Pilsen, Czech Republic,*  
*30100*  
*poorpeter@gmail.com*

J. Basl  
*Department of Industrial  
 Engineering*  
*University of West Bohemia*  
*Pilsen, Czech Republic,*  
*30100*

D. Zenisek  
*Department of Industrial  
 Engineering*  
*University of West Bohemia*  
*Pilsen, Czech Republic,*  
*30100*

### Abstract

*The aim of this article is to acquaint the reader with problematics of industrial maintenance. The first part of the article deals with the historical overview of maintenance itself. There are different definitions and approaches to maintenance, some of them are presented with the most used maintenance types of breakdown maintenance, preventive maintenance, predictive maintenance, RCM, and TPM. The paper proposes opportunities for “new” type of maintenance connected with Industry 4.0, namely Predictive Maintenance.*

**Keywords:** *Industry 4.0, predictive maintenance, industrial maintenance*

### 1. Historical overview of industrial maintenance developments

From the point of view of needs and prosperity of operational practice, machines and equipment have a central position in each enterprise. In terms of competitiveness in a global world, businesses need to synergistically develop all areas, processes and systems that are being implemented. The unstructured development of these components helps to advance and develop competition. For this reason, the problems of maintenance of machinery and equipment also appear to be important.

Maintenance history is also a history of culture and an image of industrial concept and philosophy. Effective and effective maintenance is more than ever essential for further development, especially in times of financial crisis and reduced turnover, rising costs of energy, raw materials and labor.

At present, over € 1,500 billion is spent a year on maintenance, repair and renovation (MRO) only in the EU and over € 7,000 billion globally. More than 50 million jobs are directly linked to the MRO and 150

million indirectly. However, the requirements for MROs are increasing and, especially in times of financial crisis and reduced budgets, are difficult to meet. Regulations and legislation are tightened up and meet demanding requirements, not to mention the demands of customers but the demands of the latest technology and considering economic and environmental needs and principles. However, modern industry is unable to work efficiently without fast, modern and innovative tools for maintenance of tools and equipment.

#### 1.1. Definition of maintenance

The objective of the paper is also to devise a strategy for Industry 4.0 with respect to maintenance. First of all, present theoretical aspects of machinery renewal and maintenance are presented through this paper.

In literature there are many different definitions about Maintenance: Swedish standard SS-EN 13306 define it as a “Combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function” (SS-EN 13306, 2001, p.7).

Another definition in literature about the purpose of maintenance is: “The main purpose of maintenance engineering is to reduce the adverse effects of breakdown and to increase the availability at a lower cost, in order to increase performance and improve the dependability level” (Simeu-Abazi and Sassine, 2001, p.268).

Gits (1994) thinks about industrial maintenance like a process that the production process needs; it is the primary process where input is transformed in output and Maintenance is a secondary process that helps the first to the achievement of production.

The term maintenance according to [1] means all measures that help to preserve and recreate the required state of machinery and equipment. They lead to the identification and assessment of the actual state of the technical installations as a whole and the subsequent

technical measures to restore all its functions in the required quality.

One of the most important aspects of well-organized production is the maintenance of machinery and equipment. In companies, we often meet with the approach: "Because maintenance works well when the fault is, it will always fix it quickly." If they are already planning on maintenance, it is more of a "proform" maintenance where production workers are not interested in their machinery and equipment. The error is also that economic records are not kept on the value of maintenance and repair activities and these items are expensed in overhead costs.

One of the many ways in which maintenance is divided is where we see whether maintenance is performed before or after failure. Maintenance performed prior to the occurrence of a failure prevents the occurrence of delays. We also differentiate between proactive maintenance (dependent on device status) and preventative (performed independently of the state of the device). After-fault maintenance serves mainly to return the machine to a state in which it can continue to perform the required functions. We also differentiate this kind of maintenance to be moved (moved according to the maintenance rules) and immediate (performed immediately after the failure is detected).

## 1.2. History of maintenance:

Maintenance activities are not always considered "the necessary evil". Up to a high maintenance cost in the range of 15 to 40% of manufacturing costs has led managers to raise their attention to maintenance management issues. The development of maintenance can only be recorded after World War II.

Production machines, equipment and tools will always be subjected to wear and the need for maintenance. From the first industrial revolution, the maintenance was done only when the machines stopped or they could not work, is the type of maintenance called Corrective Maintenance. A lot of parts are damaged by the application of strains and so it is possible to understand approximately their duration and so it is possible to fix, before the fault, maintenance intervals. In this way one company needs to find a tradeoff between these types of maintenance: the fault should increase costs for delayed shipment, but, at the same time, too much maintenance means also big costs. In the last years, for these two reasons, the company aim is to introduce a maintenance strategy that can understand the need of maintenance before a breakdown. The arrival of industry 4.0 introduces and develops a lot different methods to do maintenance.

In Industry 4.0, maintenance costs are considered value creator. On the contrary, the entrepreneurs, in the

past, considered them like a cost that must be reduced because the trend was to repair machine or equipment only when they stopped.

But let's start from the beginning.

### 1.2.1. Breakdown maintenance

The strategy of "letting the device work until it goes wrong" was related in the period up to 1945, and partly up to the 50s, with the problem of data collection and evaluation as well as high demand in the markets. Maintenance work was aimed at repairing damaged parts and simple activities such as cleaning and lubrication. Reactive maintenance therefore only responds to machine and machine failures - the machines work, and when a fault occurs, it interferes with the maintenance department and removes it.

In order to prolong the lifetime proposed by the designers, no special action or effort has been developed. Studies have shown how the devices are on average maintained and we see that reactive maintenance represents the largest percentage:

- > 55% reactive maintenance
- 31% preventative maintenance
- 12% predictive maintenance
- 2% others.

More than 55% of the means and maintenance activities of an average facility are reactive. The benefits of reactive maintenance can be presented as a two-sided weapon. If the equipment is new, we can expect minimal failures. If the maintenance program is purely reactive, we will not waste labor and do not incur capital costs until something goes wrong.

Since maintenance costs are zero, this period is a money saver. But reality is about something else. At a time when we are convinced that we are saving maintenance and capital costs, we actually spend much more than we would use a different approach to maintenance. Capital expenditures are higher because we expect that the equipment will fail, thus shortening its lifetime, which leads to a higher frequency of substitution. Additional costs associated with the failure of the primary device will arise, causing the secondary device to fail. These increased costs do not occur in more proactive maintenance. The wage costs associated with the repair are likely to be higher than normal, because failure will require more extensive repairs than if part of the device was not in operation until full failure. If the device is important and has to be immediately back on site, additional costs will be incurred. As we expect equipment failure, we will need large supplies of spare parts.

### 1.2.2. Preventive maintenance

The post-war period and the boom of production meant a breakthrough in monitoring and evaluating data as well as in product quality requirements. Increased pressure to reduce losses and the need for preventive maintenance of equipment, failure was no longer a natural occurrence. Reducing downtime and requirements for increasing device reliability have led to planning and scheduling of maintenance activities [17]. The so-called "planned preventive maintenance".

Preventive maintenance can be defined as follows: Action based on a specific timetable that detects, avoids or mitigates the deterioration of the component or system state in order to maintain or extend its life via controlled degradation to an acceptable level [3]. There are two basic types of preventive maintenance - maintenance in periodic cycles or maintenance based on device status. Maintenance in periodic cycles appears to be too expensive for about 92% of machine components. Device-based maintenance interchanges components and interferes with the machine only when abnormalities start to appear in its process, making it more economical.

By simply spending the necessary resources to carry out the maintenance activities of the designer, the facility's life is extended and its reliability is increased. In addition to increasing reliability, savings are also being made. This savings can be up to 12% on average to 18% of saved costs.

Depending on the current state of maintenance, device reliability and downtimes, many companies dependent on purely reactive maintenance that can save much more than 18% by starting the right preventive maintenance program.

Although preventive maintenance is not an optimal machine maintenance program, it has several advantages over purely reactive maintenance. Performing preventive maintenance on the device as designed by the designer will prolong the life of the device. This means saving money. Preventive maintenance (lubrication, filter replacement, etc.) will usually result in higher device efficiency, which will be reflected in savings. Even if we do not prevent the catastrophic scenario, the number of disorders will go down.

### 1.2.3. Predictive maintenance

Predictive maintenance is defined as follows: Measurements that detect the initiation of degradation of the mechanism and thus allow the problem originators to be removed or controlled before each significant deterioration of the device condition.

Predictive maintenance differs from preventive by staying maintenance needs on the current status of the

device rather than on a programmed schedule. Preventive maintenance is time-based.

The benefits of predictive maintenance are many. Well-organized preventive maintenance eliminates everything except catastrophic failures. It can be planned to minimize or eliminate the cost of overtime work. It is possible to minimize inventory and order the necessary parts. We can optimize device operation, save energy costs, and increase device reliability.

A well-functioning predictive maintenance program can mean savings of 8% to 12%. Depending on the equipment and material conditions, it is possible to save 30% to 40%.

The following savings resulting from the use of predictive maintenance are:

- Return on investment: 10 times
- Reduction of maintenance costs: 25% to 30%
- Troubleshooting: 70% to 75%
- Reduction of downtime: 35% to 45%
- Increased production: 20% to 25%.

But the start of predictive maintenance is not cheap. A large part of the equipment requires a cost of more than € 30,000. Staff training also requires additional funding.

### 1.2.4. Reliability centered maintenance

Reliability-based maintenance is "a procedure to establish maintenance requirements for any physical asset in its operational context." The reliability-oriented maintenance methodology addresses basic issues not covered by other maintenance programs. Recognizes that not all equipment in an enterprise has the same importance, which the construction and operation of the equipment is different and that different equipment is more likely to cause a fault for different reasons. It also takes into account the fact that the facility does not have unlimited financial and personnel resources and needs to be identified and optimized.

Reliable operational maintenance is a systematic approach to assessing equipment and resources in order to best match them, resulting in high device reliability and cost efficiencies. Reliability-based service maintenance is strongly dependent on preventive maintenance but also takes into account that maintenance activities on cheap and insignificant equipment are better left to reactive maintenance.

Since reliability-oriented maintenance is of great importance in using predictive maintenance technology, its advantages and disadvantages are reflected in those of preventive maintenance. In addition to these benefits, reliability-based maintenance will allow plants to better manage resource availability while improving reliability and lowering costs.



The RCM strategy monitors the operation of each element and defines the consequences of its failures. The RCM creates a structure of consequences in decreasing order of severity of individual disorders. When determining the consequences, all the activities of the elements of the monitored device must be specified. If the level of risk due to failure cannot be reduced by the selected maintenance mode, then it is necessary to reconstruct the element. Thus, the RCM also deals with the assessment of possible causes of device failures (e.g. neglected maintenance, wear, etc.).

The basic tools used by this strategy to assess the risks of failure include:

- FMEA - Failure Mode and Effect Analysis - Analysis of the ways and consequences of faults.
- Analysis of causes and consequences.
- Risk analysis.

### 1.2.5. Total Productive Maintenance (TPM)

Total Productive Maintenance - TPM - is the state-of-the-art philosophy, organization of machine maintenance in the production process, with a significant share in increasing productivity and eliminating costs. It is a set of maintenance activities performed throughout the lifetime - the existence of machines in order to improve their accuracy, reliability, performance and efficiency as well as reducing all possible losses. It is part of the company's production philosophy; it includes all the departments of the company and represents the interconnection of maintenance and production with the technical security of the construction and technology. [8]

TPM is a result-oriented process which provides a methodology for data collection for analysis, troubleshooting and process management. It must be managed by factory employees and is supported by process and maintenance partners so that they work together and are equal partners. It is just users and operators of machines and equipment who are able to capture information about the condition of the device and prevent its malfunction. No proper maintenance can work properly. Their activity is supported by construction and projection staff, quality, logistics, purchasing, or others who have a relationship with the production facility and the process. The TPM includes every employee from the executive to the top manager. The TPM includes a process of continuous improvement of equipment and overall quality. This dynamic process improvement has not yet been overcome. It is an enterprise-wide system that includes preventive maintenance. The TPM is based on the support of product maintenance by small group activities (production teams). The TPM is applicable everywhere; where production (operation, equipment) is

based on technological service (operators). TPM goals are [3]:

- zero downtime,
- zero errors,
- zero disturbances.

The basis of the TPM philosophy is, on one hand, to increase the efficiency of equipment and to reduce maintenance costs and losses due to downtimes, on the other hand, to retain good communication between the operator and the maintainer.

The author of the TPM is Seichi Nakajima, who in the 1950s and 60s studied Preventive Maintenance systems in the US and Europe. He worked out his knowledge in a complex design that was given a working name, (Figure 3) Total Productive Maintenance. In 1971, he introduced the system into Japanese businesses [9].

T Total - with the participation of all employees of the organization.

P Productive - the efficiency of maintenance and production equipment, constantly improving.

M Maintenance - maintenance of machinery and equipment in good technical condition.

Total productive maintenance is geared to engaging all staff in the workshop to activities that minimize downtime, minimize accidents, and minimize incidents. The TPM is about overcoming the traditional division of people into workers working on the machine, and "workers who repair it". It is based on the fact that the employee who manages the machine has the chance to first capture the abnormalities in his work and possible sources of future equipment failures. Motto of TPM is: "Protect your machine and take care of it with your own hands." Thus, the maximum diagnostic and maintenance activities in the TPM are transferred from the classical maintenance department directly to the production workers of the production sections. It usually starts with improving the workplace, cleaning machines and checking their condition. Furthermore, the operator learns to "understand his machine", to learn how to behave as his "own device". In addition to maintainers and operators, other professions are involved in the TPM, such as technical preparation of production [12].

The essence of the TPM method is to achieve better use of machinery and equipment, improving maintenance work. In practice, this system is based on the principle of early detection of abnormalities arising from accidental machine operation and removal. One of the progressive elements of the TPM is to shift part of the machine responsibility directly to the machine operator, based on the fact that operators are in daily contact with the machine and therefore can quickly identify the resulting abnormality as the potential basis for a fault or breakdown. However, the implementation of the method

itself is based on teamwork and change of staff thinking. The shift in thinking has to be done mainly in optimizing the "man-machine" relationship, where the operator performs not only in the role of the machine operator but also in the role of an active maintenance co-worker. It is absolutely essential that the entire system is covered by active management support for TPM implementation projects and the collaboration of technical staff throughout the company [12].

Main TPM Goals:

Two approaches have to be accepted to achieve the ideal conditions [13]:

1. Design of optimum conditions for human-machine system. To make the most efficient use of the device, it is necessary to know the ideal assumptions for the operation of each machine component as well as the values that represent the top performance of the machine. It is the duty of man in the man-machine system to maintain these prerequisites; the most important element in this system is still human. The system must be modified by him.

2. Improve the overall quality of the work environment:

- changing the behavior of people,
- change of equipment - the position of employees changes with it.

On the basis of the above, it can be determined that such a way of maintenance is nowadays retroactive and economically disadvantageous. Getting the so-called Greater skill in maintenance work is by creating and deploying an acceptable maintenance method that is more cost-effective, location-specific and with specific individuals in operation. These activities are being only a part of the overall TPM activity. The final state must have feedback on the construction of the equipment in the form of new maintenance requirements [3].

Training and training of employees is one of the basic elements of the TPM, which supports all the individual parts of the program. The operator is gradually engaged in machine care and takes over certain maintenance activities. The operator must gradually learn in specialized areas of machine care and become more qualified to address urgent small problems immediately.

Repairs may be involved in machine rebuilding, and its evolving skills can also lead to further self-employment, such as machine design improvements or machine-readiness arrangements. Implementation of TPM takes place in phases that are based on employee training, improvement of work, machine and workplace knowledge [9].

The first steps of the TPM are to build the basic conditions for effective detection of the specifics, namely the cleanliness of the workplace, the order and organization of the workplace. Further steps are being taken to build the ways in which the operator handles

some less complex and time-consuming machine maintenance elements, to organizing the workplace. The TPM plan must be supported by the organization of expert workshops aimed at eliminating the root causes of pollution, simplifying the cleaning and checking of the state of the machinery and equipment. The core of the program is the phase of mastering the field of stand-alone maintenance by operators, machines and devices. Implementation of the TPM must stem from a long-term project based on precisely defined annual plans and their development into smaller activities for each operation. [12].

### 1.2.6. Implementation policies

The role of the TPM team lies in specifying the individual activities that ensure the maintenance of the required ratios and the determination of specific tasks in the procedures that will be maintained by the maintainers. These tasks are the core to determine preventive maintenance. The plan of the following activities (lubrication, cleaning, etc.) must be visibly located at predetermined locations with respect to the devices it touches. The list of activities required and their timetable must be simple and clear because it is intended for people who operate, set and maintain the machines.

## 2. Influence on Industry 4.0 on maintenance strategy

Maintenance strategy is defined in literature like a choice between corrective maintenance and various kinds of preventive maintenance. Industry 4.0 is, in a simple way, the relationship between physical assets and digital assets. It is totally based on efficiency and it needs that each single assets of each machine works as to the best possible. This is the reason because a lot of companies, that have introduced Industry 4.0, are in the need of developing a maintenance strategy that ensures a high level of efficiency and a low chance that failure or breakdown happen. Therefore, this innovating company started to adapt the machine, the employees and the plants and later started to develop a maintenance strategy.

In a world where there are a lot of companies that make the same products or services, it is very important to differentiate also some little details; Industry 4.0 brought the companies to be more and more competitive and efficient and this is the main cause of firms that are doing one-third; too much maintenance. However, more often the maintenance operation can be postponed with more precise information about the component.

Next part of the article summarizes all above and tries to show an approach with our contributions and say what is presented as "new" when dealing with modern maintenance management.

## 2.1. Formulation of maintenance strategy:

This is one of the most difficult formulation because there are only few models, some of these are quite resource demanding and they are rather similar in their key elements. The common steps in the model are: identify strategic goals of all stakeholders, identify the strategic goals for the maintenance department, identify relevant KPIs, assess the current state of maintenance, set the goals for each KPI and make an action plan. These models, even though, have this common key; they have different approach to develop the strategy.

One of the first model is by Visser (1998) and it composed by 4 strategic dimension of maintenance: choice between in-house capability and outsourced service, organization of the maintenance function and the way maintenance tasks are structured, selection of maintenance policies and the last is design of the infrastructure that supports maintenance.

In the maintenance strategy it is very important, by Wilson, that maintenance objectives and strategy, align with the business goals which should reflect: customer needs to understand the performance; the resources match with the strategy objectives; what the function capabilities and the people involvement, what chance the maintenance function must make and how quickly they can be implemented etc.

The maintenance in a factory concerns each asset and each physical or digital part, for this there are a lot of kind of maintenance that can be divided in two big groups: structural decision elements and infrastructure decision elements. In the first one there are maintenance capacity, maintenance facilities, maintenance technology and also vertical integration; in the second group there are maintenance organization, maintenance policy and concepts, maintenance planning and control system, human resources, maintenance modifications and maintenance performance measurements.

After the development of maintenance strategy, it involves the identification, resourcing, execution of repair, replacing and inspection of decisions and it is concerned with formulating the best plan for each unit, formulating a maintenance schedule for the plant and establishing the organization to enable the scheduled and unscheduled maintenance work to be resourced.

In the formulation of maintenance strategy the starting point is the maintenance objective, because, according to Kelly, the formulation of them is more difficult than this in reality; then the strategy is formulated in order to achieve the specified requirements at the minimum costs. This process is useful to understand the basis for maintenance budgeting and cost control. The second steps start with stating the maintenance philosophy; after that we should consider the aims and objectives of the maintenance.

The third step is to assess and to evaluate the maintenance practices and issues; the various policy sectors; from the assessment, a maintenance program may be developed. Then tactic has to be formed in order to integrate the new practices with the existing ones. The final step is set.

## 2.2. Maintenance 4.0 as a service

In the age of Industry 4.0, maintenance managers need additional technical skills and they have to be able to manage maintenance processes, to govern the impact that maintenance has on the other sectors, on its business goals, on quality, on efficiency and also on safety.

More and more resources are dedicated to the maintenance management and this lead a lot of companies to manage maintenance not only in direct way but also with remote control: now there is the possibility to acquire and process data with software and more important with low costs. There are two new trends: the ability to analyse the huge amount of data that assets and machines produce during their life and the ability to analyse the needs of the stakeholders, to contribute at the creation of added value in line with the needs of the company. The production processes will become more and smarter in the following years cause of the new technologies. Radio Frequency Identification (RFID), Supervisory Control and Data Acquisition (SCADA), Near Field Communication (NFC), Wireless Sensor Actuator Networks (WSAN) and Wireless Sensor Networks (WSN) are some of the new innovative technologies that support the diagnosis of faults and the maintenance condition with assets applied in the machines. With the transition to digital chains, the efficiency will increase and new business models will be possible through sensors and connectivity; the industrial products will be enriched with individual services, like maintenance, or transformed into a total integrated service, like Original Equipment Manufacturer (OEM). The improvement achieved is offering customer service for each area of business. The machine manufacturer that use Industry 4.0 started in the last year to increase their business: they, in fact, do not sell only machine but they offer to the customer also the Preventive Maintenance and all kinds of maintenance that their machines needs during the production processes. The maintenance is not only one internal activity but become a service that you can sell to your customers. For example, Skoda start to offer the maintenance service to its customers. In Italy the situation is a little bit different and only the big company started to set their business in line with Industry 4.0 principles.

### 2.3. Equipment maintenance

Another important role in Industry 4.0 shall be represented by equipment; with this term we suggest the entire machine but also all the tools that are essential for the success of production processes. The company, that wants to reduce the risk of unscheduled downtime, failure or breakdown, has to keep its equipment at the optimum working condition. It is important for equipment to keep a record about the maintenance speeches and the following are some benefits about it:

**Prevent expensive repair works from happening:** all kinds of equipment are prone to wear and tear and this should become failure or breakdown; performing routine inspections allow you to see and repair small damages before they become a big problem. Documenting these inspections and small repairs help you keep track of all the maintenance work that your equipment has undertaken, ensuring that each machine is in its best shape before putting them to work. This helps you create specialised maintenance programs: working conditions are different for each equipment and with help of routine check-ups, you will be able to determine and record the differences of each individual equipment with regards to maintenance works; but the other benefit is that this information will help you in creating maintenance programs.

**Prevent problems regarding warranty claims:** keep a record of the type of maintenance, the data and the hour of this will help determinate your right for warranty claims. Some other benefits are explained further below.

**Increase operators safety:** if the equipment is well-maintained, the risk of accidents will decrease; having maintenance history documented of equipment will help you keep track of the health of equipment.

**Keep track about responsibility of a piece of equipment:** in one machine usually there are more than one worker, performing a routine inspection and recording the situation and worker will help you to understand who is responsible about some injury but also to increase operators attention for the equipment.

**Increase the resale value of the equipment:** if you record each machine problems and so the maintenance on this machine, it will be easier to resell it at a higher price.

**Fixing big problems is expensive:** the bigger, more complex repairs come with a bigger price tag, to avoid this you do not think that small repairs are loss of money and time, but they are something that makes your equipment more efficiency.

### 2.4. How to extend equipment life:

**Monitor equipment technology:** many models have a variety of sensors in place that will tell you if something isn't working properly, even if it's something small.

Make sure someone is monitoring the technology and collecting data on equipment and performance; this technology is only valuable if someone is there to watch when a sensor detects something is not working right.

**Equipment monitoring:** There are also tools you can use to supplement the technology a particular piece of equipment has. For example, vibration monitoring, thermal imaging, audio gauges and other technologies and tests can help you collect additional information on your equipment so that you have real-time information on the condition of your construction equipment.

**Train operator and technicians:** there is no substitution for the human touch; it needs a trained operator to understand the problem and a trained technician to know how to fix it or to alert someone that is able to repair it.

**Make someone accountable:** the staff structure needs someone with accountability for equipment maintenance; empowering and educating someone to take on this role is crucial to the success of preventative maintenance, otherwise it will always get pushed aside.

**Set and stick a maintenance schedule:** every piece of equipment is different. They all have their own intricacies and need a maintenance and repair schedule to match. Over time, there are some seals, belts, bolts, etc., that will need to be replaced. Rather than waiting for these parts to cause a problem, replace them when they are scheduled to be replaced.

### 2.5. Outsourcing maintenance 4.0

The Business Dictionary defines Outsourcing like “the contracting or subcontracting of noncore activities to free up cash, personnel, time and facilities for activities in which a company holds competitive advantage. Companies having strengths in other areas may contract out data processing, legal, manufacturing, marketing, payroll accounting, or other aspects of their businesses to concentrate on what they do best and thus reduce average unit cost”. In line with the definition, in the last years, cause of the importance and costs of maintenance, a lot of companies prefer to outsource the maintenance processes in order to have better results and more efficiency and also to focus on the activities that they do best. There are many levels of outsourcing maintenance work and the following are the most popular:

**Outsourcing of the whole maintenance function:** usually this is done to create a separate maintenance company and it ensures reliability and services to the manufacturing organization;

**Outsourcing of work done during scheduled shutdowns and outages:** this method is not cost effective to not bring extra resources during peak loads;

**Outsourcing of special skills:** this methodology is used when the company does not need this kind of maintenance skills;

Outsourcing of equipment or a system function: when some companies decide to sell a guaranteed function at a fixed cost including operations and maintenance; guarantee includes reliability performance; the problems of this type is caused by abusive operations practices and this can make it difficult and time consuming to decide who will pay for these incidents.

Outsourcing of spare parts store: the possibility here are two or sell the store or outsource this service to a specialized company.

There are other types of outsourcing like there are different reasons about it, but the most important is that the supplier of outsourcing promises lower cost and better reliability, but sometimes it is not the truth. When you receive something you have always to ask yourself the reasons because the supplier wants to outsource maintenance skills.

**Table 1. Correlation of industrial revolutions and maintenance.**

Industry revolution	Industry 1.0	Industry 2.0	Industry 3.0	Industry 4.0
Characteristics of the industrial revolution	Mechanization, steam power, weaving loom	Mass production, assembly lines, electrical energy	Automation, computers, electronics	Cyber Physical Systems, IoT, networks, cloud, BDA
Type of maintenance	Reactive maintenance	Planned maintenance	Productive maintenance	Predictive maintenance
Inspection	Visual inspection	Instrumental inspection	Sensor monitoring	Predictive analysis
OEE	<50%	50-75%	75-90%	>90%
Maintenance team reinforcement	Trained craftsmen	Inspectors	Reliability engineers	Data scientists

### 3. Strategies for future and conclusions

We are in the Industry 4.0 era, where the competition between different companies is very hard and also the small details should make big differences. Concept of relatively new Industry 4.0 fully used the advantaged of cyber systems, cloud storage or Internet of Things resulting in the most advanced form of maintenance “Predictive Maintenance”. Maintenance is one of the things that companies try to improve. If you improve the maintenance and so the failure or breakdown possibility

decreases. The company that developed a maintenance strategy now has a competitive advantage on the competitor, because the maintenance, if it is done in the right way, permits to produce with lower costs and better reliability.

### 4. Acknowledgements

This research was supported by the project SGS-2018-031 Optimizing the parameters of a sustainable production system.

### 5. References

[1] Bagadia, Kishan: Computerized maintenance management systems made easy: how to evaluate, select, and manage CMMS. McGraw-Hill Professional, 2006. ISBN 0071469850, 9780071469852

[2] D'armanda, Paul, "Maintenance: From Philosophy to Practice", Nigel Dann and Timothy Cantell. *Journal of Architectural Conservation*, Volume 11 Number 1, March 2005. ISSN 1355-6207

[3] Dhillon, B. S. Engineering Maintenance: A Modern Approach. Boca Raton, FL: CRC, 2002. Print.

[4] Edwards, David J. Holt, Gary D. Harris, F.C., "Predictive maintenance techniques and their relevance to construction plant", *Journal of Quality in Maintenance Engineering*, Rok: 1998, Volume: 4 Str.:25 - 37. ISSN: 1355-2511

[5] Emilia Pietrikova, "Audience Response Systems: Benefits & Utilization", *Acta Electrotechnica et Informatica*, vol. 15, no. 4, pp. 3-7, 2016, doi: 10.15546/aeci-2015-0028

[6] Geitner, Fred K. BLOCH, Heinz P., "Maximizing machinery uptime", *Gulf Professional Publishing*, Vol. 5, 2006. ISBN 0750677252, 9780750677257

[7] Ginder, A. Robinson, C, "Implementing TPM", Productivity Press, 2007, ISBN 1-56327- 386-1. 1996. ISBN 1563270803, 9781563270802

[8] Gregory P. Sullivan; Aldo P. Melendez; Ray Pugh, "FEMP'S O&M Best Practices Guide A Guide to Achieving Operational Efficiency", *Strategic Planning for Energy and the Environment*, Volume 23, Issue 4 December 2004 , pp. 40 -52.

[9] Gresham, Robert M., and George E. Totten, Lubrication and Maintenance of Industrial Machinery, Best Practices and Reliability. Boca Raton: CRC, 2009.

[10] Hathaway, Louis Buckingham, Frank - Deere & Company: Preventive maintenance, Fundamentals of machine operation. Deere & Co., 1992. ISBN 0866911715, 9780866911719

[11] Kyokai, Nihon Puranto Mentensu, TPM for every operator, Japan Institute of Plant Maintenance, Productivity Press,

[12] Mobley, R. Keith, An introduction to predictive maintenance, Butterworth-Heinemann, 2002, ISBN 0750675314, 9780750675314

- [13] Novotný, L., Abreu, H. F. G. D., Zajac, J., & Béreš, M., Imperfections generation in finite element models of welding. *Science and Technology of Welding and Joining*, 23(2), 2018, pp. 148-157.
- [14] T. Orosz, Á. Sleisz and Z. Á. Tamus, "Metaheuristic Optimization Preliminary Design Process of Core-Form Autotransformers," *IEEE Transactions on Magnetics*, vol. 52, no. 4, pp. 1-10, April 2016, Art no. 8400310. doi: 10.1109/TMAG.2015.2496905
- [15] Palner, Doc: Maintenance planning and scheduling handbook. McGraw-Hill Professional, 1999, ISBN 0070482640, 9780070482647
- [16] Peter Papcun, Erik Kajati, Jiri Koziorek, Human machine interface in concept of industry 4.0. World Symposium on Digital Intelligence for Systems and Machines (DISA), pp. 289-296, 2018, doi: 10.1109/DISA.2018.8490603.
- [17] Poór, P., Kamaryt, T., & Simon, M., "Introducing autonomous maintenance by implementing OTH hybrid positions and TPM methods in metallurgical company", *International Journal of Engineering and Technology*, 7(3), 2015, pp. 817-824.
- [18] Ruffa, Stephen A., Going Lean: How the Best Companies Apply Lean Manufacturing Principles to Shatter Uncertainty, Drive Innovation, and Maximize Profits. Publisher. AMACOM Div American Mgmt Assn, 2008, ISBN 081441057X, 9780814410578
- [19] Salokyová, Š, Krehel', R., Pollák, M. and Kočiško, M., Research on impacts of mechanical vibrations on the production machine to its rate of change of technical state. *Advances in Mechanical Engineering*, 8(7), 2016, p.168781401665577.
- [20] Toshio Nakagawa, T.: Maintenance theory of reliability. Springer, 2005. ISBN 185233939X, 9781852339395
- [21] Trebuña, P., Pástor, M., Trebuña, F., & Šimčák, F., "The analysis of failure causes of the rotor shaft of steam turbines", *Metalurgija*, 56(1-2), 2017, pp. 233-236.
- [22] Wireman, Terry, Developing performance indicators for managing maintenance. Industrial Press Inc., 2005. ISBN: 0831131845, 9780831131845



## Author Index

### A

Abeyrathna, Y.	119
Adikari, A.M.C.	63
Amalan, T. P.	63
Amarasena, P. T.	57
Anuradha, U. A. D. N.	184
Asanka, P.P.G.D.	165
Azeez, Y. R.	172

### B

Bandara, O.K.K.	190
Basl, J.	245
Basnayake, B. M. A. L.	103

### C

Charanya, R.	240
Chathumali, E. J. A. P. C.	24, 158

### D

Dayarathna, H. R. O. E.	85
Dharmasiri, S. B. D. H.	126
Dias, U.G.M.D.	212
Dissanayake, M. B.	51

### F

Fanoon, A. R. F. S.	108
---------------------	-----

### G

Ganegoda, G. U.	29
Godakandage, V. M. P.	24
Gunasekara, S. R.	51

### H

Haddela, P. S.	97
Hashan, B.	119
Hettikankanama, H. K. S. K.	196
Hewawalpita, S.	9



<b>I</b>	
Ifhaam, M. F. A.	134
Indrathilaka, L.	78
<b>J</b>	
Jayalal, S.	17, 126, 134
Jayalath, J. M. T. I.	158
<b>K</b>	
Kaldera, H. N. T. K.	51
Kaluarachchi, T. R.	220
Kaluaratchi, M.	119
Kasundi, J.	29
Kesavan, M.	240
Kothalawala, K. R. M.	24, 158
Kulathilake, E. R.	179
Kulathunge, A. S.	85
Kumara, B. T. G. S.	57, 78
Kumari, P. K. S.	97
Kurahashi, S.	1
Kuruwitaarachchi, N.	158
<b>M</b>	
Madawala, C. N.	78
Madhubhashana, A. W.	24
Madhusanka, B. G. D. A.	38, 44
Mallawaarachchi, V.	70
Meedeniya, D.	70
Moladande, M. W. C. N.	38
<b>N</b>	
Nanayakkara, J. N.	212
Nanayakkara, L. D. J. F.	220
<b>P</b>	
Pallegedera, A.	204
Pathirana, W. P. N. H.	151
Perera, A. S.	165
Perera, I.	9, 70
Piyalath, A.	119
Poor, P.	245

<b>R</b>	
Rajapaksa, S.	70
Rajapakse, C.	103, 172, 179
Ranasinghe, D. C. R.	184
Rasanjana, W.	70
<b>S</b>	
Samarasinghe, K. T. C. S.	17
Sandaruwan, S. M. D. T.	204
Shanmugalingam, K.	113
Shibata, M.	91
Sumathipala, S.	113
<b>T</b>	
Takahashi, M.	91
Tharaka, V. K.	190, 212
Thelijjagoda, S.	119
<b>U</b>	
Uwanthika, G. A. I.	108
<b>V</b>	
Vasanthapriyan, S.	196
Vithanawasam, T. M. W.	44
<b>W</b>	
Wanniarachchi, W. K. I. L.	184
Weerabahu, W. M. S. K.	220
Weerasinghe, W. P. G. M.	142
Wickramarachchi, D. N.	142, 151
Wickramarachchi, A. P. R.	190
Wickramarachchi, A.	70
<b>Z</b>	
Zenisek, D.	245

