16

Engineering and Technology

APPLYING TECHNOLOGY ACCEPTANCE MODELS FOR LOCAL TRANSPORTATION NETWORK COMPANY APPLICATION DEVELOPMENT IN CAMBODIA

Sok An Chhor

Department of International College, Panyapiwat Institute of Management, Thailand
*Corresponding author, E-mail: chhorsokan@gmail.com

ABSTRACT

The study aims to applying technology acceptance models for local transportation network company application development in Cambodia. A quantitative research has been employed in order to collect on participants' perception toward adopt intention in local transportation network company application as well as factor led to adopted intention. The results obtained five factors led to customer adopting local transportation network company application. Those factors are relative advantage, compatibility, ease of use, perceived usefulness and perceived enjoyment.

Keywords: local, transportation network company, application and development

Introduction

The last few decades are known for the technological happenings. The pace of development of new technologies has led to the development of innovative products. These technologies in themselves are innovations and have led to many new inventions and discoveries that were never thought before. Application Platform or Mobile apps have fundamentally changed almost every feature of our daily lives. Application or app are highly developed software construct to accomplish on roaming devices like a smartphone, tablet or another comparable device. Making a grocery list, ordering food, checking account balances and communicating with others. Mobile Apps are becoming part of our daily life with its useful way. This portion of the change is attribute to peer-to-peer sharing transportation mode as it led to developed an enormous factor of the newest fashion in new up-and-coming transportation notions. A TNCs or ride-sharing app, like GrabCar or Uber, is a service matching application platform that matches the driver with possible customers. The nexus of one specific consequence of peer-to-peer sharing was the formation of the mobile application, or app, geared precisely for transportation. An app is advanced software engineered to execute on roaming devices for example a smartphone or other similar devices. Apps, in general, have established productive agility and ubiquity (Federal Transit Administration, 2015). Cambodia is a developing country which is developing in reasonably high rate of growth with the average of growth rate of 7 percentage grow rate in last decades.

Cambodia also country in Southeast Asia that have young, but capable ride-hailing apps and ride-hailing market, that first appeared in Cambodia about three years ago in 2016, and then a dozen local players rapidly amassed into the sector and with more than 20,000 ride-hailing drivers work in the country of Cambodia. TNCs operate in Cambodia are WeGo, Exnet Taxi Cambodia, GO-X, ITSUMO, TOS, Well, CityApp, 71 and CamGo. But the market is directed by two major local players – PassApp and WeGo (Shaun Turton, 2019). The PassApp, the largest local TNC which also has about 10,000 drivers with two million people have downloaded the app nationwide, mostly in Phnom Penh, WeGo the second largest local company and third largest in overall market of TNC have around 7,000 drivers (Shaun Tyurton, 2019). Since TNCs began to emerge in Cambodia with ride-hailing service application the market is dominated by the local TNCs, but with the introducing of the strong and robust foreign

TNCs competitor the position of the local TNCs had been shaken. So local TNCs need to prepare for further competition from foreign. Equally Grab begin operate in Cambodia since 2017 and a year later Uber's operations in the region also been bought out by Grab. Main supporters include Microsoft, Toyota, SoftBank, and later it gathers $300 million from U.S. investment management corporation Invesco, making its $7.5 billion of the total funds. Financial strength of Grab has already assisted it piece away at PassApp's head start. The Grab had more than 10,000 drivers alongside achieve the company in 2019 with one millionth ride in Cambodia (Shaun Tyurton, 2019).

This research study will conduct on the new technology that emerge and changing our transportation industry and focus on transportation network company, an industry that new technology of application platform and transportation. This research will be significant help both global business company Cambodia local business to under Cambodia market better in the market of Cambodia because the research on the ride hailing service never been conducted in Cambodia. This study is focused on local transportation network company application development in Cambodia.

Benefit of the research

In Spite of having multiple year of experience, Cambodia still lack of the still of important and connection of application platform in our new environment of integrated digital business but with the successful of TNC' app business, the interest had beginning to emerge in this part of business:

1. This research will help to understanding the driving factor that lead to user intention of adoption of transportation network company application in Cambodia.
2. The result from this research will enable will investor to understand better of Cambodia market and potential business that integrate application platform.
3. The result from this research can be useful information for local transportation network company in Cambodia to understanding the customer' perspective toward their adoption intention of the application.
4. The research result information can be highly effective to local transportation network company in Cambodia while considering the decision making for the overall marketing strategies.
5. The reference of this research can be useful for the one who wants to further study or research in topic intention of adopting of technology for local transportation network company application in Cambodia.

Objectives

1. To explore the key variables of IDT that influence the perceived usefulness and perceived enjoyment
2. To identify the key variables of Technology Adoption Model that influence the intention to adopt TNC application.
3. To suggest managerial strategies for TNC app's service development

Literature Review

Literature review aims to gathered the related general information and insight about transportation network company and application platform market development. Follow by the Cambodia's relevant review including the whole the syndicate and modified between model of Innovation Diffusion Theory and Technology Acceptance Model that can measure the perform of transportation network company application and situation of the local transportation network company application of Cambodia.

1. Application Platform

Application Platform is digital mobile application, with most normally indicated called to as an app, is a kind of software application intended to operate on a mobile device, like a smartphone or type of tablet computer. Mobile applications often use to provide users with comparable assistances to like of those able to operate on PCs. Application are usually little, separate individual software items with purpose of limited use. The use of as software application was initially commercialized through App Store of Apple Inc, which provide millions of applications for the iPad, iPod Touch and iPhone. Mobile applications are a travel long way from the assimilated systems of software usually found on PCs. In its place, individually application offers limited and isolated functionality such as a mobile web browsing, game and calculator. While apps may have evaded multitasking since the hardware with the limited resources of the early mobile devices such as smartphone, their specificity is now portion of their attractiveness because they allow users to hand-pick what their devices are able to do (Viswanathan, 2019).

2. Innovation Diffusion Theory

Introduced in late 1962, the Innovation Diffusion Theory (IDT) was modified by Rogers (Rogers E., 2003). Innovation diffusion theory centers on knowing how, at what rate and why technologies along innovative ideas extent in a social (Rogers E. M., 1962). In positions of the theories of modification, Innovation Diffusion theory scene a different attitude to analyzing changes. In its place to concentrating on influence persons to change, it perceives change as presence mainly around the evolution or "reinvention" of products alongside behaviors so they develop into greater fits aimed at the needs of individuals person and clusters of people. In diffusion of innovations, the difference is not in person, but the innovations themselves (Robinson, 2009). Instead, diffusion is the development through which an innovation is interconnected by certain limks over time between the members of a social system (Rogers E., 2003). Fichman describes diffusion by way of the development through which a technology extent across a population of organizations (Fichman, 2000). The model of innovations diffusion generally indicates to the extent of ideas from one group to another or from a center or institution in a society to further parts of that society (Rogers E., 2003). The elements of IDT are composed by Relative Advantage, Compatibility and Ease of Use.

3. Technology Acceptance Model

Technology Acceptance Model (TAM) is an information systems theory that models in what way operators come to acknowledge alongside operate a technology, centers on the presumption lead to rational behavior concept conferring to a goal of foreseeing with user's intention to adopted regarding new technology. Davis described perceived ease of use equally to the point to a person his or her thinks that the operate of a system will be painless and described perceived usefulness as being the point to which a person think that the use of a precise system will rise his or her performance. TAM theory base on concept in which of two factors influence attitude concerning new technology, which our case TNC' app, and in return, effects user intention, and finally actual usage (Davis F. D., 1989). TAM proposes that two variables perceived usefulness and perceived ease of use, are important elements of behavioral intention to use a system/technology.

Technology acceptance model theory assumes the two factors influence attitude toward new technology, which, lead to influences usage intention, with in time lead to ultimately actual usage. Technology acceptance model had been confirmed through many researchers for its validity in studies of application platform. Later encompassed self-efficacy as a predecessor with perceived ease of use and usefulness in their online learning system study (Ong, 2004). Additional, in number of limited researches which drawn out Technology acceptance model, perceived enjoyment was use as supplementary predecessor to implementation of numerous objects (Liao, 2009).

Transportation Network Company (TNC) application platform inherently possess not only practical but hedonic characteristics that are critical for product adoption. For instance, the primary adoption reasons of Transportation Network Company (TNC) application platform like convenience. Previously, have been moderately limited studies earlier which supported perceived enjoyment in order to confirm its position in product adoption. Example, Moon research on internet adoption along the idea of perceived enjoyment, which they defined along the level of perceived enjoyment caused through the act of using the system (Moon, 2001). Likewise, Bruner attempted to handle Technology acceptance model after a hedonic perspective for the confirm of the consequences of perceived usefulness, perceived ease of use and perceived enjoyment on attitude (Bruner, 2005).

On the other hand, Transportation Network Company (TNC) application platform occupied practical values connected along their technological characters. Namely, Transportation Network Company (TNC) application platform users center their adoption of choice with the product's usefulness or effectiveness in relations with task execution. This contributes credibility with the argument of the practical values will triumph in user's definite adoption decision (Davis F. D., 1989).

The current study is based on Technology acceptance model alongside perceived usefulness and perceived enjoyment that can be considered with attitudinal indicators of practical values along with hedonic values connected with the usage of TNC's app.

4. Hypothesis

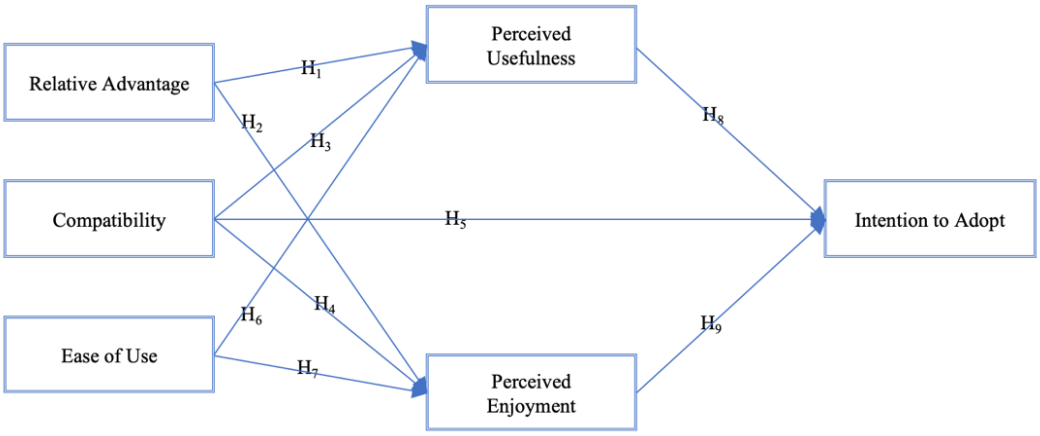


Figure 1: Conceptual Framework
Source: Davis (1989), Sung-Joon, (2014)

Hypotheses in this research are as follows:

H_1. Relative Advantage has a significant impact with perceived usefulness.
H_2. Relative Advantage has a significant impact with perceived enjoyment.
H_3. Compatibility has a significant impact with perceived usefulness.
H_4. Compatibility has a significant impact with perceived enjoyment.
H_5. Compatibility has a significant influence with intention to adopt.
H_6. Ease of use has a significant impact with perceived usefulness.
H_7. Ease of use has a significant impact with perceived enjoyment.
H_8. Perceived usefulness has a significant influence with intention to adopt.
H_9. Perceived enjoyment has a significant influence with intention to adopt.

Methods

1. Study Design

The study will be conducted by making a questionnaire. The questionnaire will be distributed in Cambodia using social media and emails. The questionnaire will be tested among 30 respondents in for checking the reliability of the questionnaire. After collecting data, the results will be combined (the response from the online survey and the response from offline survey). These data will be examined for the reliability and significance. Finally, the significant response will help the results to compare with the hypothesis of the paper and pave the way to initiate towards conclusion.

2. Sample Size

The minimum sample size was calculated according to the formula (Green, 1991). The factors to be held for the sample size calculation will be comprise of the number of categories for each variable. All 5 Dependent variables that included Ease of Use, Compatibility, Relative Advantage, Perceived Enjoyment and Perceived Usefulness. The minimum participant 90 was quoted, therefore a sample of 100 should be sufficient but to increase the reliability of the result the research sample had increase to 253. In order cost effective and time effective, Phnom Penh was selected as target area. Phnom Penh is not only Cambodia capital city, it's also the biggest commercial and economic city in Cambodia. According Cambodia's Ministry of Planning, Phnom Penh had the population of 2,129,371 (National Institute of Statistics, 2019). In order to get the proportion sample, the data was collected from 253 respondents in Phnom Penh.

Rule of thumb Formula (Green, 1991):

$$N > \text{or} = 50 + 8m$$
$$= 50 + 8*5$$
$$= 90$$

N = Sample Size
M = the number of predictors

3. Investigative Techniques

For most of the question in the questionnaire, Likert Scale will be used to analyze the respondent's perspective. Checkbox questions will be included to screen down the questions. The questionnaire will not use Direct Translation Method. The questionnaire will be using the English.

4. Research Instrumentation

This research used questionnaire survey as examination instrument to collect the research information. The questionnaire survey was separated into two section i.e. earliest section, was more about the demographic and personal inquiry such age, occupation, gender etc. whereas, the next section was aimed to model according to the hypothesis and conceptual framework of the research. A questionnaire will be created using all the variables that connect to intention adoption of TNC's app. There will be online questionnaire, the questionnaire will be created using the Google Form and this software will also be used to monitor the responses.

5. Data Collection

The data collection will mainly be conducted in the capital city of Cambodia, Phnom Penh. The process will include two steps, i.e. collecting primarily data from 30 respondents and conducting pilot test to check the reliability of the questions, collecting main data as well as record them. The Questionnaire will be allocated among the sample group using social media, i.e. Facebook, Instagram, WhatsApp, Viber, IMO and some of the questionnaire will be also distributed using email.

Primary source of data: Informative for this research is gather through questionnaires (online) from November 10, 2019 to December 05, 2019. The 253 surveys were collected from individuals who older or equal to 20 years of age. The Data collection center was Phnom Penh City, Cambodia

Secondary source of data: The Secondary source of data collected from different article based on: a) Internet Survey: Online new article and online survey questionnaire was used in the research. b) Reference Books: Journals/articles related to the transportation network company and application service, Technology Acceptance Mode and Innovation Diffusion Theory etc. to supplement theoretical background especially for the user adopted intention

6. Measurement

253 in total responses were gathered from the online study. In this study we published questionnaires on an online agency that responsible with survey to gather the experimental data. The original questionnaire contains of 2 portions. The opening portion has 8 questions to gather the simple information of the responses, such as sex, age, and use experience with TNC app. Another portion is the core section of the questionnaire and contains of 26 questions to examine the 7 factors presented in previous section with Relative Advantage, Compatibility and Ease of Use had 5 question each. Perceived Usefulness and Perceived Enjoyment had 4 question each with Intention to Adopt had 3 questions. Each question is measured on a 5-point Likert scale with the end points of “strongly agree (5)” and “strongly disagree (1)”.

7. Data Analysis Plan

The plan to use factor analysis along with confirmed item reliability by means of Cronbach’s alphas to decide the internal consistency of the measures, after data collection, Path analysis will be used for statistical research result

For the analyze the efficiency of the original data, the initial step of the test is to conduct data standardization. In this stage we analyze the average and standard deviation of each question result and also the average for each group. The outcomes are shown in Table1. Since the table we can understand that the average of all factors is larger than 3 with Relative Advantage have total average of 3.56, Compatibility have total advantage of 3.46, Ease of Use with total average of 3.82, Perceived of Usefulness with total average of 3.64, Perceived of Enjoyment with total average of 3.66 and Intention for Adoption with total average of 3.68. Which the result had advocates that the assumptive factors were distinctive.

Table 1: Question standardization and reliability analysis.

Factor	Question	Mean	SD	Total Mean
REA	REA1	3.71	0.950	3.56
	REA2	3.64	0.859	
	REA3	3.57	0.811	
	REA4	3.47	0.915	
	REA5	3.43	0.951	
CPB	CPB1	3.58	1.018	3.46
	CPB2	3.46	1.044	
	CPB3	3.40	0.969	
	CPB4	3.50	0.902	
	CPB5	3.36	1.009	

Factor	Question	Mean	SD	Total Mean
EOU	EOU1	3.59	0.977	3.82
	EOU2	3.75	1.120	
	EOU3	3.84	0.962	
	EOU4	3.98	0.865	
	EOU5	3.96	0.907	
POU	POU1	3.52	0.940	3.64
	POU2	3.60	0.846	
	POU3	3.62	1.003	
	POU4	3.83	0.919	
POE	POE1	3.62	0.906	3.66
	PEO2	3.62	0941	
	POE3	3.58	0.906	
	POE4	3.55	0.997	

The value of Cronbach Alpha is an evidence for reliability of the questionnaire and measure internal consistency. Cronbach Alpha value can be considered as test the questionnaires using a consistency (Li, 2010). The excellent alpha value the greater and equal to 0.9, the good alpha value ranges from 0.8 to 0.9, the acceptable alpha value ranges from 0.7 to 0.9, the poor alpha value ranges from 0.5 to 0.6, unacceptable alpha value ranges from 0.5 to 0.6 (Li, 2010).

Subsequently we further use Cronbach's alpha coefficient to display the convergent validity and internal reliability of the factors, which are recorded in Table 2. After Table 2 we can see that the total Cronbach's alpha coefficient is 0.964 and the coefficients of each factor are larger than 0.8 with Relative Advantage with Cronbach's alpha coefficient of 0.814, Compatibility with Cronbach's alpha coefficient of 0.880, Ease of Use with Cronbach's alpha coefficient of 0.864, Perceived of Usefulness with Cronbach's alpha coefficient of 0.849, Perceived of Enjoyment with Cronbach's alpha coefficient of 0.937 and Intention of Adoption with Cronbach's alpha coefficient of 0.894. It is then claimed that the total Cronbach's alpha coefficient is acceptable (>0.8), and the coefficients of each factor are also acceptable (>0.7) (Li, 2010). As a consequence, we determine that the data are reliable measures for their factors.

Table 2: Cronbach's alpha coefficient of each factor.

Factor	Cronbach's alpha coefficient
REA	0.814
CPB	0.880
EOU	0.864
POU	0.849
POE	0.937
ITA	0.894
Total (26 questions)	0.964

Pearson Correlation analysis is conducted to examine the relationship between the six variables. As shown, the maximum correlations among different factors are below 0.80 and indicates that all variables are related to each.

Table 3: Correlation.

Question		REA	CPB	EOU	POU	POE	ITA
REA	Pearson Correlation	1	.731**	.672**	.717**	.774**	.729**
	Sig. (2 tailed)		.000	.000	.000	.000	.000
CPB	Pearson Correlation	.731**	1	.696**	.732**	.700**	.698**
	Sig. (2 tailed)	.000		.000	.000	.000	.000
EOU	Pearson Correlation	.672**	.696**	1	.626**	.672**	.587**
	Sig. (2 tailed)	.000	.000		.000	.000	.000
POU	Pearson Correlation	.717**	.732**	.626**	1	.757**	.722**
	Sig. (2 tailed)	.000	.000	.000		.000	.000
POE	Pearson Correlation	.774**	.700**	.672**	.757**	1	.709**
	Sig. (2 tailed)	.000	.000	.000	.000		.000
ITA	Pearson Correlation	.729**	.698**	.587**	.722**	.709**	1
	Sig. (2 tailed)	.000	.000	.000	.000	.000	

Results and Discussion

1. Structural Equation Model

According to Table 4, model IDT and TAM characteristics of TNC product, compatibility (β =0.364, p = 0.000), relative advantage (β=0.300, p = 0.000) and ease of use (β = 0.260, p = 0.000) had substantial influence to perceived usefulness. Including the characteristics of TNC app product, relative advantage and ease of use had moderately greater influences, signifying people perceive higher usefulness of TNC apps since ease of use plus relative advantage are deemed instrumental to successful task completion. Additional, relative advantage (β = 0.544, p = 0.000) and ease of use (β = 0.171, p = 0.003) had significant positive impact on perceived enjoyment while compatibility (β = 0.066, p = 0.356) because of the height of p value it hard to determine the impact on perceived enjoyment. Compatibility (β = 0.202, p = 0.000) had also significant impact on intention to adopt. Final, the research discovered that perceived usefulness (β = 0.179, p = 0.000) and perceived enjoyment (β = 0.537, p = 0.000) together had substantial effects on adoption intention. Particularly important concerning this result is that perceived enjoyment had larger impact on adoption intention than perceived usefulness, which reasserts the importance of emotional attribute of Transportation Network company application.

Table 4: Result of SEM analysis.

Hypothesis	Path	β	S.E.	P	Supported
H_1	REA → POU	.300	.085	***	Accepted
H_2	REA → POE	.544	.088	***	Accepted
H_3	CPB → POU	.364	.073	***	Accepted
H_4	CPB → POE	.066	.076	.356	Rejected
H_5	CPB → ITA	.202	.066	***	Accepted
H_6	EOU → POU	.171	.057	.003	Accepted
H_7	EOU → POE	.260	.059	***	Accepted
H_8	POU → ITA	.179	.062	***	Accepted
H_9	POE → ITA	.537	.053	***	Accepted

2. Fitness Indices

Table 5 shown Df with the obtained value of 3 and X2 with the obtained value of 48.680. X2/df with the stand of obtained value of 16.227 above recommended values at equal or smaller than 5 this result in failed (Hairet al, 2006). GFI result id qualified with the obtained value is 0.944 within range of the recommended values of equal or larger than 0.90 (Chau & Hu , 2001).IFI with the obtained value of 0.965 within the range of the recommended value of equal or larger than 0.90 with in the qualified result (Chau & Hu , 2001).CFI is on qualified result with the obtain value of 0.965 within the range of the recommended value of equal or larger than 0.90 (Bagozzi and Yi, 1988). NFI had an obtain value of 0.963 that above 0.90 within the recommendation range of equal or larger than 0.90 with a qualified result (Bagozzi and Yi, 1988). RMSEA fit indices was result failed because the obtain value of 0.246 more than the acceptable range of smaller or equal to 0.08 with the result RMSEA is considered failed (Browne and Cudeck, 1993).

Table 5: Fit indices for the measurement**.**

Goodness of fit	Obtained value	Recommended values	Source	Results
Df	3			
X2	48.680			
X2/df	16.227	≤ 5.00	(Hair et al, 2006)	Failed
GFI	0.944	≥ 0.90	(Chau & Hu , 2001))	Qualified
IFI	0.965	≥ 0.90	Chau & Hu (2001)	Qualified
CFI	0.965	≥ 0.90	(Bagozzi and Yi, 1988)	Qualified
NFI	0.963	≥ 0.90	Bagozzi and Yi (1988)	Qualified
RMSEA	0.246	≤ 0.08	(Browne and Cudeck, 1993)	Failed

3. Direct, Indirect and total effects

Table 6 shows the significance direct, indirect, and total effects between the exogenous and mediating variables and the endogenous variables. The significance of indirect paths from mediating effects was calculated using Sobel's (1982) test, as displayed in Table 4.4. REA had nearly equal effect on ITA with direct effect of 0.09 and indirect of 0.08964 on ITA and with total effect of 0.17964 on ITA, REA had strong significant on influence with ITA. CPB with direct effect of 0.04 stronger than indirect effect of 0.01296 and total effect of 0.05296 had led to assuming that CPB had significant influence on ITA. EOU had seen result of significant on ITA with direct effect of 0.0676 influence stronger than indirect effect of 0.01274 and total effect of 0.08034. POU with direct of 0.0324 more influence than indirect effect of 0.01296 and total effect of 0.04536 had significant influence on ITA. POE had very strong influence on ITA with direct effect of 0.2916 much stronger effect than indirect effect of 0.00756 and total effect of 0.29916. In direct, indirect and total effects, POE had the strongest on ITA with total effect of 0.29916 follow by REA with total effect of 0.17964. CPB, EOU and POU had significant less influence on ITA than POE and REA.

Table 6: Direct effects, indirect effects and total effects (n=253)**.**

Variable	**ITA**		
Exogenous	Direct	Indirect	Total
REA	0.09	0.08964	0.17964
CPB	0.04	0.01296	0.05296
EOU	0.0676	0.01274	0.08034
POU	0.0324	0.01296	0.04536
POE	0.2916	0.00756	0.29916

Conclusions

In this study, we have collected 253 respondents with more 60% of the respondent was male, and the biggest majority of age of our respondent are between 20-29 with comprise of nearly 90% of our respondent. The majority of the respondent had bachelor degree with comprise about 80% and with 90% are had occupation of student. The majority of our respondent had income under 400,000 riel that comprise of nearly 60% of the respondent about 88% had frequency use of 1-3 per week and spend about between 2,000 riel and 10,000 riels on average spending per ride.

This study developed a theoretical framework and discussed the structural equation modelling analysis of the proposed theoretical framework for local transportation network company application development of adoption in Cambodia. A huge stream of study has been done, from the standpoint of innovation products headed by Rogers who said that compatibility, relative advantage, divisibility, complexity, and observability effect innovation adoption (Rogers E., 2003). Other stream of study led by Davis used Technology acceptance model to emphasize that perceived usefulness and perceived ease of use influence technology adoption. (Davis F. D., 1989) Contemporary studies on innovation technology acceptance tried to integrating IDT and TAM (Such, 2004). But as point previously, while meaningful in the sense that they tried to explain varied features of innovation adoption process, these earlier researches have difficulty in describing adoption decision that mirrors both product and consumer dimensions. Therefore, to fill the void in literature, this study attempted to identify local transportation network company application adoption combining present theories on innovation adoption. By combining Innovation and Diffusion Theory and Technology Adoption Model to describe TNC application.

1. Theoretical implication

Findings of this study show that Innovation Diffusion Theory (IDT) variable of ease of use, relative advantage and compatibility with Technology Acceptance Model (TAM) of perceived of usefulness and perceived of enjoyment had significant impact on customer adopted intention usage on local transportation network company application in Cambodia. Such relative advantage, compatibility and ease of use contribute to perceived of usefulness and perceived of enjoyment, in line with our conceptual framework. Sung-Joon (2014) study suggests Innovation Diffusion Theory (IDT) variable of relative advantage, compatibility and ease of use contribute to Technology Acceptance Model (TAM) of perceived of usefulness and perceived of enjoyment in the study of features affecting the adoption of Haptic Enabling Technology. In context of our study of adopted intention usage on local transportation network company application in Cambodia had seen similar effect of contribution.

Findings are consistent with the customer intention literature, suggesting that high levels of felt enjoyment and usefulness can lead to customer intention (Koufaris, 2002). Furthermore, the significance of perceived usefulness and perceived enjoyment shows that the Technology Acceptance Model can be successfully applied. When tested our results also confirmed prior TAM research that found that perceived enjoyment was a more important predictor of intended to adopted for the customer.

2. Managerial contribution

From this research it is found that perceived ease of use and perceived enjoyment are the main factors of user intention to adopted local transportation network company application. This may suggest that:

(1) Users prefer to use some easier to get started using local transportation network company application which more useful in transportation which can apply the discount point for sharing among of application scan code for social friend such Facebook, Line and WeChat to encourage spread of usage of the application among the population because with encourage of point and simply of scan code can increase number of user of the application and (2) Users regard the level of enjoyment from local transportation network company application as the most significant factor. Of these 2 factors, perceived enjoyment displays a much bigger effect than perceived ease of use, which suggests that new technology that assist in transportation will be paid much attention to on convenience and customer experience, First is payment method need to be operate with multiple channel that make can make customer can make payment in multiple way pay on cash the favorite of the older generation must be available older generation customer, pay by credit/debit card and pay application payment WeChat and Alipay the favorite among new generation of customer need to be accesses to the customer to increase the convenience of the application service and information of payment method understand need to understand multiple different discount for different method to encourage use of multiple different method among customer. Second the Map and location of transportation vehicle must easier as to understand and precise as possible location of the customer destination need and of vehicle need to be at understandable to the customer for sense of security and direction to provide the best experience to customer and strong advertisement on social media and traditional media on the convenience on the map.

From this research it is concluded that relative advantage and ease of use also does not have very big direct effect on the actual intention to adopt for the user. It may be inferred that, in the domain of local transportation network company application, users mainly want to easily get the usefulness and some enjoyment from local transportation network company application as transportation mode.

References

Bagozzi and Yi. (1988). n the Evaluation of Structural Equation Models. *Journal of the Academy of Marketing Sciences, 16*, 74-94.

Browne and Cudeck. (1993). Alternative ways of assessing model fit. In B. a. Long, *Testing structural equation models* (pp. 136-162). Newbury Park: CA: Sage.

Bruner, G. (2005, May). Explaining consumer acceptance of handheld Internet devices. , *Pages 553-558, 58*(5), 553-558.

Chau & Hu . (2001). Information Technology Acceptance by Individual Professionals: A Model of Comparison Approach. *Decision Sciences, 32*(4), 699-719.

Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: theory and results.* Massachusetts Institute of Technology, The Sloan School Of Management. Massachusetts: Massachusetts Institute of Technology.

Davis, F. D. (1989, September). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340.

Davis, V. V. (1996). A model of the antecedents of perceived ease of use: development and test. *Decision Science, 27*(3), 451-481.

Federal Transit Administration. (2015). *National Transit Database (NTD) Glossary*. Retrieved from Federal Transit Administration: https://www.transit.dot.gov/ntd/national-transit-database-ntdglossary.

Fichman, R. G. (2000). The Diffusion and Assimilation of Information Technology Innovations. In R. G. Fichman, *Framing the Domains of IT Management: Projecting the Future Through the Past.* Cincinnati, Ohio, United States of America: innaflex Educational Resources.

Green, S. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*, 499-510.

Hair et al, B. W. (2006). Multivariate data analysis. *6th Edition*.

Koufaris, M. (2002, June). Applying the Technology Acceptance Model and Flow Theory to Online Consumer Behavior. *13*(2), 205–223.

Liao, C. H. (2009, April). User acceptance of computer-mediated communication: The SkypeOut case. *Expert Systems with Applications, 36*(3), 4595-4603.

Li, Y. H. (2010). The research on index system optimization of graduation design based on cronbach coefficien. *Proceedings of the 5th International Conference on Computer Science and Education*, (pp. pp. 1843–1845). Hefei, China,.

Moon, J. W. (2001, Febury). Extending the TAM for a World-Wide-Web context. *Information & Management, 38*(4), 217-230.

National Institute of Statistics. (2019). *General Population Census of the Kingdom of Cambodia 2019.* Ministry of Planning, National Institute of Statistics. Phnom Penh: Ministry of Planning.

Ong, C. h. (2004, July). Factors affecting engineers' acceptance of asynchronous e-learning systems in high-tech companies. *Information & Management, 41*(6), 795-804.

Robinson, L. (2009). *A summary of Diffusion of Innovations.* ChangeoLogy: the Book.

Rogers, E. (2003). *Diffusion of innovations. Fifth edition* . New York: Free Press.

Rogers, E. M. (1962). *Diffusion of innovations. First editio.* New Yor: Free Press.

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology, 13*, 290-312.

Shaun Turton, B. P. (2019, July 08). *Grab takes on Cambodia's PassApp for ride-hailing supremacy*. Retrieved from Nikkei Asian Review: https://asia.nikkei.com/Business/Startups/Grab-takes-on-Cambodia-s-PassApp-for-ride-hailing-supremacy

Shaun Tyurton, B. P. (2019, July 08). *Grab takes on Cambodia's PassApp for ride-hailing supremacy*. Retrieved from Nikkei Asian Review: https://asia.nikkei.com/Business/Startups/Grab-takes-on-Cambodia-s-PassApp-for-ride-hailing-supremacy

Suh, S. (2004). Application of TAM and IDT regarding adoption of supply chain model. . *J. Ind. Econ, 17*(04), 1443–1466.

Sung-Joon, O. J. (2014). Validation of Haptic Enabling Technology Acceptance Model (HE-TAM): Integration of IDT and TAM. *Telematics and Informatics*, 585-596.

Viswanathan, P. (2019, May 28). *What Is a Mobile Application?* Retrieved from Lifewire: https://www.lifewire.com/what-is-a-mobile-application-2373354

Appendix A
The Questionnaire

Factor	Item	Measure
Relative Advantage (RAT)	RAT1	Using local transportation application had improved quality of my daily transportation compare to other type of transportation.
	RAT2	Using local transportation application had made it easier to go do my job and daily life compare to other type of transportation.
	RAT3	Using local transportation application had increased my effectiveness on my daily job and daily life activity compare to other type of transportation.
	RAT4	Using local transportation application had given me greater control over my work and daily life compare to other type of transportation.
	RAT5	Using local transportation application had enables me to accomplish tasks more quickly than other type of transportation.
Compatibility (CPB)	CPB1	Possibility for using local transportation application is to replace other transportation already in use.
	CPB2	Possible for using local transportation application is replace product's with similarity function.
	CPB3	Possible for using local transportation application is replace advanced version of product with similarity function that already in use.
	CPB4	Using local transportation network company application is compatible with all aspects of my work
	CPB5	Using local transportation application fits well with the way I like to work.
Ease of Use (EOU)	EOU1	I believe that it is easy to get local transportation application to do what I want it to do.
	EOU2	Learning to use local transportation application would be easy for me.
	EOU3	My interaction with local transportation application is clear and understandable.
	EOU4	It would be easy for me to become skillful at using local transportation application.
	EOU5	I find local transportation application easy to use.
Perceived Usefulness (PU)	PU1	Using local transportation application enhances my effectiveness in my job.
	PU2	Using local transportation application improves my performance in my job.
	PU3	Using local transportation application in my job increases my productivity.
	PU4	Find using local transportation application Useful.
Perceived Enjoyment (PE)	PE1	During my last using of local transportation application I found. using it interesting.
	PE2	During my last using of local transportation application I found using it enjoyable.
	PE3	During my last using of local transportation application I found using it exciting.
	PE4	During my last using of local transportation application I found using it fun.
Intention to Adopt (ITA)	ITA1	Have plan to sustained use of the local transportation application.
	ITA2	Have intention to use the local transportation application.
	ITA3	Have intention to use extended product of local transportation application.

AUTOMATIC INFORMATION EXTRACTION FOR THAI RESEARCHER

Nalinee Wuttivichayanan[1*], Takashi Yukawa[2], Jian Qu[3*]

[1,3]Department of Engineering of Technology, Panyapiwat Institute of Management, Thailand
[2]Nagaoka University of Technology, Japan
***Corresponding author, E-mail:** Nalineewut@gmail.com, Jianqu@pim.ac.th

ABSTRACT

Our proposed method is to implement information extraction technology and natural language processing in order to create an automated information extraction system for Thai researchers. Based on a sample of 20 Thai researchers, we tested keywords to find information related to researchers. We proposed a rules-based information extraction system and used distance between OOV and keywords and keyword list to get the most accurate information. The results showed that keywords-"graduates" with an accuracy of 66.7%; keywords-"work experience" with an accuracy of 75.0%; keywords-"research interest" with an accuracy of 71.7% accuracy; and keywords-"research area" with an accuracy of 71.4%.

Keywords: Information Extraction, Information for Researcher, Natural Language Processing.

Introduction

At present, the Internet becomes an increasingly indispensable tool for many people if information is needed. Such information can range from recreational media like cartoons, music, or series, to serious knowledge in specialized fields such as astrophysics or biology. This also applies to personal information as people would like to know more about public figures.

Despite the perceived convenience, manual information searching usually takes time especially in more obscure fields such as Thai researchers. Information such as the birth date and notable works are usually published at their respective universities' websites.

According to our observation, there was a website for finding Thai research, which is "Thailand S&T Expert Finder" (Fig.1) which is no longer active. As a result, currently there is no standard, reliable system for the collection of Thai researchers' personal profiles and works. Due to the profound lack of means to find information about Thai researchers and its potential usefulness for the researcher to query and contact each other, we plan to develop a research information extraction system.

This is a Primary study on Thai Researcher information extraction, because in the past, there has been no development of data extraction methods for Thai researchers, therefore, we aim to study researcher information extraction and build prototypes that can automatically extract the information from the Internet.

Figure 1: Website Thailand S&T Expert Finder

Objectives

Implement information extraction technology and natural language processing to create an automated information extraction system for Researcher.

Literature Review

1. Information Extraction

Wimalasuriya and Dou, (2010) defined Information Extraction (IE) as the use of computational methods to identify relevant pieces of information in a document originally generated for human use and convert this information into a representation suitable for computer-based storage, processing, and retrieval. The input to IE system is a collection of documents (email, web pages, newsgroups, news articles, business reports, research papers, blogs, resumes, proposals, and so on) and output is a representation of the relevant information from the source document according to some specific criteria. The ability of humans to effectively make use of this vast amount of information is limited as this task is quite boring, tedious and consume a lot of time. This explosion of information and the need for more sophisticated and efficient information handling tools highlighted the need for information extraction and retrieval technology (Neil et al., 1998).

2. Related Work

2.1 Jian Qu et al., (2018) built an automated system for extraction of celebrity information from free-text websites. In this work, framework of Thai actor and actress database had following modules: Web-crawling, Pattern matching and extraction and personal information selection by statistics. Tests showed that their method achieved better accuracy than Mthai.com, a manually-managed celebrity website.

2.2 Ying Chen et al., (2009) proposed a robust system which extracted lightweight features from the internet without supervision. The experiments showed that these lightweight features not only improved performances, but also increased the robustness of the disambiguation system. To extract information of the focused person, an integrated system was introduced, which could effectively re-use and combine current well-developed tools for web data, and at the same time, identify the expression properties of web data. Their flexible extraction system achieved state-of-the-art performance, especially precision, which is very important for real application.

2.3 Yaoyong Li et al., (2012) suggested that their approach needed fewer SVM classifiers to be trained than other recent SVM-based systems. The paper also compared their approach to several state-of-the-art systems (including rule learning and statistical learning algorithms) on three IE benchmark datasets: CoNLL-2003, CMU seminars, and the software jobs corpus.

2.4 Kritsada and Thanaruk (2002) studied Generalized association rule mining (GARM) as an extension of traditional association rule mining to find more information rules. GARM has many algorithms for use but recently the most efficient algorithm is Prutax. However, limitation of Prutax is high checking cost. Thus, this paper aimed to overcome this limitations by creating a new set enumeration algorithm, named SET. Experiment showed that SET outperformed the most efficient algorithm available on Prutax.

2.5 Jian Qu et al., (2016) proposed a novel cross-language definition retrieval system for OOV terms. Nevertheless, we proposed an auto re-evaluation method to evaluate the correctness of OOV translations and definitions. Our method achieve high performances against existing methods.

2.6 Jian Qu et al., (2011) proposed a rule-based method for translating English medical OOV terms to Chinese. A major problem was that English OOV terms were not perfectly translated into Chinese because English and Chinese were written differently. Thus, they used a data mining approach and used statistic-based approach to select the translation candidate. After testing with two OOV term sets ICD9-CM and ICD9, it outperformed existing approaches with a precision of 89.98% and could handle translation of Chinese sentences that also included non-Chinese characters.

2.7 Jian Qu et al., (2012) proposed an integrated ranking approach for predicting the type of OOV terms, extracting the monolingual definitions and the multilingual context information. Moreover, they proposed a novel adaptive rule-based approach with Bayesian net and Ad boost for handling hybrid translations. Experiments showed that their approach outperformed existing approaches.

2.8 Jian Qu et al., (2011) presented a novel OOV term translation mining approach, which proposes a new adaptive rules system for hybrid translations and a new recursive feature selection method for supervised machine learning. In this paper, the proposed method is for English-Chinese OOV term translation.

2.9 Pascal and Guy (2001) proposed a simple non- weighted features KNN algorithm for text categorization. The proposed KNN algorithm was efficient for classifying text documents in that context (in terms of its predictability and interpretability). Also, tests showed that its simplicity (w.r. t. its implementation and fine-tuning) became its main asset for on-the-field applications.

2.10 Zhou Yong Li Youwen and Xia Shixiong (2011) proposed an improved KNN text classification algorithm based on clustering center. Results from simulations showed that the algorithm proposed in this paper could effectively reduce the actual number of training samples and lower the calculation complexity and, also improve the accuracy of KNN text classification algorithm.

2.11 Sang-Bum Kim et al., (2006) proposed two empirical heuristics: per-document text normalization and feature weighting method. While these method were somewhat ad-hoc in nature, their proposed naive Bayes text classifier performed very well in standard benchmark collections, comparable to state-of-the-art, complex text classifiers such as SVM.

Methods

Our proposed method is to implement information extraction technology and natural language processing in order to create an automated information extraction system for Thai researchers. We have five steps to extract the correct answer information, the steps are as follows in Fig. 2

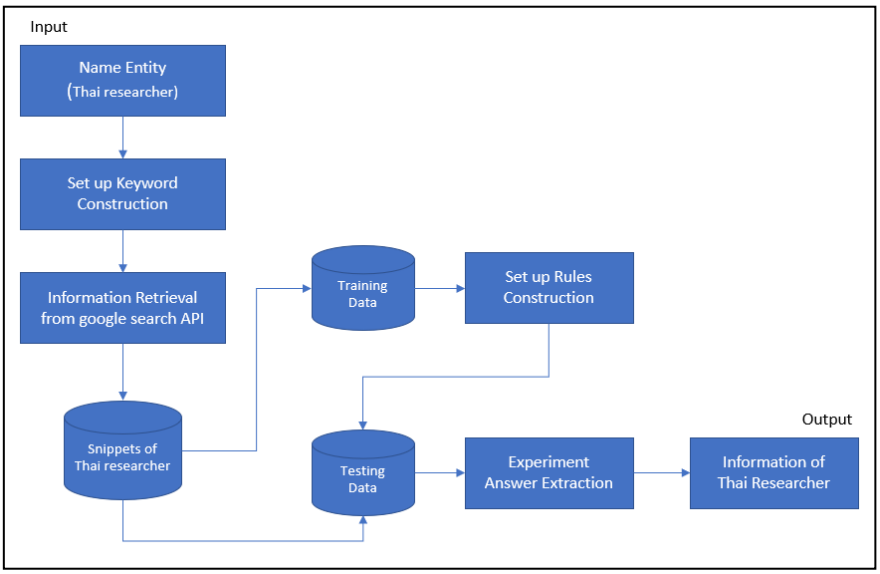


Figure 2: Flow Chart of our proposed approach

1. Get Name Entity (Thai Researcher)

In the first step, we start by selecting names of 20 Thai researchers in the fields of technology, computers, and communication. Data come from Sirindhorn International Institute of Technology (SIIT) and Panyapiwat Institute of Management (PIM) websites. The list of Thai researchers collected is shown in Table 1.

Table 1: The list of Thai researchers collected.

No	Name Entity	Type	No	Name Entity	Type
1	Gun Srijuntongsiri	SIIT	11	Itthisek Nilkhamhang	SIIT
2	Thanaruk Theeramunkong	SIIT	12	Pakinee Aimmanee	SIIT
3	Ekawit Nantajeewarawat	SIIT	13	Prapun Suksompong	SIIT
4	Chalie Charoenlarpnopparut	SIIT	14	Somsak Kittipiyakul	SIIT
5	Komwut Wipusitwarakun	SIIT	15	Teerayut Horanont	SIIT
6	Nirattaya Khamsemanan	SIIT	16	Sasiporn Usanavasin	SIIT
7	Toshiaki Kondo	SIIT	17	Virach Sornlertlamvanich	SIIT
8	Waree Kongprawechnon	SIIT	18	Jian Qu	PIM
9	Banlue Srisuchinwong	SIIT	19	Datchakorn Tancharoen	PIM
10	Cholwich Nattee	SIIT	20	Tunyawat Somjaitaweeporn	PIM

2. Set up keyword construction

After preparing the list of researchers, we started searching for the keyword that will be used in the Query data from Google search API. We randomly selected 3 researchers to experiment with 20 keywords to find the best keyword that gives us the information that we need. An element result is shown in Table 2 and Table 3.

Table 2: Query name entity each keyword.

No.	Keyword	No.	Keyword
1	research	11	work
2	researcher	12	professional
3	research interest	13	proficiency
4	research area	14	work experience
5	research area of	15	specified
6	research experience	16	relevant research
7	graduate from	17	relevant research paper
8	graduated	18	conference paper
9	experience	19	work position
10	competency	20	list of Publications

Table 3: Result of keyword construction.

No.	Keyword	Use
1	research	x
2	researcher	x
3	research interest	/
4	research area	/
5	research area of	x
6	research experience	x
7	graduate from	/
8	graduated	x
9	experience	x
10	competency	x
11	work	x
12	professional	x
13	proficiency	x
14	work experience	/
15	specified	x
16	relevant research	x
17	relevant research paper	x
18	conference paper	x
19	work position	x

3. Retrieve information using Google search API

We have the keyword that gives the desired information. The amount of 4 keywords are place of graduation, research area, research interest, and work experience. We use the aforementioned keywords in the Google search API and we got URL, Title, summary as a result.

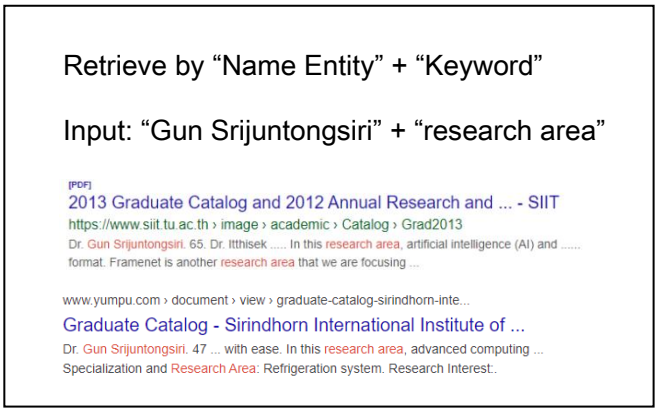


Figure 3: Example Retrieve information from Google

The acquired website summary (called “snippets” of a website) along with its title and keyword used are imported into a database. An example of web retrieved snippets of name entity is shown in Figure 4.

Id	ID_person	Fname	Lname	title	URL	summary	keyword_id
1	1	Gun	Srijuntongsiri	In Celebration of The 20th Anniversary of SIIT	https://www.siit.tu.ac.th/upfiles/publication_file...	Dr. Gun Srijuntongsiri is a full-time faculty mem...	1
2	1	Gun	Srijuntongsiri	Main Conference Proceedings - APSCE	www.apsce.net/icce/icce2017/140.115.135.84/.../ICC...	... Sciences, China. Chih-Hung CHEN, Graduate from...	1
3	1	Gun	Srijuntongsiri	Department of Mathematics - Cornell University	https://pi.math.cornell.edu/News/AnnRep/AR98-99.pd...	David Kuo and Gun Srijuntongsiri tied for first pr...	1
4	1	Gun	Srijuntongsiri	Department of Mathematics Cornell University: Annu...	educationdocbox.com/.../69563643-Department-of-mat...	David Kuo and Gun Srijuntongsiri tied for first pr...	1
5	2	Thanaruk	Theeramunkong	SIIT News, Vol. 15, No. 1, March 2010	https://www.siit.tu.ac.th/upfiles/publication_file...	Dr. Thanaruk Theeramunkong (Head of School of Info...	1
6	2	Thanaruk	Theeramunkong	SIIT News, Vol. 14, No. 3, September 2009	https://www.siit.tu.ac.th/upfiles/publication_file...	Mr. Puwanan Chumtong, a graduate from the School ...	1
7	2	Thanaruk	Theeramunkong	In Celebration of The 20th Anniversary of SIIT	https://www.siit.tu.ac.th/upfiles/publication_file...	Dr. Thanaruk Theeramunkong (Head of the School of ...	1
8	2	Thanaruk	Theeramunkong	JAIST-NECTEC-SIIT PhD. Dual Degree Program Present...	https://www.slideshare.net/.../jaistnectecsiit-phd...	The graduate from this program will receive JAIST ...	1
9	2	Thanaruk	Theeramunkong	TAIST-TOKYO tECH ANNUAL REPORT 2014 - 東京工業大学 タイオフィ...	www.titl.ipo.titech.ac.jp/Newsletter%20Vol.%2010.p...	have strong communication skills when they graduat...	1
10	2	Thanaruk	Theeramunkong	ニュースレター Vol. 11 - 東京工業大学 タイオフィス	www.titl.ipo.titech.ac.jp/Draft-Newsletter/Vol11_ak...	Dr. Thanaruk Theeramunkong, ... that scholars sh...	1
11	2	Thanaruk	Theeramunkong	JAIST-NECTEC-SIIT PhD. Dual Degree Program Present...	https://vdocuments.site/.../jaist-nectec-siit-phd-...	Steering Members Mitsuru Ikeda (JAIST) Thanaruk Th...	1
12	2	Thanaruk	Theeramunkong	Dr tan teck jack - kam training services pvt. Ltd	kamtrainingservices.com/tovlsk3kd/gos/akwj3.php?db...	Finest Soeren Is A Business Graduate From The Univ...	1
13	2	Thanaruk	Theeramunkong	Towards ontology enrichment with treatment relatio...	dl.acm.org/citation.cfm?id=2173406&preflayout=flat...	Peerasak Intarapaiboon , Ekawit Nantajeewarawat ...	1
14	2	Thanaruk	Theeramunkong	Proceedings of NCCIT2014 - Scribd	https://www.scribd.com/document/223300399/Proceedi...	School of Basic Education and the GPA to graduate ...	1
15	2	Thanaruk	Theeramunkong	INSTRUCTIONAL DESIGN ONE YEAR BY SEMESTER - PDF	https://docplayer.net/17913515-Instructional-desig...	... they specified 16 credits which are a minimum ...	1
16	2	Thanaruk	Theeramunkong	rechenschaftsbericht des rektors - Hochschule Rave...	https://www.hs-weingarten.de/cldocument_library/ge...	after graduate from the University vorgestellt. Sa...	1
17	2	Thanaruk	Theeramunkong	Download Handbook Of Pediatric Retinal Disease - C...	cigamot.com/book/download-handbook-of-pediatric-re...	After all, we look the simple graduate from our gc...	1
18	3	Ekawit	Nantajeewarawat	SIIT News, Vol. 15, No. 1, March 2010	https://www.siit.tu.ac.th/upfiles/publication_file...	Dr. Ekawit Nantajeewarawat (Executive Assistant, D...	1

Figure 4: Example Retrieve information from Google

4. Set up rule based construction

After examining the snippets, we found that in each keyword, there are 2 spaces between them for answer extraction, we called them “distance A” which is between OOV and Keyword, and “distance B” is between the keywords and answer that we need as shown in figure 5.

Table 4: Example of rule-based pattern matching

No.	Rules Base
1	OOV + Keyword + Answer
2	OOV + Keyword + Answer + *Noise*
3	OOV + Keyword + *Noise* + Answer
4	OOV + Keyword + *Noise* + Answer + *Noise*
5	OOV + *Noise* + Keyword + Answer
6	OOV + *Noise* + Keyword + Answer + *Noise*
7	OOV + *Noise* + Keyword + *Noise* + Answer + *Noise*
8	*Noise* + OOV + Keyword + Answer
9	*Noise* + OOV + Keyword + Answer + *Noise*
10	*Noise* + OOV + Keyword + *Noise* + Answer
11	*Noise* + OOV + Keyword + *Noise* + Answer + *Noise*
12	*Noise* + OOV + *Noise* + Keyword + Answer
13	*Noise* + OOV + *Noise* + Keyword + Answer + *Noise*
14	*Noise* + OOV + *Noise* + Keyword + *Noise* + Answer + *Noise*

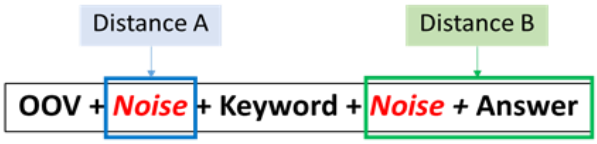


Figure 5: Rule explanation

We count the string spacing in distance A and distance B of every snippet as shown in Figure 6. Five researchers were selected as samples. In order to find the distance, mean and standard deviation are calculated so that most accurate data extraction can be achieved.

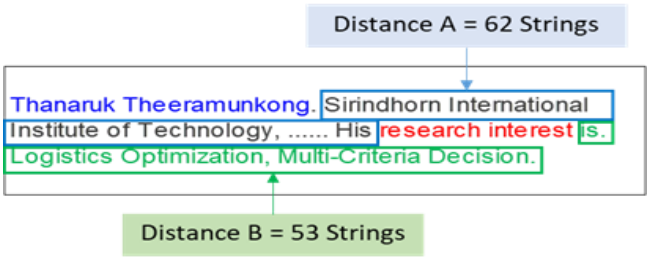


Figure 6: Example of distance A and distance B

5. Experiment rules based answer extraction

When we have the most suitable distance, we get the distance of each keyword to experiment with another 15 researchers

Results and Discussion

1. Result of keyword construction

After we have set up the keywords, we then use the four keywords to query Thai researchers from the internet. We found that in each keyword, the results will be different in the number of Snippets. Most data is the research area. It has 407 rows, but only 37 rows are correct, accounting for 9.09% of the total rows retrieved.

At the same time, a keyword with the most accurate and useful amount of data is research interest. There is a total of 320 data queries and 69 valid rows, accounting for 21.56% of the total rows retrieved.

According to table 5, if we look for research information, the term research interest should be used rather than the research area. In other words, research interest is more popular than in the research area.

Table 5: Result of retrieve data each keyword

Keyword	Total of Retrieve data	Incorrect data	Correct data	% Correct data
graduate from	90	87	3	3.33%
Work Experience	213	205	8	3.76%
research interest	320	251	69	21.56%
research area	407	370	37	9.09%

2. Result of rule-based answer extraction

We set up rules based on experiments with five researchers' data, we call them "training data". We check the data of five researchers to verify that each snippet contained information that could be used or had the exact information we needed in the OOV + Keyword + Answer structure. According to the experiment in each keyword, we get the distance that makes extract the most useful information as follows:

Table 6: Result of Distance for extract answer information

Keyword	Distance A (between OOV and Keyword)	Distance B (between Keyword + answer)
graduate from	100.97	54.03
Work Experience	77.83	64.86
research interest	67.86	64.98
research area	74.76	81.37

When we have the optimal Distance A and B, we test it with test data from 15 researchers. The results show that accuracy of the keyword "work experience" is the most accurate, 75%. Next is the research that is interested in the research area and graduate level in Table 7.

Table 7: Result of rule-based answer extraction between training data and testing data

Keyword	Total of Correct data	Training Data			Testing Data		
		Correct data	answer extraction	accuracy	Correct data	answer extraction	accuracy
graduate from	6	3	2	66.7%	3	2	66.7%
Work Experience	8	4	3	75.0%	4	3	75.0%
research interest	69	16	12	75.0%	53	38	71.7%
research area	37	9	6	66.7%	28	20	71.4%

From Figure 7, it can be concluded that the Keyword graduate from and Work Experience has the same accuracy of Training data compared to Testing data because of the same amount of data. While research interest has more accuracy of Training data than testing data at 3.3%. On the other hand, the research area has the accuracy of testing data more than training data at 4.7%.

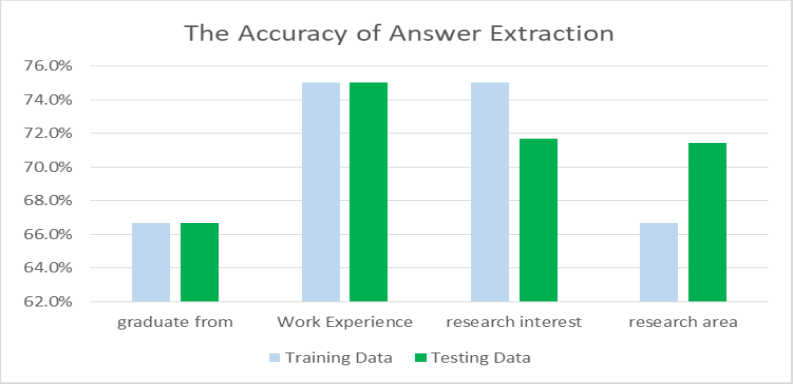


Figure 7: Result of accuracy answer extraction

Conclusions

In this article, we focus on the research of data extraction and build a prototype that can automatically retrieve data using Google APIs. Based on a sample of 20 Thai researchers, we found that the four keywords provided us with useful information: place of graduation, work experience, research interests, and research area. After that, we search for the distance between OOV and keywords, keywords and answers to get the most accurate information. The results showed that graduates with 66.7% accuracy, 75.0% accuracy in work experience, research interest with 71.7% accuracy, a research area with 71.4% of accuracy.

This research will greatly improve works in the field of Artificial Intelligence (AI), especially in the area of searching and extraction information for Thai researchers. Moreover it can be a guideline for applying rules-based pattern matching in future tasks such as Design Program Personal Learning Solution for HR.

Acknowledgments

The first author Nalinee Wuttivichayanan contributed 50% of this research, conducted the research and authored the paper. The last author Jian Qu contributed 48% for this research, guided the research and co-authored the paper.

We would like to express our sincerely thanks to Mr. Chinorot Wangtragulsang for his support in English grammar checking and guide for programing.

References

C.Daya Wimalasuriya and Dejing Dou (2010), Ontology based information extraction: An Introduction and a survey of current approaches, *Journal of Information Science*, June 2010, 36:306-323.

Paul O' Neil and Woojin Paik (1998), the Chronological Information Extraction System (CHESS), *IEEE*, 1998.

Ying Chen, Sophia Yat Mei Lee, Chu-Ren Huang (2009), Polyuhk: A robust information extraction system for web personal names, April 2009.

Ying Chen, Sophia Yat Mei Lee, Chu-Ren Huang (2012), A robust web personal name information extraction system, Expert Systems with Applications: *An International Journal, February 2012, Vol. 39*, No. 3, pp. 2690-2699.

Yaoyong Li, Kalina Bontcheva, and Hamish Cunningham (2004), SVM Based Learning System for Information Extraction, *Deterministic and Statistical Methods in Machine Learning, September 2004*, pp. 319-339.

Kritsada Sriphaew, Thanaruk Theeramunkong (2002), A New Method for Finding Generalized Frequent Itemsets in Generalized Association Rule Mining, *IEEE, Proceedings ISCC 2002 Seventh International Symposium on Computers and Communications*, Taormina-Giardini Naxos, Italy, July 2002, pp. 1040-1045.

J. Qu, L. M. Nguyen and A. Shimazu (2016), Cross-language information extraction and auto evaluation for OOV term translations, *IEEE transactions, in China Communications, December 2016, vol. 13*, no. 12, pp. 277-296.

J. Qu, T. Theeramunkong, C. Nattee, and P. Aimmanee (2011), Web Translation of English Medical OOV Terms to Chinese with Data Mining Approach, *Science & Technology Asia, vol. 16*, no. 2, pp. 26-40.

J. Qu, Shimazu A., Le Nguyen M. (2012) OOV Term Translation, Context Information and Definition Extraction Based on OOV Term Type Prediction. In: Isahara H., Kanzaki K. (eds) Advances in Natural Language Processing. *JapTAL 2012*. *Lecture Notes in Computer Science, vol 7614*. Springer, Berlin, Heidelberg

J. Qu, Nguyen Le Minh and Akira Shimazu (2011), Web based English-Chinese OOV term translation using Adaptive rules and Recursive feature selection, Institute of Digital Enhancement of Cognitive Processing, Waseda University, *Proceedings of the 25th Pacific Asia Conference on Language*, Information and Computation, pp. 1-10.

Ma Y., Zeng Y., Ren X., Zhong N. (2011) User Interests Modeling Based on Multi-source Personal Information Fusion and Semantic Reasoning. In: Zhong N., Callaghan V., Ghorbani A.A., Hu B. (eds) Active Media Technology. AMT 2011. *Lecture Notes in Computer Science, vol 6890*. Springer, Berlin, Heidelberg

P. Soucy and G. W. Mineau, "A simple KNN algorithm for text categorization," *Proceedings 2001 IEEE International Conference on Data Mining, San Jose, CA, USA, 2001*, pp. 647-648.

W. Zheng, G. Feng and Z. Nan, "An Improved KNN Text Categorization Method Based on Spanning Tree Documents Clustering," *2011 International Conference on Internet Technology and Applications, Wuhan, 2011*, pp. 1-5.

Sang-Bum Kim, Kyoung-Soo Han, Hae-Chang Rim and Sung Hyon Myaeng, "Some Effective Techniques for Naive Bayes Text Classification," in *IEEE Transactions on Knowledge and Data Engineering, vol. 18*, no. 11, pp. 1457-1466, Nov. 2006.

EXPERIMENTAL AND NUMERICAL ANALYSIS OF SMOKE CONTROL IN ATRIUM BUILDING

Sisira Jaimun[1], Supat Patvichaichod[2]

[1,2] Department of Mechanical Engineering, Faculty of Engineering at Sriracha, Kasetsart University Sriracha Campus
[*]**Corresponding author, E-mail:** sisira.ja@ku.th

ABSTRACT

In this research, the simulation of smoke control in atrium is carried out by using two different methods which are 1:28 scale model and FDS simulation of the actual building on the residential building with atrium and interior balcony to see if their results are consistent with each other; the comparison of the results show that the discrepancy of the decrease in smoke layer between them is unacceptably huge. Furthermore, the smoke movement affected by the different air velocities (0.5, 1, 1.5 m/s) of the air supply vents is later studied by using FDS simulation which tests on three areas of fire sources: at the center, at the side wall, and at the corner with 4 and 6 exhaust vents on each area. The result of the study shows that different air velocities has an effect on the fire source at the center, and the smoke ventilation by 6 exhaust vents with the velocity of 1 m/s is the most appropriate.

Keywords: atrium, smoke control, fire dynamic simulator, scale model test

Introduction

Nowadays, atrium, a large open space rising through several stories with a translucent roof, is often found featured in typically crowded large-space buildings such as residential buildings, shopping malls, school buildings, and etc. So, due to its capability of occupancy, the occurrence of fire incidents in the building with atrium can seriously harm people's lives and belongings. According to ISO 13571 (2007), there are 4 main factors obstructing a fire evacuation: asphyxiant fire gases, irritant fire gases, smoke, and heat. The death during fire incident is mainly caused by breathing in too much of asphyxiant smoke; asphyxiant gases that can lead to the unconsciousness and death are carbon monoxide (CO) resulted from an incomplete combustion and Hydrogen Cyanide (HCN) found in polyurethane (Gann et al., 2011), often used as a coating for furniture. The amount of toxic released smoke depend on kinds of fuel, oxygen, and combustion time. The smoke will move in a vertical direction from the ground by buoyant force and then cover the entire area. Therefore, the existence of smoke control system is very important for smoke exhaust in order to help occupants escape out of a fire area safely.

In the design of smoke control system for atrium, air supply vents are required for air intaking purpose as it helps the smoke to rise higher up to the ceiling. Air supply vents should be symmetrically arranged and has the velocity of air intaking at an appropriate rate (Gutiérrez-Montes et al., 2010; Kerber and Milke, 2007; Rafinazari and Hadjisophocleous, 2018). Then, the smoke will be exhausted out of the building to the atmosphere through the smoke exhaust vents installed on the ceiling. The natural ventilation is mostly used in the atrium building for both air intaking and smoke extracting (Chow and Li, 2018; Qin et al., 2009). Due to use of natural ventilation, air velocity, temperature, and the shape of the atrium hold a lot of effects to the systems (Król, 2016; Król and Król, 2017). Those effects attracts a variety of tests including full scale test, fire dynamic simulator (FDS) test, scale model test, and etc. (Capote et al., 2009; Quintiere et al., 1978). However, most

experiments has been carried out with no interior balcony in the building and often test the air velocity of the air supply vents with the same number of smoke exhaust vents. So, this research has decided to do a test on the scale modeling and FDS modeling of atrium building with interior balcony to study the consistency between the two tests and study the smoke movement affected by the different velocities of air supply that is pulled into the atrium which has an effect on the fire source and natural smoke ventilation.

Objectives

1. to compare the smoke movement in two different tests which are scale model and FDS test to investigate the consistency between the two.
2. to study the movement of the smoke that spread to each storey from the influence of velocities of air supply in which affects the fire sources and the performance of natural ventilation.
3. to study the efficiency of smoke ventilation in the atrium by 4 and 6 vents.

Literature Review

The review of literature consists of applied theories and additional information relative to the smoke control system in atrium designing as described below.

1. Plume

In the events of fire, there is a release of smoke which is solid and liquid particulates evolved from the combustion caused by the heat spreading in the air. McCaffrey (1979) has divided the range of Axisymmetric Plume ,which is explained in NFPA 92 (2018) as a plume that rises above a fire, does not come into contact with walls or other obstacles, and is not disrupted or deflected by airflow, into 3 parts as follows:

1.1 Continuous flame; it can be measured from the surface of the fire to the highest point that the fire can still be seen continuously which means the plume of this range has a high temperature.

1.2 Intermittent flame; it is the range of height where the combustion is almost complete, so the flame can be seen partially and intermittently. The heat flow of the plume in vertical direction is quite steady. The mixture of the outside air with the smoke causes the increase in plume flow rate as the height is added up.

1.3 Plume; it is the range of height where the combustion is complete. The plume in this range consists of hot smoke and added air from the outside surrounding. Since the air outside has cooler temperature, so when it mixes with the smoke, the heat and temperature of the plume is in inverse to the height causing the smoke to become cooler.

Moreover, NFPA 92 has also identify the definitions of each smoke position which can be divided into smoke layer and smoke layer interface; smoke layer is the accumulated thickness of smoke below a physical or thermal barrier. Smoke layer interface is the theoretical boundary between a smoke layer and the smoke-free air.

2. Engineering Modeling

In the field of engineering, a model is a physical representative researchers use to explain the phenomena occurring in the system they want to study by using numerical analysis technique. Due to the safety and financial reasons, the test object is scaled down to the model size relative to the size of the actual object. In this research, Dimensionless Modeling and Similitude (Munson et al., 2010) is used for the downsizing purpose as in stated equation (1) – (2) together with Froude Modeling (NFPA

92, 2018) to determine the fuel position as stated in equation (3) in order to keep the consistency between scale model test and FDS test at an appropriate level.

$$w_m/h_m = w/h \quad (1)$$

Since h is the height of the actual building (m), h_m is the height of the scale model (m). Also, w is the width of the actual building (m), and w_m is the width of the scale model (m).

$$A_m = A(l_m/l)^2 \quad (2)$$

Since A is the area of the actual fire source (m^2), A_m is the area of the model fire source (m^2). Also, l is the length of the actual building (m), and l_m is the length of the scale model (m).

$$x_m = x_F(l_m/l_F) \quad (3)$$

Since x_F is the position of the actual fire source, x_m is the position of the model fire source. Also, l_F is the length of the actual fire source (m), and l_m is the length of the scale model (m).

The numerical analysis of FDS model applies a calculating principle which separates an entire area into grids or meshes before the temperature, density, speed, and chemical components in each mesh are computed; however, the mesh size should be determined appropriately as in the equation below.

$$D^* = \left(Q/\rho C_p T_o \sqrt{g}\right)^{2/5} \quad (4)$$

D* is the characteristic fire diameter (m). Q is a heat release rate (kW). ρ is air density (kg/m^3). C_p is the specific heat (kJ/kg-K). T_o is the surrounding air temperature (K). Lastly, g is the gravitational acceleration (m/s^2).

The size of mesh can be calculated out of this D*, but the value of D* will be set differently due to the different types of mesh which are coarse (D*/dx =4), moderate (D*/dx =10), and fine (D*/dx =16) (USNRC, 2007).

Currently, the smoke control system has been widely studied and analyzed with various methods including FDS (Fire Dynamics Simulator) model, full scale test, and scale model test. Capote et al. (2009) has done the study that compares the smoke movement in FDS model with in 1:9 model based on Froude Modeling. The result indicates that the height of smoke layer depend on the temperature during the test, and the velocity of a decrease of smoke layer in scale model is slower than in FDS due to the heat loss during the test. However, Quintiere et al. (1978) also show that the scale test based on the Froude Modeling gives a result closer to the large scale test.

3. Fire Source

To select the size of fire, Yuen and Chow (2005) let the designer choose the heat release rate by considering the types of building in order to design a smoke management system; the suitable heat release rate for atrium is 7 MW. In addition, Qin et al. (2009) has studies the location of fire sources as follows: at the center, at the side wall, and at the corner, and it has been found out that the fire source at the corner has the slowest decrease, the highest density, and the lowest average temperature of smoke layer; the fire source at the center has the highest average smoke temperature; for the fire source located at the side wall, the spread of the smoke is limited causing the fluctuation in ambient air pressure.

4. Air Supply in Atrium

NFPA 92 (2018) prescribes that the make-up air velocity should not exceed 1.02 m/s to prevent smoke from being disturbed. Kerber and Milke (2007) have done the investigation on the possible effects of symmetrical make-up air supply arrangement with 4 supply vents and the velocity

at 0.5 to 3 m/s, and found out that the velocity of the symmetrical supply vents of the fire source at the center should be at 0.5 and 1m/s to keep smoke unbothered. When the air velocity of the air supply increases, the size of the vents should be smaller to keep volumetric flow rate balanced caused the heat release rate to be higher and the height of smoke layer to be lower which means the higher the velocity can get the stronger the fire resistance can be. Gutiérrez-Montes et al. (2010) also states that the use of symmetrical make-up air supply will not cause fire and smoke disturbance.

5. Smoke Exhaust Vents

Qin et al. (2009) has adopted FDS to investigate the smoke movement in atrium at the same make-up air velocity, and the result indicates that natural smoke exhaust method is necessary when the vents are located on the ceiling especially when the heat release rate hasn't been specified properly. From natural exhaust system test result, if there are more exhaust vents, the smoke pressure will be decrease due to the more amount of air being pulled into the atrium which led to more amount of smoke and the fast decrease of smoke layer corresponding with what Chow and Li (2018) has stated that natural ventilation is popular with large atrium, often installed on the ceiling and has buoyancy as the main force in smoke removal.

Methods

1. Acrylic Scale Modeling for Smoke Movement Testing

The acrylic model is scaled down to ratio size of 1:28 by Dimensionless Modeling and Similitude method; the model is constructed following the design of atrium building with interior detail as follows: the size of the first floor is 28×24×3.8 m, the size of the sond floor to the ninth floor is 24×12×24 m, each floor is 2.9 m high with the 2 m wide interior balcony on the sond floor to the ninth, and 2×1.8 m doors of both sides, and 2.5×1.25 m natural smoke exhaust vents on the ceiling. After being scaled down, the first floor is left with 0.6×0.514×0.081 m; the sond to ninth floor is 0.514×0.26×0.514 m. Each floor is 0.063 m far from one another, and the interior balcony on the sond floor to the ninth is 0.043 m. The doors of both sides has the size of 0.02×0.018 m, and the reduced size of smoke exhaust vents on the ceiling is 0.025×0.0125 m.

The smoke is simulated by lighting on a wick with Kerosene fuel. The wick, 0.1 m high with the size of 0.0021 $m^{2,}$ is placed at the center of the model. The test is done in 2 cases: 4 exhaust vents and 6 exhaust vents inside a closed room, so the wind force has no effect to the test. The smoke movement will be recorded for 180 s, and the spread and the decrease of smoke layer at 30, 60, 120, and 180 s will be compared with the FDS model later on.

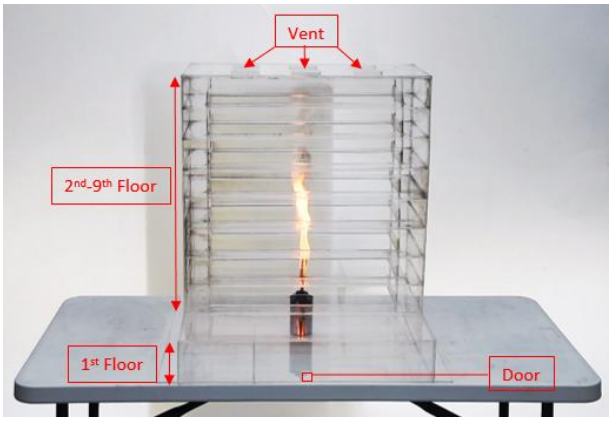


Figure 1: Scale model test

2. Conducting FDS Test to Compare with the Scale Model Test

Fire Dynamic Simulator simulates the actual building size with detailed stated in Table 1. There are 2 cases to be carried out which are 4 exhaust vents and 6 exhaust vents same as in the scale model test. FDS model is showed in Figure 2.

Table 1: Details of building components specified in FDS

Details of building components	
Atrium dimension	1st floor : 28×24×3.8 m 2nd - 9th floor : 24×12×24 m
Doors of both sides dimension	2×1.8 m
Interior balcony 2nd - 9th floor dimension	x axial : 8×2 m y axial : 24×2 m
Rail of interior balcony 2nd - 9th floor dimension	x axial : 8×1 m y axial : 20×1 m
Material	Concrete
Exhaust vents dimension	2.5×1.25 m

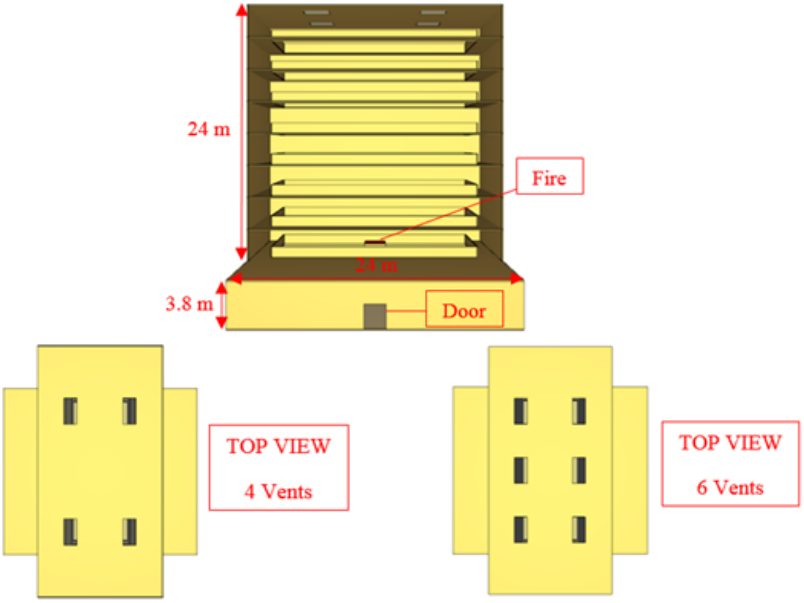


Figure 2: FDS model

Fire sources in FDS model are determined by the size of a wick by using Dimensionless Modeling and Similitude method together with Froude Modeling to find the appropriate height and size of a fire source. Therefore, the fire source at the center is at a height of 4.67 m and the size of 4.57 m^2 as showed in Figure 3 and the fuel quality is determined following the details in Table 2 for FDS calculating purpose.

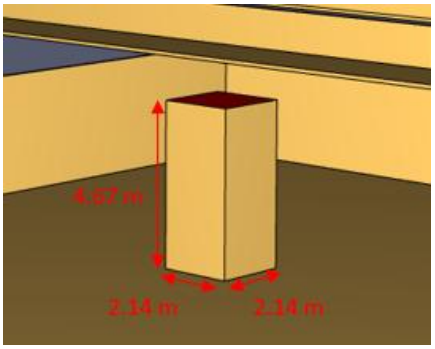


Figure 3: Fire source Kerosene fuel at the center

Table 2: Property Kerosene

	Set value
Surface (Fire)	
Heat Release Rate Per Unit Area (HRRPUA)	2300 kW/m^2
Reaction	
Fuel	Kerosene
Fuel Composition	C = 1, H = 0.011, O = 0.168 and N = 0
Fire Suppression	Critical Flame Temperature = 1027 ℃
Byproducts	Specify release per unit mass oxygen = 1.31E4 kJ/kg y_{CO} = 0.012, y_{soot} = 0.042, Hydrogen Fraction = 0.1

With heat release rate of kerosene at 2300 kW, the equation (4) can be calculated to find D* by substituting each parameter as follows: ρ = 1.204 kg/m^3, C_p = 1.005 kJ/kg-K, T_o = 305 K and g = 9.81 m/s^2 resulted in

$$D^* = \left(2300/1.204\times1.005\times305\sqrt{9.81}\right)^{2/5} = 1.32 \text{ m}$$

The mesh around fire source determined as moderate has the resolution of 0.1 m, so multiplying 0.1 m with D* = 1.32 m, the size of moderate mesh becomes 0.132 m. For coarse mesh around the building with the resolution of 0.25 m, the size of mesh can be calculated as 0.33 m by multiplying 1.32 m. Due it the different size of mesh in the same domain, the problem of mesh alignment resurfaces. To fix that, the mesh sizes are rounding up to 0.2 m and 0.4 m respectively.

3. FDS Modeling for the Study of the Smoke Movement Influenced by the Fire Source Locations, Air Velocities of Supply Vents, and the Different Exhaust Vents

To study the smoke movement in Atrium, the influence of surrounding conditions which are fire source locations, air velocity of supply vents, and the capability of smoke ventilation by different exhaust vents are considered. The details of each condition are explained as follows: three locations (at the center, at the side wall, and at the corner), symmetric 2 supply vents with three velocities (0.5, 1, and 1.5 m/s.), and two types of exhaust vents which are 4 and 6 vents. With the details earlier, the case studies can be varied as showed in table 3.

Table 3: Cases study

Case	Fire source	Air velocity of supply vents (m/s)	Air supply vents Dimension (m)	Number of air supply vent	Smoke exhaust vents Dimension (m)	Number of smoke exhaust vents
1	Center	0.5	21.6 × 4	2	2.5 × 1.25	4
2	Center	0.5	21.6 × 4	2	2.5 × 1.25	6
3	Center	1	18 × 2	2	2.5 × 1.25	4
4	Center	1	18 × 2	2	2.5 × 1.25	6
5	Center	1.5	10.8 × 2	2	2.5 × 1.25	4
6	Center	1.5	10.8 × 2	2	2.5 × 1.25	6
7	Side wall	0.5	21.6 × 4	2	2.5 × 1.25	4
8	Side wall	0.5	21.6 × 4	2	2.5 × 1.25	6
9	Side wall	1	18 × 2	2	2.5 × 1.25	4
10	Side wall	1	18 × 2	2	2.5 × 1.25	6
11	Side wall	1.5	10.8 × 2	2	2.5 × 1.25	4
12	Side wall	1.5	10.8 × 2	2	2.5 × 1.25	6
13	Corner	0.5	21.6 × 4	2	2.5 × 1.25	4
14	Corner	0.5	21.6 × 4	2	2.5 × 1.25	6
15	Corner	1	18 × 2	2	2.5 × 1.25	4
16	Corner	1	18 × 2	2	2.5 × 1.25	6
17	Corner	1.5	10.8 × 2	2	2.5 × 1.25	4
18	Corner	1.5	10.8 × 2	2	2.5 × 1.25	6

The FDS model is simulated following the building size in Table 1. Since the velocity is involved in air supplying, if the size of supply vents has to alter according to the velocity. In addition, the size of supply vents is bit by bit from the actual door size, and velocity measuring devices are installed around the supply vents on both sides as showed in Figure 4. After that, the velocity measuring result will be created into a graph average velocity as in the Figure 5, which indicates the supply vents change their size 3 times, 5 times, and 10 times bigger than the actual door size, as showed in Table 3. Due to the velocity of air supplying at 1.5 m/s, the height of the first floor should be 6.8 m which is agreeable with the supply vents height as showed in Figure 6.

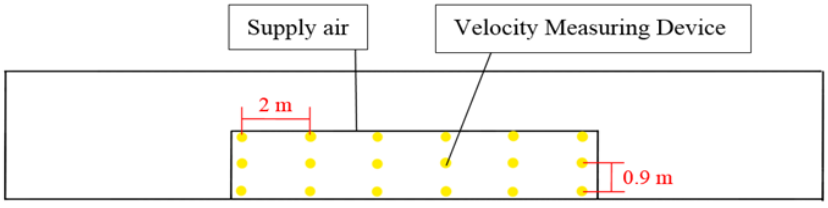


Figure 4: Velocity measuring devices arrangement

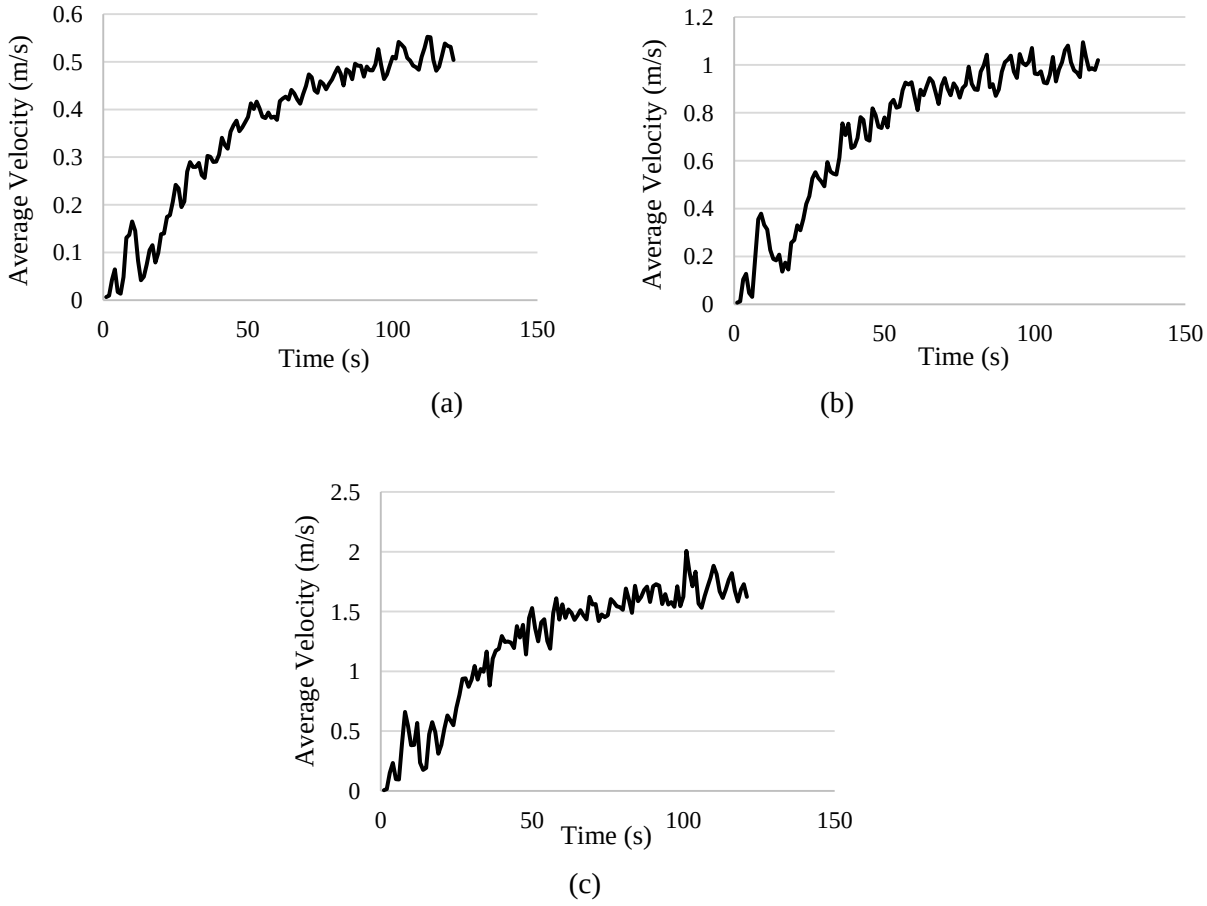


Figure 5: Average velocity of air supply vents (a) 0.5 m/s, (b) 1 m/s and (c) 1.5 m/s

In Figure 5, there are average velocity result measuring from the installed around devices air supply vents on both sides. Figure 5(a) shows the average velocity of the supply vents measuring 21.6×4 m at 0.5 m/s. Figure 5(b) shows the average velocity of 1 m/s from air supply vents measuring 18×2 m. Figure 5(c) shows the average velocity of air supply vents sized 10.8×2 m at 1.5 m/s.

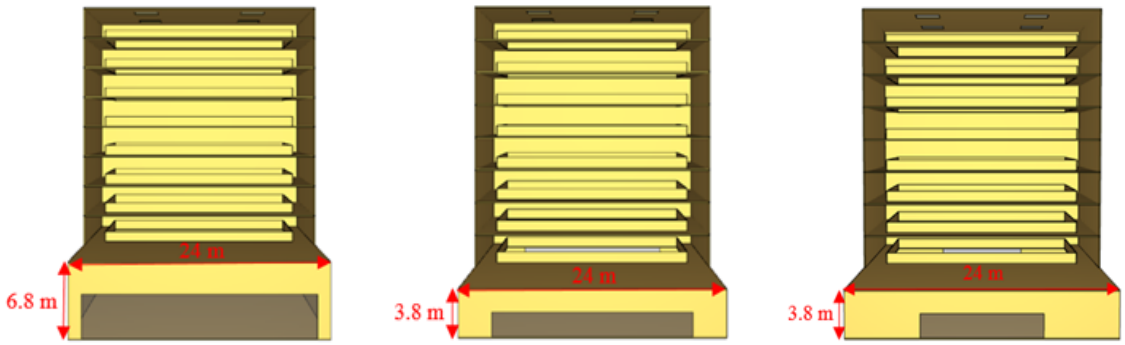


Figure 6: FDS models, air supply vents dimension 21.6×4, 18×2 and 10.8×2 m

The FDS model simulates fire sources at the center, on the side wall, and at the corner that has fire surface size set at 2×2 m on the first floor to visualize the situation where fire is on the ground as showed in Figure 7 by using Polyurethane as fuel in the test since it is the material that found the most in the residential buildings; The fuel property is elaborated in Table 4.

Table 4: Property Polyurethane

	Set value
Surface (Fire)	
Heat Release Rate Per Unit Area (HRRPUA)	7 MW = 7000 kW ∴ HRRPUA Fire area 4 m^2 = 1750 kW/m^2
Reaction	
Fuel	Polyurethane
Fuel Composition	C = 6.3, H = 7.1, O = 2.1 and N = 1
Fire Suppression	Critical Flame Temperature = 1427 ℃
Byproducts	Specify release per unit mass oxygen = 1.31E4 kJ/kg y_{CO} = 0, y_{soot} = 0.1, Hydrogen Fraction = 0.1

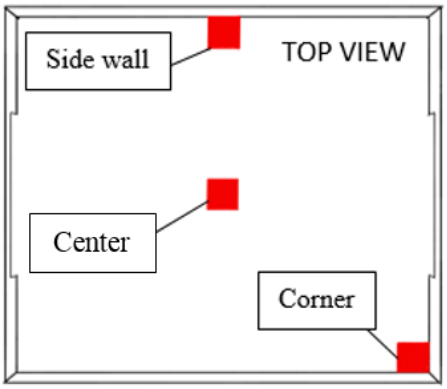


Figure 7: Fire source Polyurethane fuel at the center, side wall and corner

From the heat release rate at 7000 kW of Polyurethane, D* can be calculated by substituted parameters with the values as follows: ρ = 1.204 kg/m^3, C_p = 1.005 kJ/kg-K, T_o = 305 K and g = 9.81 m/s^2 which resulted in

$$D^* = \left(7000/1.204\times1.005\times305\sqrt{9.81}\right)^{2/5} = 2.055 \text{ m}$$

The mesh around fire source determined as moderate has the resolution of 0.1 m, so multiplying 0.1 m with D* = 2.055 m, the size of moderate mesh becomes 0.2055 m. For coarse mesh around the building with the resolution of 0.25 m, the size of mesh can be calculated as 0.5138 m by multiplying 2.055 m. Due it the different size of mesh in the same domain, the problem of mesh alignment resurfaces. To fix that, the mesh sizes are rounding up to 0.2 m and 0.4 m respectively.

After simulating all 18 cases, as showed in Table 3, the results are run with the time of 120 s to steady the test result, and then compare them with the smoke movement at 30 s, 60 s, 90 s and 120 s.

Results and Discussion

The Comparison of Smoke Movement between in Scale Model Test and in FDS Model Test

When comparing the scale model test done in a closed room to the FDS model simulating the actual building with the same conditions of the fire source located at the center, and no air supply velocities involved in 2 cases of exhaust vents which are 4 and 6 vents at 4 point in time duration of 180 s (30, 60, 120, and 180), the smoke visibility results of both cases are the same at 30 s; the smoke inside and outside the scale model ,being exhausted, can be clearly seen. At 60 s, the visible smoke starts to fade away until at 180 s that the fire dies down due to heat loss, the smoke distribution can't be seen anymore; additionally, each time the fire being ignited, the amount of smoke is unsteady as there are deviations in time and wick. For FDS simulation, the smoke is visible clearly at every point of times and the fuel package can be ignited all the time, so the fire ignition of each case can be spotted quickly without any time deviations occurred. Also, the determination of fuel properties enables the amounts of fire and smoke to stay consistent.

The smoke movement from the scale model test at 30 s in the cases of 4 and 6 exhaust vents are seen to spread around the whole area of the 6th to the 9th floor, and the smoke layer decrease down to the fifth floor. At 60 s, the smoke seems to be faded, but the spread and the smoke layer become more stable than earlier. When the time reaches 120 s and 180 s, the smoke rises to the ceiling and is exhausted out of the building so suddenly that the spread of the smoke to the lower storeys can't be observed.

The smoke movement from the FDS model test in the cases of 4 and 6 exhaust vents at the times of 30, 60, 120, and 180 s have the same result. At 30 s, the smoke spreads to the entire area of the 7th to the 9th floor and covers above the ceiling part of the 5th to the 6th floor. At 60 s, the smoke spreads to the entire area of the 5th to the 9th floor and covers the ceiling part of the 4th floor left out only a lower area of the interior balcony; a little bit of the smoke lower down to the 2nd and 3rd floor. At 120 and 180 s, the smoke movement is stable since the smoke spread to the entire area of the 4th to the 9th floor, and some of it lower down to the 3rd floor spreading just a little to both sides of atrium as shows in Figure 8.

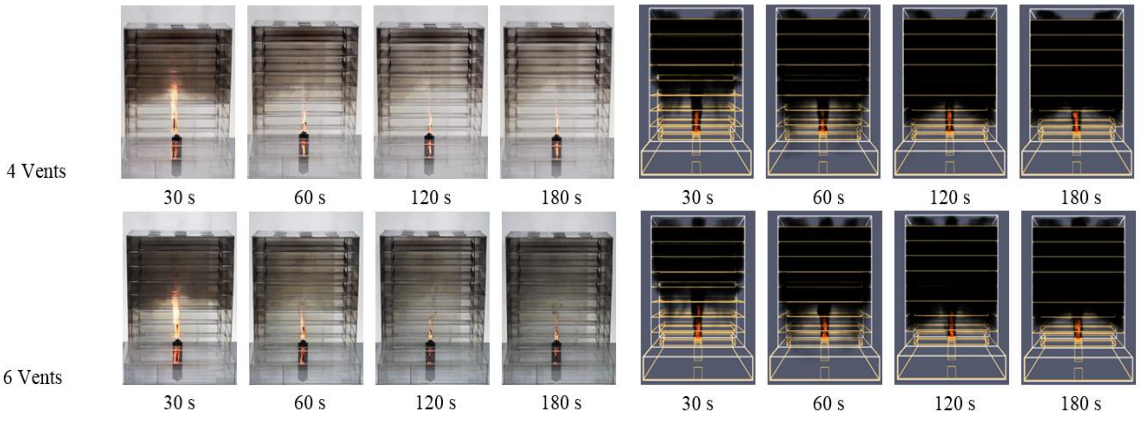


Figure 8: Smoke movement from scale tests and FDS (4 and 6 exhaust vents)

In Figure 9 showing the comparison chart between the smoke layer depth measuring from the lowest part of smoke to highest part on the ceiling of atrium from scale tests and FDS, the scale test result indicates both cases of exhaust vents have the same levels of smoke layer depth at 30 and 60 s. When the time reaches 120 s and 180 s the smoke distribution can't be seen anymore. From the FDS

test, it is found that both cases of vents have the same level of smoke layer depth regardless of a little smoke in some parts. The smoke layer is thicker in FDS model since there is no influence of the light-wick range which causes the flame to be higher in scale model than in the FDS one where the fire is set up on fire sources resulted in lower flame position. Due to the height of flame in scale model, the smoke occurs at the higher area, and the smoke layer will not lower down from the range of intermittent flame. In addition to the existence of vents in the tests, since the scale model test has been done in a closed room, the velocity of surrounding air holds no effects to the smoke movement compared to FDS test where the smoke is disturbed by the influence of air velocity causing some smoke to decrease down to the sides of fire sources. Therefore, it can be inferred that FDS test gives out a more dangerous close to the real situation and visible simulation, so the next test will be appropriately carried out by FDS method.

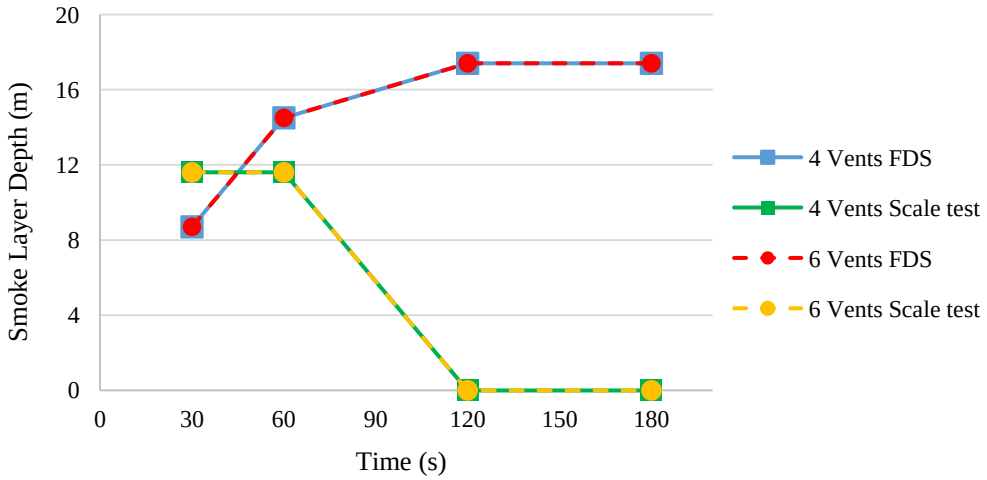


Figure 9: Comparison chart smoke layer depth between scale tests and FDS (4 and 6 exhaust vents)

The Result of FDS Test Involving Air Velocities of Supply Vents

The result of FDS test on smoke movement of fire sources located at the center, at the side wall, and at the corner with the three velocities (0.5, 1, and 1.5 m/s) of air supply vents when opening 4 and 6 exhaust vents ,as shows in Table 3, at 30, 60, 90, and 120 s is described below.

In the test of 1-6 cases of smoke movement with fire source located at the center, the smoke rises to the center of the ceiling, then spread out from the center to the 4 sides of atrium covering the 9th floor entire area. After that, the smoke eventually lowers down to cover the lower floors respectively. For 7-12 cases with fire source at the side wall, the smoke moves in a vertical direction along the wall. When meeting the ceiling of the 1st floor, the smoke divides into 2 parts which, at the same time, the first part spread to the whole area of 1st floor and the second one rises to the ceiling before spread out from the fire position in horizontal direction. For case 13-18 with the fire source at the corner, the source area is limited causing the smoke to rise to the first floor ceiling and spread out vertically along the interior balcony. Then, the smoke moves vertically through atrium. It can be concluded that with fire sources at the side wall and at the corner, the smoke can spread to cover the entire building swiftly as shows in Figure 10.

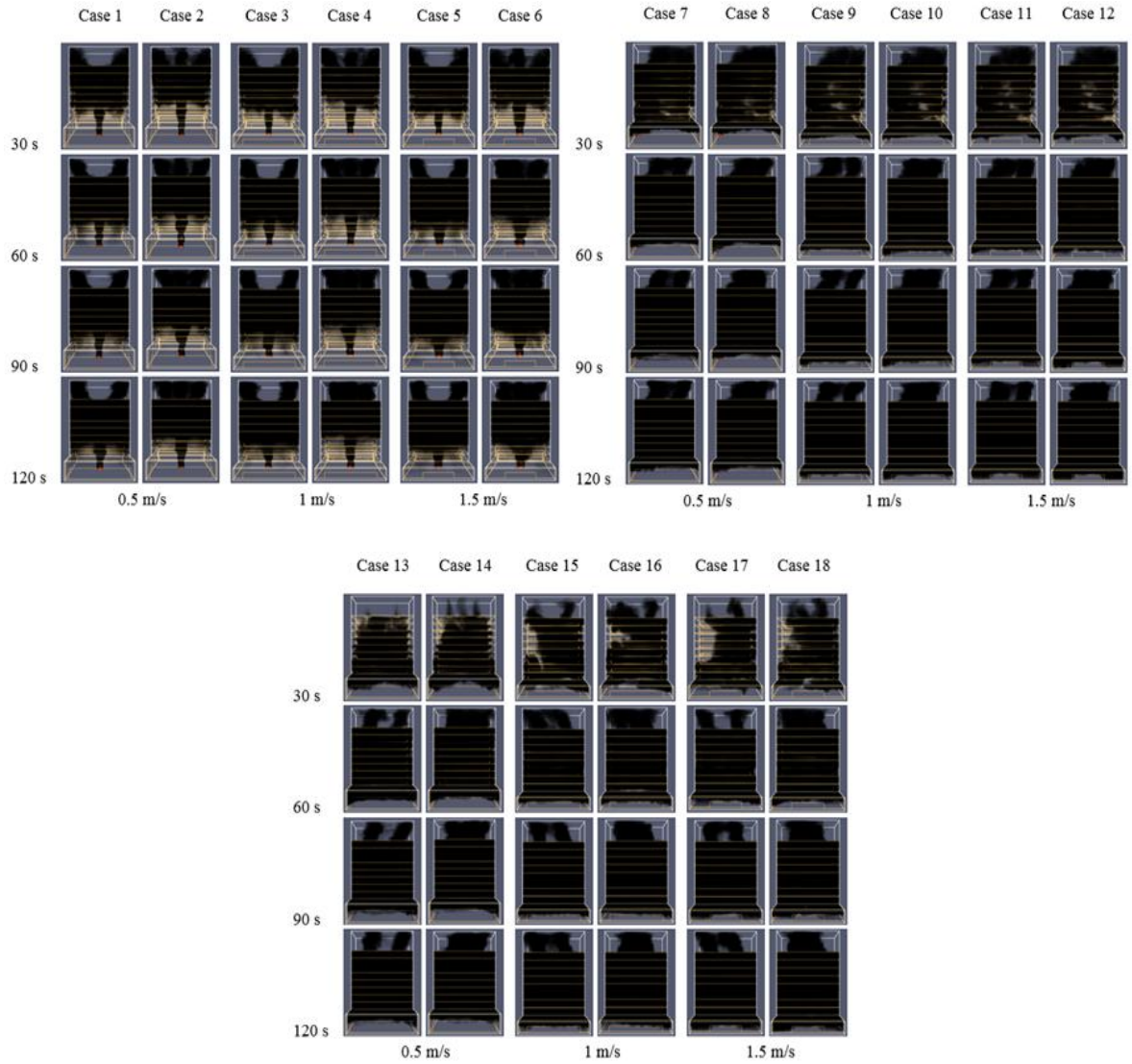


Figure 10: Smoke movement from FDS 18 cases

When creating all 18 cases into the chart comparing the smoke layer depth measuring from the lowest part of smoke to highest part on the ceiling of atrium as shows in Figure 11, it can be noticed that the smoke layer depth is equal and at its highest in both sources of fire which are at the side wall and at the corner. So, it means the influence of air velocities and air ventilation can be analyzed from the fire source at the center of atrium since the smoke layer will decrease down to the lowest which is on the 3rd floor encouraging the visibility of the increasing and decreasing smoke. With air velocities (0.5 and 1.5 m/s) of supply vents, the smoke layer is at its thickness. At 1 m/s, both 4 and 6 exhaust vents has the thickness of smoke layer at its lowest. However, the smoke layer depth in 6 exhaust vents is measured 14.5 m or equal to the smoke decreasing to the 5th floor of the building while the smoke layer depth in 4 exhaust vents is measured 17.4 or equal to the smoke decreasing to the 4th floor of the building.

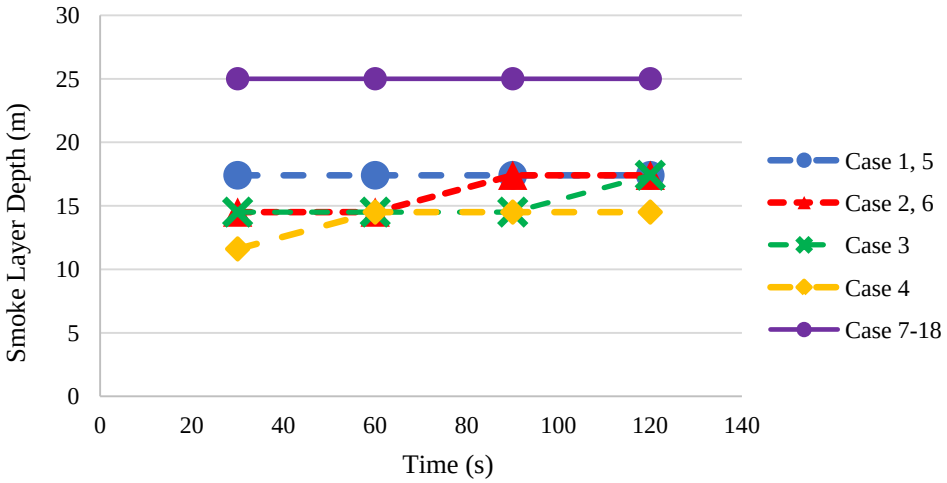


Figure 11: Comparison chart smoke layer depth from FDS 18 cases

Conclusions

The two methods used in the research which are the 1:28 scale modeling by Dimensionless Modeling and Similitude and the FDS modeling by Froude Modeling can be effectively utilized for the modeling even if the result from comparing the smoke visibility between in scale model test and in FDS test is inconsistent due to the difference in heat loss since in FDS test, the fire can be ignited all the time so that the spread and the existence of smoke can be seen more clearly opposite to scale model where heat loss occurs causing fuel package to eventually die down. The result considering the smoke layer depth of 4 and 6 exhaust vents when there is no air supply velocities involved. It has been found that the number of smoke exhaust vents has no effect on smoke ventilation.

From the model study of the influence that air velocities have on fire sources and air ventilation by 4 and 6 exhaust vents. It has been found that the velocity has no effect on fire sources at the side wall and at the corner due to the rapid decrease of smoke layer causing the smoke to spread to the entire area of the building but has the effect on the fire source at the center since the smoke can't cover the entire area of the building which gives the clearer vision to the fluctuation of smoke. After analyzing air ventilation of fire source at the center, the result shows that the smoke layer decreased down the most at the velocities of 0.5 and 1.5 m/s. The influence of air velocities from air supply vents causes the better performance of 6 exhaust vents than that of 4 vents when the velocity is at the appropriate rate of 1 m/s.

References

Capote, J., Alvear, D., Abreu, O., Lazaro, M., & Espina, P. (2009). Scale tests of smoke filling in large atria. *Fire technology, 45*(2), 201.

Chow, W. K., & Li, J. (2018). Wind action on natural smoke exhaust in atria. *Journal of computational science, 28*, 140-147.

Gann, R. G., Babrauskas, V., Grayson, S. J., & Marsh, N. D. (2011). Hazards of combustion products: Toxicity, opacity, corrosivity, and heat release: The experts' views on capability and issues. *Fire and Materials, 35*(2), 115-127.

Gutiérrez-Montes, C., Sanmiguel-Rojas, E., & Viedma, A. (2010). Influence of different make-up air configurations on the fire-induced conditions in an atrium. *Building and Environment, 45*(11), 2458-2472.

ISO 13571. (2007). *Life-threatening components of Fire-Guidelines for the estimation of time available for escape using fire data*.

Kerber, S., & Milke, J. A. (2007). Using FDS to simulate smoke layer interface height in a simple atrium. *Fire technology, 43*(1), 45-75.

Król, M. (2016). Numerical studies on the wind effects on natural smoke venting of atria. *International Journal of Ventilation, 15*(1), 67-78.

Król, M., & Król, A. (2017). Multi-criteria numerical analysis of factors influencing the efficiency of natural smoke venting of atria. *Journal of Wind Engineering and Industrial Aerodynamics, 170*, 149-161.

McCaffrey, B. J. (1979). *Purely Buoyant Diffusion: Some Experimental Results*: National Bureau of Standards, Gaithersburg, MD.

Munson, B. R., Okiishi, T. H., Huebsch, W. W., & Rothmayer, A. P. (2010). *Fluid Mechanic Seventh edition*: John Wiley & Sons Singapore Pte. Ltd.

NFPA 92. (2018). *Standard for smoke control systems*: National Fire Protection Association.

Qin, T., Guo, Y., Chan, C., & Lin, W. (2009). Numerical simulation of the spread of smoke in an atrium under fire scenario. *Building and Environment, 44*(1), 56-65.

Quintiere, J., McCaffrey, B. J., & Kashiwagi, T. (1978). A scaling study of a corridor subject to a room fire. *Combustion Science and Technology, 18*(1-2), 1-19.

Rafinazari, A., & Hadjisophocleous, G. (2018). A study of the effect of make-up air velocity on the smoke layer height with symmetric openings in atrium fires. *Fire technology, 54*(1), 229-253.

USNRC. (2007). Verification and Validation of Selected Fire Models for Nuclear Power Plant Applications: Fire Dynamics Simulator (FDS)(NUREG-1824, Volume 7). *United States Nuclear Regulatory Commission, USA*.

Yuen, W. W., & Chow, W. K. (2005). A new method for selecting the design fire for safety provision. *Fire Science and Technology, 24*(3), 133-149.

ONLINE EXAMINATION IMPLEMENTED ON GOOGLE CLOUD USING GOOGLE APPS SCRIPT

Vasutan Tunbunheng[1], Adisorn Kheaksong[2], Jukrapun Chitpong[3], Pakpoom Patompak[4], Sehyun Jang[5]

[1-4] Faculty of Engineering and Technology, Panyapiwat Institute of Managment
[5]Major of Computer Science, Kyonggi University
***Corresponding author, E-mail:** adisornkhe@pim.ac.th

ABSTRACT

Online learning is an opportunity for students to study in their location. But one thing that the instructor must consider is the evaluation of the learners. In this research, we have presented an online GUI examination to provide different and random questions for students simultaneously. This includes summarizing scores for management and can save teachers time.

Keywords: online teaching tool, examination, student evaluation, Google apps script

Introduction

In 2019, the world has faced a widespread disease, Corona Virus Disease (COVID-19), an emerging disease. COVID-19 is a contagious disease caused by a new strain of the virus. This disease causes illness in the respiratory system as well as influenza with symptoms. Such as coughing, fever, and difficulty breathing in patients with severe symptoms. The way to prevent this virus infection is to wash your hands frequently for 20 seconds with soap and water or an alcohol-based hand sanitizer gels. Use tissue or a crook of your arm or inner elbow to cover the mouth and nose while coughing or sneezing. Avoid close contact with sick people within 1 meter or 3 feet. Stay at home and keep yourself away from others in the house if you feel unwell. In May 2020, there were 4 million people infected and 300,000 deaths (Control, Corona Virus Disease (COVID-19) report in 2020, 2020), making measures to prevent the spread of COVID-19, such as social distancing, lockdown in various areas which is a gathering together (Control, Corona Virus Disease (COVID-19) report in 2020, 2020)

There are many organizations that are affected by this situation, educational institutions are one organization that needs to change the method of learning and teaching to handle uncontrollable situations. In general, the instructor has to design the teaching to cover the contents. But in such a situation that students are unable to study in the institution. Therefore, online teaching is inevitable. In addition, what the instructor has to consider is the measurement process that must be adjusted. The test is the type of assessment that is commonly used to test the knowledge of students after completing all courses. But cheating in online learning is widespread. The teachers have considered this problem and try to find ways to prevent cheating that may occur.

Online quizzes has been developed extensively and used in various exams. However, when the student is taking the exam, the student may be close or talk to a friend and may easily cheat.

To prevent the exam from being copied, an online test has been developed to provide capabilities such as Lockdown screen — the screen locks down when the examination starts thus preventing the students from visiting any other website or pages. This provides a secure and cheating free environment for the exams. (SoftwareSuggest, 2020), Supports 14 different types of questions. Generates graphically analyzed reports and results. (Softwaresuggest, 2020) In (Jun, 2009), present architecture of online examination with a structure consisting of C # and developed using web service

and COM LMS. In (Golden and Kohlbeck, 2020), has shown that when students are unable to find information on the internet, therefore their results are lower than those in the internet. In(Nagal, Ramesh & Nemkul, Prakash & Kumar, Dispesh & Kumar, Naveen & Joseph, Asha., 2017), is a mobile development application to check who come to take the exam and prevent others from taking the exam instead.

In this research, we create web application is for creating exams from the manuscript to prevent learners from copying, that build from Google Cloud using Google apps script since it is free web hosting. We can reduce the hardware cost of organizing solution that can give automation and performance in low cost. The rest of the paper is organized as follow. Section II Basic Ideas and Implementation. Section III details the Using guideline and the paper is concluded in Section III.

Basic Ideas and Implementation

Since the students have to do examinations from home, it is easy to cheat by copying answers from their friends. There is a strategy to prevent cheating by giving different questions to students, but it becomes hard work for the teacher to generate the different questions separately. Moreover, those questions should have the same difficulty to be fair for all students.

This research shows a possible way of generating many different questions from a based question. Teacher has to define which part of the base question could be changed and our developed application will assign it randomly for each different question.

By developing our application as a web app, students could join the examination easily. The application is implemented on Google Cloud using Google apps script since it is free web hosting. The examination question and answer from students are stored on Google sheets. Thus, teachers who would like to arrange examinations using our application have to set security of the Google sheet to be editable by everyone.

The process of generating questions, taking answers from students, and evaluating scores are shown step by step in the next section.

Using guideline

Teachers can access the application from URL : https://random-exam-online.web.app/ as shown in Figure 1. After inputting the Google sheet URL into the input textbox, the "SUBMIT" button below becomes enabled and the teacher can enter to compose the question. The Google sheet must be shared for everyone to edit, thus, the application could write data.

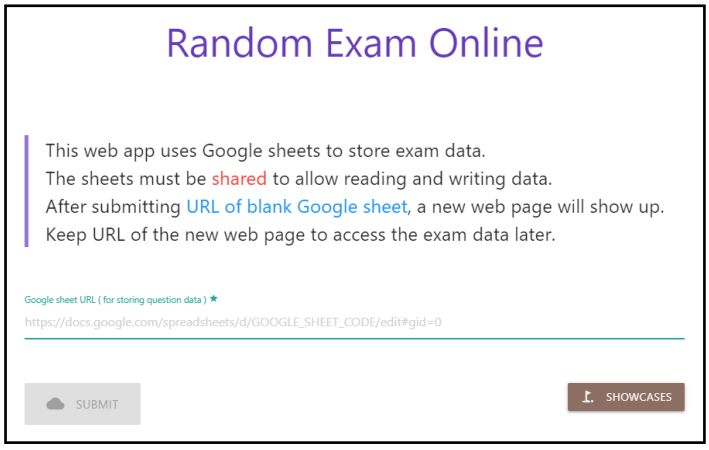


Figure 1: Random exam online web page

After the teacher clicks the "Submit" button to enter Google sheet URL, the web page will show compose tools for editing questions, export tool, and import tool as shown in Figure 2, 3, and 4 . Teachers could save the current URL for accessing the web page later. Or, teachers may input the same Google sheet URL at the first page to access again.

The purple tab in Figure 2 is corresponding to a question. The tab is collapsible, thus, users could collapse it and select the other tabs easily. There are 3 parts inside each question consisting of 8 small editing question buttons in row on top, question composing tool (blue tab), and solution scoring tool (red tab).

The buttons with up/down arrow and bar icon (black buttons) are used to add a copied question tab next to the current tab, in order to increase the number of questions. After inserting a new copied question on top or below, teachers will modify content inside by using the question composing tool (blue tab), and solution scoring tool (red tab) later.

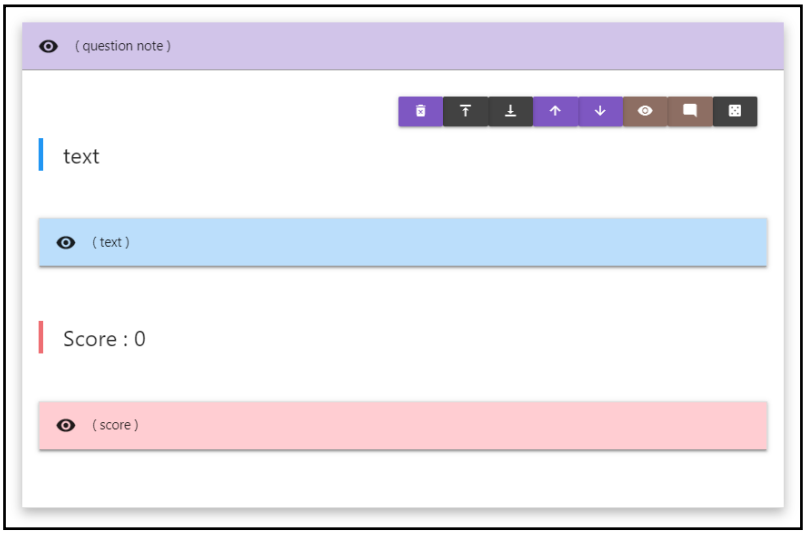


Figure 2: Compose tools for editing question

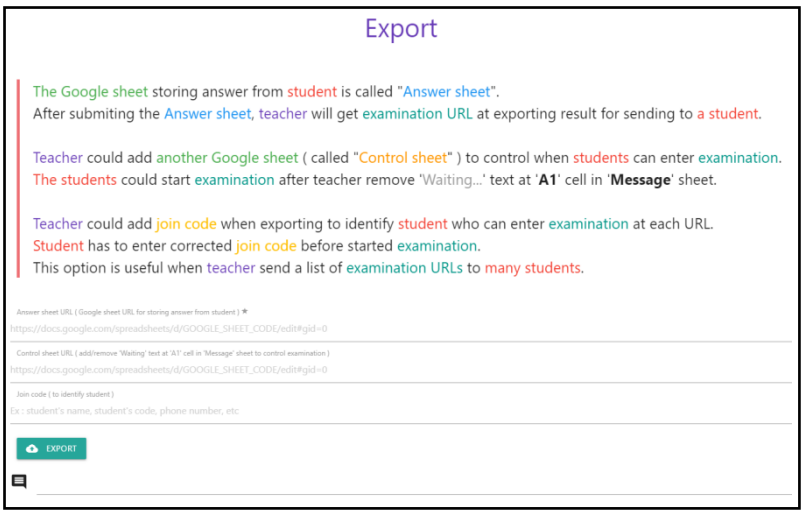


Figure 3: Export tool for generating question

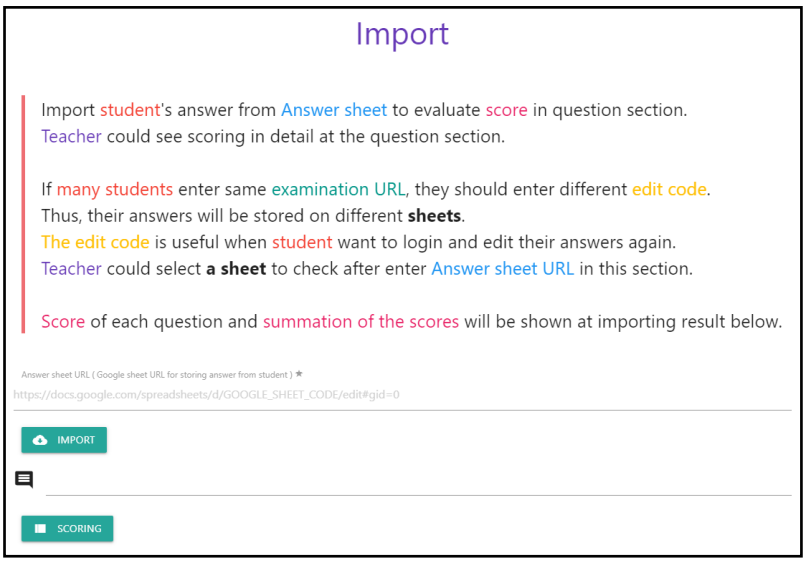


Figure 4: Import tool for checking answer from students

The buttons with up/down arrow (without bar) icon (purple buttons) could be used to reorder the question list by moving the current question tab up or down. The button with the eye icon (brown button) could be clicked to show or hide the current question tab from students. It is useful when teachers create many versions of the same topic and select one of them to test students on different academy years. The button with a comment icon (brown button) could be clicked to edit text displayed on the question tab, thus, the students select question tabs to answer easily. The example could be shown as Figure 5. This example comes from the final examination of class "Computer Architecture (CE59310)". The question list consists of 2 parts. The first part has 5 multiple choice questions, and the second part has 19 text questions. The text displayed on the question tab should include score points and testing topics as a good practice.

The leftmost button with the bin icon (purple button) could be used to remove the current question tab. The rightmost button with dice icon (black button) could be used to random the composing question to be different for each student. This button could be used to change random text generated from "Random-Text" type items to see different possibilities of composing questions. Teachers could random all items at once or change an item individually.

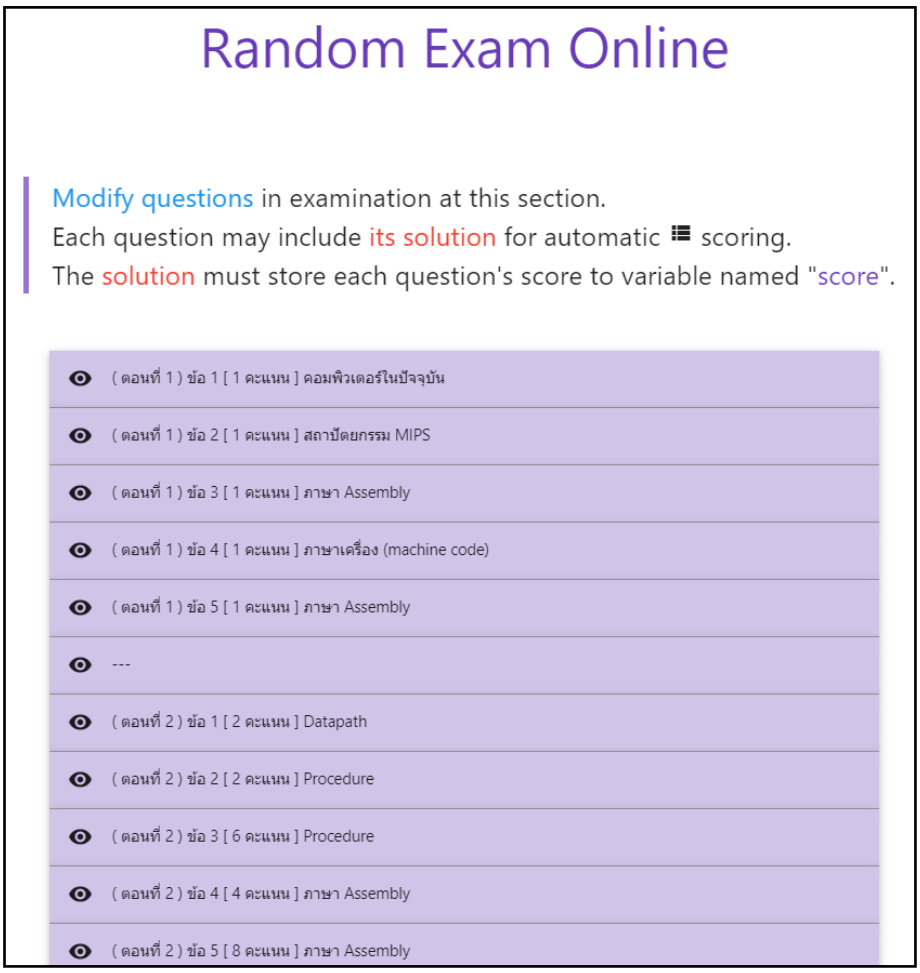


Figure 5: Example of question list from the final examination of class CE59310

Question text inside each tab could be composed by joining many items together. Those items are displayed as blue tabs as shown in Figure 6. In this example, teachers want to test the additional skill of students. Each student has to add a different pair of numbers, thus, the first 2 items are "Random Text" type for randoming 2 numbers, then stored into variable "a" and "b". The configuration of the first "Random Text" type item could be shown as Figure 7. User could change both item type and configuration by clicking on the rightmost button with a wrench icon (blue button under the blue item tab) to show the setting tool as Figure 8.

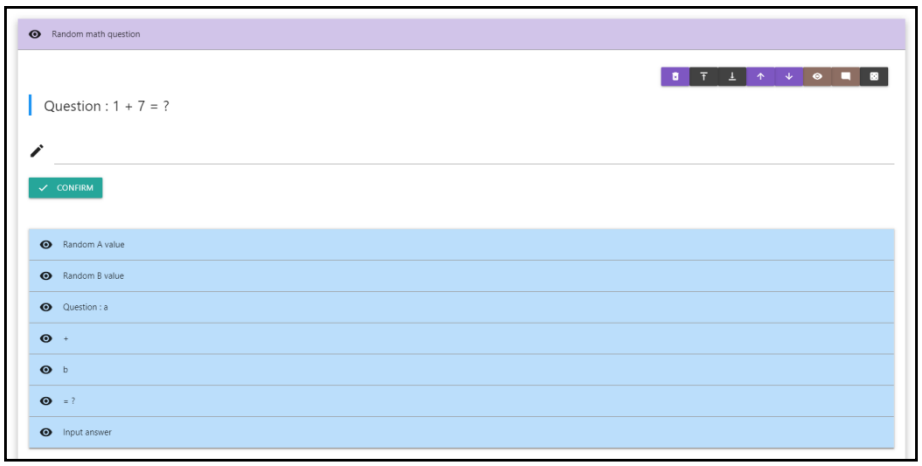


Figure 6: Example of composing question

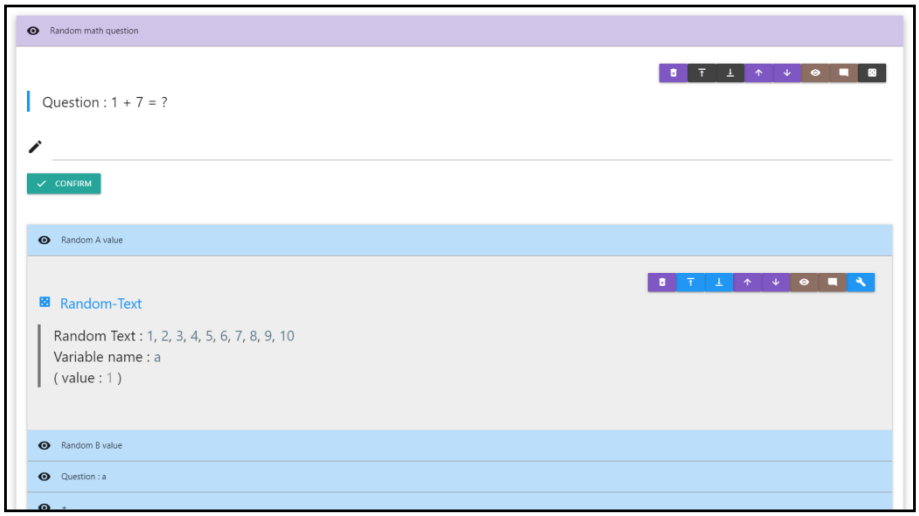


Figure 7: Example of configuration in "Random-Text" type item to random a number from 1 to 10

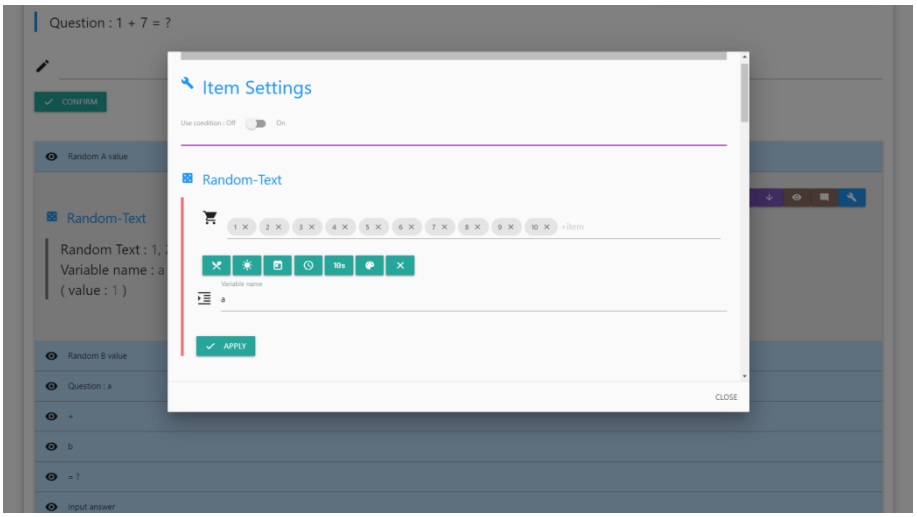


Figure 8: Setting dialog of "Random-Text" type item to random a number at variable "a"

Users could scroll down and click the "APPLY" button to change the item to another type after copying an item, in order to add more items to the list. The following items are "Print Text" type items for displaying numbers stored in variable "a" and "b" with plus, equal signs, and question mark to compose the question (the text "1 + 7 = ?" in Figure 6 and 7). The last item is "Input Text" type for displaying input textbox to obtain an answer from a student, then store the answer into variable "answer".

The solution scoring tool (red tab in Figure 2) could be used to evaluate scores corresponding to the answer. The "Java Script" type item executes the statement "if (answer == (parseInt(a) + parseFloat(b))) score = 1;" to evaluate score, is shown as Figure 9. Since the values stored in variable "a" and "b" are text, the function "parseInt()" or "parseFloat()" is required to turn the text value to number for calculation (adding). Then, the next "Show Score" item displays the "score" value of this question.

The "Input Radio" type item could be used to compose multiple choice questions as shown in Figure 10, while the "Show Image" type item could be used to display image in the question. Students could zoom to see large images by clicking on the image.

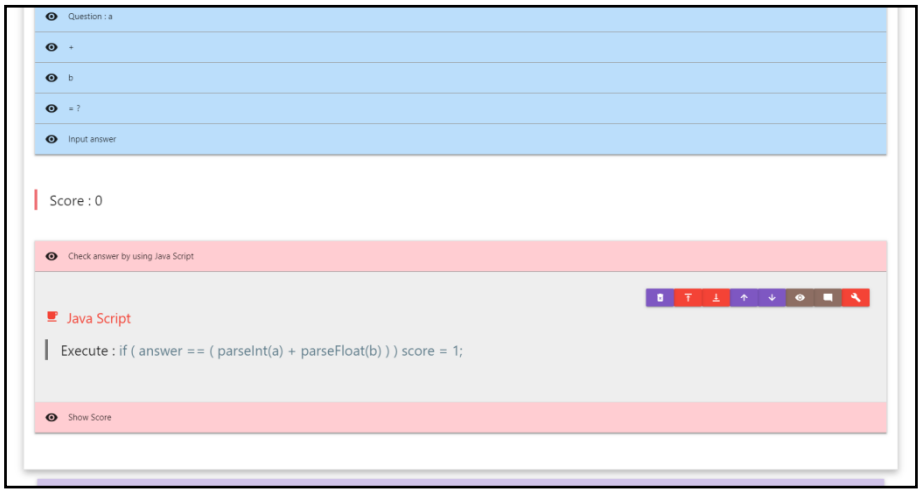


Figure 9: Java Script for evaluating score for the "answer = a + b" question

Figure 10: Example of multiple choice question

After the teacher composes a list of questions for examination, the teacher could generate a URL for each student to join the examination by using the exporting tool (shown in Figure 3). Another Google sheet URL for storing answers from a student is entered at the first input box (this Google sheet is called "Answer" sheet). The examination URL could be used for only a student, thus, the teacher must generate the examination equal to the number of students. The second and third input boxes are options to control the examination. They could be empty if not needed. After teachers press the "EXPORT" button, an URL for sending to a student will appear at the textbox below.

The optional second input box could be filled with another Google sheet URL to control start and stop time of examination (this Google sheet is called "Control" sheet). When the control sheet option was used, students could not see the question list until the teacher removed "waiting..." text from the "A1" cell inside the "Message" tab of the control sheet. Many students could share the same Control sheet, thus, all students could start examinations at the same time. After the examination finishes, teachers will add "waiting..." text at the "A1" cell inside "Message" tab of the control sheet again. This action will force all students to stop answering questions.

The optional third input box could be used as a password to join the examination. This password (called "Join" code) is set independent for each URL. Students have to input the correct password in order to enter the examination. This option is useful to protect students from entering wrong URL since teachers usually show a list of URLs to all students. If this option is used, there will be an input box to check "Join code" as shown in Figure 11. Another "Edit code" input box is used to identify students when students accidentally close the web page and would like to continue editing the last answers. This option provides additional security to the student who does not want bad friends to erase or change his or hers answer, since, teacher may give the Join code of every student with the list of examination URLs.

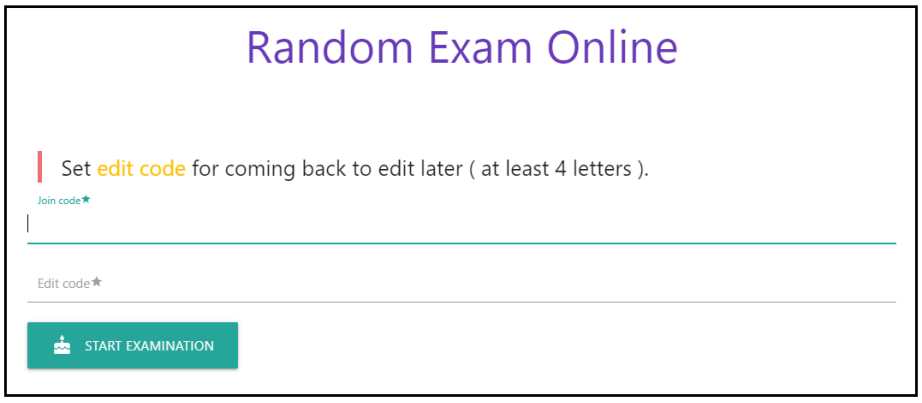


Figure 11: Web page waiting for "Join" and "Edit" code before students enter examination

After students finish answering the examination, the teacher could use an importing tool, shown in Figure 4, to obtain an answer from every student. By entering the Answer sheet URL (Google sheet URL which is used to generate examination URL when exporting) into the input box, the answer from the sheet will transfer into the current web page and application calculate summation score from every question to show at the textbox below.

Evaluation

Since the random exam online system has been tested on the final examination of class "Computer Architecture (CE59310)", teachers also make questionnaires to ask students in order to evaluate the system. There are 26 students and the questionnaires is consisting of 5 questions as followed: (1) rating between 1 to 5 for satisfaction, (2) rating between 1 to 5 for comprehension how to use, (3) short answer about answer copying protection system, (4) short answer of problem occurred during examination, and (5) short answer for future development. The first and second questions are required, while the rest questions could be blanked on students' demand.

The average satisfaction value is 4.12 and the average comprehension value is 4.00 (rating 1 for lowest value and rating 5 for highest value). The answer copying protection system has been implemented to reject copy and paste operations when students type answers at the input text box. Teachers explain the existence of the protection system before the examination starts for experiment. There are 13 students who agree that implementing the protection system is a good practice to protect cheating at online examination. There are 2 students who disagree because the protection system slows down their answering speed and the examination is not enough. There are 5 students who both agree and disagree at the same time and the rest of the students have no comment. Even the answer copying protection system could protect most students from cheating, however, some students who have Javascript programming skills could pass the protection. Teachers are aware of this fact because there are 3 mistakes: (1) the cheating answers consist of the same misprint or misspelling, (2) all words are exactly the same, and (3) the height value of the input box is not saved.

The comments from students about problems occurring during examination are mostly about the requirement of pressing the confirm button or enter key to save answers. Students report that they have to spend extra examination time to type the same answer again. Students suggest adding an auto save system, global saving button, or save warning icon in order to solve the problem. The other problem is that students lose focus on questions which have been answered when closing the question tab to answer the next question since opening another question tab causes the previous tab to close. Students suggest adding a mark icon to indicate which questions have been completed as the future development.

Discussion

This section explains how this research is different from other online examinations. The main feature of "random exam online" is generating different questions for each student. This feature could reduce cheating examinations by copying answers from other students since students could use social media to communicate with the other student easily. The examination time could be controlled in real time since giving too much examination time will cause students who finish fast to help the slow students. Teachers could peek at the student process on Google sheet of each answer sheet since answer data of different questions is saved on different rows. The empty row means there is no answer at the question while encoding text data on the row means a student has submitted an answer at least one time.

When a student joins an examination, a student has the ability to define personal edit code to protect his or her answer from editing by the other bad students. The edit code is also useful when students close the web browser and come back to re-edit their answer again. If students want to get a new answer sheet or clear all answers to start from beginning, the students could simply re-join the examination by using another edit code.

The answer copying protection system, which rejects copy and paste operation, is also rarely found on the other online examinations. Most students have positive sentiments about this feature. The random exam online in this research allows executing Javascript to evaluate scores while most of the

other online examinations only compare to a corrected answer when scoring. The Javascript executing allows teachers to check answers of math questions with random value or evaluate the score of short answer questions.

The random exam online requires the teacher to create his or her own empty Google sheet for storing examination questions and answers from students. This feature becomes extra work to teachers who use the random exam online. However, the teacher should feel safe to store question and solution data on their own Google drive rather than storing this important data on an unknown server. Moreover, the random exam online developper does not need to maintain a server to store enormous data from users around the world.

The random exam online in the current state could be used on real examination, however, there are still many things to be improved based on suggestions from students shown in the previous section. Based on the experience of teachers, the application is not convenient enough for heavy use. The answer sheet must be checked and score one by one manually in the current version. There are many room to update in the next version for example: (1) collect score from a group of students by clicking only a button, (2) import score from other activity such as class participation score, group report score, and homework score, then, find summation and automatic grading (A to F), (3) add another type of GUI for drawing or hand writing answer, (4) graph plotting of overall student's scores, (5) report feedback to each student after checking their answer, (6) allow student to submit file data, and (7) mechanism to evaluate score on different types of question with less using Javascript as possible.

Conclusions

The application developed in this research provides a rich Graphics User Interface (GUI) to generate different questions for students. The examination time could be controlled by the teacher. After the examination finishes, teachers can import answers from students to evaluate their score. The application stores examination data on Google sheet. It is available for everyone who wants to arrange online examinations, however, the teachers must have basic Java-Script programming skill in order to control randoming of questions and automatic scoring. Since this research is just starting, GUI will be improved to be more comfortable in the future.

References

Control, D. o. (2020, May). Corona Virus Disease (COVID-19) report in 2020. Retrieved from https://ddc.moph.go.th: https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no130-120563.pdf

Control, D. o. (2020, May 12). Corona Virus Disease (COVID-19) report in 2020. Retrieved from https://ddc.moph.go.th: https://ddc.moph.go.th/viralpneumonia/eng/announcement.php

Jun, L. (2009). Design of Online Examination System Based on Web Service and COM. First International Conference on Information Science and Engineering, (pp. 3276-3279). Nanjing.

J. Golden and M. Kohlbeck. (2020). Addressing cheating when using test bank questions in online Classes. Journal of Accounting Education, 100671.

Nagal, Ramesh & Nemkul, Prakash & Kumar, Dispesh & Kumar, Naveen & Joseph, Asha. (2017). Android based Secure Exam Management System to Prevent Impersonation. International Journal of Latest Technology in Engineering, Management & Applied Science. VI. , 46-49.

Softwaresuggest. (2020). Retrieved from Online Assessment Software: https://www.softwaresuggest.com/virtualx

SoftwareSuggest. (2020). Retrieved from Online Assessment Software: https://www.softwaresuggest.com/edbase

SMART APPLIANCE HOME CONTROL AND MONITORING ON SMARTPHONE USING INTERNET OF THINGS (SmartAppHome IoTs)

Watanyou Yodthong[1], Pruegsa Duangphasuk[2*]

[1,2] Faculty of Information Science and Technology, Mahanakorn University of Technology
[*]**Corresponding author, E-mail:** pruegsa@mut.ac.th

ABSTRACT

Smart Home or Home Automation is a modern home system, that can control appliance such as devices and systems in the home then people prefer to stay home happily. However, in order to meet current lifestyles with staying outside, they cannot control the home system but at the present, it is easily to access internet via Smartphone application and can also connect to security systems to increase home security. This research proposed the "Smart Appliance Home Control and Monitoring on Smartphone using Internet of Things", called SmartAppHome IoTs. With Internet of Things (IoT) technology and equipment, the development on mobile application is combination in the lifestyle of controlling electrical appliances to comfort facilitate the use of appliance control and record the historical user behavior, electrical appliances usage, with home control and monitoring. There are three phases; (1) Analysis and designing phase, the proposed system is designed with UML Diagrams and then (2) Production phase, by bringing the various sensors connected to the NodeMCU ESP8266 to take the values from the sensor to process and save. After that (3) Test phase, data into the database and developed mobile application are used to display the various values sent by the sensor to support the operation of various devices from users and displays user history of various devices. The experimental results showed that, on mobile application, it can display values from sensors and it supports the control panel operations and it records various information. Accordingly, the application can assist users to check the working status of various electrical devices via Smartphone. The system can support the connection of power plug. The application can analyse the usage of user and presentation in the form of graphs as well.

Keywords: Internet of Things (IoT), Smart Home, Appliance Control, Home Control and Monitoring, NodeMCU, Mobile Application, Smartphone.

Introduction

Internet of Things (IoT) means that things are linked together. In the internet world, there allow humans to control various devices via network such as command to turn on-off electrical appliances, mobile phone office, supplies agricultural tools, industrial machinery, houses, including appliances through the internet. In the future, smart home will become the truth in technology that can help users easily access and monitor homes remotely. Using computer control system, it can notify users by sending an SMS or using Smart Home Applications installed on a smartphone.

Switching on-off electrical appliances in the room is an important key for safety and saving the cost of electricity. The benefit is cost reduction such as reducing costs for electricity and also the manpower of staff to check the electricity meter every month. If the idea of Internet of Things is implemented, it will be able to reduce costs because the meter can send the data to the system to keep records of electricity usage automatically. It also helps to reduce electricity consumption as well.

Objectives

1. To provide equipment in the room that be able to control using electrical appliances via the mobile application on a smartphone, also control of equipment in the room with more convenient.
2. To apply modern technology on equipment in the room and be compatible with using IoT devices.
3. To increase room safety by controlling accidents that can be the incidents of electrical equipment in the room with notification on application to the user in advance.

Research Contributions

1. Create system models using the Internet of Things. There are six models as following
    1) Room model
    2) Model of the power switch relay module
    3) Light sensor model
    4) Doors and windows opening and closing detection sensor model
    5) Sensor button alert buzzer model
    6) Infrared sensor model
2. Create the control panel of NodeMCU into five systems as following
    1) The control panel to turn off-turn on the power plug
    2) The motor control panel closing-opening curtains
    3) The control panel detects the closing - opening doors and windows
    4) The doorbell control panel
    5) The emergency lamp control panel
3. Create an application on the Android operating system to display values from sensors as following
    1) Value from the relay module
    2) Value from the light sensor
    3) Values from sensors, magnetic switches detect the closing - opening doors and windows
    4) Values from the push sensor
    5) Value from the ultrasonic sensor
4. To increase room safety by controlling accidents that can be the incidents of electrical equipment in the room with notification on application to the user in advance.

Literature Review

To reduce forgetfulness or lack of time to take care of the house, we are ready to improve the quality of human's life with various innovative control styles. For examples;

- Home controls use clock lights to adjust the area of the room while in the house. Presented in Budianto, Dimas & Nurmaini, Siti & Tutuko, Bambang & Raflesia, Sarifah. (2018), and Alaa, Musaab & Zaidan, A. & Bahaa, Bilal & Talal, Mohammed & Mat Kiah, Miss Laiha. (2017).
- Home monitoring remotes control via coil circuit camera, called smart home discussed in Dener, Murat & Özkök, Yunus & Bostancıoğlu, Cevat. (2015).
- Media Center brings together entertainment content for everyone in the house to play movies anywhere in the house. Described in Mendes, T.D.P. & Godina, Radu & Rodrigues, Eduardo & Matias, João & Catalão, João. (2015).

- Smart Kits, like smart kits, as well as kitchens that can communicate directly with you or communicate with technology under your heart. Appeared in M. Ständer, A. Hadjakos, N. Lochschmidt, C. Klos, B. Renner and M. Mühlhäuser. (2012) and A. Arya, A. Taliyan, P. Chauhan and A. Gautam. (2019).
- Electronic control devices via the application and notification of operation on a smartphone such as an automatic vacuuming robot, Moedor Cleaning Robot. As described in A. Taneja, G. Bansal, R. Setia and N. Hema. (2018).
- Landscape Control to turn on-off watering trees instantly when needed while we are being not at home. As mention in K. Lekjaroen, R. Ponganantayotin, A. Charoenrat, S. Funilkul, U. Supasitthimethee and T. Triyason. (2016) and S. Aygün, E. O. Güneş, M. A. Subaşı and S. Alkan. (2019).

Theories and Methods of the Proposed System

The Internet of Things (IoT) is a device connection or things with the network or the internet to enable humans to operate and command control operations with more conveniently. IoT was invented by Kevin Ashton in 1999. IOT is also known as M2M.

C ++ is a computer language that can be written in both object and structured programming. It consists of the Free Software Foundation, LLVM, Microsoft Intel, and IBM. C ++ is set to be the standard language by ISO.

JAVA is an object-oriented programming language developed by James Gosling and the engineering team. The strength of the JAVA language is that programmers can use the principles of Object-Oriented Programming to develop programs. The program written is created within the class. Therefore, classes are methods for storing methods or behaviors that have states and identities in behavior.

Firebase is designed to be an API and Cloud Storage for Real-time Application development. It supports many platforms including IOS, Android, and Web app. The company was founded in 2011 by Andrew Lee and James Tamplin. Firebase's initial product is Real-time Database.

Micro Controller is an electronic device. It is used to control electrical devices or electronic systems, which the Micro Controller is like a small computer and collected in the same Chipset.

NodeMCU is an Arduino-like device that can connect to WiFi. It can be programmed with the Arduino IDE, just like Arduino, and the board is cheap. It consists of ESP8266 (Micro Controller that can connect to WiFi) with facilities such as Port Micro USB for power and uploading programs.

The proposed system overview of Smart Appliance Home Control and Monitoring on Smartphone using Internet of Things is divided into 3 parts which are IoT, Google Firebase Real-time Database and Mobile Application as described in Figure 1. The first part is “IoT”, a sensor that is used to monitor various status, including closing-opening the door in the room, the amount of sunlight outside the room, the distance of the user and the sensor while sleeping in bed. It will send data to Google Firebase Real-time Database via HTTP protocol to store data and further processing at NodeMCU ESP8266 V.2 to enable devices. The second part is “Google Firebase Real-time Database”, server that is responsible for storing various data sent by sensor for NodeMCU ESP8266 V.2 which can use data to operate various devices and send to application that use the data stored and be able to analyse data. The third part is “Mobile Application”, application IoT for smart homes’ users, need to login to authorise and authenticate before using Smart Appliance Home Control and Monitoring. After the user has logged in, it can browse the information from sensors and operate various devices following:

(1) The power switch to turn on-off the power plug
(2) The light sensor for closing-opening curtains
(3) The doors and windows opening and closing detection sensor
(4) The sensor button alert buzzer to control doorbell
(5) The infrared sensor to alarm an emergency lamp on the head-bed light.

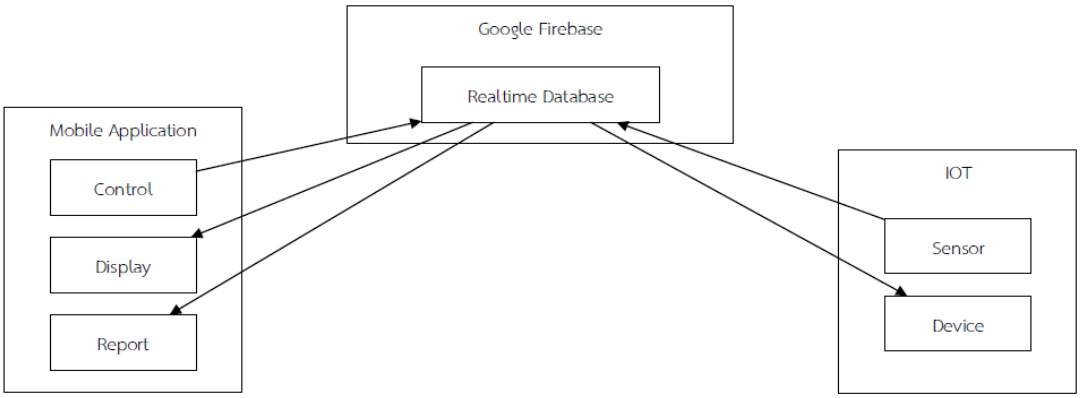


Figure 1: The system overview of the proposed system

In the analysis and design phase, this research has been studied on UML consisting of use case diagram of IoT for Smart Room (Mobile Application) as shown in Figure 2 and use case diagram of IoT for Smart Room (Internet of Things) as shown in Figure 3. Also, there are class diagram as shown in Figure 4 and activity diagram as shown in Figure 5.

In the figure 2, there are eleven use cases as describe in below.

1) Use Case Name: Register. The system is registered with 2 related actors, namely User and Real-time Database. When receiving data from User via Mobile. Application will have to check the username, the password and the password confirmation. The registered information will be examined which username must be matched with username in Real-time Database and username is unique.
2) Use Case Name: Login. User must log in to start using the application. The user must enter the user id and password to log in. The application will verify from the information in Real-time Database using included use case Authentication.
3) Use Case Name: Authentication. After user enters the username and password then the system will confirm the identity for accessing the system with the database in order to access the mobile application. In case of entering incorrect personal ID or password. The system will alert in order to fill correctly.
4) Use Case Name: Controlling. The devices, which have five devices such as the door, the doorbell, the head-bed light, the power plug and bamboo curtains, are controlled via mobile application by connecting to Real-time database.
5) Use Case Name: Control curtain. Data are received through the light exposure circuitry meter. After that it will send the value to Real-time database with NodeMCU ESP8266 v.2 device. The device will update the data via WiFi and update the data to the Mobile Application page.

6) Use Case Name: Control on/off outlet. The control is turned on-off via Mobile Application. When the power button is pressed on the Application, the data are sent to the database. NodeMCU ESP8266 v.2 will update the database from the command to enable the device to turn on / off.
7) Use Case Name: Control unlock door. There are two controls which are the door unlocking system and the doorbell to alert. The door unlocking system can be controlled via mobile application and by pressing the release switch next to the door. The buzzer will send the notification to the mobile application. Both of controls will send the usage data through the NodeMCU ESP8266 v.2 device to the database to update and return the value to NodeMCU ESP8266 v.2 for further execution.
8) Use Case Name: Control light on bed. Data are received through the ultrasonic sensor. When moving through the sensor, the lamp will turn on automatically. The values are sent to the database.
9) Use Case Name: Monitoring. Device usage data is checked and sent to database to update the Mobile Application page.
10) Use Case Name: Show status all devices. There is a status display on the Mobile Application page whether the device is active or not. The data are sent from NodeMCU ESP 8266 V.2 to the database and displayed on the Mobile Application page.
11) Use Case Name: Show device processing report. Device usage data is displayed on the Mobile Application page. Device usage data is sent from the NodeMCU ESP 8266 V.2 to the database and displayed on the Mobile Application page.

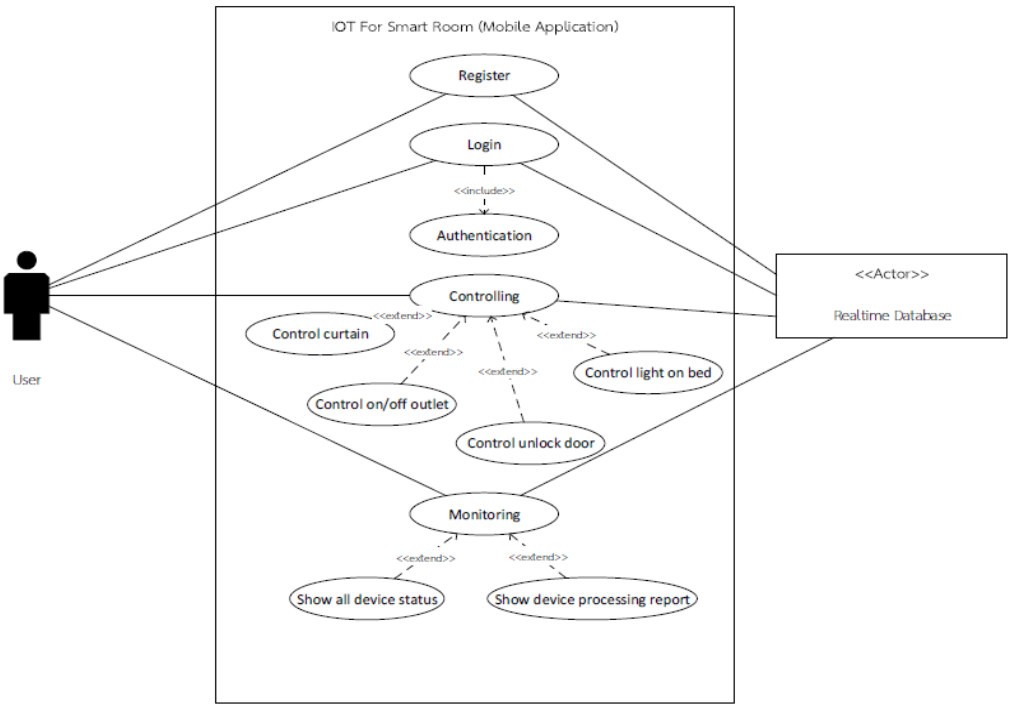


Figure 2: Use Case Diagram of IoT for Smart Room (Mobile Application)

In the figure 3, there are ten use cases as describe in below.

(1) Use Case Name: Send lock status. The status of the door is read whether it is locked or not with a magnetic door switch sensor and sent to the database.

(2) Use Case Name: Send value light. There are reading and transmitting light input values with a light sensor and sending to the database.

(3) Use Case Name: Send detection status. The distance reading of the ultrasonic sensor is read and the value is sent to the database.

(4) Use Case Name: Receive press a button unlock. The value of pressing the button to unlock the door is received from the user and its status updates to the database.

(5) Use Case Name: Receive press a button alarm. The ringing value from visitor is received and its status updates to the database.

(6) Use Case Name: Receive lock status. The door lock status is obtained from the database.

(7) Use Case Name: Control on/off light. The system receives the lamp status from the database and controls the operation according to the lamp's on-off status.

(8) Use Case Name: Control curtain. The system controls the curtain blinds with adjusting the blinds. When a large amount of light intensity is detected, which means that there is sunlight entering through the window a lot, thus this will be causing the curtain to rotate or control the adjustment of the blinds by reading from the database for light reduction.

(9) Use Case Name: Control on/off outlet. The system controls the on / off of the power plug. By controlling via mobile application, it will send values to database and execute to NodeMCU ESP8266 V.2 to enable the power plug to turn off.

(10) Use Case Name: Control door. The system has control to unlock the door. By controlling via mobile application, it will send values to the database and activate to NodeMCU ESP8266 V.2 to lock or unlock the door.

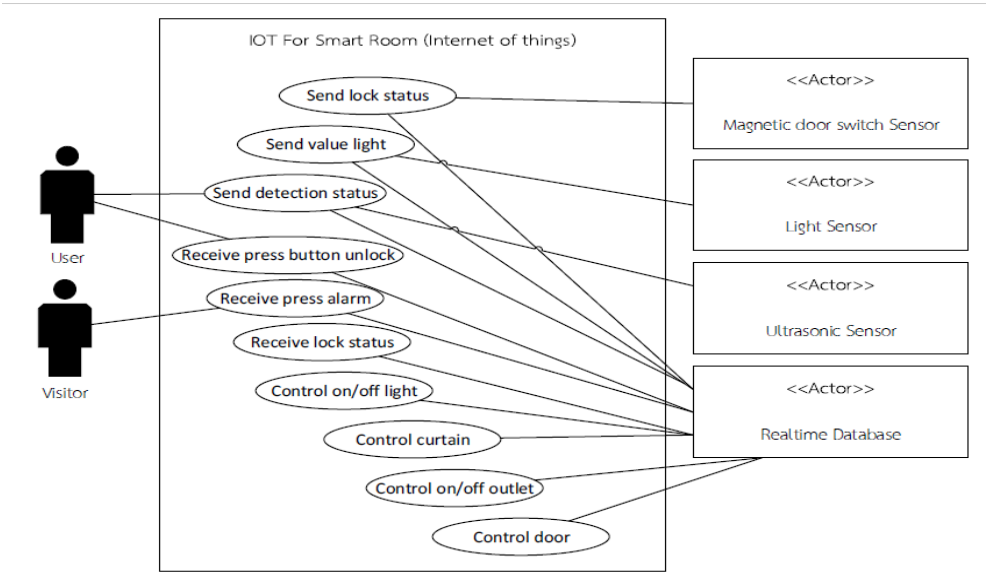


Figure 3: Use Case Diagram of IoT for Smart Room (Internet of Things)

In the figure 4, there are nine classes as described in below.

(1) Class PathRealtimeDatabase contains one operation, setPath(String). To define the data address.

(2) Class Plug consists of three attributes; path, status and mac. There are two operations; setPlug(boolean) and getPlug(). To define plug status and read its status.

(3) Class Light consists of three attributes; database (JSON format), status, path. There are two operations; getStatus() and setStatus(boolean). To define light tube status and read its status.

(4) Class Curtain consists of four attributes; database (JSON format), Degree, Value_Light and path. There are three operations; getDegree(), setDegree(int) and getValueLight(). To define/update and read the degree of curtain angles, also read the light intensity.

(5) Class Door consists of four attributes; database (JSON format), Lock, Switch and path. There are three operations; getLock(), setLock(Boolean) and getSwitch(). To define and read door status, also read the status of pressing the door release button.

(6) Class Alarm consists of three attributes; database (JSON format), status and path. There are one operation, getStatus(). To read the buzzer status.

(7) Class Controlling consists of four attributes; Plug (List of Class Plug), light (object of Class Light), curtain (object of Class Curtain) and door (object of Class Door). There are four operations; controlPlug(boolean), controlLight(boolean), controlCurtain(int) and controlDoor(boolean). To turn off-turn on the power plug, alarm emergency lamp on the head-bed light, close or open curtains with appropriated degree for property light through inside room, also alert buzzer to control doorbell and open or close the doors and windows.

(8) Class Monitoring consists of five attributes; Plug (List of Class Plug), light (object of Class Light), curtain (object of Class Curtain), door (object of Class Door) and alarm (object of Class Alarm). There are five operations; showPlug(), showLight(), showCurtain(), showDoor() and showAlram(). To display theirs' status.

(9) Class Reporting consists of five attributes; Plug (List of Class Plug), light (object of Class Light), curtain (object of Class Curtain), door (object of Class Door) and alarm (object of Class Alarm). There are five operations; reportPlug(), reportLight(), reportCurtain(), reportDoor() and reportAlram(). To report theirs' historical data on working.

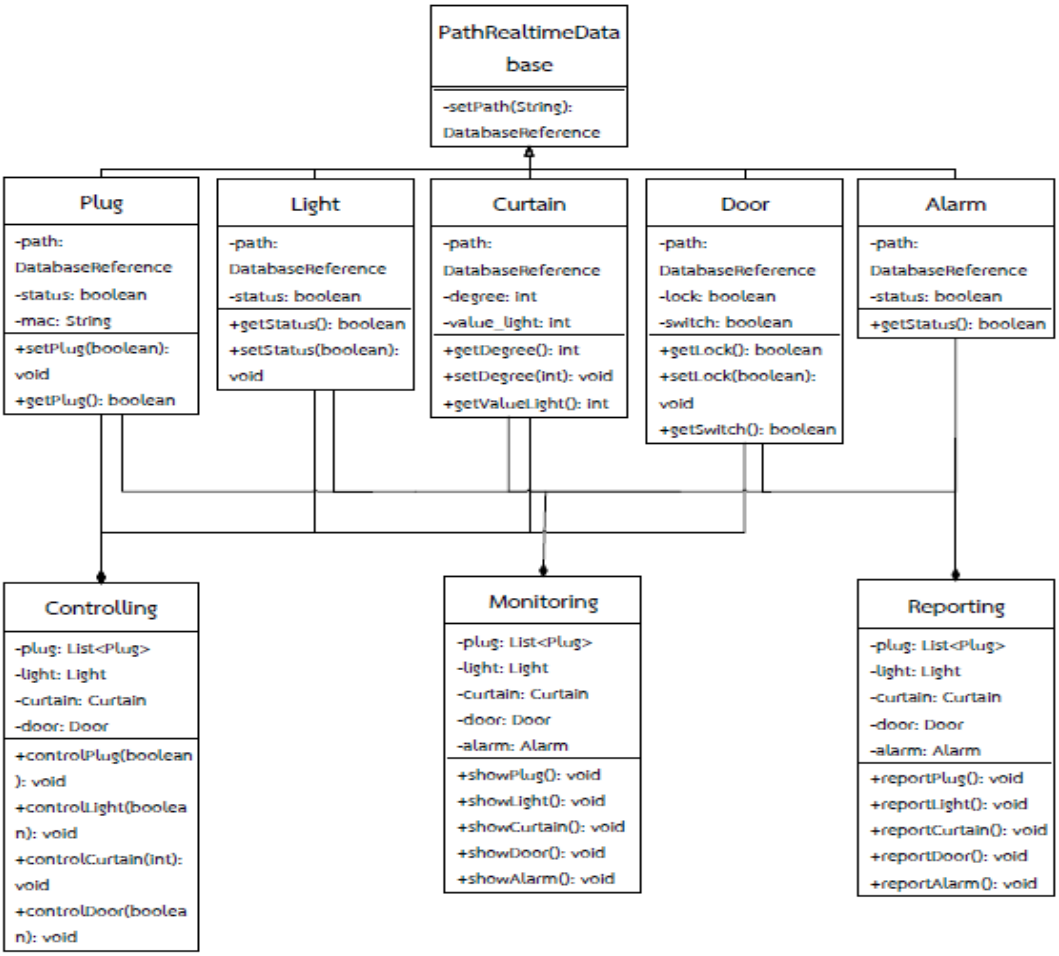


Figure 4: Class Diagram of the proposed system

In the figure 5, there are eight activity diagrams, main activities, as described in below.

(1) Activity Diagram Register as shown in figure 5 (a) can be described step-by-step as follows:

1. Users press the Register button.

2. Connect to the database.

3. Check the connection to the database.

3.1 If yes, continue with step 4.

3.2 If not, the system will connect again.

4. User to fill in User ID, Password and Confirm Password.

5. Check if the User ID is duplicated with the information in the system or not.

5.1 If yes, display the message that the account name has already been used. Then allow the user to fill in again.

5.2 If yes, show the registration message successfully.

(2) Activity Diagram Login as shown in figure 5 (b) can be described step-by-step as follows:

1. Users press the Login button.
2. Connect to the database.
3. Check the connection to the database.
    3.1 If yes, continue with step 4.
    3.2 If not, the system will connect again.
4. User to enter User ID, Password.
5. Check that the User ID is correct or not.
    5.1 If yes, display the message Access successful.
    5.2 If yes, displays the message User ID or Password incorrectly, re-enter.
6. Import all device data.
7. Show all device statuses.
8. Users can access all device statuses within the intelligent room.

(3) Activity Diagram Monitoring as shown in figure 5 (c) can be described step-by-step as follows:

1. User login successfully.
2. Connect to the database
3. Check the connection to the database.
    3.1 If yes, continue with step 4.
    3.2 If not, the system will connect again.
4. Show status values of all devices by receiving data from the database.

(4) Activity Diagram Control Curtain as shown in figure 5 (d) can be described step-by-step as follows:

1. The user selects the curtain control menu.
2. Connect to the database
3. Check the connection to the database.
    3.1 If yes, continue with step 4.
    3.2 If not, the system will connect again.
4. Check the information and show the status of the blinds.
5. Receive status information on/off.
6. When the command to enable/disable the status on the Application will change.
7. Keep the status to the database.

(5) Activity Diagram Control outlet as shown in figure 5 (e) can be described step-by-step as follows:

1. Users choose the plug control menu.
2. Connect to the database.
3. Check the connection to the database.
    3.1 If yes, continue with step 4.
    3.2 If not, the system will connect again.
4. Check the information and display the status of the power plug.
5. Receive status information on/off.
6. When the command to enable/disable the status on the application will change.
7. Save the status to the database.

(6) Activity Diagram Control door as shown in figure 5 (f) can be described step-by-step as follows:

1. The user selects the door control menu.
2. Connect to the database.
3. Check the connection to the database.
    3.1 If yes, continue with step 4.
    3.2 If not, the system will connect again.
4. Check the information and show the status of the door
5. Receive status information on/off
6. When the command to enable/disable the status on the application will change.
7. Keep the status to the database.

(7) Activity Diagram Control light as shown in figure 5 (g) can be described step-by-step as follows:

1. Users choose the lamp control menu.
2. Connect to the database.
3. Check the connection to the database.
    3.1 If yes, continue with step 4.
    3.2 If not, the system will connect again.
4. Check the information and display the status of the bed headlights.
5. Receive status information on/off.
6. When the command to enable/disable the status on the application will change.
7. Keep the status to the database.

(8) Activity Diagram Show device processing report as shown in figure 5 (h) can be described step-by-step as follows:

1. Users choose the report menu.

2. Connect to the database.

3. Check the connection to the database.

3.1 If yes, continue with step 4.

3.2 If not, the system will connect again.

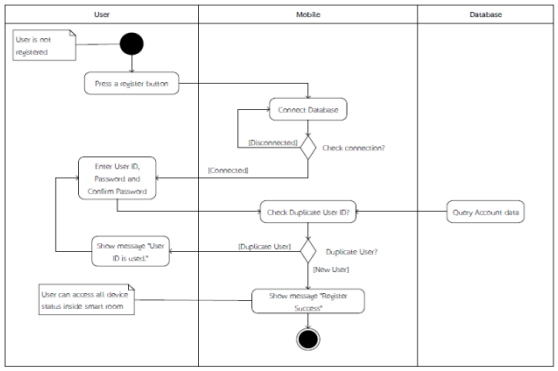
4. Users specify the day they wish to view the report.

5. Get all the information of the device from the database to analyze what the data in each day.

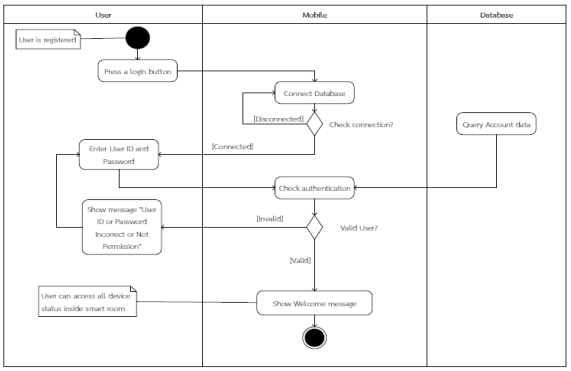
6. Display a graph of device usage data, such as the opening / closing of the curtains for each day, the

On / off of the power plug, unlocking the door each day, checking that the lights of the bed turn on. How many times each day, to check how many times to ring the bell.

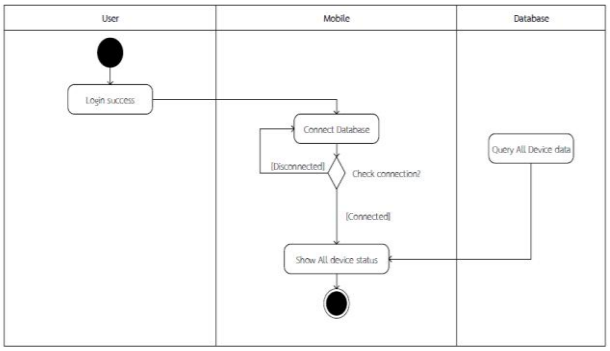
In the circuit board design phase, this research has been designed into four circuit boards such as open-close circuit, lock the door and alert the doorbell as shown in Figure 6, Open-close circuit of the power plug as shown in Figure 7, Circuit - turn off the light on the head of the bed as shown in Figure 8 and Light sensing and adjusting circuit for curtain as shown in Figure 9.



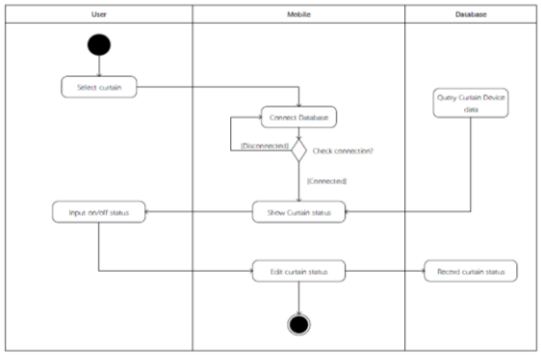
(a)



(b)



(c)



(d)

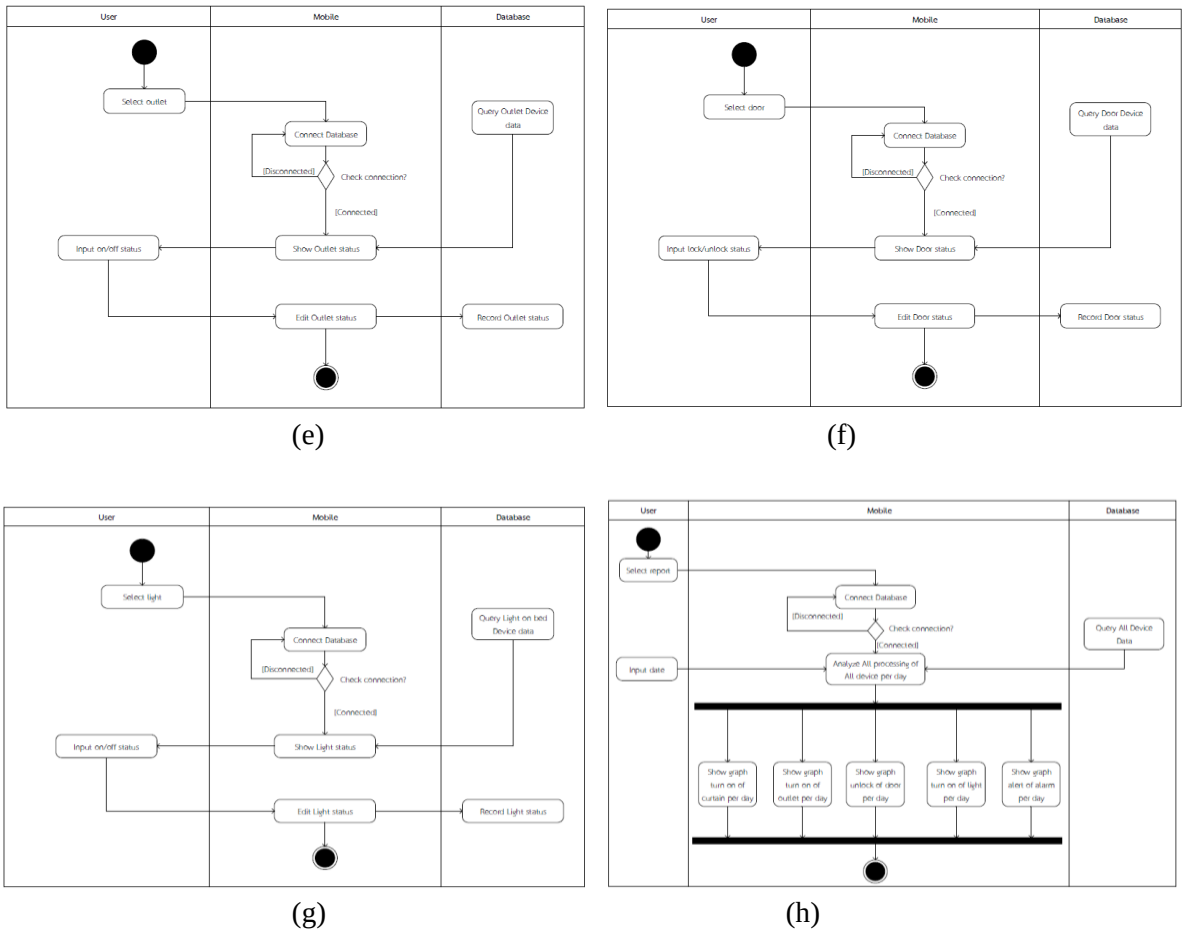


Figure 5: Activity Diagram of the proposed system

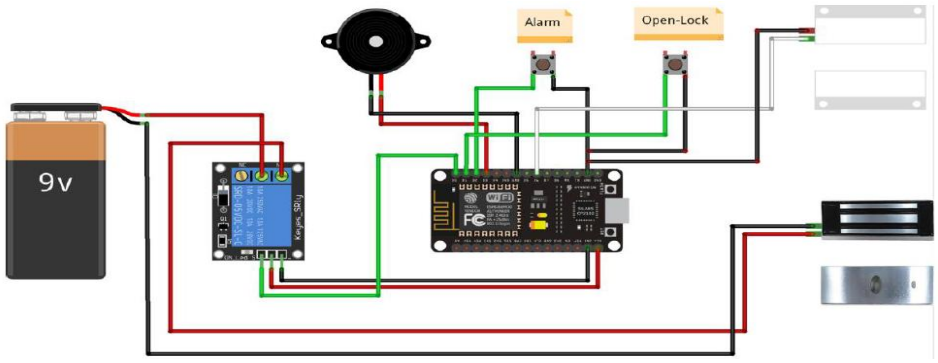


Figure 6: Open-close circuit to lock the door and alert the doorbell.

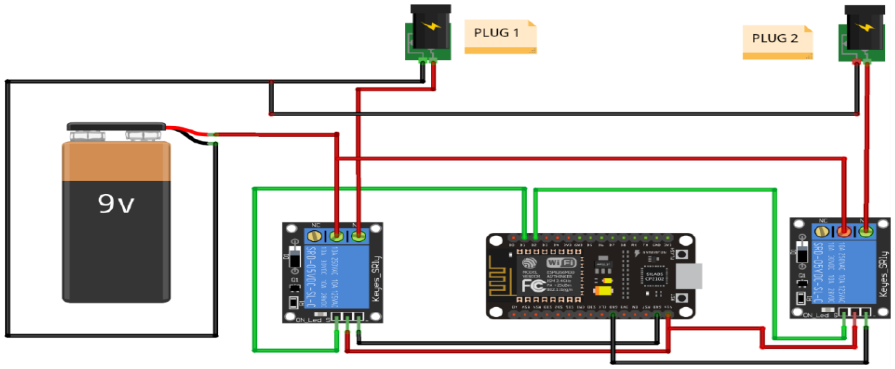


Figure 7: Open-close circuit of the power plug

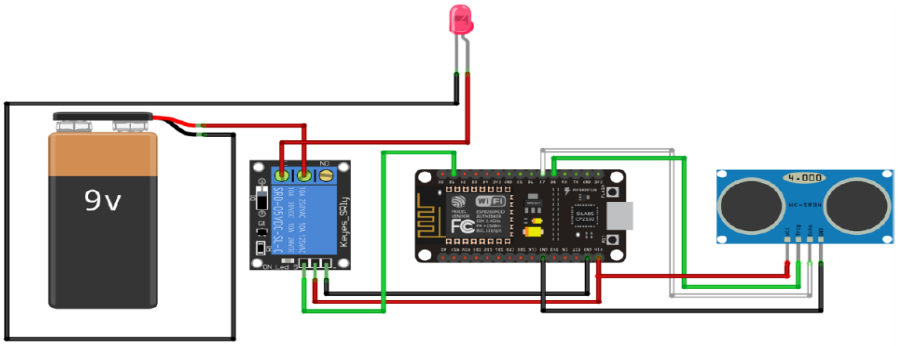


Figure 8: Circuit turn off the light on the head of the bed

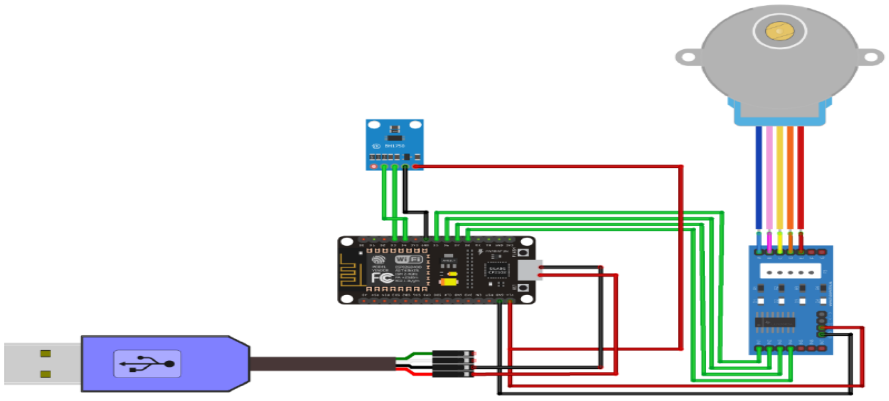


Figure 9: Light sensing and adjusting circuit for curtain

Results and Discussion

In designing a system testing plan, we have been deployed and tested with the mobile – based application on Android operating system. Then we began to produce the work by bringing the various sensors connected to NodeMCU ESP8266 to take the values from the sensor to process and store the data in the database and develop mobile application used to display the values that Sent from a sensor to support various devices from users. Also it displays user historical command and results of various devices. There are four models as shown in figure 10, (a) The door model, (b) Light bulb model (in the head of bed), (c) Model of electrical plug and (d) Curtain Blinds Model.

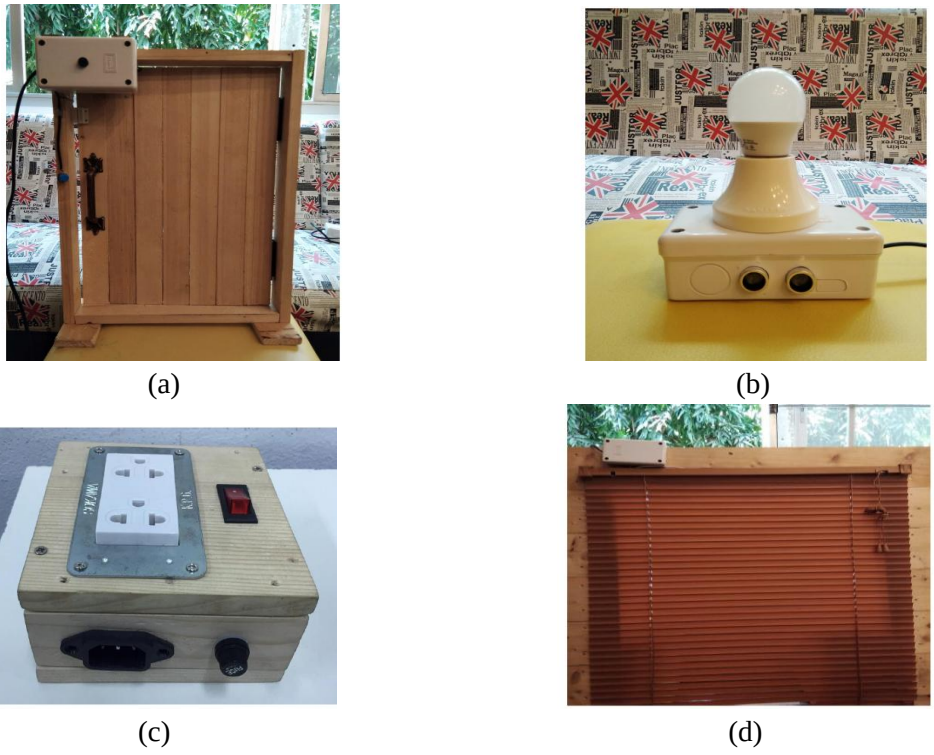
The test phase, for an example, in case of the power switch to turn on-off the power plug as shown in Figure 11 (a) show setting screen on mobile application to turn on plug1 and (b) show the result of "turn on" command to plug1.

To discuss about performance of this proposed system, the application is able to allow users to check various values received by the sensor, support the operation of various devices from users, allow users to look back on their devices and also notify users.

There are some problems regarding product creation. We have chosen wood to produce the models and wood helps to prevent faults caused by electric leakage. So we really need the rely heavily on expertise to build the models. Another one is that Firebase is a library chosen by the developer to use in the development. The NoSQL database is used, which makes the storage design complicated.

For the system capabilities, the application can assist users to check the working status of various electrical devices via smartphone. The system can support the connection of power plug. The application can analyse the usage of user and presentation in the form of graphs.

With the limitations of the proposed system, the application needs to use the internet all the time. The application uses the Firebase free package database system, which has quota or restrictions on its use. Now, the application can be used only in the Android operating system.



(a) (b) (c) (d)

Figure 10: Four Models Built-in Smart Appliance Home Control and Monitoring on Smartphone using Internet of Things

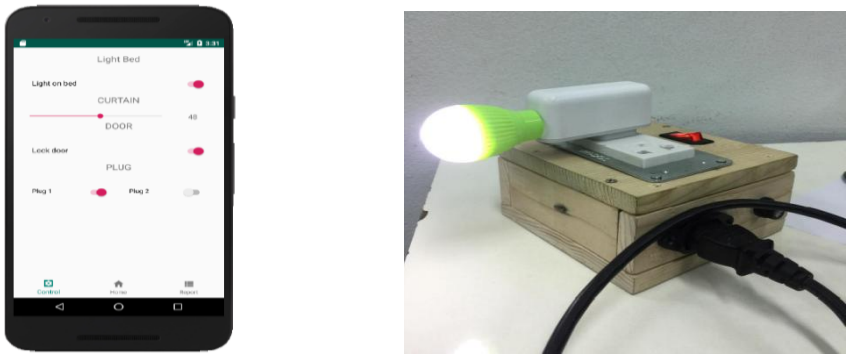
As described in Figure 12 (a) Main menu of mobile application, (b) The control command page on mobile application. The light on bed is set to be turn on, the curtain has been set the light intensity equal to 48 lumens, the lock door is enabled to be "Locked", also there are two plugs, such as plug1 and plug2, are set to "Turn off". There are the default values at the beginning step then we can test as seen in Figure 11.

Conclusions

The proposed system can implement on mobile application, the Android operating system, to display values from sensors such as (1) the relay module, (2) the light sensor, (3) sensors of magnetic switches detect, (4) the push sensor and (5) the ultrasonic sensor.

The application can support users' access to the Android operating system, working in conjunction with the Firebase functions. Also, it supports the control panel operation such as (1) closing - opening of the power plug with specifying the time period for automatic turn-off or manually closed-on by users, (2) opening and closing curtain blinds with receiving the value from the light sensor to turn on-off or turn off-on by the user, (3) closing-opening detection Door or window with a magnetic switch by attaching a sensor for turning off-on the security alert mode from users within the condition that when the magnets that are attached to a door or window are separated and it will send a vibrating notification to the user's smartphone, (4) the buzzer notification, when the doorbell is pressed and there will be a notification to the user's smartphone and (5) intelligent lamp on the head of the bed. When the users turn off all lights in the room. The sensor will then start to work. When the user gets out of bed at night. The control panel will open the emergency lamp on the head of the bed.

The control panel can record various information, for the result of the number of power-ups, the number of closing-opening the door or window, the number of closing-opening curtain blinds, the number of emergency lights on-off and the number of buzzer presses. Therefore, the application can display the usage history of users for the specified period; (1) To turn off-turn on the power plug (2) Closing-opening doors or windows (3) To close-open the curtain blinds (4) Switching off-on emergency lights as well as (5) Ring the bell.



(a) (b)

Figure 11: Testing result (a) setting screen on mobile application to turn on plug 1, (b) result of “turn on” plug1

(a) (b)

Figure 12: Screens (a) Main menu of mobile application, (b) The control command page on mobile application

References

Budianto, Dimas & Nurmaini, Siti & Tutuko, Bambang & Raflesia, Sarifah. (2018). Real-Time Lighting Control System for Smart Home Applications. Computer Engineering and Applications Journal. 7. 10.18495/comengapp.v7i3.267.

Alaa, Musaab & Zaidan, A. & Bahaa, Bilal & Talal, Mohammed & Mat Kiah, Miss Laiha. (2017). A Review of Smart Home Applications based on Internet of Things. Journal of Network and Computer Applications. 97. 10.1016/j.jnca.2017.08.017.

Dener, Murat & Özkök, Yunus & Bostancıoğlu, Cevat. (2015). Creating Scenarios and Using for Smart Systems Generated with Wireless Sensor Networks. Procedia - Social and Behavioral Sciences. 195. 1898-1904. 10.1016/j.sbspro.2015.06.193.

Mendes, T.D.P. & Godina, Radu & Rodrigues, Eduardo & Matias, João & Catalão, João. (2015). Smart Home Communication Technologies and Applications: Wireless Protocol Assessment for Home Area Network Resources. Energies. 8. 7279-7311. 10.3390/en8077279.

M. Ständer, A. Hadjakos, N. Lochschmidt, C. Klos, B. Renner and M. Mühlhäuser. (2012). A Smart Kitchen Infrastructure: 2012 IEEE International Symposium on Multimedia. Irvine. CA. pp. 96-99.

A. Arya, A. Taliyan, P. Chauhan and A. Gautam. (2019). Smart Kitchen with New Measurement. Web and Application Based with Affordable Design. 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU). Ghaziabad. India. pp. 1-6.

A. Taneja, G. Bansal, R. Setia and N. Hema. (2018). Moedor Cleaning Robot. 2018 Eleventh International Conference on Contemporary Computing (IC3). Noida. pp. 1-5.

K. Lekjaroen, R. Ponganantayotin, A. Charoenrat, S. Funilkul, U. Supasitthimethee and T. Triyason. (2016). IoT Planting: Watering system using mobile application for the elderly. 2016 International Computer Science and Engineering Conference (ICSEC). Chiang Mai. pp. 1-6.

S. Aygün, E. O. Güneş, M. A. Subaşı and S. Alkan. (2019). Sensor Fusion for IoT-based Intelligent Agriculture System. 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics). Istanbul. Turkey. pp. 1-5.

THAI AVIATION AMIDST THE COVID-19 PANDEMIC

Nick Alfred Trinidad Umadhay[1], Piyanat Sumetawenunt[2], Jureca Failanga Nacionales[3]

[1,3] Faculty of Education, Iloilo Science and Technology University, Philippines
[2] Faculty of Management Sciences, Panyapiwat Institute of Management, Thailand
Corresponding author, E-mail: piyanatsume@pim.ac.th

Abstract

The global airline industry is facing huge financial losses as Corona Virus-19 (COVID-19) pandemic spreads all over the world, which force the growing number of carriers to cancel flights to countries with a high risk of COVID-19 infection. The International Air Transport Association (IATA) analysis on the financial impact of the novel coronavirus (COVID-19) public health emergency on the global air transport industry issued on 20 February 2020 stated that global revenue losses between $63 billion passenger businesses (in a scenario where COVID-19 is contained in current markets with over 100 cases as of 2 March 2020) and $113 billion (in a scenario with a broader spreading of COVID-19). IATA identified the impact of COVID-19 largely confined to markets associated with China lost revenues at $29.3 Billion, which was equivalent to almost 943 Million Baht. The passenger numbers dropped to -23% in Australia, China, Japan, Malaysia, Singapore, South Korea, Thailand, and Vietnam, indicating a loss of $49.7 Million in revenues. The rest of Asia Pacific dropped 9% to passenger numbers and $7.6 Billion in revenues. This occurred to all markets that had ten or more confirmed COVID-19 cases (as of 2 March 2020). These numbers are worse compared with $7 Billion revenue loss of the SARS epidemic in 2003.

As a result, Thailand's Aviation Industry has faced a number of challenges, such as revenue loss, a low number of tourist arrivals, lay off and voluntary leaves of Aviation employees, stricter monitoring procedure, and additional requirements (Fit to Fly Certificate) to both Domestic and International travels.

Keywords: Thai Aviation, Pandemic, COVID-19, Coronavirus, SARS

Introduction

According to Hannah Jeon of Good Housekeeping Website (2020), Thailand is one of the ten best places to travel in 2020, which means that Thailand expects several travelers from all around the world. Phuket, Chiang Mai, Bangkok, or Krabi are the top tourist destinations in Thailand. Moreover, those provinces have airports to support all prospective travelers who wish to travel by plane as air travels are more comfortable faster regarding travel time and distance, and safe if compared with any other modes of transportation.

Yet, in a current pandemic situation or simply called COVID-19, several reasons forbid people to use air transportation. Among them, the gap between passenger seats in economy class or the filtration system within the plane itself does not prove that it can prevent all kinds of viruses. These bring fear to every passenger, especially in this kind of situation, when there is no medicine or vaccine to cure COVID 19 as of now.

This is not the first time when a virus or germ causes a problem to the Aviation Industry. Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) that happened in 2003 caused a similar problem to the industry.

Impacts of SARS-CoV and COVID-19 to Aviation Industry

The Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) identified in 2003 as an epidemic affected 26 countries and resulted in more than 8,000 cases that appeared in Guangdong Province of Southern China in November 2002. This epidemic believed to be transmitted through person to person or animal to humans, which resulted in symptoms like fever, malaise, influenza-like, headache, myalgia, diarrhea and shivering. (World Health Organization, 2020)

In the same period of April 2002, tourist arrivals in April 2003 declined from 407,917 to 241,351 due to the SARS-CoV epidemic, which is equivalent to 41% decrease of International tourist arrivals at Bangkok International Airport. This was considered as the biggest fall ever in the history of Thai tourism. The figure showed an unprecedented decline of tourist arrivals to Thailand: From China -58%; Singapore -72%; Japan -57% from East Asia. (Imtiaz Moqbuil, 2003)

The Corona Virus -19 (COVID-19) was first reported to the World Health Organization (WHO) Country Office in December 2019 as pneumonia of unknown cause. WHO announced a name for the new coronavirus disease: COVID-19 on 11 February 2020. On 11 March 2020, COVID 19 was characterized as a pandemic based on the assessment of WHO. Symptoms may appear from 2-14 days after exposure to the virus-like cough, shortness of breath or breathing difficulty, fever, chills, muscle pain, sore throat, loss of taste or smell, and other less common symptoms reported including gastrointestinal like nausea, vomiting or diarrhea. (World Health Organization, 2020)

The Tourism Authority of Thailand (TAT) estimates only 14 to 16 million or may fall 65% tourist arrival this year from 33.8 million projected in March due to COVID-19 crisis, according to Panarat Thepgumphanat (2020)

Corona Viruses cause both COVID-19 and SARS CoV. According to Meredith Goodwin, the virus that causes SARS is known as SARS-CoV, while the virus that causes COVID-19 is known as SARS-CoV-2. Corona viruses have spiky projections on their surface that look like crowns. Corona means “crown” in Latin — and that’s how this family of viruses got their name. COVID-19 and SARS are similar in many ways; for example, both are respiratory illnesses caused by coronaviruses. COVID-19 believed to originate in bats, jumping to humans via an intermediate animal host and among humans by respiratory droplets produced when a person with the virus coughs or sneezes, or by contact with contaminated objects or surfaces, with similar stability in the air and on various surfaces, which can lead to potentially serious illness, sometimes requiring oxygen or mechanical ventilation, can have worsening Trusted Source symptoms later on in the illness, have similar at-risk groups, such as older adults and those with underlying health conditions, have no specific treatments or vaccines as of writing.

Overall, the symptoms of COVID-19 and SARS are similar, but there are some subtle differences. Common symptoms for COVID-19 are fever, cough, fatigue, and shortness of breath while SARS are fever, cough, malaise, body aches and pains, headache, shortness of breath. COVID-19’s less common symptoms are runny or stuffy nose, headache, muscle aches and sore throat nausea, diarrhea, chills (with or without repeated shaking), loss of taste, loss of smell unlike SARS diarrhea and chills are the two less common symptoms. (Jill Seladi-Schulman, 2020)

The best way to reduce COVID-19 is by limiting face to face contact. This is where Social Distancing or commonly called as Physical Distancing, which means keeping space between you and other people outside your house at least 6 feet or 1 meter. This practice avoids the large crowd or mass gathering. Social distancing is also important, especially for people with a higher risk of contaminating severe illness from COVID-19. Spread happens when a disease carrier coughs, sneezes, talks, or spread droplets from the mouth or nose to other people’s mouths, eyes and nose and can be inhaled into the lungs. (National Center for Immunization and Respiratory Diseases, 2020)

Impacts of COVID-19 Pandemic to the World/Asia and Pacific Aviation Industry

The possible COVID-19 impact on world scheduled passenger flights estimates a huge reduction, ranging from 32% to 59% of seats by airlines and an overall decrease of 1,815 to 3,213 million passengers, which is equivalent to approximately $236 to $419 billion loss of gross operating revenues of airlines. In Asia and Pacific International Passenger traffic, -38% to -70%, which is reduced to the passenger seat capacity, which is equivalent to -233 to -414 million passengers and revenue loss of -$48 to -$86 Billion. While the estimated impact on domestic passenger traffic is reduced -28% to -50% in capacity and -430 to -756 million passengers with a revenue loss of -$32 to -$56 Billion. (Air Transport Bureau, ICAO, Effects of Novel Coronavirus (COVID-19) on Civil Aviation: Economic Impact Analysis, 12 May 2020)

Thai Airways (TG) requires foreign nationals transiting Thailand and Thai Nationals returning to Thailand to present documents at the check-in counter. This requirement includes foreign passengers transiting Thailand to the third country with the transit time of no more than 24 hours to present a "Fit to Fly" health certificate (can be in English or the official language of the country of departure). The exempted passengers will be required to be in the designated transit areas and strictly comply with the disease control measures applicable at the airport of transit. However, for Thai Nationals returning to Thailand, they are required to present a "fit to fly" health certificate and a certification letter from the Royal Thai Embassy, Thai Consulate Office or Ministry of Foreign Affairs, Kingdom of Thailand. (Thai Airways Travel Advisory, 2020)

The Beginning of COVID-19 in Thailand

According to Thairath website published (2020), first COVID-19 patient in Thailand on 20 January 2020. She is a 61-year-old Chinese tourist who came from Wuhan, China. During the time Thailand was dealing with high-risk air pollution; thus, people did not pay much attention to COVID-19. Also, the government still mentioned that all masks N95 that was widely used in Thailand could protect against dust and virus risks. Furthermore, as of 5 February 2020, reported that the Chinese Woman was cured of COVID-19, which made people thought that the virus was not that bad. This led people to care less about wearing a mask until the situation has gotten worse.

Since Wuhan China was the epicenter of COVID-19, the Thai Nationals in Wuhan requested the government to bring them back to Thailand. In response, the government repatriated a number of Thai Nationals on 4 February 2020 and required to comply with 14-day quarantine as recommended by medical experts.

At that time, there was a rapid spike of local transmission on 10 February 2020, with 32 confirmed COVID-19 patients. The situation got worse when an infected succumbed to death. Thailand's Prime Minister Prayuth Chan-o-cha announced that anyone who got infected with COVID 19 was considered as an urgent patient (Thairath, 2020).

Government declared the State of Emergency and launched Curfew Law.

Thai people learned and understood the words "Super Spreader" when the case surged on after 100 people on 6 March 2020 patients who have attended an event at the Boxing Stadium. Some patients who had the virus were not aware that they were already infected because they had no symptoms or signs of the disease. Moreover, the production of masks was low at that time, and the COVID-19 swab test is very expensive.

On 21 March 2020, the government ordered to close all department stores and Markets in Bangkok. Subsequently, the State of Emergency was declared on 26 March 2020, which prompted public places like department stores, schools, or universities to close temporarily. The government hoped to reduce the number of infected patients, but the figure remained high, so they finally imposed a curfew on 3 April 2020 that limited to leave the house or living place from 10 pm until 4 am. Also, the government issued another law that authorities must quarantine any passengers who came back from outside of the country for 14 days.

Conflict between arrival passengers and authorities after curfew law has launched.

The imposition of curfew did not reach the people outside the country, creating conflicts of incoming passengers. Despite conflicts, the government required them for a 14-day quarantine for public safety and safety assurance. However, some of the passengers did not follow the authority's order and left the airport instead. Those illegal passengers who broke the law were out of a couple of days; thus, the government decided to release warming and demand them to comply with the requirement.

Airline take action with COVID 19

After imposing a curfew, it definitely affected the Airline Industry in one of its major problems was the disruption of airline operations as people had to observe the curfew.

There was one airline that was allowed to operate during the COVID pandemic, but the uniform of the flight crew was changed. From very light, modern, and well designed to big face shield with face mask, and personal protective equipment suit (PPE). In Asian Countries, flight attendants or cabin crews are expected to put on a nice uniform.

Biggest Concern for Airline Industry

This pandemic situation has prompted people to live a different lifestyle and brought the biggest concern to Airline Companies.

First, passengers are unable to sit next to each other because of the social distancing rule. Even if this pandemic would disappear, any humans will learn to keep themselves safe from any viruses. A two meter gap between each other to prevent virus spreading. This new rule remains a big challenge for airline companies as they are required to provide a proper response to the matter.

Second, further monitoring will be provided to passengers and flight crews before security checks were limited to weapons, liquids, drugs, or anything that will harm yourself, other passengers, or cabin crews. Now, airlines would require health measures to battle germs and viruses. After this situation, people are going to observe further personal sanitation like wearing gloves, etc. In addition, airline authorities may also modify the safety procedures to make sure that passengers are fit to fly. The downside of this will lead to dissatisfaction with travelers as they may not have the patience to follow the additional procedures.

Third, is the affordability of the air ticket. The newly imposed guidelines will affect airfare rates. The passengers are always hunting for a cheaper plane ticket. If we consider all factors that make passengers consider using the airline, the price would be the top of the list.

Lastly, the onboard selling of Duty-Free and Souvenir Items will be affected due to the newly imposed guidelines.

The "New Normal" Practices of different Airline Companies

Cabin Crews and Airline Employees from different Airline Companies participated through interviews and online messaging and discussed the "New Normal" practices.

In this time of uncertainty, Delta Air Lines Inc. (Delta) meets with the COVID-19 Disease Control recommendations to its employees. They demonstrate their commitment to safety at every turn. One of their practices is to wear a face mask whenever they are unable to maintain at least 6 feet away from another person. This is also a requirement at crew rooms (while deadheading or commuting) crew hotels (while checking in and out) and ground transportations, including employee buses and airport shuttles. The airlines introduced a new standard of cleanliness for all the aircraft, which called electrostatic spraying. Cleaners come onboard after customers deplane to spray down the cabin with a diluted form of the same cleaning product used to wipe down surfaces. In addition, the spraying cleaners come on board to clean the cabin at every turn. Gate agents are responsible for verifying the full completion of the cleaners before boarding customers. Delay may take place if cleaning standards are not entirely achieved. They also provide masks to customers at the ticket counters and boarding gates. Gate agents are responsible for ensuring customers to procure masks before their boarding.

The Onboard hygiene announcement informs passengers about the new standards of cleanliness and asking everyone to keep their masks on as well for everyone's safety. The service has been changed to care packages that contain a bottle of water, some snacks, and disinfecting wipes. Those packages are available by requests. As an additional effort for the social distancing, metered boarding, metered deplaning are the new practice as well as jet bridge distancing and blocked customer seats.

Qatar Airways, as 5th-time winner of the World's Best Airline in 2019 World Best Airline Award. (Miller**,** 2019) trained its cabin crew on how to minimize their chances of contacting or spreading the COVID-19 infection. Cabin crews are authorized to wear surgical masks, disposable gloves and PPE in selected flights. The crew will be thermally screened before and after the flight and would be quarantined if any colleagues or passengers on a flight show any symptoms or tests positive of COVID-19. It also encourages passengers to follow hygiene practices recommended by the World Health Organization, such as regular hand washing and refraining from touching their face.

The Philippine Airlines (PR), Thai Lion Air (SL), Bangkok Airways (PG) and Cebu Pacific (5J) requires their Cabin Crew and Airport personnel to wear PPE suit for international flights, especially to high-risk cities by wearing mask and gloves, stepping on the foot bath, and going through temperature scanning in every entry points, social distancing is strictly implemented in common areas. Rapid antibody testing is required for all 5J Cabin Crew and Pilot, while PPEs are provided depending on employee function. Air Asia (XJ) introduces red-hot PPE suits designed by a Filipino Designer Puey Quinones come in XJ signature bright red and were used in the recent repatriation flight from Bangkok to Manila.

Conclusions

Based on the findings of the study, the COVID-19 pandemic, Thai Aviation Industry, faced the worst crisis that resulted in revenue loss. A huge decrease in Thailand's tourist arrivals was recorded amidst the COVID-19 pandemic. There is a high aviation employee lay-off, and employees filed voluntary leaves to lessen the expenses of the Airline companies. The Thai government, in coordination with the Airport of Thailand (AOT), imposed stricter monitoring and security procedures to both Inbound and Outbound passengers. Additional requirements (Fit to Fly Certificate) for International and Domestic passengers were implemented.

Suggestions

In order to recover from the COVID-19 crisis, the airline company needs to cut costs in order to return to profits and, at the same time, seek to increase flexibility to cope with fluctuations in demand. The onboard services and amenities should also be reduced to its food and beverage service and blankets to be distributed to customers only by request. The Airline company can also freeze hiring new staff, adjust bonuses and incentives as well as the salary and any significant management pay for the meantime until revenue will be recovered. The company can also offer voluntary leave to its employees to retain jobs. They should include COVID-19 free in their campaigns and marketing ads by showcasing their safety measures to give assurance to the passengers of their safety and convenience while flying with them. Onboard items to sell might include safety kits such as hygiene wipes, masks, alcohol spray, sanitizer, and PPEs. All Airline Companies are also encouraged to open new flight routes to low-risk cities.

References

Airports of Thailand, April 2020, https://www.airportthai.co.th Civil Aviation: Economic Impact Analysis, May 2020, https://www.icao.int

International Civil Aviation Organization (ICAO), Effects of Novel Corona Virus (COVID19) on Thai Airways Travel Advisory, March 2020, https://www.thaiairways.com

International Air Transport Association, COVID-19 Impact on Asia-Pacific Aviation Worsens, April 2020. https://www.iata.org

Imtiaz Moqbuil, April 2003, https://www.travel-impact-newswire.com

Jeon, H. (2020). The 10 Hottest Travel Destinations for 2020. Retrieved from https://www.goodhousekeeping.com/life/travel/g30335840/best-places-to-travel-2020

National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases, 6 May 2020, https://www.cdc.gov)

Jill Seladi-Schulman, PhD, 2 April 2020, https://www.healthline.com

Panarat Thepgumphanat, Reuters Business News, May 2020, https://www.reuters.com/

World Health Organization (WHO), Disease Information, May 2020, https://www.who.int

THE REVIEW OF RFID TECHNOLOGY FOR GRID-BASED STORAGE MOBILE ROBOT LOCALIZATION

Vikrom Maikaensarn*

Department of Logistics and Transportation Management, Panyapiwat Institute of Management
***Corresponding author, Email:** vikrommai@pim.ac.th

ABSTRACT

This paper focused on the review of RFID technology to supports the order picking activities of grid-based storage, one of the vertical storages designed which incorporated grid rail for the mobile robots to move around and pick up the orders. However, with the lack of literature in the grid-based storage operation, the proper controls protocol and wireless technologies used in navigating the robots on it are yet to be verified. For the system to work effectively; the robot required a module to help localized its position such as RFID technology, which offers the flexibility of implementation. The selection of RFID is based on the implementation purpose and the properties of different types of RFID, which required the understanding of RFID characteristics and properties. Thus, this paper is consisted of the review of RFID technologies and the previous works of RFID in robot navigation, to support the technology selection for a localization system for the grid-based storage mobile robot.

Keywords: Passive RFID, Grid-based Storage, RFID for localization, Wireless Sensor, Automated Warehouse

Introduction

Warehouse operation is one of the main activities in supply chain management. It is involved in both inbound logistics and outbound logistics. In inbound logistics, transported goods and materials are stored inside the warehouse for either distribution or manufacturing purpose. On the other hand, warehouse activities also include retrieves, sorts and prepared the goods for distribution process in outbound logistics activity which uses wireless technologies in activities management and asset flows tracking. However, wireless technologies implementation requires the user to understand the system and process of their model to select the ideal wireless technology for later implementation.

There are many products stored within the warehouse; therefore, to improve the efficiency in storing and retrieving activities, automated technologies were incorporated into warehouse operation processes to reduce the time used in the activity and error from inattention. In warehouse operation, technologies, such as automated technology are used to help reduce the time the operators spend in both inbound and outbound logistics activities. These technologies use sensors, such as RFID to help keep track and verify whether the operations are executed effectively.

In warehouse operation, RFID technologies are used to identify and to keep track of the movement of the products inside the warehouse via IoT sensory in automated technologies and material handling equipment (Gandino, Montrucchio, Rebaudengo, & Sanchez, 2007; Hashimoto, Isshiki, Iguchi, Morisaki, & Ishii, 2006; Kwok & Wu, 2009; Lim, Bahr, & Leung, 2013; Lu, Xu, Zhong, & Wang, 2017; Lu, Xu, Zhong, & Wang, 2018; Mo & Pearson, 2011; Wamba, Lefebvre, & Lefebvre, 2006). There are many reviews related to automated warehouse technology; however, there is one particular system in automated warehouse technology, such as grid-based storage, that was left without any operation processes literature.

Grid-based storage or grid storage is vertical storage designed as a hive grid to store goods, and the grid rail, used by mobile robots for picking activity. However, the grid-based storage model is lacked in technologies implementation and operations literature (Beckschäfer, Malberg, Tierney, & Weskamp, 2017). Nonetheless, as RFID technology progress, many warehouse operations also begin to adopt its technology. Thus, this paper aims to find a suitable RFID technology for grid-based storage model for future development and implementation in both education and warehouse operations.

Reviews of literature

There are numerous applications which use an RFID system as a mean to manage the activities and collect information (Jia, Feng, Fan, & Lei, 2012). However, one of the main usages of RFID is indoor localization. In warehouse operation, RFID technologies are presented in different applications ranged from receiving, storing, picking, and shipping.

Based on the mentioned activities, RFID applications can be grouped into three: products verification, storing location, and navigation. First, in product verification, RFID is used similarly to the barcode in representing goods and inventories (Kaur, Sandhu, Mohan, & Sandhu, 2011; Loebbecke, 2005), and the interrogator is used, not only limited to a handheld device, but also with robotic technology. With the RFID technology, checking and order pick up from the stored location can be done with ease (Chen et al., 2013; Lim et al., 2013). Second, in storing location, RFID is used as a locator to indicate where each product is stored on the shelves. Lastly, for navigation, similar to storage application, RFID can be used to represent the current location of automated technology, such as AGV, and to keep track of the current location of goods when passed through the warehouse facility.

There are many applications of RFID in warehouse management using varieties of RFID technologies (Lim et al., 2013). Nonetheless, it is necessary to understand the fundamental properties of RFID to utilize them effectively. In the following section, the review of RFID implementation in the contexts of AGV and robot navigation to find the suitable RFID to implement in the grid-based, followed by the type of RFID and different frequencies of RFID will be discussed to provide more background information regarding their properties.

RFID for AGV and robot navigation

To find out which RFID technology fitted with the grid-based storage or to be more precise, GSR (grid system robotic) model; it is important to evaluate previous studies relating to RFID for AGV navigation system. AGV is part of automated technology founded in the warehouse, which used different sensors to navigate throughout the facility. Several papers had depicted how RFID technology is used as part of the AGV navigation system. For example, the RFID tags deployment in a grid shape to navigates the AGV (Lu et al., 2017; Lu et al., 2018; Park & Hashimoto, 2009). The AGV used the RFID tags determines its travel path from its starting position and then verifies the final destination based on the following workflow shown in Figure 1 below. This flowchart can also serve as a basis for the navigation system using RFID, which can be modified to suit different implementations.

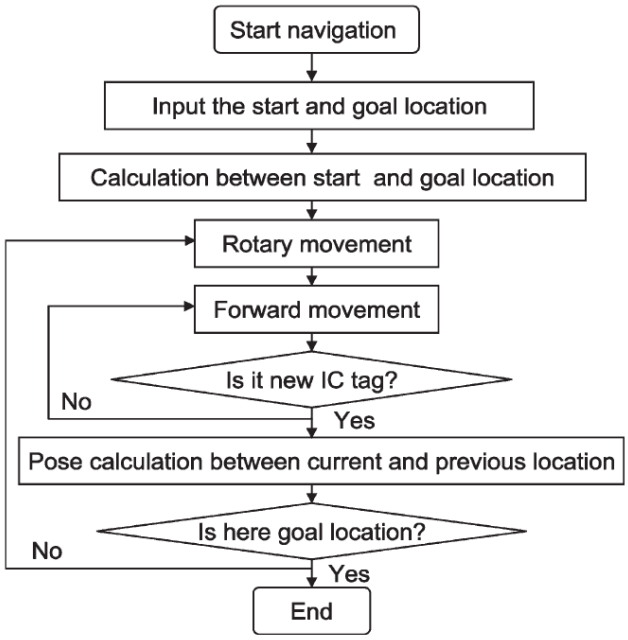


Figure 1: Navigation procedure for RFID based AGV (Park & Hashimoto, 2009)

Another navigation example using RFID is the localization using a weighted average of RFID tags position with the weight variables, which depends on the signal strength of the RFID (Mo & Pearson, 2011). Another case is the used of multi-sensory, including RFID, on the robot to enhance its navigation system (Xiao-Long, Chun-Fu, Guo-Dong, & Qing-Xie, 2017). Although there are many forms of RFID implementation, their operation is very similar to each other where RFID tags are placed and used as a representation of specific locations. In short, RFID tags can be used to either represent the particular location or provides signal strength information to estimate the robot position within the facility.

Type of RFID

In general, RFID technology has different properties based on their designs. Nonetheless, RFID technology can be classified into two types, which are based on the communication between the reader (interrogator) and the tag (transponder): near-field RFID and far-field RFID. The difference between these two is the range of communication. The near-field RFID uses the principle of magnetic induction for the reader to communicate with the tag, which results in short read range. On the other hand, far-field RFID technology communicates with each other using far-field emission capture of electromagnetic (EM) wave (Kaur et al., 2011) with a backscattering technique which yields longer read range, as compared to the near-field RFID. Figure 2 below shows how both near-field and far-field RFID reader communicates with the tag.

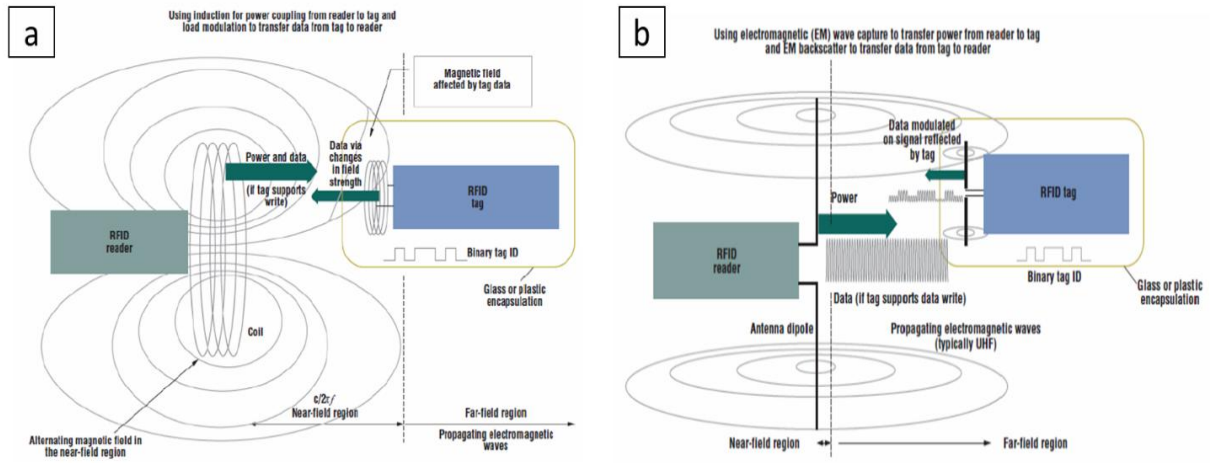


Figure 2: (a) Near-field and (b) far-field RFID communication between the reader and passive tag (Kaur et al., 2011)

Passive and Active RFID tags

Another important component of RFID technology is the tag. RFID tags are grouped into three based on their designs: active RFID tag, semi-passive RFID tag, and passive RFID tag. Both semi-passive RFID and passive RFID are similar to each other, in which they both lacked signal transmission, whereas semi-passive RFID design has the battery included for additional features, which can be applied to the tag. However, RFID is intended for the navigation for AGV and mobile robots only, making the semi-passive RFID and its additional functions superfluous for the grid-based storage model. Thus, the comparison between RFID tags is simplified to the comparison of active and passive tags only, as shown in Table 1.

Table 1: The difference between active and passive RFID tag

Type of RFID tag	Advantage	Disadvantage	References
Active	• Required no energy from the RFID reader to be operable / Always transmitting RF signal • Large data storage • Long detection range • Faster data transfer • Enabled multiple RFID tags detection simultaneously	• Larger size to hold the battery • Sensitive to the surrounding environment • Creates radio noises • Duration of operation depends on battery life • High cost • Required RFID reader to read/write the information	• (Park & Hashimoto, 2009) • (ELAInovation, 2019) • (Ray, 2018) • (RFID4U, 2020a) • (Smiley, 2019)
Passive	• Smaller designs • Low cost • Long life span • Flexible deployment • Require no battery to be operable	• Required the presence of RFID reader to be operable • Small data storage • Small detection range • Inferior in term of multiple tags detection compared to active RFID	

Frequencies of RFID

Even though understanding the difference between types of tags is essential; nevertheless, to pick the right RFID for the grid-based storage design, it is also important to consider the frequency used in the selected RFID. Hence, the following section will discuss the difference between different RFID frequencies to provide necessary information regarding how to choose the right RFID for the grid-based storage. RFID technology, reader and tag, communicate with other using a specific radio frequency which can be categorized into three groups: low frequency (LF), high frequency (HF), and ultra-high frequency (UHF). Each frequency range has unique properties, as shown in Table 2, which is suitable for different applications.

Table 2: Properties of different frequencies used in RFID technology

Frequency	Advantages	Disadvantages	Read range	References
LF **125-134 kHz**	• Low cost • Not sensitive to another wave interference • Works with metallic and liquid materials or surfaces	• Slow reading/writing speed	• Read range of the tag < 10cm depending on the antenna configuration	• (Klair, Chin, & Raad, 2010) • AB&R® (2019) • RFID4U (2020b) • ResourceLabelGroup (2018) • TSL (2020)
HF **13.56 MHz**	• Moderate performance on metallic and liquid materials or surfaces • Faster read/write rates compared to LF • Low-cost tag	• Sensitive to another radio wave interference • Fewer tags simultaneously detected when scanned	• Read range of the tag < 30cm depending on the antenna configuration	
UHF **860-960 MHz**	• Multiple tags detection supported • Faster data transfer rate as compared to HF	• High cost (equipment) • Performance dropped when placed on metallic or liquid materials/ surfaces	• Read range of the tag < 100m depending on the antenna configuration	

Theoretical Background

In the process of RFID selection; first, it is important to understand the concept of grid-based storage and its functions in the warehouse or the distribution center. Grid storage is similar to that of the vertical storage in the warehouse. Nevertheless, the main difference between grid-based storage and regular vertical storage is its designs, which not only saved a considerable amount of space in the warehouse by omitting the aisles spacing but also incorporates a grid-rail on the top for the mobile robots to move around. The basic design of the rail of the storage is as shown in Figure 3 below. Second, with the designs of grid-based storage and its function in mind, the selection of RFID must also consider different types of RFID technologies and its limitation. Finally, the preliminary test is a necessity to fully understand how well the selected RFID or any other wireless technologies function when being implemented in the model.

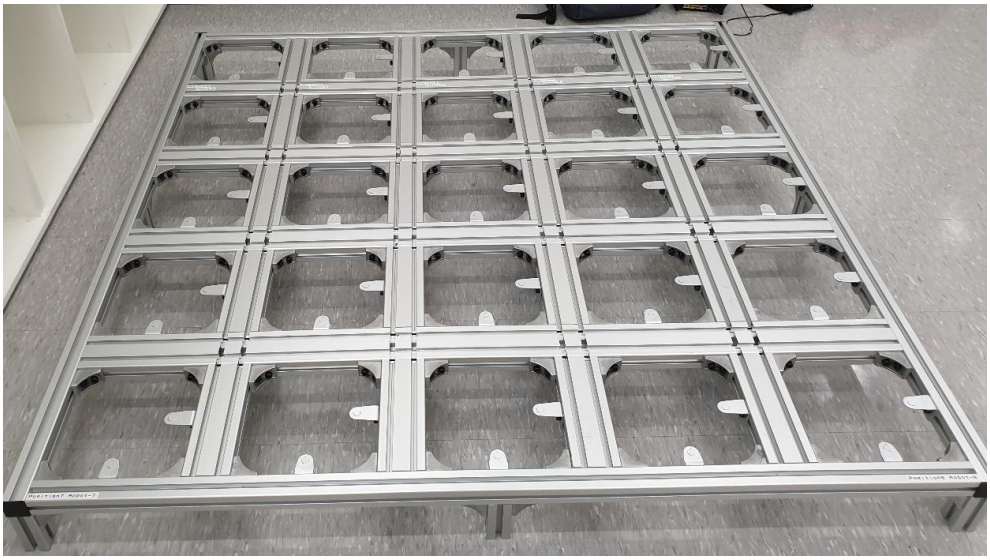


Figure 3: The design of grid rail for grid-based storage

Limitation of RFID

RFID may have the advantage much like other wireless technologies, which do not require to be within line of sight. Nonetheless, there are points, which must be considered: the effects of the surrounding environment to the performance of wireless technologies; the read range of the selected wireless technologies and their modules; and the effects of the tag SOAP.

Read range

Several factors are affecting the read range of RFID technologies: the antenna's sizes; type of communication (near-field and far-field); active or passive; and RFID frequencies. As said, the device read range plays an important role in determining the accuracy of the localization system. Long read range could easily trigger multiple tags detection, which lowered the localization accuracy depending on the method used to determine the location leading to a more complex algorithm to determine the asset's position. Also, far-field RFID tends to be more expensive as compared to near-field RFID. Therefore, near-field technologies, which used a passive RFID system, are more suitable to the system as compared to the active or far-field RFID. Furthermore, this also narrows down the range of RFID frequencies, since near-field RFID, which operates using frequency less than 100MHz (Kaur et al., 2011) is either LF or HF RFID.

Effects of the surrounding environment

As for the environmental factor, the users must note that RFID technologies used magnetic induction and EM wave propagation to communicate with the near-field and far-field RFID, respectively. Therefore, the environment in which RFID tags are deployed, and the position between the tags and the readers can affect RFID communication (Klair et al., 2010). In this case, the main issue is what is surface RFID tags are placed, where the metal surface caused a disturbance in both magnetic induction and signal backscattering, as shown in Table 3, which also summarized the properties of the types of RFID technology.

Table 3: Summary of different types of RFID technology (Klair et al., 2010)

	LF	HF	UHF
Frequency	< 135 kHz	13.56 MHz	860 – 930 MHz
Tag type	Passive		Active, Semi-passive, Passive
Physical coupling	Magnetic induction (inductive coupling systems)		Backscatter systems
Communication Range	Near-field communication		Far-field communication
Metal	Magnetic induction disturbance		Signal backscatter disturbance
Water/human body	None	Attenuation	Attenuation

In this case, the grid storage is manufactured using metal. Thus, the metal surface in which RFID tags are placed on could affect the inductive coupling between the reader and tag. The example of how metal surface affects the efficiency of RFID magnetic field is as shown in Figure 4 below.

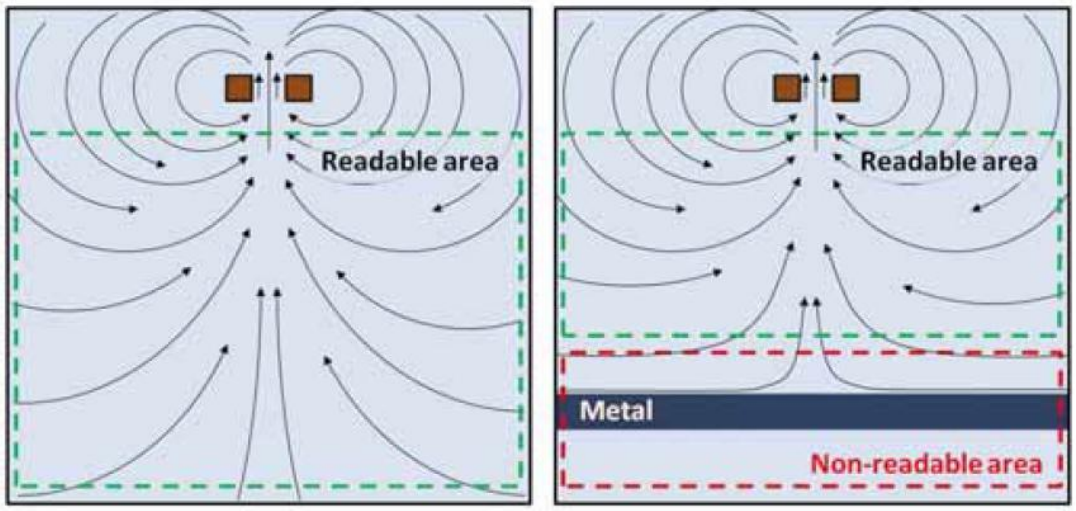


Figure 4: The distortion of the magnetic field caused by metal surface (Ciudad, Cobos, Sanchez, & Aroca, 2010)

Although metal surface caused a disturbance to the inductive coupling between an RFID reader and tag, the distance between RFID tag and reader is also a part of the communication efficiency. Even if the disturbance occurred, but the reader and tag are close to each other, the passive tag would be able to receive enough energy to transmit back its information to the reader. To solve this problem, in the event where the grid-storage model has the significant effects to the RFID communication, a flexible absorbent material that has the property to neglect electromagnetic wave interference could be applied to recover the inductive coupling efficiency between a reader and tag. Otherwise, unique RFID tags designed for the metal surface, which are more expensive, could be used for this model instead of the regular RFID tag to make the navigation system functional. While on the other hand, the metal surface not only causing a blind-spot by reflecting the transmitted signal but also can amplify the reflected signal from the far-field RFID tag, by acting as a repeater (Sattlegger & Denk, 2014); results in greater read range, which are not always the desired outcome and increased the complication in implementation design.

RFID tag SOAP

Therefore, to effectively utilize the selected RFID technology; other than the specification of RFID technologies and its antenna, RFID tag SOAP is another general indicator of RFID communication efficiency (Armstrong, 2020). The word SOAP is derived from factors which affect RFID performance: size, orientation, angle, and placement. These factors could affect both near-field and far-field RFID. However, it tends to have more impact on the far-field RFID, which is more sensitive than the near-field RFID.

- Size: Larger tag size increase the communication range as it contains a larger antenna. This factor is applied to both near-field and far-field RFID.
- Orientation: Tag orientation is how the tags are placed on the surface vertically or horizontally, as shown in Figure 5 below. In general, different tag orientation could result in lower read range for those with orientation sensitive design or the RFID using backscatter systems. On the other hand, RFID tags using inductive coupling systems are less sensitive to the tag orientation.

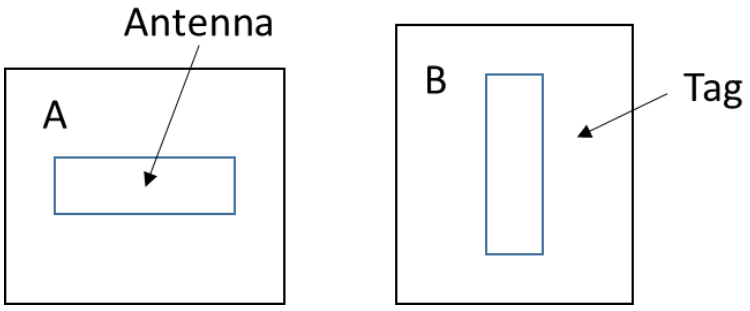


Figure 5: RFID tag orientation; (A) horizontal placement, and (B) vertical placement

- Angle: Tag's angle to the RFID reader could affect the read range. Thus, it is recommended to place the tag parallel, or on the same plane, to the reader to maximize the read range (Armstrong, 2020; Sanghera et al., 2007).
- Placement: Tag's placement is related to the angle and the orientation of the tag. If the tag is placed perpendicular to the reader, the read range could be reduced or, in some case, cannot detects the tag. Another point to consider is the surface where the RFID tag is placed on as mentioned in the previous section, effects of the surrounding environment, where either metal surface or objects with high water content could affect the inductive coupling systems.

Analysis

Initially, without the RFID technology to help localized the grid-based storage mobile robot, the robot can only follow basic commands, using manual input, to move only one step at a time. To eliminates user input process, the smart algorithm comes in to help analyze user's requirements and controls robots' activities. However, for the smart algorithm to work, the information about the robots' current location is required for the algorithm to process and initiates smart actions. Therefore a localization technology suitable for the grid storage design is necessary for the model to be functional.

RFID selection

Technology selection must cohere with the user purpose or objective of technology implementation, which allows the user to setup system requirement to match up with the properties of the RFID technologies, as shown in Figure 6. Therefore, there are three elements to be considered when selecting the RFID for grid storage model: the RFID-based application for the model, the properties of grid storage model, and the understanding of RFID limitation.

The first element is the desired application, or the purpose of implementing RFID technology, for the model. In this study, RFID is intended for localization system for a mobile robot to navigate throughout the rail on the grid storage. When considering this application, both far-field and near-field RFID technologies are still viable for this model. For the far-field, the ability to detect multiple RFID tags enabled location estimation for the robot. However, multiple RFID detection for navigation is more suitable for the robot without movement restriction. On the other hand, near-field RFID can be used to represent a specific location due to having short communication. Thus, near-field RFID not only yield higher accuracy but also allow less complicated localization algorithm.

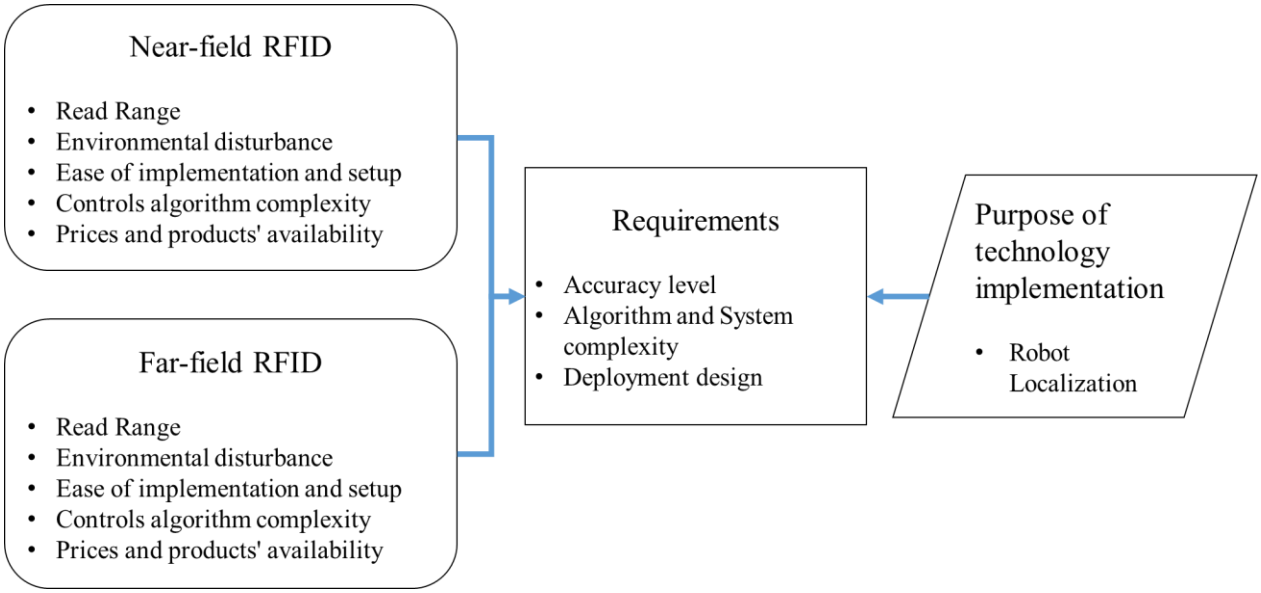


Figure 6: The RFID requirements matching process

Next is the basic properties of the grid storage model: surface material, and the distance from the surface to the bottom end of the robot. Based on grid storage design, the distance between the lower ends of the robot to the surface of the rail is less than 5cm, depending on the robot design. Furthermore, grid storage also has a movement constraint to the robot movement. Since the robots can only move in four directions, the far-field RFID is not necessary, since the robot only required a verification point to determine its travel path. Therefore, the use of near-field RFID should be enough for the localization system.

Finally, when it comes to the environmental factor, both far-field and near-field technologies could be affected by the metallic environment on the grid storage. As referred to Table 3, the metal surface caused a disturbance to communication between RFID tag and RFID reader. Since the purpose of the RFID implementation, the localization system required high accuracy module to locate the mobile robot on the grid storage. Therefore, lesser sensitive technology, which allows flexible RFID deployment, is more suitable than the sensitive technology. Thus, making near-field RFID more preferable than far-field RFID.

Implementation Design

The implementation design of near-field RFID can be based on the far-field RFID setup, as demonstrated by Mo and Pearson (2011), as shown in Figure 7, since the RFID grid in this model is similar to the grid storage design as shown in Figure 3.

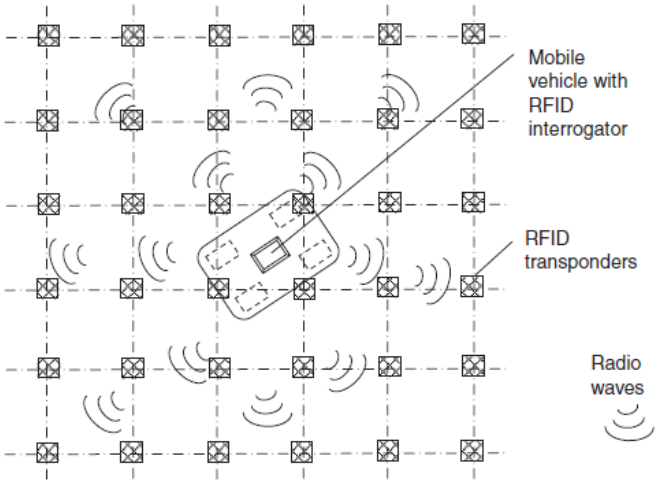


Figure 7: RFID grid and mobile vehicle setup using UHF RFID (Mo & Pearson, 2011)

With fixed dimension and limited robot movement of the grid storage, near-field RFID tags can be deployed on the grid as shown in Figure 8, where each block in the grid storage only requires one RFID tag to represent its location. Furthermore, a single point representation of near-field RFID also offers high precision and accuracy as compared to the multiple tags detection. Thus, offer an effective localization system for the smart algorithm to operate the grid storage robot model. RFID tag SOAP is also considered in the design. To represent a specific location, the read range of the near-field RFID tag, which neglects the effect of tag orientation, must be very small, thus resulted in a smaller tag which also enabled flexible deployment. Moreover, based on the implementation design, the RFID reader and the RFID tag are parallel to each other, which is a recommended angle for the RFID tag to receive the maximum energy transfer via magnetic induction (Sanghera et al., 2007).

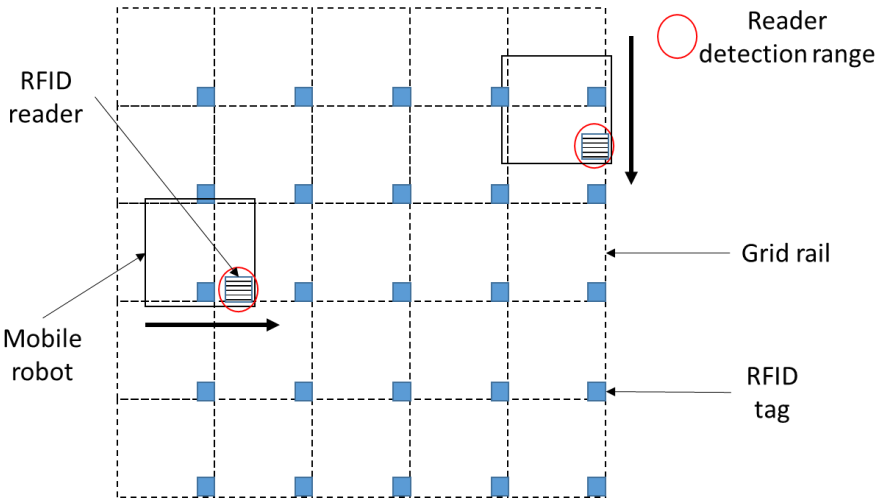


Figure 8: RFID grid and mobile robots setup using near-field RFID

Nonetheless, to fully grasp the RFID performance and the extent of the signal disturbance when being deployed in a metal environment, it is recommended to thoroughly test the selected technology to check how well it performs.

Conclusion

The smart robot used a smart algorithm to initiate an intelligent action. However, for the algorithm to work, the location of the robots within the facility is required. However, with the lack of literature in the context of grid storage control systems and the RFID module used to localized the robots in grid storage model; therefore the investigation of different RFID technologies is necessary to supports the localization module selection for the future development of the grid storage model. Through reviewing the related literature and properties of RFID technologies, it is founded the ideal type of RFID is the near-field RFID, which has a short communication range and can be used to represent each box in the grid storage. Nonetheless, the testing of RFID performance is required to understand how well the selected RFID module performs when applied to the grid storage and mobile robot. Nevertheless, even with some limitations, the near-field RFID would offer high precision and accuracy in localizing the robot; allowing an ease of implementation, and robot calibration.

Suggestion

In addition to the robot localization, RFID technology can also be used in application such as population monitoring, indoor navigation, assets tracking, and storage verification. Furthermore, due to the flexibility and reliable performance of the RFID technology, it can also be used in employees' training as checkpoints component during the training sessions. Therefore, not only that RFID is reliable for the localization system, but this technology also provides a mean to improve the performance within the workplace and helps with the assessment of the facilities for better operations and services.

References

AB&R®. (2019). What is RFID and How Does RFID Work? Retrieved from https://www.abr.com/what-is-rfid-how-does-rfid-work/

Armstrong, S. (2020). 6 Factors that Affect RFID Read Range. Retrieved from https://www.atlasrfidstore.com/rfid-insider/improve-rfid-read-range

Beckschäfer, M., Malberg, S., Tierney, K., & Weskamp, C. (2017). *Simulating Storage Policies for an Automated Grid-Based Warehouse System*, Cham.

Chen, J. C., Cheng, C.-H., Huang, P. B., Wang, K.-J., Huang, C.-J., & Ting, T.-C. (2013). Warehouse management with lean and RFID application: a case study. *The International Journal of Advanced Manufacturing Technology*, *69*(1-4), 531-542.

Ciudad, D., Cobos, P., Sanchez, P., & Aroca, C. (2010). RFID in Metal Environments: An Overview on Low (LF) and Ultra-Low (ULF) Frequency Systemsled.

ELAInovation. (2019). Active RFID vs. Passive RFID: which differences? Retrieved from https://elainnovation.com/active-rfid-vs-passive-rfid-which-differences.html

Gandino, F., Montrucchio, B., Rebaudengo, M., & Sanchez, E. (2007). *Analysis of an RFID-based information system for tracking and tracing in an agri-food chain.* Paper presented at the 2007 1st Annual RFID Eurasia.

Hashimoto, N., Isshiki, N., Iguchi, M., Morisaki, M., & Ishii, K. (2006). Assets location management solution based on the combination of SmartLocator and RFID. *NEC technical journal, 1*(2), 92-96.

Jia, X., Feng, Q., Fan, T., & Lei, Q. (2012). *RFID technology and its applications in Internet of Things (IoT).* Paper presented at the 2012 2nd international conference on consumer electronics, communications and networks (CECNet).

Kaur, M., Sandhu, M., Mohan, N., & Sandhu, P. S. (2011). RFID technology principles, advantages, limitations & its applications. *International Journal of Computer and Electrical Engineering, 3*(1), 151.

Klair, D. K., Chin, K.-W., & Raad, R. (2010). A survey and tutorial of RFID anti-collision protocols. *IEEE Communications Surveys & Tutorials, 12*(3), 400-421.

Kwok, S. K., & Wu, K. K. (2009). RFID-based intra-supply chain in textile industry. *Industrial Management & Data Systems*.

Lim, M. K., Bahr, W., & Leung, S. C. (2013). RFID in the warehouse: A literature analysis (1995–2010) of its applications, benefits, challenges and future trends. *International Journal of Production Economics, 145*(1), 409-430.

Loebbecke, C. (2005). RFID technology and applications in the retail supply chain: The early metro group pilot. *BLED 2005 Proceedings*, 42.

Lu, S., Xu, C., Zhong, R. Y., & Wang, L. (2017). A RFID-enabled positioning system in automated guided vehicle for smart factories. *Journal of Manufacturing Systems, 44*, 179-190. doi:https://doi.org/10.1016/j.jmsy.2017.03.009

Lu, S., Xu, C., Zhong, R. Y., & Wang, L. (2018). A passive RFID tag-based locating and navigating approach for automated guided vehicle. *Computers & Industrial Engineering, 125*, 628-636.

Mo, J., & Pearson, D. (2011). Localization of position using radio frequency identification array. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 225*(5), 675-684.

Park, S., & Hashimoto, S. (2009). Autonomous Mobile Robot Navigation Using Passive RFID in Indoor Environment. *IEEE Transactions on Industrial Electronics, 56*(7), 2366-2373. doi:10.1109/TIE.2009.2013690

Ray, B. (2018). Active Vs. Passive RFID For Location Tracking. Retrieved from https://www.airfinder.com/blog/active-vs-passive-rfid

ResourceLabelGroup. (2018). Comparing different types of RFID tags. Retrieved from https://www.resourcelabel.com/comparing-different-types-of-rfid-tags/

RFID4U. (2020a). How to Select a Correct RFID Tag – Passive vs. Active. Retrieved from https://rfid4u.com/rfid-basics-resources/how-to-select-a-correct-rfid-tag-passive-vs-active/

RFID4U. (2020b). How to Select a Correct Tag – Frequency. Retrieved from https://rfid4u.com/rfid-basics-resources/how-to-select-a-correct-tag-frequency/

Sanghera, P., Thornton, F., Haines, B., Fung, F. K. M., Kleinschmidt, J., Das, A. M., . . . Campbell, A. (2007). *How to cheat at deploying and securing RFID*: Elsevier.

Sattlegger, K., & Denk, U. (2014). Navigating your way through the RFID jungle. *White paper, Texas Instruments*.

Smiley, S. (2019). Active RFID vs. Passive RFID: What's the Difference? Retrieved from https://blog.atlasrfidstore.com/active-rfid-vs-passive-rfid

TSL. (2020). RFID Comparison. Retrieved from https://www.tsl.com/support/rfid-comparison/

Wamba, S. F., Lefebvre, L. A., & Lefebvre, E. (2006). *Enabling intelligent B-to-B eCommerce supply chain management using RFID and the EPC network: a case study in the retail industry.* Paper presented at the Proceedings of the 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet.

Xiao-Long, W., Chun-Fu, W., Guo-Dong, L., & Qing-Xie, C. (2017, 20-22 Oct. 2017). *A robot navigation method based on RFID and QR code in the warehouse.* Paper presented at the 2017 Chinese Automation Congress (CAC).

THREE-STAGE STOCHASTIC PROGRAMMING APPROACH TO SOFTWARE DEFINED NETWORK AND CLOUD RESOURCE PROVISIONING FOR INTERNET OF VEHICULAR NETWORK

Nay Myo Sandar

Faculty of International College, Panyapiwat Institute of Management, Thailand
Corresponding author, E-mail: naymyo@pim.ac.th

ABSTRACT

This paper conducts the research for Internet of Vehicles (IoV) in which a group of vehicles are equipped with wireless sensors. Such sensors monitor data about traffic jam, accident, location to provide safety and comfort to drivers. To provide useful information to users, data readings from sensors need to be processed. In this case, cloud computing is a convenient model which can provide abundant resources of storage and processing power to data from numerous sensors. However, it can encounter bandwidth constraint and also bottleneck when numerous sensors transfer data to the cloud over long range communication. Therefore, cloud-based internet of vehicular network framework designed by software-defined networking (SDN) is proposed to address network scalability problem when data from sensors are transferred to the cloud services. The SDN approach can achieve the adaptive routing control which improves the network performance between vehicular network and cloud services. In addition, three-stage stochastic programming is applied to minimize the total cost of provisioning SDN bandwidth and cloud resources under uncertainty scenario of data demand. The numerical results can provide the minimum total cost for implementing proposed framework.

Keywords: Internet of Vehicles, Cloud Computing, Software-Defined Networking, Three Stage Stochastic Programming

Introduction

Today, the number of vehicles used in many countries is increasingly higher than 1 billion and is expected to reach 2 billion by 2035 (Sperling & Gordon, 2008). Due to increasing number of vehicles, it can raise traffic congestion and accidents on the roads. Until recently, Internet of Vehicles (IoV) is emerged to solve the various driving and traffic problems in current transportation systems (Yang, Li, Lei, & Wang, 2017: 1-17). In IoV system, the intelligent sensor devices are installed in vehicles and traffic lights which can be able to absorb data from the environment and exchange information between vehicles to vehicles and vehicles to roads (Sadiku, Tembely, & Musa, 2018: 11-13). The use of IoV system can provide vehicle transportation safety and improve users' convenience (Buyya & Dastjerdi, 2016). However, data readings from sensors installed in vehicles and traffic lights needs to be processed to become useful and meaningful information. In the past, data stream management system (DSMS) is applied to store and process data streams. However, we need to take into account the weakness of DSMS because it has limited amount of memory and CPU to store and process data streams. Therefore, DSMS can be overloaded when the data streams from sensors is increasingly arriving. Under overload, DSMS can experience excessive data loss and processing delay.

As a result, cloud computing has been emerged to provide greater storage capacity and processing power for achieving value-added services such as data analytics, backup, database management, respectively (RajkumarBuyya, Yeo, SrikumarVenugopal, Broberg, & Brandic, 2009: 599-616). Therefore, data streamed from sensors can be transferred to the cloud for storing and processing in order to utilize value-added services. Then, the data processed or stored by the cloud services can be remotely monitored and managed anywhere and anytime by the users through their Internet-connected devices such as mobile phones, tablets, and other smart devices.

However, although the cloud computing provides a large pool of resources to use cloud services, serious network latency could encounter when sensors with limited bandwidth transfer data to the cloud services over long distance. Moreover, network bottleneck can be experienced for data transmission between several sensors and cloud. As a result, we propose a cloud-based internet of vehicular network framework using software-defined networking (SDN). In the framework, sensor data from vehicular network is sent to cloud server with better SDN bandwidth. By using SDN approach, it can flexibly manage network with the use of programmability for traffic scheduling, congestion, and routing (Xia, Wen, Foh, Niyato, & Xie, 2014: 27-51). In this case, ISP service providers can provision the sufficient amount of network bandwidth between vehicular network and cloud services.

Furthermore, we also applied three-stage stochastic programming to minimize the total cost of provisioning SDN bandwidth and cloud resources under uncertainty of data demand. Then, the optimal solution obtained from three stage stochastic programming is also evaluated by expected value formulation approach. The numerical result shows that using optimization approach of three stage stochastic programming can provide the minimum total cost of provisioning SDN bandwidth and cloud resources for data demand uncertainty.

The rest of this paper is organized as follows. Literature review is presented in Section 2. The proposed framework is described in Section 3. Then, the problem formulation with parameter setting and numerical results are discussed in Section 4. Finally, the paper is concluded in Section 5.

Objectives

The main objective of this research is to explore three-stage stochastic programming approach to minimize the total cost for software defined network and cloud resources provisioning for internet of vehicular network.

Literature Review

Recently, earlier researchers have studied the combination of cloud computing and internet of vehicles. For example, a novel multilayered vehicular data cloud platform was proposed in (Ashokkumar, Sam, Arshadprabhu, & Britto, 2015: 58-63) to solve the transportation issues. In (Bitam & Mellouk, 2012: 1-6), a vehicular cloud architecture called ITS-Cloud was proposed to improve vehicle-to-vehicle communication and road safety. The authors in (Sasikala & Deepti, 2013: 8-14) proposed vehicle cloud architecture to provide services for road traffic monitoring and healthcare monitoring of drivers. The article (Meneguette, 2016: 1-11) proposed a vehicular cloud-based framework to provide intelligent transportation management for big cities.

Based on the aforementioned papers, cloud computing has integrated with internet of vehicles because it can provide benefits of increased amount of storage and computational capacities to provide services for data analytics, backup, database management. Moreover, data stored in cloud can be retrieved from anytime, anywhere, and shared among the vehicles through internet connected devices. However, critical issues of bandwidth constraint and bottleneck need to be taken into account when a number of sensors from vehicles transfer data to the cloud services over long range communication.

Dealing with this issue, the concept of software defined network approach (SDN) can be applied to achieve better network performance. This network technology can not only provide large amount of bandwidth but also allows application developers to manage network path programmatically for traffic management, congestion control, and packet routing control. Earlier research articles also applied SDN technology for cloud-based vehicular network. In (Wang, Gao, & Zhu, 2019: 1819-1828), cloud enabled software defined vehicular network architecture is developed to improve network performance. The authors in (Truong, Lee, & Ghamri-Doudane, 2015: 1202-1207) proposed SDN-based VANET architecture to provide efficient network management for cloud based vehicular network. Software defined network is introduced in (Nkenyereye, Nkenyereye, Tama, Reddy, & Song, 2020: 1-26) to improve connectivity and bandwidth through programmable network for vehicular cloud network.

However, although the earlier authors conducted research about the combination of software defined network and cloud computing for vehicular network, they did not consider the cost of SDN bandwidth and cloud resources provisioning. In practical scenario, SDN bandwidth has to be purchased from Internet Service Provider (ISP) and cloud resources of storage and processing power have to be purchased from public cloud providers. Therefore, it can be costly to apply the SDN bandwidth and cloud resources in real situation.

As a result, in this paper, we propose a cloud-based internet of vehicular network framework by applying SDN approach which can provide the flexible network management to minimize the network latency between vehicles and cloud computing. Then, we further considered minimizing the total cost for provisioning of SDN bandwidth and cloud resources to implement the proposed framework. In this case, it is not easy to provision exactly the amount of SDN bandwidth and cloud resources because the sensing data from vehicles is in uncertainty. Therefore, we applied three stage stochastic programming which is a powerful optimization technique to solve the SDN bandwidth and cloud resources provisioning under uncertainty of data demand from vehicles in order to minimize the total cost of proposed framework for long term plan. As far as we are aware, none of the research works has applied three stage stochastic programming approach for SDN bandwidth and cloud resources provisioning in the field of internet of vehicular network.

Cloud-based Internet of Vehicular Network Framework

In this section, the proposed cloud-based internet of vehicular network framework is presented as shown in Fig. 1. In the framework, the bottom layer is Vehicular Ad Hoc Network (VANET) which is composed of vehicles and traffic lights installed with buffer. In VANET, vehicles are equipped with wireless sensors in order to collect road data such as traffic jam, accident detection, post-accident investigation, and intersection collision, respectively. Then, the collection of road data from sensors is sent to the traffic lights with buffers. The buffer from each traffic light sends the data from vehicles to the cloud for further storing and processing to provide useful information. Then, the vehicles can receive the information from the cloud via internet connected devices and use that information for transport safety, traffic management, convenient driving, etc. However, several network related problems can arise when real-time road data from traffic lights with buffers are uploaded to the cloud.

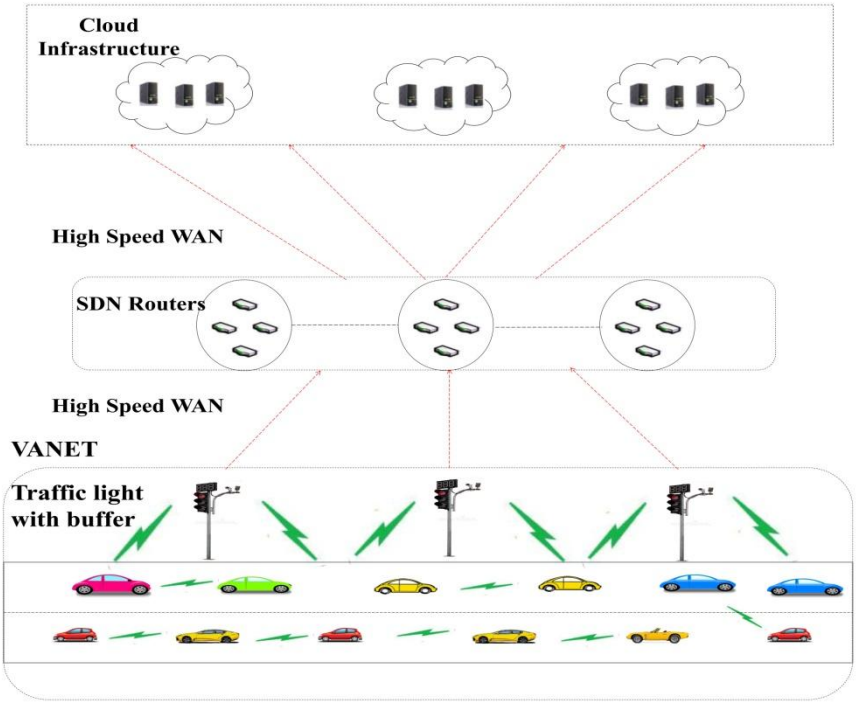


Figure 1: Cloud-based internet of vehicular network framework

Due to limited wireless bandwidth by sensors from vehicles, undesirable network latency can be experienced for uploading data to the cloud. In addition, a serious network bottleneck can be encountered when numerous vehicles are connected to the cloud in a parallel manner. For this issue, software defined network (SDN) approach is applied between traffic lights with buffers and cloud. Specifically, SDN is an emerging network technique which can provide high speed WAN and also allows controller to write programmable network management for traffic control.

Since the framework is organized by three models, i.e., traffic lights with buffers, SDN bandwidth, and cloud resources, we also proposed optimization approach of three stage stochastic programming to minimize the total cost for implementing the framework. SDN bandwidth has to be purchased from ISP provider and cloud resources have to be purchased from public cloud providers. Both ISP provider and cloud providers can provide two options, i.e., reservation and on-demand. With on-demand option, SDN bandwidth and cloud resources can be provisioned anytime but it is expensive and can have significant impact on the monetary cost. Otherwise, the reservation plan can be used by subscribing the certain amount of SDN bandwidth and cloud resources with cheaper price in advance. However, since the vehicles are mobile in nature, the data demand gathered from vehicles is unpredictable. As a result, the reservation plan with advance resource subscription may not be able to meet the data demand and result in underprovisioning and overprovisioning. The underprovisioning happens when the reserved amount is insufficient for data demand so that the on-demand option must be used for the required amount and the corresponding cost leads to higher on-demand cost. In contrast, the overprovisioning happens when the reserved amount is exceeded than the data demand and incurs to oversubscribed cost. Hence, it is untrivial to provision the required amount of SDN bandwidth and cloud resources to adjust between reservation plan and on-demand plan. To address this problem, three-stage stochastic programming is applied to minimize total provisioning cost of SDN bandwidth and cloud resources.

Stochastic Programming Model with Three Stages to Provisioning of SDN Bandwidth and Cloud Resources

In this section, the stochastic programming with three stages is presented in order to achieve optimal provisioning of SDN bandwidth and cloud resources for longer plan. Fig. 2 depicts the three stage scenario tree for provisioning of SDN bandwidth and cloud resources.

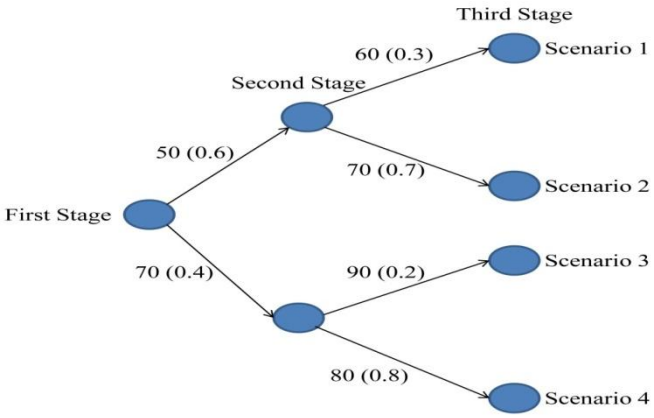


Figure 2: Three stage scenario tree for provisioning of SDN bandwidth and cloud resources

In Fig. 2, it is considered two periods for provisioning SDN bandwidth and cloud resources. At the first stage, the decision is made about how much the amount of SDN bandwidth and cloud resources needs to be provisioned in advance for the first period (e.g., 1 year) before the data demand is known. Here, for simplicity, we just consider two possible realizations of the data demand at the second stage and third stage. At the second stage, data demand will be realized and then decision will be made how much the amount of SDN bandwidth and cloud resources needs to be provisioned for the data demand in the next period (e.g., 1 year). At the third stage, one of the two possible realizations of data demand will be observed depending upon the realization at the second stage. Then, the provisioning phases for SDN bandwidth and cloud resources are illustrated in Fig. 3.

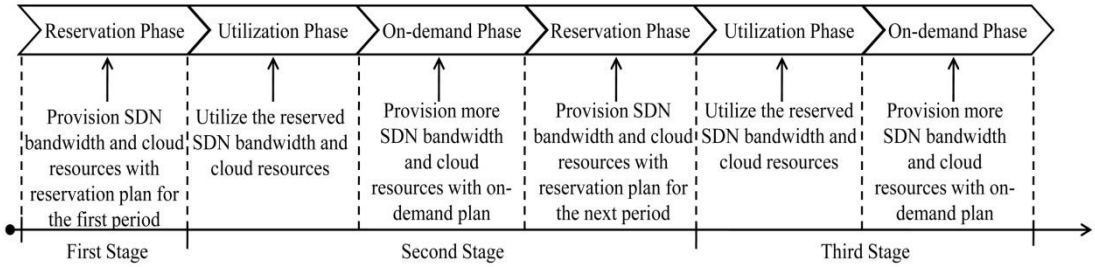


Figure 3: Provisioning phases for SDN bandwidth and cloud resources

At the first stage, the amount of SDN bandwidth and cloud resources is provisioned with reservation plan for the first period before the data demand is known. At the second stage, the data demand will be realized. Based on realized data demand, the reserved amount of SDN bandwidth and cloud resources are utilized. If the data demand exceeds the reserved amount, the on-demand phase starts to provision the required amount to meet the data demand. Then, the amount of SDN bandwidth and cloud resources will be provisioned with reservation plan before the data demand for the next

period is known. At the third stage, the realization of data demand will be observed. According to data demand, the reserved amount of SDN bandwidth and cloud resources will be utilized. If the reserved amount is insufficient and cannot fully meet the data demand, the required amount needs to be provisioned with on-demand plan. Since it is considered reservation phase, utilization phase, and on-demand phase in Fig. 3, there are three types of reservation cost, utilization cost, and on-demand cost for SDN bandwidth and cloud resources provisioning. To minimize total provisioning costs in three stages, the stochastic programming with three stages for provisioning of SDN bandwidth and cloud resources is formulated as follow.

Minimize:

$$\sum_{r\in R} C_r^{(re)} x_{rt_1}^{(re)} + \sum_{a\in A}\sum_{j\in J} C_{aj}^{(re)} y_{ajt_1}^{(re)} + \sum_{\omega\in\Omega}\sum_{p\in P} p(\omega)\left\{\left[\sum_{r\in R}\left(C_r^{(u)} x_{rt_2}^{(u)}(\omega) + C_r^{(o)} x_{rt_2}^{(o)}(\omega)\right) + \sum_{a\in A}\sum_{j\in J}\left(C_{aj}^{(u)} y_{ajt_2}^{(u)}(\omega) + C_{aj}^{(o)} y_{ajt_2}^{(o)}(\omega)\right)\right] + \left[\sum_{r\in R}\left(C_r^{(re)} x_{rt_2}^{(re)}(\omega)\right) + \sum_{a\in A}\sum_{j\in J}\left(C_{aj}^{(re)} y_{ajt_2}^{(re)}(\omega)\right)\right] + \left[\sum_{r\in R}\left(C_r^{(u)} x_{rt_3}^{(u)}(\omega) + C_r^{(o)} x_{rt_3}^{(o)}(\omega)\right) + \sum_{a\in A}\sum_{j\in J}\left(C_{aj}^{(u)} y_{ajt_3}^{(u)}(\omega) + C_{aj}^{(o)} y_{ajt_3}^{(o)}(\omega)\right)\right]\right\} \quad (1)$$

Subject to:

$$x_{rt_2}^{(u)}(\omega) \le x_{rt_1}^{(re)}, \forall r \in R, \omega \in \Omega \quad (2)$$

$$x_{rt_3}^{(u)}(\omega) \le x_{rt_2}^{(re)}(\omega), \forall r \in R, \omega \in \Omega \quad (3)$$

$$y_{ajt_2}^{(u)}(\omega) \le y_{ajt_1}^{(re)}, \forall a \in A, j \in J, \omega \in \Omega \quad (4)$$

$$y_{ajt_3}^{(u)}(\omega) \le y_{ajt_2}^{(re)}(\omega), \forall a \in A, j \in J, \omega \in \Omega \quad (5)$$

$$\sum_{j\in J}\left(y_{ajt_2}^{(u)}(\omega) + y_{ajt_2}^{(o)}(\omega)\right) \ge D_1(\omega), \forall a \in A, \omega \in \Omega \quad (6)$$

$$\sum_{j\in J}\left(y_{ajt_2}^{(u)}(\omega) + y_{ajt_2}^{(o)}(\omega) - D_1(\omega) + y_{ajt_3}^{(u)}(\omega) + y_{ajt_3}^{(o)}(\omega)\right) \ge D_2(\omega), \forall a \in A, \omega \in \Omega \quad (7)$$

$$\sum_{a\in A}\left(G_a^{(h)}\left(y_{ajt_2}^{(u)}(\omega) + y_{ajt_2}^{(o)}(\omega)\right)\right) \le E_j^{(h)}, \forall j \in J, \omega \in \Omega \quad (8)$$

$$\sum_{a\in A}\left(G_a^{(h)}\left(y_{ajt_3}^{(u)}(\omega) + y_{ajt_3}^{(o)}(\omega)\right)\right) \le E_j^{(h)}, \forall j \in J, \omega \in \Omega \quad (9)$$

$$\sum_{a\in A}\left(G_a^{(s)}\left(y_{ajt_2}^{(u)}(\omega) + y_{ajt_2}^{(o)}(\omega)\right)\right) \le E_j^{(s)}, \forall j \in J, \omega \in \Omega \quad (10)$$

$$\sum_{a\in A}\left(G_a^{(s)}\left(y_{ajt_3}^{(u)}(\omega) + y_{ajt_3}^{(o)}(\omega)\right)\right) \le E_j^{(s)}, \forall j \in J, \omega \in \Omega \quad (11)$$

$$x_{rt_2}^{(u)}(\omega) + x_{rt_2}^{(o)}(\omega) \le M_r, \forall r \in R, \omega \in \Omega \quad (12)$$

$$x_{rt_3}^{(u)}(\omega) + x_{rt_3}^{(o)}(\omega) \le M_r, \forall r \in R, \omega \in \Omega \quad (13)$$

$$\left(x_{rt_2}^{(u)}(\omega) + x_{rt_2}^{(o)}(\omega)\right) \ge \sum_{j\in J}\left(y_{ajt_2}^{(u)}(\omega) + y_{ajt_2}^{(o)}(\omega)\right), \forall r \in R, a \in A, \omega \in \Omega \quad (14)$$

$$\left(x_{rt_3}^{(u)}(\omega) + x_{rt_3}^{(o)}(\omega)\right) \ge \sum_{j\in J}\left(y_{ajt_3}^{(u)}(\omega) + y_{ajt_3}^{(o)}(\omega)\right), \forall r \in R, a \in A, \omega \in \Omega \quad (15)$$

$$\left(x_{rt_2}^{(u)}(\omega) + x_{rt_2}^{(o)}(\omega)\right) \ge \sum_{a\in A}\left(y_{ajt_2}^{(u)}(\omega) + y_{ajt_2}^{(o)}(\omega)\right), \forall r \in R, j \in J, \omega \in \Omega \quad (16)$$

$$\left(x_{rt_3}^{(u)}(\omega) + x_{rt_3}^{(o)}(\omega)\right) \ge \sum_{a\in A}\left(y_{ajt_3}^{(u)}(\omega) + y_{ajt_3}^{(o)}(\omega)\right), \forall r \in R, j \in J, \omega \in \Omega \quad (17)$$

$$x_{rt_1}^{(re)}, x_{rt_2}^{(re)}(\omega), x_{rt_2}^{(u)}(\omega), x_{rt_2}^{(o)}(\omega), x_{rt_3}^{(u)}(\omega), x_{rt_3}^{(o)}(\omega) > 0, \forall r \in R, \omega \in \Omega \quad (18)$$

$$y_{ajt_1}^{(re)}, y_{ajt_2}^{(re)}(\omega), y_{ajt_2}^{(u)}(\omega), y_{ajt_2}^{(o)}(\omega), y_{ajt_3}^{(u)}(\omega), y_{ajt_3}^{(o)}(\omega) \in \{0,1,\dots\}, \forall a \in A, j \in J, \omega \in \Omega \quad (19)$$

The objective function in (1) is to minimize the reservation cost, utilization cost, and on-demand cost of SDN bandwidth (i.e., $C_r^{(re)}$, $C_r^{(u)}$, and $C_r^{(o)}$) and cloud resources (i.e., $C_{aj}^{(re)}$, $C_{aj}^{(u)}$, and $C_{aj}^{(o)}$) in three decision stages. Let decision variables $x_{rt_1}^{(re)}$ and $y_{ajt_1}^{(re)}$ denote the reservation amount of SDN bandwidth and cloud resources at the first stage. Let decision variables $x_{rt_2}^{(u)}(\omega)$, $x_{rt_2}^{(o)}(\omega)$, $y_{ajt_2}^{(u)}(\omega)$, and $y_{ajt_2}^{(o)}(\omega)$ denote the utilization and on-demand amount of SDN bandwidth and cloud resources at the second stage. Let decision variables $x_{rt_2}^{(re)}(\omega)$ and $y_{ajt_2}^{(re)}(\omega)$ denote the reservation amount of SDN bandwidth and cloud resources at the second stage. Let decision variables $x_{rt_3}^{(u)}(\omega)$, $x_{rt_3}^{(o)}(\omega)$, $y_{ajt_3}^{(u)}(\omega)$, and $y_{ajt_3}^{(o)}(\omega)$ denote the utilization and on-demand amount of SDN bandwidth and cloud resources at the third stage. Let $\omega \in \Omega$ denotes a set of scenarios and $p \in P$ denotes a set of probabilities. Let $p(\omega)$ denotes the probability of data demand at a certain scenario will be realized.

The constraint in (2) ensures that the utilization amount of SDN bandwidth at the second stage must not exceed the reserved amount of SDN bandwidth at the first stage while the constraint in (3) ensures that the utilization amount of SDN bandwidth at the third stage must not exceed the reserved amount of SDN bandwidth at the second stage. The constraint in (4) ensures that the utilization amount of cloud resources at the second stage must not exceed the reserved amount of cloud resources at the first stage while the constraint in (5) ensures that the utilization amount of cloud resources at the third stage must not exceed the reserved amount of cloud resources at the second stage.

The constraint in (6) ensures that the data demand for the first period (i.e., $D_1(\omega)$) is met at the second stage and the constraint in (7) ensures that the data demand for the second perioed (i.e., $D_2(\omega)$) is met at the third stage. In (7), the left over units from the constraint in (6) will help make decision to meet future data demand for the second period at the third stage. The constraint in (8) ensures that the utilization and on-demand amount of processing power required by data demand at the second stage must not exceed the maximum capacity of processing power offered by each cloud provider. The constraint in (9) ensures that the utilization and on-demand amount of processing power required by data demand at the third stage must not exceed the maximum capacity of processing power offered by each cloud provider. The constraint in (10) ensures that the utilization and on-demand amount of storage capacity required by data demand at the second stage must not exceed the maximum capacity of storage offered by each cloud provider. The constraint in (11) ensures that utilization and on-demand amount of storage capacity required by data demand at the third stage must not exceed the maximum capacity of storage offered by each cloud provider. The constraint in (12) ensures that the utilization and on-demand amount of network bandwidth at the second stage must not exceed the maximum bandwidth offered by each router. The constraint in (13) ensures that the utilization and on-demand amount of network bandwidth at the third stage must not exceed the maximum bandwidth offered by each router.

The constraint in (14) ensures that the total bandwidth supply is provided to the traffic lights with buffers by the routers at the second stage. The constraint in (15) ensures that the total bandwidth supply is provided to the traffic lights with buffers by the routers at the third stage. The constraint in (16) ensures that the total bandwidth supply is provided to the cloud providers by the routers at the second stage. The constraint in (17) ensures that the total bandwidth supply is provided to the cloud providers by the routers at the third stage.

The constraint in (18) indicates that the decision variables take positive non-integer amount of SDN bandwidth. The constraint in (19) mentions that the decision variables take positive integer number of data packets from traffic lights with buffers allocated to cloud providers.

In order to solve the formulations in (1)-(19), the parameters are defined below. The assumption of some parameters is defined as follow. It is assumed that there are three traffic lights with buffers, denoted by $|R|$= 3. Then, each traffic light with buffer collects the data from sensors, packages the same size, and transfers to 3 cloud providers (i.e., $|J|$ = 3). The same size of each data packet is 20 (GB) and the same processing power of each data packet is 15 (CPU-hours). The maximum amount of cloud resources for storage and processing power from each cloud provider is assumed as unlimited. The same maximum capacity of bandwidth offered by routers r_1 to r_3 is assumed as 1250 (MB/sec) (Singtel, 1879).

Table 1 describes the prices of cloud resources per day in each provisioning plan defined by cloud providers (Chaisiri, Lee, & Niyato, 2009: 103-110). Table 2 describes the prices of SDN bandwidth through routers per day in each provisioning plan defined by ISP provider (Chase, Kaewpuang, Yonggang, & Niyato, 2014: 2969-2974).

Table 1: Prices of cloud resources per day in each provisioning plan defined by cloud providers

Provider	Reservation		Utilization		On-demand	
	Processing power	Storage	Processing power	Storage	Processing power	Storage
j_1	0.20	0.30	0.10	0.20	0.40	0.60
j_2	0.20	0.30	0.10	0.20	0.50	0.70
j_3	0.10	0.20	0.10	0.10	0.50	0.70
j_4	0.10	0.20	0.10	0.10	0.60	0.80

Table 2: Prices of SDN bandwidth through routers per day in each provisioning plan defined by ISP provider

Routers	Reservation	Utilization	On-demand
r_1	0.17	0.02	0.30
r_2	0.17	0.02	0.30
r_3	0.17	0.03	0.50

In Fig. 2, there are four possible scenarios. To know the probability of each scenario, it can be calculated below.

Scenario	Second stage	Third stage	Probability
ω_1	50	60	0.6 (0.3) = 0.18 (p_1)
ω_2	50	70	0.6 (0.7) = 0.42 (p_2)
ω_3	70	90	0.4 (0.2) = 0.08 (p_3)
ω_4	70	80	0.4 (0.8) = 0.32 (p_4)

Among those probabilities, one of them will be realized for data demand at a certain scenario. Then, the formulations in (1)-(19) are submitted in GAMS/CPLEX solver (GAMS Solver, 1987) and generated the numerical results below.

Table 3: Numerical result for minimum total cost of cloud resource provisioning and SDN bandwidth provisioning by stochastic programming three stages

Decision variables for cloud resource provisioning and SDN bandwidth provisioning											
First Stage Reservation		Second Stage						Third Stage			
		Utilization	On-demand	Utilization	On-demand	Reservation	Reservation	Utilization	On-demand	Utilization	On-demand
Cloud	SDN	Cloud	Cloud	SDN	SDN	Cloud	SDN	Cloud	Cloud	SDN	SDN
$y_{ajt_1}^{(re)}$	$x_{rt_1}^{(re)}$	$y_{ajt_2}^{(u)}(\omega)$	$y_{ajt_2}^{(o)}(\omega)$	$x_{rt_2}^{(u)}(\omega)$	$x_{rt_2}^{(o)}(\omega)$	$y_{ajt_2}^{(re)}(\omega)$	$x_{rt_2}^{(re)}(\omega)$	$y_{ajt_3}^{(u)}(\omega)$	$y_{ajt_3}^{(o)}(\omega)$	$x_{rt_3}^{(u)}(\omega)$	$x_{rt_3}^{(o)}(\omega)$
$y_{a_1j_1t_1}^{(re)}$ 70	$x_{r_1t_1}^{(re)}$ 100	$y_{a_1j_1t_2}^{(u)}(\omega)$ 70	$y_{a_1j_1t_2}^{(o)}(\omega)$ 0	$x_{r_1t_2}^{(u)}(\omega)$ 100	$x_{r_1t_2}^{(o)}(\omega)$ 0	$y_{a_1j_1t_2}^{(re)}(\omega)$ 0	$x_{r_1t_2}^{(re)}(\omega)$ 100	$y_{a_1j_1t_3}^{(u)}(\omega)$ 0	$y_{a_1j_1t_3}^{(o)}(\omega)$ 10	$x_{r_1t_3}^{(u)}(\omega)$ 100	$x_{r_1t_3}^{(o)}(\omega)$ 0
$y_{a_1j_2t_1}^{(re)}$ 0	$x_{r_2t_1}^{(re)}$ 100	$y_{a_1j_2t_2}^{(u)}(\omega)$ 0	$y_{a_1j_2t_2}^{(o)}(\omega)$ 0	$x_{r_2t_2}^{(u)}(\omega)$ 100	$x_{r_2t_2}^{(o)}(\omega)$ 0	$y_{a_1j_2t_2}^{(re)}(\omega)$ 0	$x_{r_2t_2}^{(re)}(\omega)$ 100	$y_{a_1j_2t_3}^{(u)}(\omega)$ 0	$y_{a_1j_2t_3}^{(o)}(\omega)$ 0	$x_{r_2t_3}^{(u)}(\omega)$ 100	$x_{r_2t_3}^{(o)}(\omega)$ 0
$y_{a_1j_3t_1}^{(re)}$ 0	$x_{r_3t_1}^{(re)}$ 100	$y_{a_1j_3t_2}^{(u)}(\omega)$ 0	$y_{a_1j_3t_2}^{(o)}(\omega)$ 0	$x_{r_3t_2}^{(u)}(\omega)$ 100	$x_{r_3t_2}^{(o)}(\omega)$ 0	$y_{a_1j_3t_2}^{(re)}(\omega)$ 80	$x_{r_3t_2}^{(re)}(\omega)$ 100	$y_{a_1j_3t_3}^{(u)}(\omega)$ 80	$y_{a_1j_3t_3}^{(o)}(\omega)$ 0	$x_{r_3t_3}^{(u)}(\omega)$ 100	$x_{r_3t_3}^{(o)}(\omega)$ 0
$y_{a_2j_1t_1}^{(re)}$ 30		$y_{a_2j_1t_2}^{(u)}(\omega)$ 30	$y_{a_2j_1t_2}^{(o)}(\omega)$ 0			$y_{a_2j_1t_2}^{(re)}(\omega)$ 0		$y_{a_2j_1t_3}^{(u)}(\omega)$ 0	$y_{a_2j_1t_3}^{(o)}(\omega)$ 10		
$y_{a_2j_2t_1}^{(re)}$ 10		$y_{a_2j_2t_2}^{(u)}(\omega)$ 10	$y_{a_2j_2t_2}^{(o)}(\omega)$ 0			$y_{a_2j_2t_2}^{(re)}(\omega)$ 80		$y_{a_2j_2t_3}^{(u)}(\omega)$ 80	$y_{a_2j_2t_3}^{(o)}(\omega)$ 0		
$y_{a_2j_3t_1}^{(re)}$ 30		$y_{a_2j_3t_2}^{(u)}(\omega)$ 30	$y_{a_2j_3t_2}^{(o)}(\omega)$ 0			$y_{a_2j_3t_2}^{(re)}(\omega)$ 0		$y_{a_2j_3t_3}^{(u)}(\omega)$ 0	$y_{a_2j_3t_3}^{(o)}(\omega)$ 0		
$y_{a_3j_1t_1}^{(re)}$ 0		$y_{a_3j_1t_2}^{(u)}(\omega)$ 0	$y_{a_3j_1t_2}^{(o)}(\omega)$ 0			$y_{a_3j_1t_2}^{(re)}(\omega)$ 60		$y_{a_3j_1t_3}^{(u)}(\omega)$ 60	$y_{a_3j_1t_3}^{(o)}(\omega)$ 10		
$y_{a_3j_2t_1}^{(re)}$ 0		$y_{a_3j_2t_2}^{(u)}(\omega)$ 0	$y_{a_3j_2t_2}^{(o)}(\omega)$ 0			$y_{a_3j_2t_2}^{(re)}(\omega)$ 0		$y_{a_3j_2t_3}^{(u)}(\omega)$ 0	$y_{a_3j_2t_3}^{(o)}(\omega)$ 0		
$y_{a_3j_3t_1}^{(re)}$ 70		$y_{a_3j_3t_2}^{(u)}(\omega)$ 70	$y_{a_3j_3t_2}^{(o)}(\omega)$ 0			$y_{a_3j_3t_2}^{(re)}(\omega)$ 20		$y_{a_3j_3t_3}^{(u)}(\omega)$ 20	$y_{a_3j_3t_3}^{(o)}(\omega)$ 0		
Total Cost = $4767											

Table 3 shows the numerical result for minimum total cost of cloud resource provisioning and SDN bandwidth provisioning by stochastic programming three stages. According to the numerical results, it can be analyzed that the probability 3 for data demand at scenario 3 (i.e., $p_3(\omega_3)$) is realized. In Table 3, at the first stage of cloud resource reservation, traffic lights with buffers a_1 reserves 70 data packets in cloud provider j_1; traffic lights with buffers a_2 reserves 30 data packets in cloud provider j_1, 10 data packets in cloud provider j_2, 30 data packets in cloud provider j_3; and traffic lights with buffers a_3 reserves 70 data packets in cloud provider j_3. At the first stage of SDN bandwidth reservation, 100 (MB/sec) is reserved in each router r_1, r_2, and r_3.

Since each traffic light with buffer reserves 70 data packets in cloud providers at the first stage, the data demand at the second stage for scenario 3 (i.e. 70) is met. Therefore, at the second stage of cloud resource utilization, the exact number of 70 data packets reserved by each traffic light with buffer in each cloud provider is utilized and at the second stage of cloud resource on-demand, there is no extra number of data packets provisioned. At the second stage of SDN bandwidth utilization, the reserved amount of 100 MB/sec bandwidth is utilized in each router and at the second stage of SDN bandwidth on-demand, there is no additional amount of bandwidth provisioned. Then, at the second stage of cloud resource reservation, traffic light with buffer a_1 reserves 80 data packets in cloud provider j_3; traffic light with buffer a_2 reserves 80 data packets in cloud provider j_2; and traffic light with buffer a_3 reserves 60 data packets in cloud provider j_1 and 20 data packets in cloud provider j_3. At the second stage of SDN bandwidth reservation, 100 (MB/sec) is reserved in each router r_1, r_2, and r_3.

Since each traffic light with buffer reserves 80 data packets in cloud providers, the data demand at the third stage for scenario 3 (i.e. 90) is not met. Therefore, at the third stage of cloud resource utilization, the exact number of 80 data packets reserved by each traffic light with buffer in each cloud provider is utilized and at the third stage of cloud resource on-demand, 10 more data packets from each traffic light with buffer are provisioned. At the third stage of SDN bandwidth utilization, the reserved amount of 100 MB/sec bandwidth is utilized in each router and at the third stage of SDN bandwidth on-demand, there is no additional amount of bandwidth provisioned.

Finally, the minimum total cost of cloud resource provisioning and SDN bandwidth provisioning in three stages that is achieved as $4767.

Expected Value Formulation Approach

In this section, the expected value formulation (EVF) approach is presented to evaluate the total cost obtained in stochastic programming with three stages. In this approach, it calculates the average value of data demands at the second stage (i.e., (50+70)/2 = 60) and applies in the data demand from each traffic light with buffer (i.e., B_a) as defined in deterministic integer program. Then, the deterministic integer program formulations in (5.25)-(5.33) are solved by GAMS/CPLEX solver. Next, the optimal solution for SDN bandwidth provisioning and cloud resource provisioning obtained by deterministic integer program is fixed to the reservation phase of SDN bandwidth provisioning and cloud resource provisioning at the first stage in stochastic programming with three stages. Finally, the numerical result for minimum total cost of cloud resource provisioning and SDN bandwidth provisioning by expected value formulation approach is shown in Table 4 on the next page.

In Table 4, at the first stage of cloud resource reservation, traffic light with buffer a_1 reserves 60 data packets in cloud provider j_1; traffic light with buffer a_2 reserves 60 data packets in cloud provider j_3; and traffic light with buffer a_3 reserves 20 data packets in cloud provider j_1 and 40 data packets in cloud provider j_3. At the first stage of SDN bandwidth reservation, 100 (MB/sec) is reserved in each router r_1, r_2, and r_3.

Since each traffic light with buffer reserves 60 data packets in cloud providers at the first stage, the data demand at the second stage for scenario 3 (i.e. 70) is not met. Therefore, at the second stage of cloud resource utilization, the exact number of 60 data packets reserved by each traffic light with buffer in each cloud provider is utilized and at the second stage of cloud resource on-demand, traffic light with buffer a_1 allocates 10 data packets in in cloud provider j_2, and traffic light with buffer a_2 and a_3 allocate 10 data packets in cloud provider j_1. At the second stage of SDN bandwidth utilization, the reserved amount of 100 MB/sec bandwidth is utilized in each router and at the second stage of SDN bandwidth on-demand, there is no additional amount of

Table 4: Numerical result for minimum total cost of cloud resource provisioning and SDN bandwidth provisioning by expected value formulation approach

Decision variables for cloud resource provisioning and SDN bandwidth provisioning											
First Stage (Reservation)		Second Stage						Third Stage			
		Utilization	On-demand	Utilization	On-demand	Reservation	Reservation	Utilization	On-demand	Utilization	On-demand
Cloud	SDN	Cloud	Cloud	SDN	SDN	Cloud	SDN	Cloud	Cloud	SDN	SDN
$y_{aj t_1}^{(re)}$	$x_{rt_1}^{(re)}$	$y_{aj t_2}^{(u)}(\omega)$	$y_{aj t_2}^{(o)}(\omega)$	$x_{rt_2}^{(u)}(\omega)$	$x_{rt_2}^{(o)}(\omega)$	$y_{aj t_2}^{(re)}(\omega)$	$x_{rt_2}^{(re)}(\omega)$	$y_{aj t_3}^{(u)}(\omega)$	$y_{aj t_3}^{(o)}(\omega)$	$x_{rt_3}^{(u)}(\omega)$	$x_{rt_3}^{(o)}(\omega)$
$y_{a_1 j_1 t_1}^{(re)}$ 60	$x_{r_1 t_1}^{(re)}$ 100	$y_{a_1 j_1 t_2}^{(u)}(\omega)$ 60	$y_{a_1 j_1 t_2}^{(o)}(\omega)$ 0	$x_{r_1 t_2}^{(u)}(\omega)$ 100	$x_{r_1 t_2}^{(o)}(\omega)$ 0	$y_{a_1 j_1 t_2}^{(re)}(\omega)$ 70	$x_{r_1 t_2}^{(re)}(\omega)$ 100	$y_{a_1 j_1 t_3}^{(u)}(\omega)$ 70	$y_{a_1 j_1 t_3}^{(o)}(\omega)$ 10	$x_{r_1 t_3}^{(u)}(\omega)$ 100	$x_{r_1 t_3}^{(o)}(\omega)$ 0
$y_{a_1 j_2 t_1}^{(re)}$ 0	$x_{r_2 t_1}^{(re)}$ 100	$y_{a_1 j_2 t_2}^{(u)}(\omega)$ 0	$y_{a_1 j_2 t_2}^{(o)}(\omega)$ 10	$x_{r_2 t_2}^{(u)}(\omega)$ 100	$x_{r_2 t_2}^{(o)}(\omega)$ 0	$y_{a_1 j_2 t_2}^{(re)}(\omega)$ 10	$x_{r_2 t_2}^{(re)}(\omega)$ 100	$y_{a_1 j_2 t_3}^{(u)}(\omega)$ 10	$y_{a_1 j_2 t_3}^{(o)}(\omega)$ 0	$x_{r_2 t_3}^{(u)}(\omega)$ 100	$x_{r_2 t_3}^{(o)}(\omega)$ 0
$y_{a_1 j_3 t_1}^{(re)}$ 0	$x_{r_3 t_1}^{(re)}$ 100	$y_{a_1 j_3 t_2}^{(u)}(\omega)$ 0	$y_{a_1 j_3 t_2}^{(o)}(\omega)$ 0	$x_{r_3 t_2}^{(u)}(\omega)$ 100	$x_{r_3 t_2}^{(o)}(\omega)$ 0	$y_{a_1 j_3 t_2}^{(re)}(\omega)$ 0	$x_{r_3 t_2}^{(re)}(\omega)$ 100	$y_{a_1 j_3 t_3}^{(u)}(\omega)$ 0	$y_{a_1 j_3 t_3}^{(o)}(\omega)$ 0	$x_{r_3 t_3}^{(u)}(\omega)$ 100	$x_{r_3 t_3}^{(o)}(\omega)$ 0
$y_{a_2 j_1 t_1}^{(re)}$ 0		$y_{a_2 j_1 t_2}^{(u)}(\omega)$ 0	$y_{a_2 j_1 t_2}^{(o)}(\omega)$ 10			$y_{a_2 j_1 t_2}^{(re)}(\omega)$ 0		$y_{a_2 j_1 t_3}^{(u)}(\omega)$ 0	$y_{a_2 j_1 t_3}^{(o)}(\omega)$ 10		
$y_{a_2 j_2 t_1}^{(re)}$ 0		$y_{a_2 j_2 t_2}^{(u)}(\omega)$ 0	$y_{a_2 j_2 t_2}^{(o)}(\omega)$ 0			$y_{a_2 j_2 t_2}^{(re)}(\omega)$ 0		$y_{a_2 j_2 t_3}^{(u)}(\omega)$ 0	$y_{a_2 j_2 t_3}^{(o)}(\omega)$ 0		
$y_{a_2 j_3 t_1}^{(re)}$ 60		$y_{a_2 j_3 t_2}^{(u)}(\omega)$ 60	$y_{a_2 j_3 t_2}^{(o)}(\omega)$ 0			$y_{a_2 j_3 t_2}^{(re)}(\omega)$ 80		$y_{a_2 j_3 t_3}^{(u)}(\omega)$ 80	$y_{a_2 j_3 t_3}^{(o)}(\omega)$ 0		
$y_{a_3 j_1 t_1}^{(re)}$ 20		$y_{a_3 j_1 t_2}^{(u)}(\omega)$ 20	$y_{a_3 j_1 t_2}^{(o)}(\omega)$ 10			$y_{a_3 j_1 t_2}^{(re)}(\omega)$ 0		$y_{a_3 j_1 t_3}^{(u)}(\omega)$ 0	$y_{a_3 j_1 t_3}^{(o)}(\omega)$ 10		
$y_{a_3 j_2 t_1}^{(re)}$ 0		$y_{a_3 j_2 t_2}^{(u)}(\omega)$ 0	$y_{a_3 j_2 t_2}^{(o)}(\omega)$ 0			$y_{a_3 j_2 t_2}^{(re)}(\omega)$ 60		$y_{a_3 j_2 t_3}^{(u)}(\omega)$ 60	$y_{a_3 j_2 t_3}^{(o)}(\omega)$ 0		
$y_{a_3 j_3 t_1}^{(re)}$ 40		$y_{a_3 j_3 t_2}^{(u)}(\omega)$ 40	$y_{a_3 j_3 t_2}^{(o)}(\omega)$ 0			$y_{a_3 j_3 t_2}^{(re)}(\omega)$ 20		$y_{a_3 j_3 t_3}^{(u)}(\omega)$ 20	$y_{a_3 j_3 t_3}^{(o)}(\omega)$ 0		
Total Cost = \$4823											

bandwidth provisioned. Then, at the second stage of cloud resource reservation, traffic light with buffer a_1 reserves 70 data packets in cloud provider j_1 and 10 data packets in cloud provider j_2; traffic light with buffer a_2 reserves 80 data packets in cloud provider j_3; and traffic light with buffer a_3 reserves 60 data packets in cloud provider $j_2$and 20 data packets in cloud provider j_3. At the second stage of SDN bandwidth reservation, 100 (MB/sec) is reserved in each router r_1, r_2, and r_3.

Since each traffic light with buffer reserves 80 data packets in cloud providers, the data demand at the third stage for scenario 3 (i.e. 90) is not met. Therefore, at the third stage of cloud resource utilization, the exact number of 80 data packets reserved by each traffic light with buffer in each cloud provider is utilized and at the third stage of cloud resource on-demand, 10 more data packets from each traffic light with buffer are provisioned. At the third stage of SDN bandwidth utilization, the reserved amount of 100 MB/sec bandwidth is utilized in each router and at the third stage of SDN bandwidth on-demand, there is no additional amount of bandwidth provisioned.

Finally, the minimum total cost of cloud resource provisioning and SDN bandwidth provisioning in three stages that is achieved as \$4823. Compared to the total cost of stochastic programming with three stages, the total cost of EVF approach is higher than that. Since EVF approach approximates the average values of data packet demands, it cannot guarantee to achieve the minimum total cost.

Conclusion

Cloud computing has applied to a vehicular network in order to provide a large amount of resources such as storage and processing power for storing and processing data captured by sensors from vehicles. However, it can encounter the network latency and also bottleneck when numerous sensors with bandwidth limitation transfer data to the cloud. To overcome this problem, this paper proposed the cloud based internet of vehicular network framework using software-defined networking (SDN). The SDN approach provides flexible network management to improve the network connectivity between vehicles with sensors and cloud services. By utilizing SDN bandwidth, the network performance of the proposed system can be significantly upgraded. In addition, we also extended this work by developing three stage stochastic programming for the proposed system to minimize the total cost of provisioning SDN bandwidth and cloud resources under uncertainty of data demand from sensors. Moreover, expected value formulation approach is applied to evaluate the optimal solution obtained by three stage stochastic programming. Compared to the previous works, our proposed framework can be implemented in real scenario with the minimum total cost for optimal provisioning of SDN bandwidth and cloud resources under data demand uncertainty for internet ofs vehicular network.

For the future work, we will develop multiple stages of stochastic programming for provisioning of SDN bandwidth and cloud resources to plan for longer term.

References

Ashokkumar, K., Sam, B., Arshadprabhu, R., & Britto. (2015). Cloud based Intelligent Transport System. *Procedia Computer Science, 50*, 58-63.

Bitam, S., & Mellouk, A. (n.d.). ITS-Cloud: Cloud Computing for Intelligent Transportation System. *IEEE Global Communications Conference (GLOBECOM)* (pp. 1-6). Anaheim, California: IEEE.

Buyya, R., & Dastjerdi, A. V. (2016). *Internet of Things.* Burlington: Morgan Kaufmann.

Chaisiri, S., Lee, B.-S., & Niyato, D. (2009). Optimal Virtual Machine Placement across Multiple Cloud Providers. *IEEE Asia-Pacific Services Computing Conference (APSCC)* (pp. 103-110). Singapore: IEEE.

Chase, J., Kaewpuang, R., Yonggang, W., & Niyato, D. (2014). Joint Virtual Machine and Bandwidth Allocation in Software Defined Network (SDN) and Cloud Computing environments. *IEEE International Conference on Communications (ICC)* (pp. 2969-2974). Sydney: IEEE.

GAMS Solver. (1987). Retrieved April 29, 2020, from GAMS: https://www.gams.com

Meneguette, R. I. (2016). A Vehicular Cloud-based Framework for the Intelligent Transport Management of Big Cities . *International Journal of Distributed Sensor Networks*, 1-11.

Nkenyereye, L., Nkenyereye, L., Tama, B. A., Reddy, A. G., & Song, J. (2020). Software-Defined Vehicular Cloud Networks: Architecture, Applications and Virtual Machine Migration. *Sensors, 20*(4), 1-26.

RajkumarBuyya, Yeo, C. S., SrikumarVenugopal, Broberg, J., & Brandic, I. (2009). Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility. *Future Generation Computer Systems, 25*(6), 599-616.

Sadiku, M. N., Tembely, M., & Musa, S. M. (2018). Internet of Vehicles: An Introduction. *International Journal of Advanced Research in Computer Science and Software Engineering, 8*(1), 11-13.

Sasikala, G., & Deepti, A. (2013). Real Time Services for CloudComputing Enabled Vehicle Networks. *IOSR Journal of Computer Engineering (IOSR-JCE), 11*(1), 8-14.

SDN Bandwidth. (1879). Retrieved April 29, 2020, from Singtel: https://www.singtel.com/personal/i/internet/broadband-at-home/fibre-broadband-plans

Sperling, D., & Gordon, D. (2008). *Two Billion Cars Transforming a Culture*. Retrieved April 29, 2020, from http://www.trb.org/Environment/Blurbs/160606.aspx.

Truong, N. B., Lee, G. M., & Ghamri-Doudane, Y. (2015). Software Defined Networking Based Vehicular Adhoc Network with Fog Computing. *IFIP/IEEE International Symposium on Integrated Network Management (IM)*, 1202-1207.

Wang, Q., Gao, D., & Zhu, W. (2019). Cloud-enabled Software Defined Vehicular Networks: Architecture, Applications and Challenges. *Journal of Internet Technology, 20*(6), 1819-1828.

Xia, W., Wen, Y., Foh, C. H., Niyato, D., & Xie, H. (2014). A Survey on Software-Defined Networking. *IEEE Communications Surveys & Tutorials, 17*(1), 27-51.

Yang, F., Li, J., Lei, T., & Wang, S. (2017). Architecture and Key Technologies for Internet of Vehicles: A Survey. *Journal of Communications and Information Networks, 2*(2), 1-17.

基于 ARMA 模型的协作通信中继选择及功率分配算法
ARMA MODEL BASED RELAY SELECTION AND POWER ALLOCATION ALGORITHM IN COOPERATIVE COMMUNICATION

王树军 [1*]，徐平平 [2]，杨小琴 [3]
Shujun Wang[1*], Pingping Xu[2], Xiaoqin Yang[3]

南京工业大学浦江学院
Nanjing Tech University Pujiang Institute
[*]**Corresponding author, E-mail:** shujun.wang@qq.com

摘要

协作通信能够有效提升网络性能。对于AF网络，以优化网络能量效率为目标，提出一种中继选择算法，并在此基础上，提出源节点和中继节点的最优功率分配算法。利用某一时间段内各中继节点历史能量信息值通过ARMA模型对各节点下一时刻能量信息值预测，选择出最优的中继，再利用Lagrange Multiplier对源节点和选择的中继节点之间进行最优功率分配，以使系统满足需求的QoS。理论分析和仿真结果表明，同等条件下，与传统随机选择中继算法和功率分配方法相比，所提算法能够有效降低网络能量消耗，延长网络生命周期。

关键词：协作通信，中继选择，功率分配，能量效率

Abstract

Cooperative communication can effectively improve network performance. A relay selection and power allocation algorithm was also proposed in amplify-and-forward cooperative network aiming at optimizing network energy efficiency. In this scheme, based on historical energy information of each relay nodes in a certain period of time we adopt ARMA model to predict the next time energy information of each nodes, and to select the optimal relay set. Further, the Lagrange Multiplier is used to optimize the power allocation between source node and selected relay nodes, and hence to make the system meet the demands of QoS. Theoretical analysis and simulation results show that, the proposed relay selection and power allocation algorithm, under the same conditions, compared in the traditional relay selection and power allocation algorithm, effectively reduce network energy consumption and prolong the network life cycle.

Keywords: Cooperative Communication, Relay selection, Power allocation, Energy efficiency

引言

随着通信行业的快速发展，未来网络对于通信技术提出了更高的技术挑战，节约资源与提高效率成为了未来网络发展的主题。绿色通信在这种背景下应运而生。绿色通信(Green Communication)指节能减排、减少环境污染和资源浪费以及对人体和环境危害的新一代通信理念，主要采用创新的高效功效、多载波、分布式、智能温控等技术，配合灵活的站点场景模型，对基站进行积极改造，以达到降低能耗的目的，最终实现人于自然和谐相处，实现可持续发展。传统的无线网络由于其所处外界环境或自身设备硬件条件等限制，引起信道衰落大、系统功耗大和接收信号质量差等弊端，协作通信技术可以充分利用无线设备终端在收发端建立多条通信

链路，通过传统多天线技术中的空间分集克服无线信道的衰落，从而提高无线网络系统的性能和鲁棒性[1~4]。

协作通信技术还可以在不影响用户服务质量的前提下有效减少网络系统功率能量消耗，提高能量利用效率[5]。文献[6][7]研究了低复杂度情况下的经典协作分集协议，有效抵抗无线网络中多径引起的信道衰落，并分析了高信噪比条件下的中断概率性能。Hu等人研究在多个天线分散覆盖整个小区的情况下，用户就近接入基站，通过分散天线间的协作，提高频谱效率，减少发射功率，从而提高功率效率[8]。文献[9]首次研究机会中继及在此基础上的两种中继选择方法，在不知道网络拓扑情况下，机会中继方案通过利用发收两端之间传输准备发送(RTS，Ready to send)和清除发送(CTS，Clear to send)分组同时对信道的状态信息进行检测。Si等人[10]在DF协作协议的基础上，提出了一种中继选择和功率分配的算法，该算法要求系统满足一定的中断概率，从而最小化系统传输所需要的总功率。而对于采用AF (Amplify and Forward)协作协议的中继选择和功率分配算法，相关的研究还比较少。

本文在上述研究的基础上，针对AF协作通信网络，提出一种基于ARMA模型的中继选择和功率分配算法。所提算法要求系统消耗功率最小，降低系统传输所需要的总功率。第2节建立网络系统模型。第3节详细描述ARMA能量预测模型下的中继选择和功率分配算法。第4节对所提算法与普通中继选择和功率分配的算法进行系统能耗的对比，并对结果进行分析。第5节对全文进行总结并指出进一步的研究方向。

系统模型

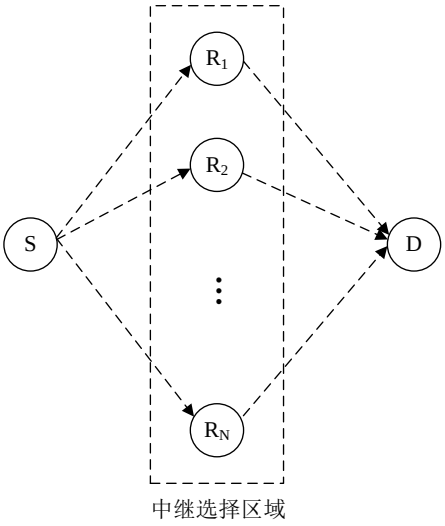


图1: 协作通信网络模型

协作网络通信模型如图1所示。网络中S为源节点，中继节点为集合 $\Re$ ， $\Re=\{R_i\}$ ，其中i={1,2,...,n}，D为目的节点。网络中中继节点可以预先部署或者为暂时无信息需要发送的空闲节点。当源S需要中继传输信息时，需要从中继选择区域选择中继节点中继其所要发送的信息至目的节点D。传输过程可以分为两个阶段进行：第I阶段，源在正交信道(例如TDMA信道)向所有中继节点传输所要发送的信息；第II阶段，完成中继选择后，中继协助源节点向目的地转发信息。目的节点将接收到的信息进行信号处理。第I阶段中继i，($i \in R$)接收到的信号和第II阶段目的地接收到中继发出的信号分别为

$$y_{si} = \sqrt{P_s} h_{si} x + n_{si},\ \mathrm{i} \in \mathrm{R} \tag{1}$$

$$y_d = \sqrt{P_i} h_{id} x^{'} + n_{id},\ \mathrm{i} \in \mathrm{R} \tag{2}$$

其中，x 和 $x^{'}$ 为源节点和中继节点功率归一化后的发射信号，$E[|x|^2] = E[|x^{'}|^2] = 1$，$\mathrm{P}_s$ 和 P_i 为源节点和中继节点i的发射功率，h_{si} 和 h_{id} 分别为源节点与中继i之间和中继i与目的节点直接的信道衰落系数，它们都是均值为0的相互独立的循环对称复高斯随机变量，方差分别为 σ_{si}^2 和 σ_{id}^2。n_{si} 和 n_{id} 分别是对应信道的独立零均值加性高斯白噪声，且方差都为 σ^2。

采用AF协议，在中继节点处接收信号放大增益为

$$G = 1 \Big/ \sqrt{P_i |h_{si}|^2 + \sigma_{id}^2}$$

$$x^{'} = G y_{si} \tag{3}$$

此时，目的节点接收到的信号(2)式变为

$$y_d = \sqrt{P_i} h_{id} x^{'} + n_{id} = \sqrt{\frac{P_s P_i}{p_s |h_{si}|^2 + \sigma_{id}^2}} h_{si} h_{id} x + \sqrt{\frac{P_i}{p_i |h_{si}|^2 + \sigma_{id}^2}} h_{id} n_{si} + n_{id},\ \mathrm{i} \in \mathrm{R} \tag{4}$$

目的节点接收信号的信噪比为

$$\gamma = \frac{\gamma_{si} \gamma_{id}}{\gamma_{si} + \gamma_{id} + 1},\ \mathrm{i} \in \mathrm{R} \tag{5}$$

其中，$\gamma_{si} = P_s |h_{si}|^2 / \sigma^2$ 和 $\gamma_{id} = P_i |h_{id}|^2 / \sigma^2$，为计算简便，本文假定信道方差 $\sigma^2 = 1$。

算法描述

1. ARMA 模型

ARMA 模型 (Auto-Regressive and Moving Average Model) 适用于短相关预测，算法复杂度简单，用于预测短相关流量时精度较高，适用于在线预测或能量有限的场景。比 AR 模型具有更小的预测误差方差。ARMA 模型对平稳性数据序列能够有效地分析其数据序列相关性。

定义[11]: 设 $\{\varepsilon_t\}$ 是均值为 0，方差为 σ^2 的白噪声，实系数多项式 A(Z)和 B(Z)没有公共根，满足 $b_0 = 1$，$a_p b_q \neq 0$ 和

$$A(z) = 1 - \sum_{j=1}^{p} a_j z^j \neq 0,\ |z| \leq 1,$$
$$B(z) = \sum_{j=0}^{q} b_j z^j \neq 0,\ |z| \geq 1, \tag{6}$$

称差分方程

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \sum_{j=0}^{q} b_j \varepsilon_{t-j},\ \mathrm{t} \in \not{\phi} \tag{7}$$

是一个自回归滑动平均模型，简称 ARMA(p,q)模型，简称满足式(7)的平稳序列$\{X_t\}$为平稳解或 ARMA(p,q)序列。对于采集的数据序列假设为$s^{'}=[x_0^{'},x_1^{'},\mathrm{L}\ ,x_n^{'}]$，利用平稳化序列方法对其进行对数处理，得$s=[x_0,x_1,\mathrm{L}\ ,x_n]$。判定 s 时间序列的平稳性，计算 s 的自相关函数和偏相关函数，依据呈现出的拖尾现象，断定为 ARMA 序列[12]。

本文采用文献[13]中的 AIC 定阶法对该序列 ARMA 定阶。AIC 准则函数为：

$$AIC=-2L(\hat{\beta})+2k \tag{8}$$

其中，$\hat{\beta}$为参数的最大似然估计值，$L(\mathrm{g})$为似然函数，k 为独立参数个数。

样本长度 n 足够大时，ARMA(p,q)模型的似然函数近似为

$$L(\hat{\beta})=-\frac{n}{2}\lg 2\pi-\frac{n}{2}\lg\hat{\sigma}^{2}-\frac{S(\hat{\beta})}{2\hat{\sigma}^{2}} \tag{9}$$

其中，$S(\hat{\beta})=n\hat{\sigma}^{2}$，$\hat{\beta}=(\hat{\varphi},\theta)^{T}=(\varphi_1,\varphi_2,\varphi_3,\mathrm{L}\ ,\varphi_p,\theta_1,\theta_2,\theta_3,\mathrm{L}\ ,\theta_q)^{T}$。将式(9)代入式(8)可得 n 充分大但给定前提下，ARMA(p,q)模型拟合的最小信息准则使下式最小

$$AIC(p,q)=n\lg\hat{\sigma}^{2}+2(p+q+1) \tag{10}$$

通过分析,本文选用 ARMA(2,1)[12~14]进行预测，模型为

$$(1-\varphi_1B-\varphi_2B^{2})x_i=\theta(B)a_i \tag{11}$$

式中，B 为后移算子，a_i为白噪声。$\varphi_1,\varphi_2,\theta_1,\sigma^{2}$为估计参数。利用最小二乘估计方法求解参数$\hat{\varphi}_1,\hat{\varphi}_2,\hat{\theta}_1,\hat{\sigma}^{2}$。对于该时间序列的稳定性判断，依据稳定性条件$\hat{\varphi}_1+\hat{\varphi}_2\ \mathrm{p}\ 1$，$\hat{\varphi}_2-\hat{\varphi}_1\ \mathrm{p}\ 1$，$|\hat{\varphi}_2|\mathrm{p}\ 1$，判断为平稳序列，得出 ARMA 拟合模型

$$x_t=\hat{\varphi}_1x_{t-1}+\hat{\varphi}_2x_{t-2}+a_t-\hat{\theta}_1a_{t-1} \tag{12}$$

随后利用逆函数法，得出进一步 ARMA 预测模型为[12]

$$\hat{x}_t(1)=\sum_{j=1}^{m}I_jx_{t+1-j} \tag{13}$$

式中，I_j为 ARMA 逆函数，m 为x_t之前的 m 次观测值，可由预测精度确定。

相应多次预测模型为

$$\hat{x}_t(l)=\hat{\varphi}_1\hat{x}_t(l-1)+\hat{\varphi}_2\hat{x}_t(l-2) \tag{14}$$

2. 中继选择算法

如何选择中继是协作通信中的关键问题，针对不同的目标具有不同的中继寻找方法。已有研究中出现了大部分的单中继选择方法[15~17]。本文提出一种基于能量预测的单中继选择算法，并分析该算法对于网络能量性能的影响。

设定候选中继集合 $\Re$，对所有 $R_i \in \Re, i \in \{1,2,\mathrm{L}\ ,N\}$ 按 ARMA 预测能量值进行降序排列，在单中继节点协作通信过程中，源节点 S 选择能量值最大的中继节点进行协作通信，用公式描述为

$$R_S = \underset{R_i \in \Re, i \in \{1,2,\mathrm{L}\ ,N\}}{\arg\max}\ E_{R_i} \tag{15}$$

式中，E_{R_i} 为中继 R_i 的预测能量值。

根据上述描述，单中继选择算法过程如下

步骤 1 判断源节点是否需要中继，若是，进行步骤 2，否则，直接传输；

步骤 2 搜集候选中继集合 中各中继历史能量信息值，依照 3.1 节中方法确定 ARMA 模型；

步骤 3 分别对于中继集合 $\Re$ 中的所有节点，依据 t 时刻收集能量数据 $s' = [x_0', x_1', \mathrm{L}\ , x_n']$ 及历史能量信息，将其代入 ARMA 模型，求出 t+1 时刻的能量值，并按降序排列中继节点为序列集合 L；

步骤 4 源节点 S 中继其信息时，选择序列集合 L 中的第一个中继节点传输其数据；

步骤 5 结束。

该算法在中继选择时，保证始终是能量值最大的中继节点中继源消息，既减轻了低能量节点中继信息时能量迅速耗尽而“死亡”，同时保证避免节点的通信传输范围的缩小。

3. 功率分配算法

在上述中继选择的基础上，为保证网络的通信 QoS(Quality of Service，服务质量)，需要再满足一定 QoS 的条件下，对中继网络进行合理的功率分配，并且保证网络功率最小。本文根据链路信道状态信息满足 QoS 条件下最优分配网络中节点的功率，采用文献[18]这种情况可以描述为

$$\begin{aligned} &\min_{R_i \in \Re} P = P_S + P_{SR_i} \\ &s.t.\ \ \gamma \geq \mathrm{D} \end{aligned} \tag{16}$$

式中，P 是中继网络的总功率，P_{SR_i} 是所选中继 R_i 的功率，γ 是目的节点接收信噪比，D是满足 QoS 的最小信噪比。

结合式(5)、(15)和(16)，为解决式(16)中的凸优化问题，优化公式中的变量 P_S 和 P_{SR_i}，利用拉格朗日乘子法(Lagrange Multiplier)可以求得满足式(16)的最优功率表达式为

$$P_S = \frac{\sqrt{\mathrm{D}(\mathrm{D}+1)}\,|h_{si}| + \mathrm{D}|h_{id}|}{|h_{id}|\,|h_{si}|^2},\quad P_{SR_i} = \frac{\mathrm{D}|h_{si}| + \sqrt{\mathrm{D}(\mathrm{D}+1)}\,|h_{id}|}{|h_{id}|^2\,|h_{si}|} \tag{17}$$

此时，最优中继情况下的网络系统最小总功率消耗为

$$P=\frac{(\sqrt{\mathrm{D}}|h_{si}|+\sqrt{\mathrm{D}+1}|h_{id}|)^2-|h_{id}|^2}{|h_{si}|^2|h_{id}|^2} \tag{18}$$

由式(17)和式(18)可知，当在图 1 所示模型下，由 ARMA(2,1)能量预测模型确定中继节点 R_i 后，按照式(17)对网络中的源和确定性单一中继进行功率分配时网络总能耗最小。

仿真结果及分析

本节将对本文 ARMA 能量预测模型的确定进行仿真分析对比，在此基础上，通过选定的 ARMA 能量预测模型，对比采用提出的算法与传统的随机中继选择算法在中继节点采用 AF 协议时网络能量方面进行性能对比。

1. ARMA(2,1)能量预测模型确定

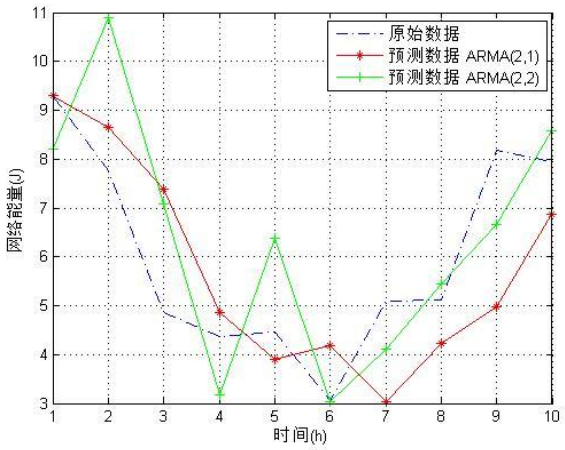
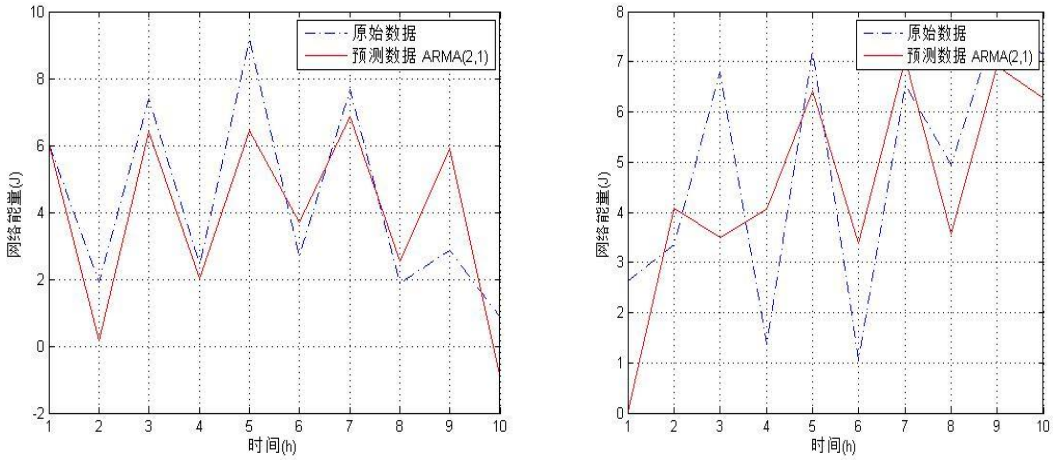


图 2: ARMA(2,1)模型与 ARMA(2,2)模型比较

为简便处理，不失一般性，本文仿真过程中，验证确定模型所需的能量数据采用随机值。仿真过程中，经过多次反复执行 ARMA 能量预测模型，随机产生 10 组实验数据验证模型的正确性，通过 3.1 节分析过程及选择对比 ARMA(2,1)和 ARMA(2,2)能量模型，结果如图 2 所示。

由图 2 可知，在同样场景下，对于一系列数据，仿真受所建场景模型及搜集数据的影响，选用 ARMA(2,1)能量预测模型和 ARMA(2,2)能量预测模型较原始数据都存在误差，但是，从图中可知，相较于 ARMA(2,2)能量预测模型，选用 ARMA(2,1)能量预测模型能够更好的反映原始数据的变化状态情况。

对于确定性的 ARMA(2,1)能量预测模型，对比采用此模型进行一步预测和两步预测的情况，如图 3 所示。



(a)一步预测模型　　(b)两步预测模型

图 3: ARMA(2,1)一步预测及二步预测

由图 3 可知，当 ARMA(2,1)能量预测模型采用一步预测时对于数据信息能够有着更好的预测，因此本文仿真过程中选择 ARMA(2,1)能量预测模型一步预测的方式进行数据信息的预测分析。

2. 网络能量分析

考虑单个源节点、单个所选中继节点和单个目的节点的情况，选择单个中继节点的方法分为随机选择和文中所提方法两种情况。为方便描述，本文对网络能量进行了简单处理，重点关注中继节点的能量消耗。随机生成历史能量数值，在此情况下，对比采用随机选择中继节点后进行功率分配和本文提出方案两种情况下的网络能量对比，如图 4 所示。

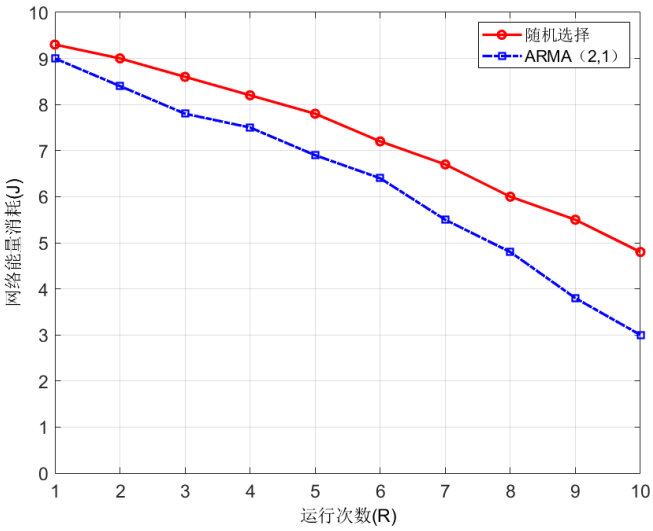


图 4: 不同方案下的网络能量对比

由图 4 可知，采用本文所提方案部署分配网络，能够更加有效的利用网络节点能量，降低网络能量消耗，有效的延长网络生命周期。

总结与讨论

本文分别提出了采用放大转发协议下的协作中继网络中的中继选择算法与功率分配算法。采用ARMA能量预测模型的中继选择算法可以有效均衡放大转发协议下的网络中各终端节点的能量消耗，降低协作网络总能量消耗，在此基础上，对源节点、确定的中继节点和目的节点在满足需求QoS的情况下运用拉格朗日乘法进行功率分配，求得最优解。仿真结果表明采用ARMA(2,1)一步能量预测模型的方法来选择中继节点，从而有效分配网络功率，能够降低网络的能量消耗，延长网络生命周期。

资助

本文受到以下项目资助，南京工业大学浦江学院2019教改项目：《通信原理》课程双对标的研究与实践（项目编号2019JG003Y）；南京工业大学浦江学院2019人才培养工程计划项目课题：ICT人才科学思维能力培养体系研究与实践（项目编号njpj2019-01-03），特向南京工业大学浦江学院表示感谢。

参考文献

[1] Laneman J N, Tse D N C, Wornell G W. Cooperative diversity in wireless networks: Efficient protocols and outage behavior[J]. Information Theory, IEEE Transactions on, 2004, 50(12): 3062-3080.

[2] Nosratinia A, Hunter T E, Hedayat A. Cooperative communication in wireless networks [J]. Communications Magazine, IEEE, 2004, 42(10): 74-80.

[3] Sendonaris A, Erkip E, Aazhang B. User cooperation diversity. Part I. System description[J]. Communications, IEEE Transactions on, 2003, 51(11): 1927-1938.

[4] Sendonaris A, Erkip E, Aazhang B. User cooperation diversity. Part II. Implementation aspects and performance analysis[J]. Communications, IEEE Transactions on, 2003, 51(11): 1939-1948.

[5] 白智全, 贾建岚, 杨祥来, 等. 协作通信网络中基于 Lagrange 算法的中继选择和功率优化方案[J]. 山东大学学报 (工学版), 2012, 42(3): 25-30.

[6] Laneman J N, Wornell G W. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks[J]. Information Theory, IEEE Transactions on, 2003, 49(10): 2415-2425.7.

[7] Laneman J N, Tse D N C, Wornell G W. Cooperative diversity in wireless networks: Efficient protocols and outage behavior[J]. Information Theory, IEEE Transactions on, 2004, 50(12): 3062-3080.

[8] Hu H,Zhang Y,Luo J.Distributed Antenna Systems:Architecture for Future Wireless Communications[M].Boca Raton,FL:Auerbach Publications,2007.

[9] Bletsas A, Khisti A, Reed D P, et al. A simple cooperative diversity method based on network path selection[J]. Selected Areas in Communications, IEEE Journal on, 2006, 24(3): 659-672.

[10] Si Jiang-bo, Li Zan, Dang Lan-jun, and Liu zeng-ji, Joint optimization of relay selection and power allocation in cooperative wireless networks[C].International Conference on Communication Systems,Guangzhou,China.Nov.19-21,2008:1264-1268.

[11] 何书元. 应用时间序列分析[M]. 北京大学出版社, 2004.

[12] 李捷,刘先省,韩志杰.基于ARMA的无线传感器网络流量预测模型的研究[J].电子与信息学报,2007,29(5):1224-1227.

[13] 张树京,齐立心.时间序列分析简明教程[M].北京:清华大学出版社。2003,9:98-122.

[14] Deng J, Han R, Mishra S. Defending against path-based DoS attacks in wireless sensor networks[C]//Proceedings of the 3rd ACM workshop on Security of ad hoc and sensor networks. ACM, 2005: 89-96.

[15] 张静美, 邵春菊, 王莹, 等. 基于不同信道状态信息的两跳中继系统最优功率分配[J]. 北京邮电大学学报, 2005, 28(4): 58-61.

[16] 张静美, 王莹, 张平. 多天线协作传输系统的容量分析与功率分配[J]. 北京邮电大学学报, 2005, 28(5): 11-15.

[17] Sreng V, Yanikomeroglu H, Falconer D D. Relayer selection strategies in cellular networks with peer-to-peer relaying[C]//Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th. IEEE, 2003, 3: 1949-1953.

[18] 龚渝钧, 朱宇. 最小功耗的中继选择与功率分配的联合优化[J]. 信息与电子工程, 2012, 10(1): 7-12.

基于视觉感知的实验室出入库零件识别软件

RECOGNITION SOFTWARE OF PARTS IN AND OUT OF LABORATORY BAS ED ON VISUAL PERCEPTION

潘时祥 [1]，徐平平 [2]
Shixiang Pan[1], Pingping Xu[2]

南京工业大学浦江学院 计算机与通信工程学院
Nanjing Tech University Pujiang Institute
*Corresponding author: 潘时祥, 徐平平, E-mail: 1976408867@qq.com, xpp@seu.edu.cn

摘要

随着时代发展，人工记录实验室物品的出库入库都会随之被取代。本文介绍基于视觉感知的实验室出入库识别软件系统用于实现实验室仓库零件的快速入库出库，实现高效快捷的管理物品流程，对指定需要的零件进行准确的管理，提高实验室对零件管理的效率与准确性，从而达成基于无人管理理念下零件为中心的技术支持平台。

关键词：视觉感知、移动出入库、无人管理

Abstract

With the development of the times, the manual recording of laboratory items will be replaced. This paper introduces a software system based on visual perception, which is used to realize the quic k warehousing and ex warehouse of parts in the laboratory, to realize the efficient and quick managem ent of goods flow, to carry out the accurate management of the required parts, to improve the efficienc y and accuracy of parts management in the laboratory, so as to achieve the parts centered technical su pport level based on the concept of unmanned management Set.

Keywords: Visual perception, mobile in and out, unmanned management

1 引言

随着互联网的发展与普及，目前大多数学校或其他地区的仓库零件已开始采用技术手段进行管理，虽然如此，但是还是需要使用纸张进行记录之后、才能使用手工输入计算机的方式进行采集和统计整理。这种方法造成了大量人力物力的浪费，在管理期间因人为管理的因素，会使零件出库入库效率低、速度慢等问题。改革开放以来，国家越来越注重学校实验室的发展，所以随着实验室的发展，学校实验室零件的物资种类机数量随着新科技的发展在不断增加。由于二十一世纪各国对学生科研大力支持，学生或老师科研积极性也在剧增，所以导致实验室的物品出库出库频率太高，仓库管理作业也已十分复杂和多样化，因此传统的人工管理以及新型二维码出入库早已不满足各类实验室的物品管理，严重影响了实验室的运行工作效率，成为制约学校实验室或企业发展的一大障碍。

2 研究目的

通过图片对比，拍摄需要出库或入库的器材图片，与所在的实验室仓库图片库进行对比，来完成实验室仓库物品的管理，解决实验室传统表格出入库需要专业专职仓库管理员操作记录而带来的人力物力的过于浪费以及新型二维码出入库记录方式中归还入库时由于二维码损坏产生的各种弊端造成数据丢失的后果。

本软件通过执行下位机拍照，自动上传至指定文件夹，在web端执行新建数据库，图片库，识别成功后的出库入库等所有操作，以达到简单方便，易操作，无需专业人员操作的目的。

3 文献综述

改革开放以来，我国国力综合水平逐渐上升，因此对各种实验室的管理也更加规范。但是现阶段依旧是人工管理实验室出入库系统，或者新型二维码扫描物品进行零件的出入库的管理。现阶段出入库几种管理方法都存在一定的弊端：

1. 传统实验室器件出入库以表格记录，过于浪费人力物力。
2. 新型二维码出入库，用户使用完实验器材后导致二维码丢失损坏，归还实验室器材时，无法做到正常入库，以及贴二维码浪费时间精力，并且需要每个物品都携带，如果贴错则或导致实验室器材的损失与数据库的混乱。
3. 若遇到无法入库的物品时，需要去查询该物品在数据库是什么名称，非转职人员无法做到认识所有的物品，并且及其浪费时间。
4. 需要转职录用一位仓库记录管理员，并且遇到记录员不在，或离职导致数据链断开，导致实验室无法正常使用，延误上课进度。

4 研究方法

现国外内实验室出入库主要分两种方法：一种为传统模式记录，一种为二维码扫码出入库。

1. 其中大多数都以传统出入库为主，传统模式弊端：传统的实验室出入库管理大多采用如图1所示表格纸质化记录或是EXCEL表格记录的形式，需要手动填写数据，效率低、存在重复冗余工作的情况，且容易出错，本地保存还有丢失的风险。

入库情况							出库情况					归还情况			库存量
日期	规格型号	有效期至	入库量	入库编号	生产厂家	管理员	日期	领取量	领用用途	领取人	管理员	日期	归还人	管理员	

图1: 传统出入库记录表

2. 利用扫描二维码的模式对产品出入库进行管理，可以大大提高管理效率，也保证了数据的准确度，降低失误率。扫码出入库流程如表1所示:

表1: 扫码流程图

步骤1	建立商品基础信息表
步骤2	建立入库表，商品编码关联商品基础信息表，并勾选扫码输入
步骤3	入库填表时点击“商品编码”扫描二维码读取编码进行入库操作

不过经过市场调研发现新型二维码出入库也存在弊端，用户使用完实验器材后导致二维码丢失损坏，归还实验室器材时，无法做到正常入库。

3. 针对该弊端设计出一款基于轻量级设计的理念通过手机app实现扫描二维码或条形码做到出库、入库、建库的，并且基于视觉感知，建立通过扫描货物图片进行出入库统计与新建图片库，在服务器云端，建立图片库与数据库，实时可对货物进行扫描比对。

5 研究结果

1、图2为系统架构实现图：

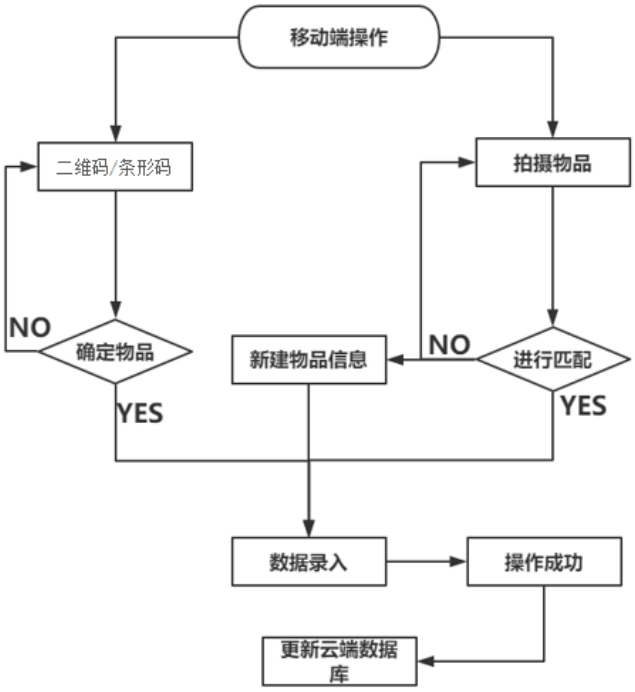


图 2: 系统架构图

2、图3为后台实现过程图：

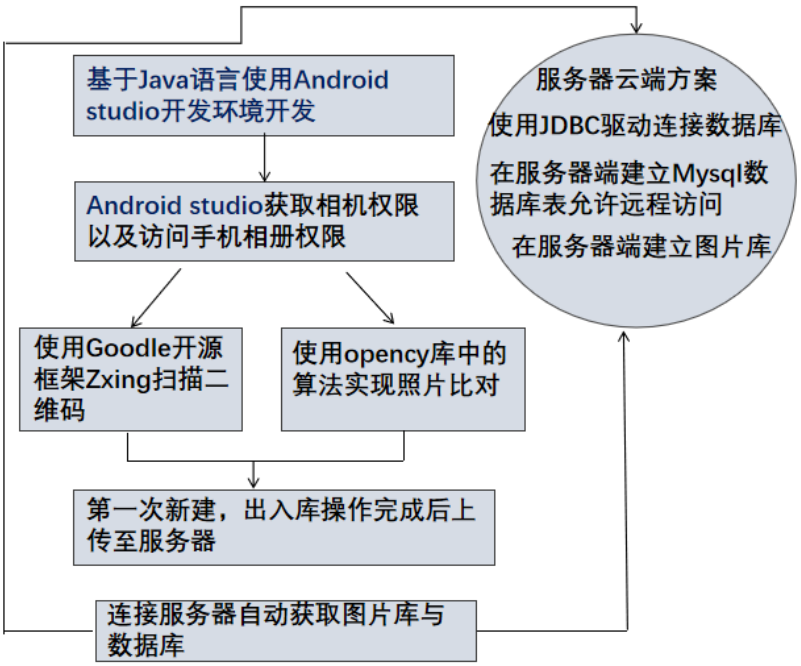


图 3: 后台实现流程图

3、图4为前后端技术路线图：

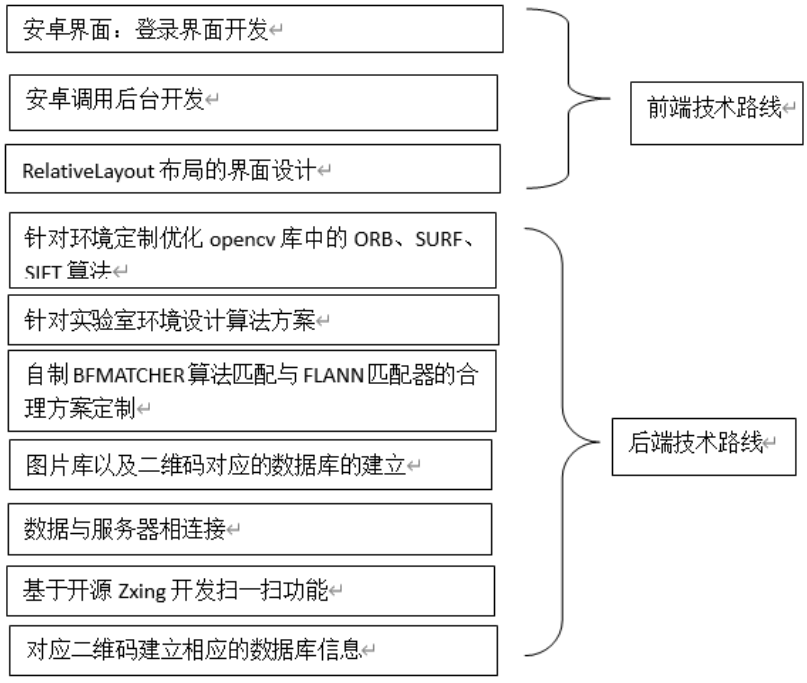


图4: 技术路线图

4、 图3为后台算法整体实现图：

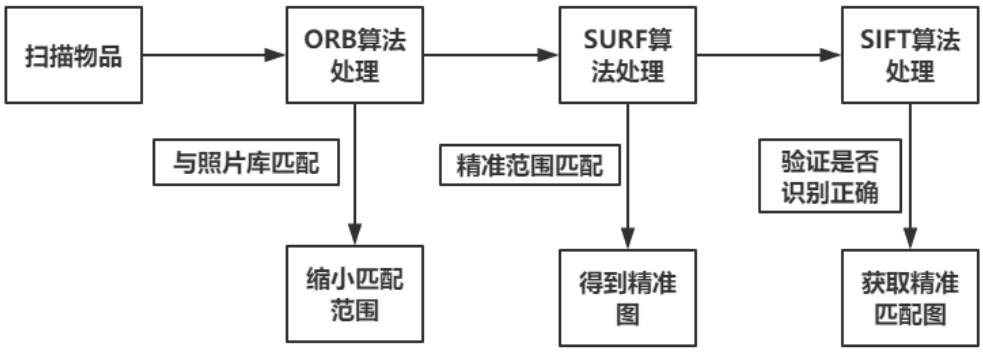


图 5: 后台算法实现

① ORB 特征+BFMATCHER 算法匹配（速度极快是后两种的 100 倍为了先缩小范围）；
② SURF 特征算法精确到图（FLANN 匹配器，k-NearestNeighbor 分类算法);
③ SIFT 特征算法进行验证;

5、ORB 算法分析，图 6 为 ORB 算法实现过程。这里主要是对特征点的检测。特征点是指相机在运动过程中，某一区域中，能反应自身特点与其他区域区别开，而且运动过程中未发生明显化的点或者区块。[1]

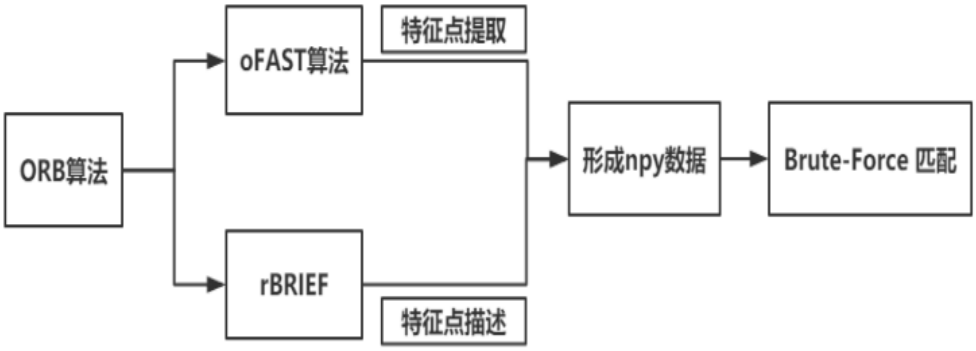


图 6: ORB 算法实现

①特征提取是由 FAST（Features from Accelerated Segment Test）算法发展来的。

②特征点描述是通过改进 BRIEF（Binary Robust IndependentElementary Features）特征描述算法实现的。

③步骤：首先，利用 FAST 特征点检测算法实现特征点检测，其次利用 Harris 角点的度量方法，通过 FAST 特征点算法中选出的 Harris 角点响应值最大的 N 个特征点。其中 Harris 角点的响应函数定义为：

$$R=detM-\alpha(traceM)^2$$

6、SIFT 算法分析：

SIFT 算法具有许多有点，包括特征稳定，尺度变换、亮度保持不变性，对视角变换等，同时在噪声处理上也有一定程度的稳定性，图 7 为 SIFT 的实现流程图。

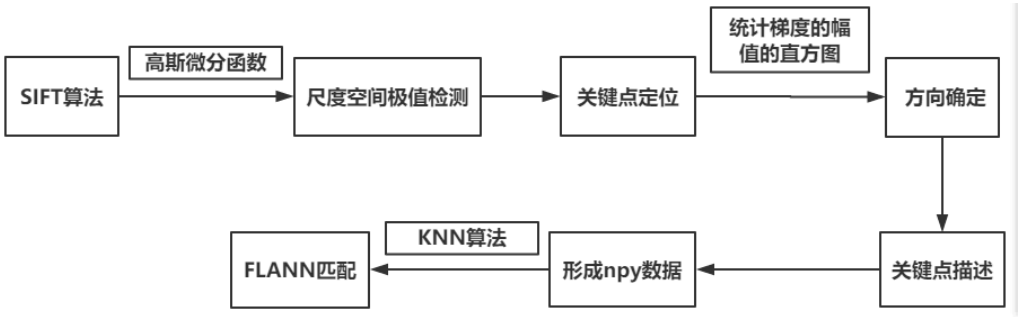


图 7: SIFT 算法实现

①尺度空间极值检测：

通过图像识别查找图像位置。通过高斯微分函数识别潜在的尺度和固定不旋转的关键点；

②关键点定位：

在每个待测点的位置上，通过一个计算精细的模型确定该点位置以及尺度。关键点的选择依据在于他们稳定程度；

③方向确定

基于图像局部的阶层方向，分配给每个关键点位置，包括一个或多个方向。所有后面对图像数据的变换都是相对于关键点的方向、尺度以及位置，从而保证这些变换的不变性；

④ 关键点描述

在每个关键点周围范围中，测量选定的尺度上的图像梯度。这些梯度被变换成一种表示，这种表示允许比较大的局部形状的变形和光照变化。

7、SURF（Speeded Up Robust Features）通过对 SIFT 特征的提取和描述方式进行一定程度上的改进，达到特征提取以及描述更加的高效，图 8 为实现 SURF 流程图。

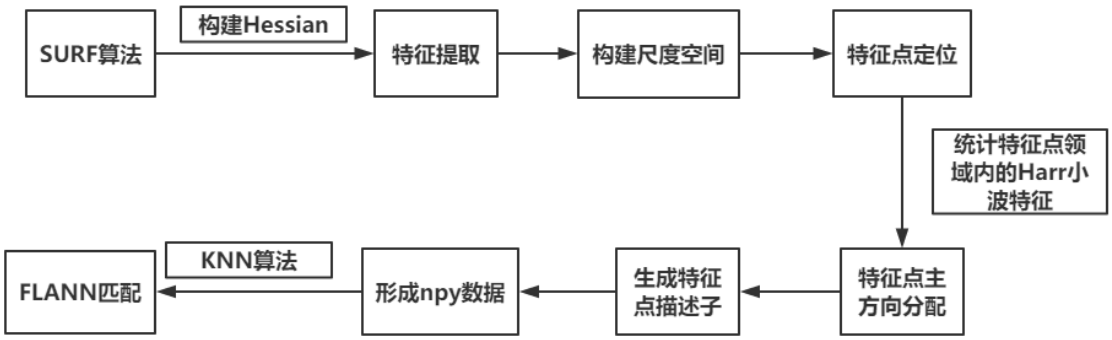


图 8: SURF 算法实现

① 构建 Hessian（黑塞矩阵），生成所有的关键点，用于特征点的提取；
② 构建相对应的尺度空间；
③ 进行特征点的定位；
④ 定位成功进行特征点主方向匹配；
⑤ 匹配完毕，生成特征点描述子；
⑥ 特征点匹配。

8、系统软件实现

①图 9 为系统登陆界面，logo 为自己创办团队大学生科创中心标志。输入账号密码登录软件首界面。

图 9: 登录界面

②图 10 为系统主页面效果图，展示出入库操作界面，方便管理员查看物品历史出入库具体时间。

图 10 系统操作主页面

③11 为仓库物品出入库情况以及库存查询效果图。

图 11 系统库存功能界面图

④点击某个物品的详情按钮，得到如图 12 所示效果图，展示此物品的出入库历史记录，包括历史时间、出入库数量等。

图 12: 物品详情按钮效果图

⑤点击某个物品的操作按钮，得到如图 13 所示效果图，可以对此物品进行出库、入库以及出入库数量等操作。

图 13: 物品操作按钮效果图

⑥搜索页面可扫码或直接图片拍摄识别进行出入库的操作。点击扫码按钮进行物品扫码识别，扫码过程图如图 14 所示。经过后台算法运算后得到如图 15 所示效果图，展示出扫码后软件对比数据库得到此物品的出入库数量与剩余量情况。

图 14: 扫码中效果图

图 15: 扫码后展示图

扫描物品与扫描二维码同样原理，图 16 和图 17 分别为物品扫描中与扫描后效果图。

图 16: 实物拍摄中效果图

图 17: 扫码后展示图

5 后续补充研究

本文研究的出入库识别系统现阶段算法只能对实验室适用，在今后会继续对此软件进行算法优化以及界面优化。使本研究对象可以对广泛区域进行使用，比如以下几点：

1、 学校或公司实验室，食堂，小卖部，医务室等地方都需要去实施出入库的管理。解决学校或公司需聘请社会人员来整理物件，浪费人力物力。

2、 社会上的小型超市，及小乡镇上的小卖部使用本产品可以清楚的知道库存里的物品存放我们下一个版本会更新警告功能，如果是发现这个商品少于你设定的数量，那么软件将会发出警示，尽快补货。

基于以上研究可以解决实验室或其他地方出入库麻烦等问题。

参考文献

[1]殷新凯,茅健,周玉凤,陈晓平.基于改进 ORB 算法的视觉里程计特征匹配方法[J]，2020,41(04):57-62.

基金资助项目：

[1]南京工业大学浦江学院 2019 年大学生创新创业训练计划项目：基于移动平台的实验室零件出入库识别系统。（PJ20191390507）

作者简介：

潘时祥 男，南京工业大学浦江学院计算机与通信工程学院信息管理 1801 班本科生，计通学院大学生科创中心主席， 2019 全国大学生电子设计竞赛省二等奖，全国大学生数模竞赛参赛奖，华为云计算 HCIA-Cloud 证书。

徐平平 女，南京工业大学浦江学院计算机与通信工程学院院长，东南大学移动通信国家重点实验室通信与信息系统学科教授博导，主要研究方向，5G 通信及移动云计算，移动自组织网络，车联网通信，物联网应用等

基于语音信号分析的婴儿情感识别与研究
RECOGNITION AND RESEARCH OF INFANT EMONTION BASED ON SPEECH SIGNAL ANALYSIS

李琪 [1]，徐平平 [2]，冯倩倩 [3]
Qi Li[1], Pingping Xu[2], Qianqian Feng[3]

南京工业大学浦江学院 计算机与通信工程学院
Nanjing Tech University Pujiang Institute
*Corresponding author: 李琪，徐平平，冯倩倩
E-mail: 1668202408@qq.com, xpp@seu.edu.cn, 974742953@qq.com

摘要

在时代的不断地变迁之下，成人面临生活和工作的压力不减反增，诸多成人在育婴育儿方面理论知识不足且不具有实践经历，因此利用语音情感识别分类技术研究基于语音信号的婴儿情感识别与分析，帮助人们解决此类问题更显得迫在眉睫。为解决该问题，本文提出从婴儿本身需求出发，帮助成人解决在育婴方面的知识匮乏和经验不足问题。本文采用MATLAB语音信号处理技术，主要针对婴儿的哭声进行一系列预处理，提取12阶梅尔频率倒谱系数，利用KNN算法对婴儿情感需求进行识别和分类学习，帮助成人快速了解并满足婴儿的情感需求，缓解成人在孕婴方面的压力，降低其在生活上、工作上的压力。本文的研究方法和技术能够有效改善婴儿情感需求无法被准确识别以及快速获取的问题，可以提高成人在解决育婴方面问题的能力。

关键词：MATLAB语音信号处理，梅尔频率倒谱系数，情感识别，KNN算法

Abstract

With the continuous changes of the times, adults are facing increasing pressure in life and work, and many adults lack theoretical knowledge and practical experience in infant care. Therefore, it is even more urgent to use voice emotion recognition classification technology to study infant emotion recognition and analysis based on voice signals to help people solve such problems. In order to solve this problem, this paper puts forward to help adults to solve the problems of lack of knowledge and experience in nursing from the needs of infants themselves. In this paper, MATLAB voice signal processing technology is used, mainly for a series of pre-processing of infant crying, 12-order mel-frequency cepstral coefficients is extracted, and KNN algorithm is used to identify and classify the emotional needs of infants, so as to help adults quickly understand and meet the emotional needs of infants, relieve the pressure of adults on pregnancy and infants, and reduce the pressure on life and work. The research methods and techniques in this paper can effectively improve the problem that the emotional needs of infants cannot be accurately identified and quickly acquired, and can improve the ability of adults to solve the problems in nursing.

Keywords: MATLAB Speech Signal Processing, mel-frequency cepstral coefficients, Emotion Recognition, KNN Algorithm

引言

当代社会，每个家庭对于育婴工作都十分重视，婴儿的各方面开支在整个家庭中占比越来越大。随着育婴工作不断被重视，很多年轻女性甚至是男性在承受着社会工作重大压力的同时，育婴方面的压力也不断加重。新手父母对于育婴的经验和精力的不足，致使他们不能够很好的照顾婴儿，有效的陪伴婴儿的成长，习得婴儿的需求。弱小的婴儿在繁杂的世界中，只有依靠成年的人精心呵护，才能获得基本的生理需求和情感满足。Sauter（2013）等人认为婴儿一般通过哭声来向成人表达自己的情感需求，因此提出基于婴儿语音信号的情感识别研究显得格外重要。

Reyes Galaviz（2009）等人在文章中提出初生婴儿不具备语言表达能力，哭是婴儿的基础性生理反应，也是他们表现情感状态和对反应外界刺激的重要方式这一观点。所以本文的研究是基于婴儿的啼哭信号分析进行情感识别和分类的。

1.语音情感识别分类

1.1 原理和方法

Steidl（2010）在撰写的文章中提出语音信号情感识别分类的基础原理是一种模式识别的过程，根据信息与特征值之间的匹配程度，实现情感的分类识别，推测出情感情绪状态，达到情感识别的目的。主要过程涵盖了语音信号的获取、预处理、情感特征提取、情感分类几个过程。

在语音信号获取方面Batliner（2008）等人发布了一个最具代表性的最为著名的是FAU AIBO语料库，该语料库是少有的纯自然语音库，包含了婴儿的愤怒、强调、中性、积极（母爱和快乐组成）和休息（不属于其他类别的情感组成，如无助、无聊等 ）五个种类。Orlandi（2012）等人发布来自意大利的6-24周婴儿的啼哭特性数据库（D1），包括正常性和反抗性（身体冷/热、衣物或和尿不湿不舒服啼哭）特性的啼哭。Sheinkopf（2012）等人发布来自美国罗德岛州的6个月左右的婴儿异常性患有疾病的啼哭特性数据库（D2）。Zabidi（2009）等人发0-7个月大的婴儿异常性的啼哭特性数据库（D3），包括肺炎和口腔炎症。Lederman（2010）等人发布健康的具有生理性需求的婴儿啼哭特性数据库（D4），包括口渴、饥饿和困倦特性的啼哭。Okada（2011）等人发布健康的具有心理性需求的婴儿啼哭特性的数据库（D5），包括需要抱、陪伴和闹脾气特性的啼哭。Kheddache（2014）发布来自古巴婴儿的正常性和异常性啼哭数据库（D6），其中异常性包括头疼胀痛和睡不沉特性的啼哭。Patil HA（2009）发布来自印度的小于6个月的婴儿正常性和异常性啼哭数据库（D7），其中异常性啼哭包括哮喘、口腔、头疼、中耳炎和疝气特性的啼哭。Messaound（2010）等人发布来自加拿大的小于6个月的婴儿健康的生理性啼哭数据库（D8），包括饥饿、口渴和困倦性的啼哭。Petroni（1999）等人发布来自加拿大的2-6个月的婴儿反抗性啼哭数据库（D9），包括痛苦、惊恐、生气和异物刺痛特性的啼哭。由于以上的数据库不是可以自由评估供公众使用的，并且语音情感类别不全。因此作者考虑参照以上数据库的情感分类标准建立更加符合本实验需求的语料库。

Linhui Sun（2019）等人在语音情感识别分类的特征选择阶段，提出采用Fisher准则和MFCC系数两种滤除区分能力较强的方法提取特征参数。刘振焘（2010）等人在特征参数提取和降维方法上主要论述了从基频特征提取、共振峰特征提取、梅尔-倒谱频率系数提取、基于导数的非个性语音情感特征提取、基于Teager能量算子（TEO）非线性特征提取和基于深度学习的特征提取等方法来进行提取特征参数。沈再阳(2017)在出版物介绍了高斯混合模型(GMM),

隐马尔可夫模型（HMM），支持向量机，神经网络，混合设计模型几种特征参数提取方法和模型。

Stefano（2019）等人撰写的文章表明在情感分类算法中决策树算法，支持向量机，随机森林算法比较具有优越性；在模型匹配中朴素贝叶斯、线性回归几种算法机制比较具有优越性。林涛（2019）等人提出利用优化的K近邻算法进行情感分类识别。本文采用KNN算法进行婴儿啼哭情感分类识别。

1.2 语音信号的预处理

婴儿哭声的预处理主要是指在对于婴儿哭声正式进行情感识别和特征提取处理过程之前进行一些语音信号的处理和操作，赵力（2009）在出版物中提出语音信号处理的具体操作过程包括:婴儿哭声的采样及格式化、婴儿哭声的降噪、预加重、分帧、加窗和端点检测等操作。

1.2.1 婴儿啼哭信号的采样及格式化

本文的自然语音信号采集是利用手机移动端声卡，其中手机移动端声卡系统包括安卓和IOS。采集语音信号之后，建立语音库，由于采集语音信号来源不同导致音频格式不同，有m4a格式、mp3格式和wav格式，因而要对语音库进行格式化处理。本文利用专业的语音格式转换器对语音库语音格式进行转换。

1.2.2 婴儿啼哭信号的降噪

本文选用递推最小二乘法（RLS）进行语音信号降噪处理。RLS降噪算法的具体数学过程和计算公式如公式1和2所示。

1.输入信号和输出信号之间关系：

$$q(n)=\sum_{i=1}^{M} w_i(n)p(n-i+1)=w^T(n)p(n) \qquad M\le n \qquad \text{公式（1）}$$

其中$\{p(n)\}$为输入信号，含有样本$\{p(1),p(2),p(3),\mathrm{K},p(N)\}$；$\{d(n)\}$为期望信号序列，含样本$\{d(1),d(2),d(3),\mathrm{K},d(N)\}$。$q(n)$为滤波器输出的信号，$W_i(n)$为冲击响应的滤波系数，含样本$\{w_1(n),w_2(n),w_3(n),\mathrm{K},w_m(n)\}$。M（$M\le n$）为滤波器长度，n为语音片段的数据长度。

2.RLS递推计算公式为：

$$\hat{\mathrm{w}}(n)=\hat{\mathrm{w}}(n-1)+K(n)[d(n)-x^T(n)\hat{w}(n-1)] \qquad \text{公式（2）}$$

其中$K(n)$为增益矢量，与输入阵$p(n)$和相关矩阵的逆矩阵的乘积数相等，$d(n)-x^T(n)\hat{\mathrm{w}}(n-1)$是真正估计误差。

1.2.3 婴儿啼哭信号预加重

在人体发生系统的发声过程中，声带和嘴唇会对生成的语音有所影响，对采集的语音信号进行预加重处理，就可以补偿在发声过程中语音信号受到抑制的高频部分。赵力（2009）提出将采集的婴儿啼哭语音信号片段通过一个高通滤波器这样的过程就是预加重过程。式中μ的值介于0.9-1.0之间。

$$H(Z)=1-\mu z^{-1} \qquad \text{公式（3）}$$

1.2.4 婴儿啼哭信号的分帧

Pramanik（2012）等人提出语音信号是根据时间变化的信号，而且不是具有稳定状态的信号，但在一定时间内可以看作是不变的即相对稳定的，语音信号的这一特性也称为短时平稳性。在婴儿啼哭语音信号短时的基础上对其进行分析和处理，这个过程就是将婴儿啼哭信号进行分段，进而分析每个分段的特征参数，啼哭信号中的每一分段称为“帧”。

1.2.5 婴儿啼哭信号的加窗

婴儿啼哭信号具有短时平稳性，对婴儿啼哭信号进行分帧操作之后，为了对抽取样本N附近的波形进行加强，对其余部分的波形进行弱化，还要对其进行加窗操作。Batliner（2008）等人提出加窗操作的原理就是对婴儿啼哭信号的各个短片段进行变换运算处理。目前研究中用得最多的变换运算处理就是窗函数运算，矩形窗函数公式如下。

$$\mathrm{w}(n)=\begin{cases}1, 0<n\le N-1\\ 0, \text{其他}\end{cases} \qquad \text{公式（4）}$$

1.2.6 婴儿啼哭信号的端点检测

1.短时能量

随着时间的改变婴儿啼哭信号的振幅或能量会发生缓慢的变化，这个过程中的时间、振幅以及能量之间的关系变化规律可以通过短时能量序列来反映，其计算公式如 5 所示。

$$E_n=\sum_{\mathrm{m}=-\infty}^{\infty}[d(m)w(n-m)]^2=d(n)^2*h(n) \qquad \text{公式（5）}$$

其中 $h(n)=w(n)^2$， $w(n)$ 是矩形窗口函数， $h(n)$ 是汉明窗口函数。

2.过零率

每秒内婴儿啼哭信号通过零值的次数就是过零率，其计算公式如 6 所示。

$$Z_{\mathrm{n}}=\sum_{m=-\infty}^{\infty}|\mathrm{sgn}[s(m)]-\mathrm{sgn}[s(m-1)]|\,w(n-m) \qquad \text{公式（6）}$$

其中式中 $\mathrm{sgn}[]$ 是符号函数，即 $\mathrm{sgn}[s]=\begin{cases}1, s\ge 0\\ -1, s<0\end{cases}$。

2. 特征参数提取

2.1 特征参数提取过程

2.1.1 快速傅里叶变换（FFT）

语音片段经过预处理的一系列加工操作之后，得到预处理后语音信号记为 $x(n)$。在对语音信号 $x(n)$ 施以快速傅里叶变换（是一种DFT的高效算法）得到离散谱 $X(k)$，公式如7所示。

$$X[\mathrm{k}]=\sum_{n=0}^{N-1}x[n]e^{-j\frac{2\pi}{N}nk} \qquad k=0,1,2,\mathrm{L}\ ,\ N\text{-}1 \qquad \text{公式（7）}$$

2.1.2 取绝对值数

语音信号在经过快速傅里叶变换处理（FFT）后，语音信号从时域信号变换到了频域信号，此时能够轻易看出语音信号的特征，再对 $X[k]$ 进行取绝对值操作，得到信号幅度谱 $\left|X[k]\right|$。

2.1.3 Mel滤波器

将幅度谱输入滤波器组进行处理。一组有 Mel 刻度的三角形滤波器组组成 Mel 滤波器组，1000Hz以下的10个滤波器线性相隔，1000Hz以上的剩余滤波器对数相隔。设定一个滤波器组，含有 m 个三角滤波器，其中 $m=1,2,\mathrm{K},M$，M 通常取22-26。各 $f(m)$ 之间的间隔随着 m 值的减少而变小，随着 m 值的增加而变大。

Step1 将实际频率尺度转换为 Mel 频率尺度：

$$\mathrm{Mel}(f)=2597\lg(1+\frac{f}{700}) \qquad \text{公式（8）}$$

$Mel(f)$ 是 Mel 频率，f 是实际频率，单位为 Hz。

Step2 设置三角形滤波器组并计算婴儿啼哭信号幅度谱滤波经过每一个三角形滤波器后的输出：

$$H_{\mathrm{m}}(k)=\begin{cases}0 & ,k<f(m-1)\\ \dfrac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m)-f(m-1))} & ,f(m-1)\le k\le f(m)\\ \dfrac{2(f(m+1)-k)}{(f(m+1)-f(m-1))(f(m)-f(m-1))} & ,f(m)\le k\le f(m+1)\\ 0 & ,k\ge f(m+1)\end{cases} \qquad \text{公式（9）}$$

$$\sum_{m=0}^{M-1}H_m(k)=1,\quad 0\le \mathrm{m}\le M \qquad \text{公式（10）}$$

$$f(m)=(\frac{N}{f_{\mathrm{s}}})Mel^{-1}(Mel(f_l)+m\frac{Mel(f_h)-Mel(f_l)}{M+1}) \qquad \text{公式（11）}$$

其中 $H_m(k)$ 为带通滤波器，M 是滤波器个数，$f(m)$ 是中心频率。N 为帧长，f_{s} 是采样频率，Mel^{-1} 为函数的逆函数，f_l 为频率范围中的最低频率，f_h 为频率范围中的最高频率。

Zabidi（2009）等人提出语音片段的音调或音高，不会反应在MFCC参数中也不受其影响，所以将MFCC作为声学特征，即使输入的语音信号音调不同，识别结果也不会有所影响，还可以降低运算量。

2.1.4 取对数

在婴儿啼哭信号经过 *Mel* 滤波器组处理之后，计算出由滤波器组输出的对数频谱。每个滤波器输出的对数频谱计算公式如12所示。

$$s(\mathrm{m}) = \ln(\sum_{k=0}^{N-1} |X(k)|^2 H_m(k)) \qquad 0 \le \mathrm{m} \le M \qquad \text{公式（12）}$$

2.1.5 倒谱/离散余弦变换（DCT）

Zabidi（2009）等人提出将通过傅里叶变换、取对数处理后的婴儿啼哭信号进行傅里叶反变换处理获得就是倒谱。在波形上，倒谱与频谱有类似的波形，假如频谱在低频处有个峰值，相应的倒谱在低倒谱系数上也有峰值，假如频谱在高频处有个峰值，相应的倒谱在高倒谱系数上也有峰值。

因此用低倒谱系数检测音元，用高倒谱系数检测音高。倒谱不同系数之间的变化是不相关的，这就是倒谱系数的优点之一，但是由于倒谱处理的计算过程较为复杂，所以本文利用离散余弦变换来替代反傅里叶变换来获取倒谱，其计算公式如13所示。

$$C(n) = \sum_{m=0}^{N-1} s(m) \cos(\frac{\pi n(m-0.5)}{M}) \qquad \mathrm{n} = 1, 2, \mathrm{K}\ , L \qquad \text{公式（13）}$$

2.2 特征参数提取结果

特征参数表是采集的每个婴儿啼哭语音片段的MFCC特征参数生成n行12阶（12是在MFCC特征参数提取过程中设置的计算离散变换系数的循环次数）的特征参数表，根据采集的语音信号的长度不同所产生的特征参数库的大小不同，即语音信号时间长，提取的特征参数行数越多，总体特征参数值越多。语音信号的部分特征参数列表如表1所示。

表 1: 部分特征参数表

	1	2	3	4	5	6	…	11	12
1	-48.8801	-11.0906	-12.7076	-7.3457	-4.8806	-1.3332	…	-2.5104	-1.9091
2	-15.3736	2.5819	1.5870	1.9563	0.8714	3.8903	…	0.1763	0.4948
3	-21.0135	-0.3970	1.7782	1.9672	1.4289	0.9452	…	-0.2060	-0.6611
4	-21.6849	0.6917	1.8563	2.2196	0.5167	3.8772	…	1.0285	0.5832
5	-23.7903	0.8539	1.3888	0.9888	3.8656	1.8110	…	0.1714	-0.2115
6	-20.6227	-2.6140	-2.9618	0.8486	0.1695	0.7874	…	-0.5225	0.0552
7	-22.6453	-1.9163	2.6347	4.5175	0.9778	1.9095	…	0.2987	1.8762
8	-25.0978	-28.4126	-6.9069	-0.0688	0.8022	-0.1438	…	9.5628	-1.7696
9	-22.8519	-27.5008	-13.0717	-4.8269	1.0623	-0.3336	…	9.4556	-2.6813
10	-16.3079	-22.2102	-15.4183	-6.8997	-2.1323	-1.2850	…	10.2719	-2.4514
11	-21.3297	-24.7422	-11.9526	-6.3537	-0.1667	-0.2081	…	8.2317	-5.3475
12	-20.1133	-14.8399	-12.3021	-2.6666	-0.3118	-1.0917	…	3.8637	-6.4228
…	…	…	…	…	…	…	…	…	…

3. 识别分类算法

本文将婴儿啼哭信号经过特征参数提取之后，提取出其特征向量，进而利用KNN算法进行分类学习，划分婴儿哭声信号种类，识别出的情感需求种类。KNN分类算法是指K近邻算法（K-Nearest Neighbor），一种分类算法，是数据挖掘分类算法中最简单的一种。

3.1 算法的数学原理

KNN算法用数学表达式来表示如下：

Step1假设簇划分为(C_1,C_2,K,C_K)，最小化平方误差E公式如14所示。

$$E=\sum_{i=1}^{k}\sum_{x\in C_i}\|x-\lambda_i\|_2^2 \qquad \text{公式（14）}$$

其中λ_i是簇C_i的均值相邻，也称为质心，数学表达式15所示。

$$\lambda_i=\frac{1}{|C_i|}\sum_{x\in C_i}x \qquad \text{公式（15）}$$

Step2计算距离公式，二维空间的欧氏距离公式如公式16所示。

$$\rho=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2} \qquad \text{公式（16）}$$

拓展到多维空间距离公式则如17所示。

$$d(x,y)=\sqrt{\sum_{i=0}^{n}(x_i-y_i)^2} \qquad \text{公式（17）}$$

3.2 算法的伪代码

Step1导入N个训练样本的特征，选择K个近邻数，导入X个测试样本特征；

Step2取K个近邻数作为中心点；

Step3计算测试样本与中心点之间的距离；

Step4给距离排序，取离中心点最近的测试样本，归为中心点所在的聚类中；

Step5计算聚类新的中心点，判断中心点是否稳定，若稳定则结束循环；若不稳定则继续循环以上步骤。

4. 仿真实验结果

在已划分的五大情感类中（正常性，异常性，心理性，生理性，反抗性），训练语音数据集情感划分较为准确，82.25%的测试语音数据都能够分类到准备的聚类中。对于情感库中的语音数据集（即训练集），各类情感语音中识别最为精确的是正常性哭泣，在理论上正常性哭泣时婴儿的哭声声音响亮具有节奏感，实验分析中婴儿正常性哭泣节奏感强，频幅高即响亮，如图1所示。其次是心理性啼哭、异常性啼哭的识别率高于生理性啼哭和反抗性啼哭。

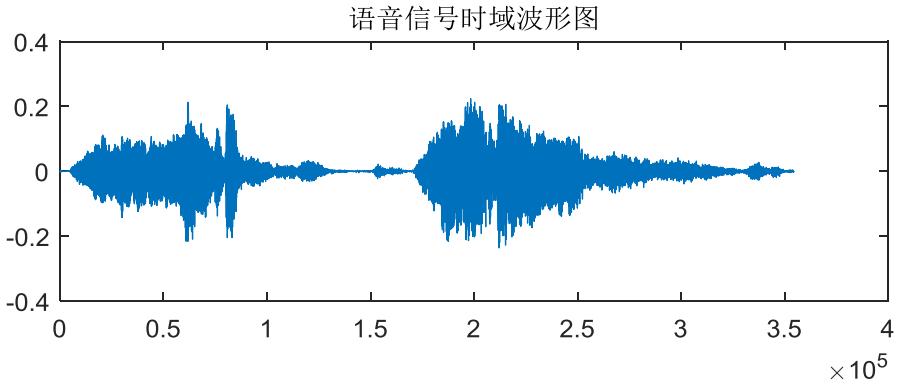


图 1: 正常性啼哭的样本语音时域波形图

为了促进婴儿情感需求的识别，建立一种语音信号情感识别的分类模型，即基于语音信号识别的KNN分类模型。在该模型的基础上，本文通过比对标准库的情感划分标准，参照FAU AIBO语音库（一种自然语音数据库）的建立方法，构建婴儿啼哭的语料库和情感库。在数据处理流程上，本文首先利用MATLAB对于语音数据进行预处理，其次利用MFCC法提取婴儿啼哭信号特征参数，然后采用KNN分类算法对于语音信号进行情感识别分类，继而得出实验结论并对实验结论进行分析。

5. 结论

本文采用较为基础的处理方法，基于语音信号识别对婴儿情感需求做出分析和研究。本文的实验是从育婴事业理论基础和实践方面出发，改善父母在育婴上的经验不足和精力有限的问题。该研究有助于成人提高工作效率，开展正确良好的育婴事项，及时满足婴儿的情感需求，获得满足的情感体验。

本文的实验在很多方面有很多不足和缺点。在分类算法方面，本文采用的KNN算法具有一定的缺陷，速度较慢，效率不高。后期可以采用最先进的深度学习算法进行分类学习，或者可以考虑减少特征值个数来提高分类速度。在选择特征参数时，只保留区别性能强的特征作为特征参数；或者可以考虑使用KD树来进行情感识别的分类学习，提高分类器的效率。在婴儿啼哭情感库和语音库的建立方面，语音的采集样本数量不是大样本，会产生数据差异。婴儿啼哭语音片段的采样范围有限，会对婴儿啼哭情感识别准确率有所影响。要扩大语音库的采样范围，提高语音信号的涵盖率从而提高语音识别的准确率。

参考文献

Batliner, Anton & Steidl, Stefan & Noeth, Elmar.(2008).Releasing a thoroughly annotated and processed spontaneous emotional database: the FAU Aibo Emotion Corpus. Proc. of a Satellite Workshop of LREC 2008 on Corpora for Research on Emotion and AffectMarrakesh

Kheddache T. (2014).Resonance frequencies behavior in pathologic cries of newborns. J Voice.29:1-12.

Lederman D.(2010). Estimation of infant's cry fundamental frequency using a modified SIFT algorithm. arXiv: 1009.2796v1 [cs. SD],1-5.

Messaound A, &Tadj C.(2010). A cry-based babies identification system. Lect Notes Comput Sci.6134:192-9.

Orlandi S, Manfredi C, Bocchi1 L, &Scattoni ML. (2012).Automatic new- born cry analysis: a non-invasive tool to help autism early diag-nosis. Proceedings of the 34th Annual International Conference of the IEEE EMBS San Diego, California USA,2953-2956.

Okada Y, Fukuta K, &Nagashima T. (2011).Iterative forward on cross- validation approach and its application to infant cry classification. Proceedings of the International Multi Conference of Engineers and Computer Scientists.

Pramanik, A. , & Raha, R.. (2012). De-noising/noise cancellation mechanism for sampled speech/voice signal. International Conference on Wireless & Optical Communications Networks. IEEE.

Patil HA.(2009).Infant identification from their cry. Proceedings of the 7th International Conference on Advances in Pattern Recognition,107-10.

Petroni M, Malowany AS, Johnston CC, &Stevens BJ.(1995). Classification of infant cry vocalization using artificial neural net- works (ANNs). Proceedings of the International Conference on Acoustics, Speech and Signal Processing, 3475-3478.

Petroni M, Malowany AS, Johnston CC, &Stevens BJ.(1995). Identification of pain from infant cry vocalizations using artificial neural networks (ANNs). Proceedings of the SPIE's Symposium on OE/Aerospace Sensing and Dual Use Photonics,729–738.

Reyes-Galaviz OF, Cano-Ortiz S, &Reyes-Garcia C. (2009).Evolutionary- neural system to classify infant cry units for pathologies identification in recently born babie. In: Proceedings of the 8th Mexican International Conference on Artificial Intelligence (MICAI-2009), Guanajuato, Mexico,330–335.

Sun,Sheng Fu,&Fu Wang. (2019).Decision tree SVM model with Fisher feature selection for speech emotion recognition. EURASIP Journal on Audio, Speech, and Music Processing,2019(1).

Stefano Rovetta,Zied Mnasri,Francesco Masulli,&Alberto Cabri. (2019).Emotion recognition from speech signal using fuzzy clustering. Proceedings of the 11th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2019).

Sauter, D. A. (2013).Children's recognition of emotions from vocal cues.British Journal of Developmental Psychology.

Sheinkopf SJ, Iverson JM, Rinaldi ML, &Lester BM.(2012). A typical cry acoustics in 6-month-old infants at risk for autism spectrum disor- der. INSAR Research Article, Autism Research,1-9.

Steidl, Stefan & Batliner, Anton & Seppi, Dino & Schuller,& Björn. (2010).On the Impact of Children's Emotional Speech on Acoustic and Language Models. EURASIP Journal on Audio, Speech, and Music Processing. 783954. 10.1186/1687-4722-2010-783954.

Zabidi A, Mansor W, Khuan LY, Sahak R, &Rahman YA. (2009).Mel- frequency cepstrum coefficient analysis of infant cry with hypo- thyroidism. Proceedings of the5th International Colloquium on signal processing & its applications,204-208.

刘振焘,徐建平,吴敏,曹卫华,陈略峰,丁学文,郝曼,谢桥.(2018).语音情感特征提取及其降维方法综述.计算机学报,41(12):2833-2851.

沈再阳.(2017,9 月).MATLAB 信号处理.北京:清华大学出版社.

赵力.(2009,6 月)语音信号处理.北京:机械工业出版社.

作者简介：

	**李琪** 女，南京工业大学浦江学院计算机与通信工程学院信息管理与信息系统专业1601班本科生，获2019全国大学生创意创新大赛初赛一等奖，主要研究方向，算法识别和机器学习等。
	**徐平平** 女，南京工业大学浦江学院计算机与通信工程学院院长，东南大学移动通信国家重点实验室通信与信息系统学科教授博导，主要研究方向，5G通信及移动云计算，移动自组织网络，车联网通信，物联网应用等。
	**冯倩倩** 女，南京工业大学浦江学院计算机与通信工程学院信息管理与信息系统专业1601班本科生，主要研究方向，数据分析等。

基金资助项目：

[1]南京工业大学浦江学院 2019 教改课题：基于 OBE 理念的“云大物智”应用型人才培养模式研究(No:2019JG002Z)
[2]南京工业大学浦江学院2019人才培养工程计划项目课题：ICT人才科学思维能力培养体系研究与实践(njpj2019-01-03)

基于云服务的桥梁微变感知系统研究
RESEARCH ON BRIDGE MICRO CHANGE SENSING SYSTEM BASED ON CLOUD SERVICE

申悦，徐平平
Yue Shen, Pingping Xu

南京工业大学浦江学院 计算机与通信工程学院
Nanjing Tech University Pujiang Institute
*Corresponding author: 申悦、徐平平
E-mail: 1055355163@qq.com, xpp@seu.edu.cn

摘要

桥梁作为交通工程的咽喉，是社会经济命脉的重要环节。任何一座桥梁失去其原有效用，都将导致整条公路的交通瘫痪。随着我国综合国力的增长，桥梁工程作为基建的重要部分，近年桥梁工程遍地开花，桥梁安全问题成为我国需要考虑的重要问题之一。本文就桥梁安全问题介绍了基于云服务的桥梁感知系统，用于监测桥梁健康状态。系统采用机器视觉技术法包括前端数据获取、数据单元传输、云端处理、终端查看等模块。云端对微变位移成像使用机器视觉算法处理，解算出桥梁相对偏移与绝对距离，绘制出桥梁微变的轨迹曲线，并建立预警系统与预测模型，以微服务为架构的后台将所得结果发送监控中心的 web 端以及手机端。若超出桥梁设定安全范围，则将预警信息发送至 web 端和手机端。

关键词：云服务，桥梁健康，监测系统，视觉感知

Abstract

As the throat of traffic engineering, bridge is an important link of social and economic lifeline. If any bridge loses its original effective use, the traffic of the whole road will be paralyzed. With the growth of China's comprehensive national strength, bridge engineering as an important part of infrastructure construction, bridge engineering is blooming everywhere in recent years, and bridge safety has become one of the important issues to be considered in China. This paper introduces a bridge awareness system based on cloud service for monitoring the health status of bridges. The system includes front-end data acquisition, data unit transmission, cloud processing, terminal viewing and other modules. The service end adopts the micro service architecture, with a variety of data analysis, early warning, trend prediction and other functions. Under the Internet of things, the bridge is monitored in an integrated way to ensure the operation safety of the bridge. It meets the requirements of the bridge detection system, such as short-term high efficiency, reliable evaluation, simple equipment, fast installation, etc.

Keywords: Cloud service, bridge health, monitoring system, micro service

1 引言

2020 年 10 月 10 晚，无锡高架桥侧倾，除影响交通以外，该桥覆压 3 部车辆，造成了 3 死 2 伤的悲痛惨剧。这一事件引起了全国各地的广泛关注，并就“桥梁安全”问题进行了激烈讨论。

2020 年 5 月 5 日下午，广东虎门大桥异常晃动。有关部门第一时间进行交通管制，大桥的封锁严了影响交通，同时引起了民众的担忧。5 月 7 日晚，大桥仍在晃动，且原因仍无法定论。5 月 8 日早，新京报发文：建议完善监测体系。

以上可知，近年来国内外经常发生桥梁坍塌等事故，一方面是一些不可控因素；还有一方面，是因为随着时代的进步，桥梁建设的结构也逐渐复杂，桥梁众多导致监测情况不及时，不准确，从而为了防止此等情况经常发生，影响国民经济发展，本文介绍一个基于云服务的桥梁微变感知系统，通过云端通信实时监控保证桥梁的健康状况与安全。因此，为能够在桥梁使用中发现异常现象，排除存在安全隐患的因素，并为桥梁的维修养护提供科学依据，以适时采取合理的维修养护方法，延长桥梁的使用寿命、提高其承载能力，降低桥梁的维护费用，或避免拆除重建导致的资源浪费；考察桥梁是否能满足将来运输量的要求；为下一步桥梁设计、规范修订和完善等提供数据支撑，需设计一个实时高效、精密可靠、设备简便、安装快速等要求的桥梁安全健康检测系统。

2 研究目的

开发桥梁安全健康感知系统，具备远程监测、实时预警、数据分析预测、桥梁管理、安装快速等功能。实现在桥梁使用中发现异常现象，排除存在安全隐患等因素，除此之外为桥梁的维修提供科学分析依据，以及时安排桥梁维修养护，使桥梁的使用寿命更加持久、提高其承载能力，降低桥梁的维护费用，或避免拆除重建导致的资源浪费；考察桥梁是否能满足将来运输量的要求；为下一步桥梁设计、规范修订和完善等提供数据支撑。

3 文献综述

改革开放以来，我国经济实力以及科研技术水平得到了质的飞跃，综合国力与国际地位也在不断提高。公路交通是我国经济发展以及联系各个地区的重要经济命脉和基础设施，随着我国社会主义经济的不断发展，使得道路工程的规模和数量都在与日俱增,而在此各界对桥梁工程的建设安全有着越来越高的重视。[1] 所以研究桥梁健康感知系统是必要的工作。

本文对桥梁感知系统进行研究，实现基于云服务的桥梁微变感知系统，采用视觉感知，解算出桥梁相对偏移与绝对距离，绘制出桥梁微变的轨迹曲线，并建立预警系统与预测模型，后台将所得结果发送监控中心的 web 终端以及手机端。

4 研究方法

现存在的桥梁监测方法存在很多种，大类为接触式测量法、非接触式测量法。

4.1 接触式监测方法

4.1.1 百分表法

百分表法，是用百分表的仪器去测量桥梁微变的方法。它是通用长度测量工具，是由精确度很高的齿条或者齿轮机构构成的构成的。其具体操作通常是使用坚硬的连接杆，与桥梁待测点和桥底的不动层面进行连接，这时将百分表仪器放入连接杆中间，从而可以达到测

量桥梁微变结构的目的。百分表测量法的测量偏差本身无法完全消除，其因设备结构形式、设备精度、电机转速、功率以及联轴器等形式不同而存在差异。[2]这种测量方式只适合测量干枯河流或小型桥梁中，并且测量之时必须在桥下搭设固定支架，这对测量工作造成了巨大的行动困难。

4.1.2 倾角仪法

倾角仪法指用倾角仪测量仪器测量桥梁微变的方法。利用倾角仪测试方法，能监测桥梁变形情况，并且测量时不受桥墩变形或其他环境的影响。从而可为桥梁的健康监测提供可靠的基础数据。[3]

4.2 非接触式微变测量方法

4.2.1 精密水准仪法

数字水准仪，又称为电子水准仪，是一种集光学、机械、电子、数字图像处理和计算机等多种技术于一体的高科技测绘仪器，具有测量精度高、速度快、操作简便、人工干预少、劳动强度低、自动化程度高和内外业一体化等优点。[4]

精密水准仪测量桥梁微变，是将精密的水准仪器安装在桥墩或者桥头上，准备工作完毕后，将水准塔尺安置在需待测点位，然后通过光直线原理进行微变测试。此方法较为简单精度较高，在桥梁安全检测里面是较为普遍的方法，但此方法不能远程实时监测并且只能人为现场测量，同时人为或者环境因素对桥梁安全检测存在很大影响，因此此方法不适合使用在需长期使用的健康监测系统中使用。

4.2.2 GPS 卫星定位系统

GPS 卫星定位系统是利用卫星定位待测点，直接读出测点的坐标，测量所得坐标与初始坐标之差即为所求得微变。此方法可以同时测定点的三维坐标，测量时无需人为现场监控，并且监测网布设自由、方便，可以进行全天观测，但其方法测量精度不为准确，不适用小型桥梁，只能用于一线城市大型桥梁，并且此方法的成本极高，短时间内很难全面普及，不适用小型企业使用。

4.2.3 机器视觉技术法

机器视觉技术法是目前许多行业测量各种物品的流行技术，尤其是桥梁微变的测量中。该方法是在测点上用 LED 灯进行标记，然后在桥头其他固定位置运用高像素的工业数码相机对测点进行拍照，通过相关软件处理照片上 LED 等的位置变化得到桥梁的微变值。该测量方法集成了无需人工现场监控、高精度、实时监测、无线连接以及成本较低等优势。本文章研究校企合作项目《基于云服务的桥梁微变感知系统》这个实例，通过对整个系统的研究，一起对机器视觉技术法的其中一种算法进行解说。

5 研究结果

5.1 系统设计

图1为系统框架图；

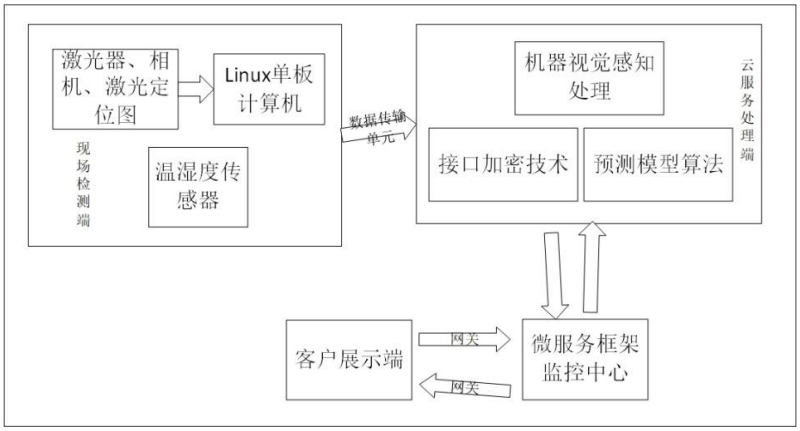


图 1: 系统框架

硬件系统设计上，在现场监测端引入高速Linux单板计算机机为核心的控制系统，摄像头进行视频流收集图像。同时，基于DHT22温湿度传感器，设计温湿度检测模块获取现场的温湿度数据，通过数据传输单元到云端服务器，从而保证采集的视频图像与温湿度数据实时传送到控制中心。

在软件设计上，云端获取光斑彩色图像后，使用滤波、去噪、图像彩色模型等算法进行机器视觉预处理，采用二维定位技术获得精确的桥梁微动信息。温湿度数据和微动信息输入到服务端，服务端以微服务框架为架构。监控中向服务端发送数据请求后，该请求通过网关到达微服务应用。接着微服务会根据对应请求调用数据库中的相应数据，得到数据后将数据通过前端传输给用户。图2为总体流程图。

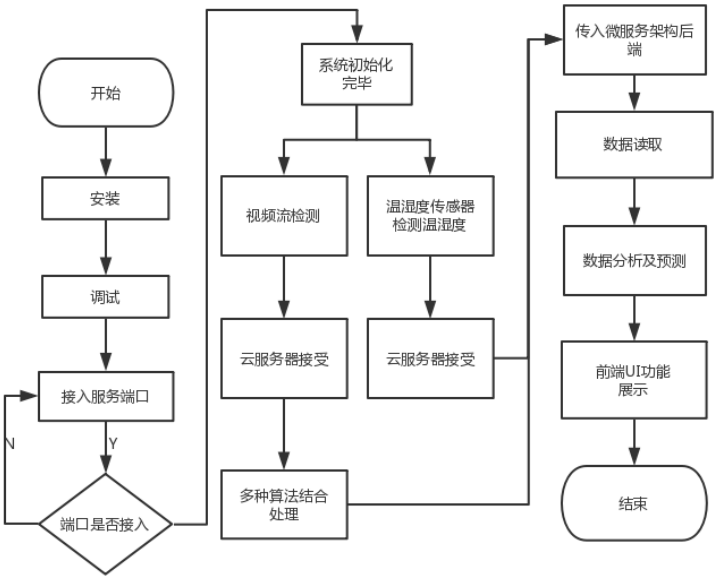


图 2: 系统设计流程框图

5.2 系统功能设计

本文章提到的桥梁健康监测系统将系统分为如图3所示的四大模块，并且将这四部分的子系统细分为各子功能模块以及具体的系统功能结构。

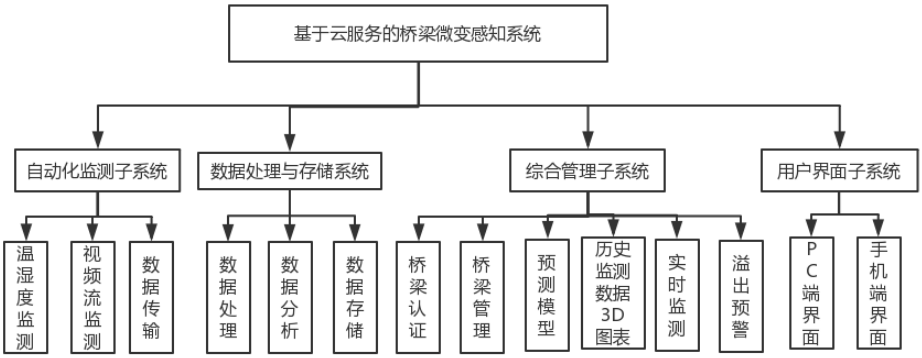


图 3: 系统功能结构图

5.3 测量坐标图校准

本作品考虑到因激光对准中心点带来的困难性以及误差性较高，因此系统采用的是中心点设计，所以安装设备时无需对其进行激光中心点校准，既用基准点代替中心点。启动激光后，通过100组数据测试中心点位置，当激光点与中心点的真实长度误差不超过0.001mm，激光点所在位置就为基准点。

5.4 温度传感器设计

温湿度传感器采用单总线通信，实时收集和处理外部环境的温湿度数据，通过HTTP协议上传监控中心，并通过前端展示给用户。图4温湿度监测流程图：

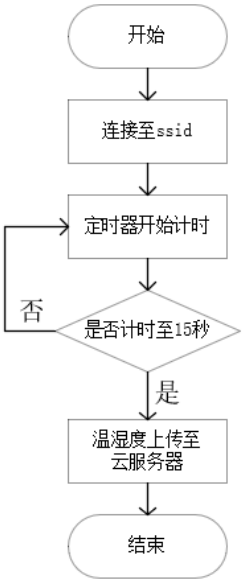


图 4: 温湿度检测流程图

5.6 机器视觉感知

文章研究的检测系统重点使用机器视觉感知算法，视频流采集图像，传入云服务器，云端采用机器视觉算法对图像识别方面进行计算，如色彩空间转换、高斯滤波去噪、Fast算法特征检测等。设计的激光提取技术与图像预处理流程如图5所示。

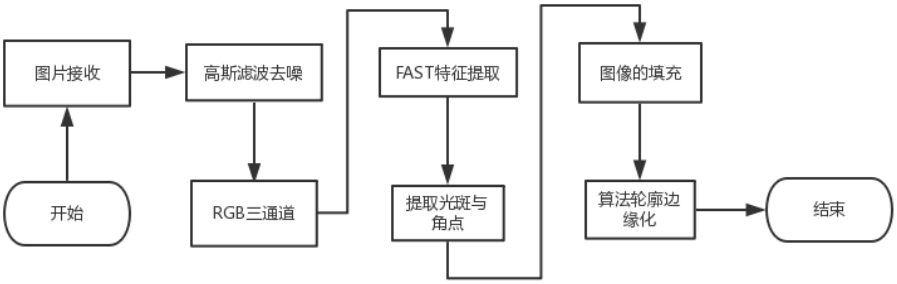


图 5: 机器视觉感知法流程图

5.7 趋势预测基于时间预测(ARIMA 模型)的建立

ARIMA模型，称为单整自回归移动平均模型，又称B-J模型，是由博克斯（Box）、詹金斯（Jenkins）于1970年首次提出的，它是拟合非平稳时间序列的常用随机时序模型，是一种精度较高的时序短期预测法。[5] 桥梁意外事故有时是瞬间即逝，因此本作品建立基于时间预测模型（ARIMA模型）实现桥梁的趋势预测。

5.8 后台与接口设计

为了增强桥梁监测系统的安全性与扩展性以及容错率。基于云服务的桥梁微变感知系统后台采用如图6设计的微服务微架构，解决了平台功能使用重复率低、容错率低、健壮性低以及模块化部署困难等问题。并在此基础上使用OAuth2.0安全协议增加了系统网关的安全性。在接口设计上采用RSA和AES混合加密算法实现前后端数据的安全传输。

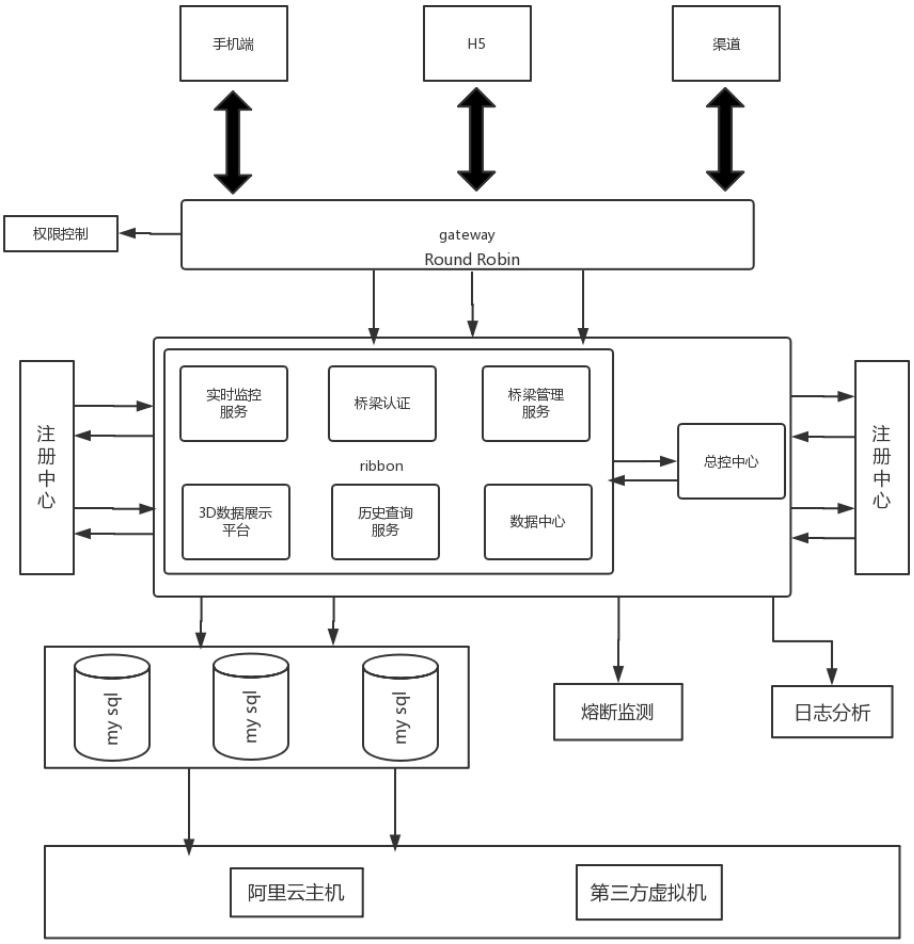


图 6: 微服务架构设计图

5.9 负载均衡算法

负载均衡算法包括动态、静态两种。静态负载均衡算法按照固定比例分配任务，如轮询算法、加权轮询算法[6]；本文采用的是完全轮询的算法基础上进行扩展，响应速度越快的服务器选择权重越大，越容易被选中。图7为系统负载均衡算法方案图。

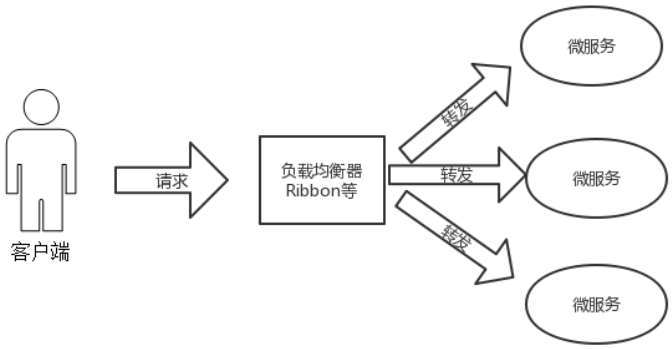


图 7: 系统负载均衡方案

5.10 系统界面实现图

5.10.1 用户登录页面设计

图8为用户登录界面，Web端用户需在登录界面输入对应用户名及密码，经过后台验证通过后，方可进入系统。

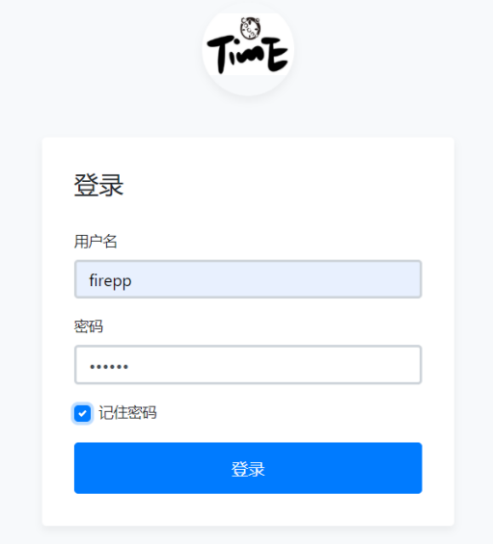


图 8: Web 登录页面

5.10.2 系统主界面设计

用户登陆后进入系统主界面，系统主界面模拟地图方式展示了所有接入系统的桥梁分布信息图，用户可以在桥梁分布信息图上点击所需查看桥梁的具体信息。图9为Web端主页面效果图。

图 9: Web 系统主界面

5.10.3 实时预览

图10为实时预览模块效果图，在实时预览这个模块上，用户可以自行选择需检查的桥梁进行实时预览，系统会根据处理流程实时上传图片以及桥梁数据上传至此界面。

图 10: 实时预览界面

图11为管理员接收预警信息画面，当监测数据超过桥梁设定级别范围时，管理员手机会实时接收预警信息对其提醒。

图 11: 桥梁预警信息功能

5.10.4 历史画面查询

在历史画面查询模块中，可以自定义是否需要打开时间范围区域选择功能。窗口可以选择单、双、四窗口模式进行历史画面查询同时窗口下有一个时间轴功能，用于进行具体时间具体画面的跳转，拖动时间轴，窗口画面随时间跳转对应历史画面与数据，图12为历史画面功能界面。

图 12: 历史画面查询界面

5.10.5 历史监测模块

在历史监测模块，主要展示监测设备的历史数据，可以通过选择桥梁、时间以及展示方式来展示历史监测数据。图13、图14、图15分别为表格、3D折线、饼图方式展现的数据及数据分析。

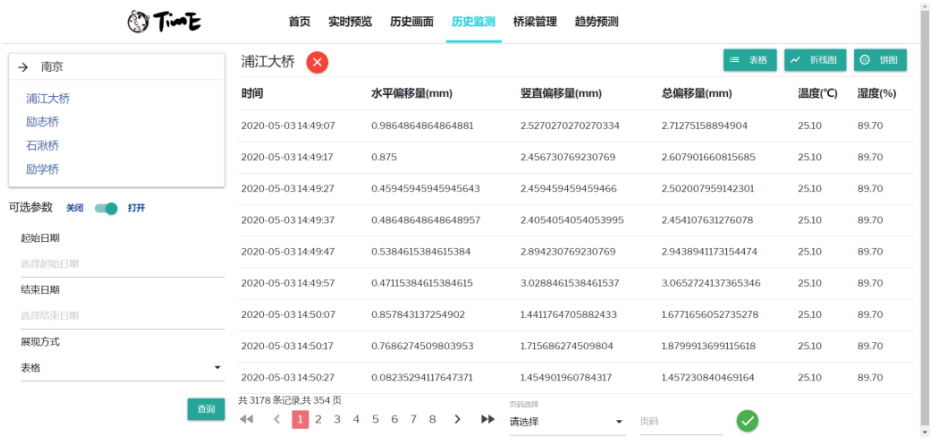


图 13: 历史表格展示界面

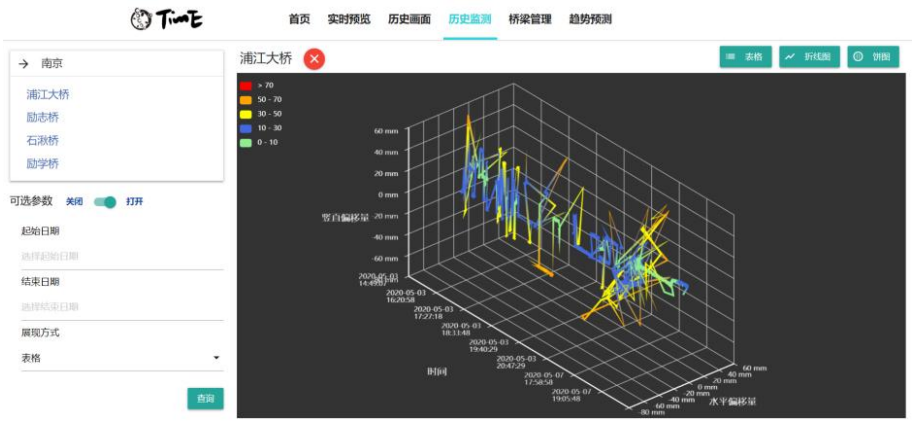


图 14: 3D 历史折线数据图

图 15: 监测圆饼图分析界面

5.10.6 桥梁管理模块

用户在桥梁管理模块可以进行桥梁的增删改查，并且可以自定义桥梁预警级别、级别范围、颜色、预警方式。图16、图17、图18、图19分别为桥梁管理浏览界面、修改桥梁界面、预警管理界面、桥梁预警设置修改界面。

图 16: 桥梁管理模块界面

图 17: 修改桥梁界面

图 18: 预警管理界面

图 19: 桥梁预警信息修改界面

5.10.7 趋势预测模块

图20为趋势预测模块，用来展现通过收集的历史数据经过算法处理得到的未来趋势变化折线图。

图 20: 桥梁变化趋势界面

5.11 手机端设计

在本次设计中，鉴于实时监控与预警功能需要的必要性，在技术上解决了图像传送与接收后设计了手机网页端来重点来实时跟踪显示桥梁监测结果。从手机端中可以实时查看桥梁健康情况，从现场监测端实时传回监测数据，对桥梁进行实时监控。管理人员在现场安装设备或维修设备时可以实时添加或修改设备信息。

(1) 用户登录成功后，进入手机首页面，如图21所示。

图 21: 手机端首页界面

(2)设计手机端最重要目的为实时预览，解决web端有时不便于实时监测，图22为实时预览界面，展现了现场实时上传的图片以及二维偏移量和温湿度数据。图23为历史监测界面,手机端以圆饼图的方式展现指定桥梁预警级别比例分析。

图 22: 手机实时预览界面图

图 23: 手机端历史监测界面

(3)为了更加方便管理员对桥梁进行管理，手机端设置桥梁管理功能，对桥梁随时进行增删改查的操作，以及预警管理的操作，图24、图25、图26分别为桥梁管理中的主页面、桥梁信息修改页面以及桥梁预警设置信息展示页面。

图 24: 桥梁管理主页面图

图 25: 桥梁信息修改图

图 26: 桥梁预警展现界面

6 项目总结

随着我国在基本建设领域的迅速发展，桥梁工程作为基建的重要组成部分，建设数量越来越大，桥梁结构变得越来越复杂，怎样保证桥梁的安全、监测桥梁的健康状态成为了突出而又严峻的问题。因此本作品设计和实现了一套基于云服务的桥梁微变感知系统，应用现代化的结构分析技术、感知技术、视觉感知处理技术、数理统计分析技术、现代网络通讯通信技术，建立集环境监测、结构监测、数据分析处理、预警与预测等为一体的桥梁健康监测系统。本作品测试通过，作品稳定性高，安全性高等性能稳定，表明作品实用性高，方案完整，相比于现有桥梁检测系统，优势较多。

根据对本文章所提系统的研究介绍，也可以将此研究方法运用于其他行业或者其他应用领域。

参考文献

[1]王丽博.变形监测技术在桥梁监测中的应用[J].工程技术研究,2019,4(04):93-94.
[2]曾程.百分表对中找正表架垂直挠度的控制方法[J].现代制造技术与装备,2018(12):115-117.
[3]兰日清,丰彪,匙庆磊.倾角仪法测量高铁桥梁动挠度研究[J].世界地震工程,2020,36(01):56-62.
[4]吴迪军.基于数字水准仪的桥上精密水准测量方法[J].铁道勘察,2019,45(01):5-8.
[5]王爽,汪海飞.基于 ARIMA 模型的海南省国内生产总值预测[J].对外经贸,2020(04):44-46.
[6]郭秀才,张悦,贺耀宜.负载均衡算法在智慧矿山软件平台中的应用[J/OL].工矿自动化:1-5[2020-05-19].

基金资助项目：

[1]南京工业大学浦江学院 2019 校企签约项目：于视频分析的目标微变监测物联网应用系统研制（PJXQ-2019-01）

作者简介：

申悦 女，南京工业大学浦江学院-计算机与通信工程学院-软件工程（嵌入式） 1901 班本科生，计通学院大学生科创成员，获得华为云计算 HCIA-Cloud 证书。

徐平平 女，南京工业大学浦江学院计算机与通信工程学院院长，东南大学移动通信国家重点实验室通信与信息系统学科教授博导，主要研究方向，5G 通信及移动云计算，移动自组织网络，车联网通信，物联网应用等

人脸识别算法在情绪管理中的应用研究
RESEARSH ON THE APPLICATION OF FACE RECOGNITION ALGORITHM IN EMOTION MANAGEMENT

刘轩绮 [1]，徐平平 [2]，王树军 [3]
Xuanqi Liu[1], Pingping Xu[2], Shujun Wang[3]

南京工业大学浦江学院 计算机与通信工程学院
Nanjing Tech University Pujiang Institute
[*]**Corresponding author:** 刘轩绮，徐平平，王树军
E-mail: 1138732172@qq.com, xpp@seu.edu.cn, 176594841@qq.com

摘要

近年来人脸识别技术已经发展成熟，并且广泛应用于各种领域。随着移动通信技术的不断进步，5G时代已经到来。这就意味着车联网技术与远程医疗技术将逐步走进人类的世界，而这两项技术的发展都离不开对人类情绪的研究。因此，将机器赋予人类的感情色彩成为了不可或缺的研究方向。本文主要结合人脸识别算法，通过对人脸静态图像的表情分析来判断此人当下的情绪状态。本文使用MATLAB软件，首先对图像进行灰度化和灰度化调整，并找出人脸区域，从而实现人脸检测。其次，采用主成分分析法（Principal Component Analysis，PCA）对检测出的人脸图像进行特征提取，最后分析人脸图像的表情所反映的情绪。为了整个系统操作的完整性，我们还结合MATLAB的图形用户接口（Graphic User Interface，GUI）设计了人脸表情分析的演示系统。系统的测试结果表明，本算法能够根据人脸静态图像识别出情绪状态。

关键词：主成分分析法，特征提取，表情分析，情绪判断，GUI

Abstract

In recent years, face recognition technology has been developed and widely applied in various fields. With the continuous progress of mobile communication technology, the era of 5G has arrived. This means that the technology of Internet of vehicles and telemedicine technology will gradually enter the world of human beings. The development of these two technologies is inseparable from the study of human emotions. Giving the human emotion to the machine has become an indispensable research direction. In this paper, the face recognition algorithm is used to judge the person's emotional state through facial expression analysis. In this paper, we use MATLAB software to first adjust the grayscale and gray degree of the image, and find out the face area, so as to achieve face detection. Principal Component Analysis (PCA) is used to extract the features of the detected face images. Finally, the emotions reflected by the facial expression are analyzed. For the integrity of the whole system operation, we also design a facial expression analysis demo system based on Graphic User Interface (GUI) of MATLAB. The test results of the system show that the algorithm can recognize the emotional state from the static face images.

Keywords: Principal Component Analysis, feature extraction, expression analysis, emotion judgment, GUI

引言

社会在进步，科技在发展。随着人工智能时代的到来，人类的生活模式发生了翻天覆地的变化，智能机器力量正不断取代简单的劳动力，人机交互模式的需求也在不断壮大。目前，人工智能的研究主要还是集中在模拟人的智能，而人工心理理论则是对人的心理活动更全面地再一次的人工智能的实现。通过识别人脸表情进行情绪分析就是人工心理理论的重要组成部分。如果我们能够将机器赋予人的情感，使机器人能够识别人类的情绪发展状况，自觉处理与人类相处的环境变化，那将是人类与机器在共处关系上的质的飞跃，也使得机器能够更好地服务于人类。

人脸表情分析实际上是结合图像处理，机器视觉，心理学，生理学，神经网络，人工智能等多方面学科于一体的强大研究体系。简单来说就是通过机器获得人的面部特征，根据人的思考方式对人脸表情加以识别与分类，如高兴、伤心、讨厌、惊讶、愤怒、恐慌等。面部表情分类在机器视觉分析中具有相当关键的位置。而这些表情都从一定程度上反映了人的情绪，如快乐、悲伤、恐惧、厌恶、惊讶等情绪。随着5G时代的到来，未来5G的商业化运用将主要集中在远程医疗和车联网行业，这就更是加强了机器对人类情绪分析的要求。目前5G技术的低时延和高速率已经在一定程度上满足了车联网和智慧医疗的需求，如果我们可以将人脸表情分析技术引用其中，比如无人驾驶时，摄像头可以根据人脸表情的变化及时改变车速，从而避免危险的发生；比如在车内音箱系统里添加人脸表情分析软件，可以根据人类的情绪变化随时转换音乐的播放风格；比如在远程医疗中，医生可以根据机器内反映出的病人情绪变化及时进行治疗；比如远程手术中，机器可以及时识别医生或者病人的表情变化，让远程的专家可以及时改变手术方案等等。这些应用都对未来的移动通信技术的发展提供了明确的发展方向，更是极大地促进社会的进步，真正的实现4G改变生活，5G改变社会。因此，人脸表情分析课题的研究具有深远意义。

1. 人脸表情分析方法

人脸表情分析系统主要分为三个步骤：1、通过对获取的图像进行灰度化处理，从而进行人脸检测，检测出图像中人脸的区域；2、对检测出的人脸图像进行脸部特征提取；3、将提取后的结果与相关MATLAB算法结合，进行表情分析，进而判断该表情所反映的情绪。

整个人脸表情分析系统如图1所示：

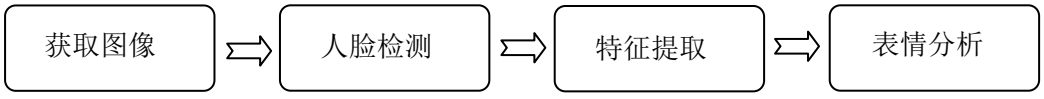


图1：人脸表情分析系统

1.1 人脸表情识别一般方法

在对人脸进行特征提取和表情分类之前，需要将图像进行灰度化和几何归一化处理，灰度化是通过将图像灰度化和灰度化增强或者减弱来减少光照等因素对图像的影响，使得特征提取与表情分类可以更加清晰准确。几何归一化则是由于分割后的图像大小不一致，会对特征的提取产生一定的影响，所以为了提高识别识别率，采用缩放变换将所有的图像变换为大小一致的图像。人脸表情识别一般有如下三种方法：

（1）基于几何的识别方法：这种方法主要是对人脸表情的显著特征部位进行定位，从而进行表情识别。比如对眼睛进行定位，我们可以根据已经规则好的确定人眼位置算法判断人眼大概位置，再使用二值化的人眼定位算法得到人眼准确位置，最后将分割后的图像采用双线性插值方法，将表情图像大小统一化，从而完成表情识别。

（2）基于整体的识别方法：这种方法是通过对整个人脸图像中特别的区域进行变换，获取各种表情的特征来进行识别。最主要的应用就是主成分分析法（PCA 算法），PCA 算法也是本文使用的方法。

（3）基于模型的识别算法：这种方法主要是建立精确的物理模型，根据解剖学知识确定关键特征，并且通过这些特征变化进行人脸识别。

1.2 人脸特征提取方法

人脸表情分析系统中，最主要的步骤就是对人脸图像的特征提取。因此研究特征提取算法是至关重要的一步。目前已经广泛应用的特征提取法如下：

（1）LBP 算法（Local Binary Pattern）:LBP 算法首次提出于 1994 年，这是一种描述图像特征像素点与各个像素点之间的灰度关系的局部特征的非参数算法，同时也是一种高效的纹理算法，纹理是物体表面的自然特性，它描述图像像素点与图像领域之间的灰度空间的分布关系，不会因为光照强弱而改变图像的视觉变化。

（2）主成分分析（Principal Component Analysis）:也叫主元分析。这是一种掌握事物主要矛盾的统计方法，它可以从多元事物中解析出主要影响因素，揭示事物的本质，简化复杂的问题。其最主要的过程就是将多维的数据降维，从而达到简化的效果。

1.3 表情分类方法

表情分类是整个系统最后一步中的核心内容，它的分类方式直接代表了最后的测试结果，也是影响系统能否成功运行的关键。以下是几种常用的分类方法：

（1）最近邻方法

最近邻法是一种把数据之间的间隔作为分类依据的识别方法。 1968 年，科弗和哈特最早提出了最近邻法,它在分类算法中存在十分关键的用途。较为常用的度量标准有 L1 范式，L2 范式，欧氏距离等。

（2）基于模板的匹配方法

这种方法为每个等待分类的样本创建一个模板，将等待测量的样本与每一个训练模板进行比较和分类，相似度越高，等待测量的样本与这种样本的相同度越高。它的优势之处在于运算量不大，简单并且容易实现，缺点是识别率较低。

（3）基于人工神经网络的方法

神经网络有效地使用人类分析问题和解决问题的能力，改进不足之处，并最终可以得到较好的分类结果。它的弊端是计算耗时很长，实时分类不容易达到，而且输入变化时必须再次训练神经网络，具有较差的扩充能力。BP 神经网络、多层感知器、RBF 神经网络等是表情识别中比较常用的神经网络分类器。

2. 图像处理技术

图像处理技术是人脸表情分析系统的最基础部分，图像处理的好坏会影响人脸区域的识别能力，也会影响表情分析结果的正误。本文主要利用MATLAB图形处理系统对图像进行灰度化并且进行灰度化增强。然后对处理好的人脸图像采用PCA算法，得到特征提取的数据，最后结合提取出的数据和表情分类方法进行表情识别，从而判断出录入的图像中人脸所反映的情绪。

2.1 人脸检测

本文主要研究在静态图像下的表情识别，所以首先我们需要录入一张图像，然后用 MATLAB 软件对所得图像进行灰度化和人脸位置检测，并且根据灰度化的效果采用图像调整技术使得图像更加清晰。

2.1.1 人脸位置检测

当系统获得一张静态图像时，它可能会只有一张人脸，但是更多情况下，人脸只是整个图像的一部分，甚至一张图像里有很多的人脸。因此我们在获得图像后，首先要提取出图像中人脸的位置，即人脸位置检测，从而缩小图像特征提取的范围。简单来说，人脸位置检测就相当于使用手机照相机来拍照时，手指点击某个区域时出现的红框里的对象。在这里我们是将图像灰度化并将图像中的人脸用红框标注出来。具体效果如图2：

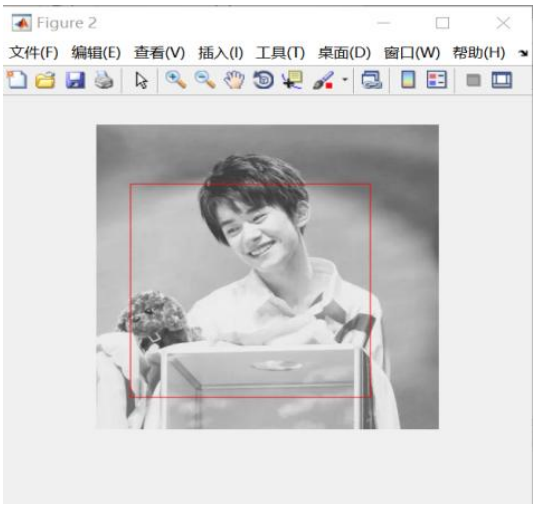


图2：人脸位置检测效果图

2.1.2 图像灰度化增强

本文主要研究在静态图像下的表情识别，所以首先我们需要录入一张图像，然后用MATLAB软件对所得图像进行灰度化，并且根据灰度化的效果采用图像调整技术。在这里我们主要使用图像灰度调整，即图像增强。

图像增强是增强图像中的有用信息，它可以是一个失真的过程，其目的是要改善图像的视觉效果，针对给定图像的应用场合。有目的地强调图像的整体或局部特性，将原来不清晰的图像变得清晰获强调某些感兴趣的特征，扩大图像中不同物体特征之间的差别，抑制不感兴趣的特征，满足某些特殊分析的需要。在这里，就是可以满足我们对人脸面部特征清晰度的要求。

图像增强的算法和效果图如下所示：

（1）图像增强算法如图3：

```
P=imread('qxi1.png');
I=rgb2gray(P);
subplot(231);imshow(I,[]);title('原始图像');
subplot(234);imhist(I);title('由原始图像得到的直方图');
J=imadjust(I,[0.15 0.9],[0 1]);
subplot(232);imshow(J);title('由新直方图得到的新图像');
subplot(235);imhist(J);title('利用imadjust调节后的新直方图');
K=histeq(I);
subplot(233);imshow(K);title('由均衡化后的直方图得到的均衡后的图像');
subplot(236);imhist(K);title('均衡化后的直方图');
```

图3：图像增强算法

（2）图像增强算法效果图如图4：

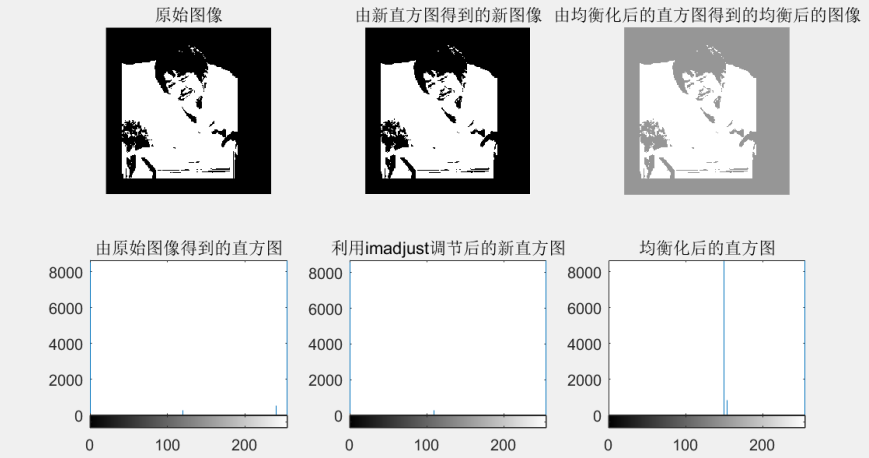


图4：图像增强效果图

由于在进行图像增强之前我们已经对录入的图像进行了人脸位置的检测和图像灰度化，所以图4中的原始图像为已经处理后的图像。

2.2 基于PCA的特征提取算法

2.2.1 PCA算法原理

PCA（Principal Component Analysis）是一种常用的数据分析方法。PCA通过线性变换将原始数据变换为一组各维度线性无关的表示，可用于提取数据的主要特征分量，常用于高维数据的降维。也就是说，使用PCA算法时，我们需要处理的数据都将会被抽象为一组数据。因此我们可以从向量内积法理解PCA算法。

内积运算将两个向量映射为一个实数。其计算方式非常容易理解，但是其意义并不明显。我们可以从内积的几何意义出发，假设A和B是两个n维向量，我们知道n维向量可以等价表示为n维空间中的一条从原点发射的有向线段，为了简单起见我们假设A和B均为二维向量，则在二维平面上A和B可以用两条发自原点的有向线段表示，见下图5：

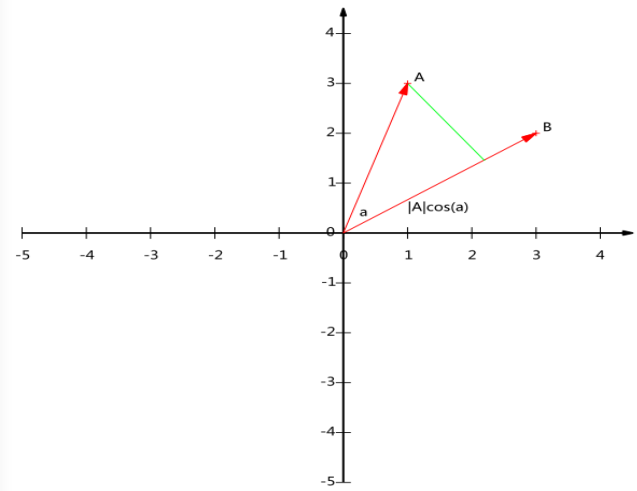


图5：内积的几何意义

现在我们从A点向B所在的直线作一条垂线，那么垂线与B的交点叫做A在B上的投影，再设A与B的夹角是a，则投影的矢量长度为$|A|COS(a)$，其中$|A|$是向量A的模，也就是A线段的标量长度。而A与B的内积等于A到B的投影长度乘以B的模。在此基础上，我们再从矩阵的角度上理解PCA算法就容易很多了。对于图像分析来说，就是将图像的几个维度降低为一维矩阵来处理。

注意，在这里我们专门区分了矢量长度和标量长度，标量长度总是大于等于0，值就是线段的长度；而矢量长度可能为负，其绝对值是线段长度，而符号则取决于其方向与标准方向相同或相反。

2.2.2 PCA算法步骤

PCA算法的主要目的就是将高维数据进行降维，而实现数据降维的步骤如下：

（1）将原始数据中的每一个样本用向量表示，把所有样本组合起来构成一个矩阵，通常

需要对样本矩阵进行处理，得到中心化样本矩阵。

（2）求样本矩阵的协方差矩阵。

（3）求协方差矩阵的特征值和特征向量。

（4）将求出的特征向量按照特征值的大小进行组合形成一个映射矩阵，并根据指定的PCA保留的特征个数取出映射矩阵的前n行或者前n列作为最终的映射矩阵。

（5）用映射矩阵对数据进行映射，达到数据降维的目的。

其中在步骤（1）中有提到中心化样本矩阵，那么什么是中心化样本矩阵呢？

中心化样本矩阵：先让样本矩阵中心化，即每一维度减去该维度的均值，然后直接用新得到的样本矩阵乘上它的转置，然后除以(N-1)即可，具体操作如下：

1 % 中心化样本矩阵，样本矩阵MySample是10*3大小的矩阵,
repmat(mean{MySample},10,1)产生一个每行都是mean(Mysample),共10行
2 X = MySample–repmat(mean(MySample),10,1); %每一维度减去该维度的均值
3 %X=X-ones(size(X,1),1)*mean(X);这样同样可以达到每一维度减去该维度的均值的效果，相比上面那种，这种更通用，因为不须每次改写repmat中的第二个参数
4 C = (X'*X)/(size(X,1)-1) %新得到的矩阵X乘上的它的转置

2.2.3 PCA实际应用：特征脸

现在，假设我们要执行面部识别，即基于带有标记的面部图像训练数据集来确定人的身份。其中一个方法就是把图像上每个像素的亮度作为特征。如果输入图像的大小是32×32，这意味着该特征向量包含1024个特征值。判断新的图像通过计算这1024维矢量与我们训练数据集中特征向量之间的欧氏距离完成。然后最小距离告诉我们正在寻找的那个人。然而，如果我们只有几百训练样本，在1024维空间中运行会变得有问题。此外，欧氏距离在高维空间中行为比较奇怪。因此，我们可以利用PCA方法通过计算1024维特征向量协方差矩阵的特征向量来降低特征空间的维数，然后映射每个特征向量到最大特征向量。因为2D数据的特征向量是2维的，三维数据的特征向量是3维的，1024维数据的特征向量是1024维。换句话说，为了可视化，我们可以重塑每个1024维特征向量到一个32×32的图像。图6展示了由剑桥人脸数据集的特征分解获得的前四个特征向量：

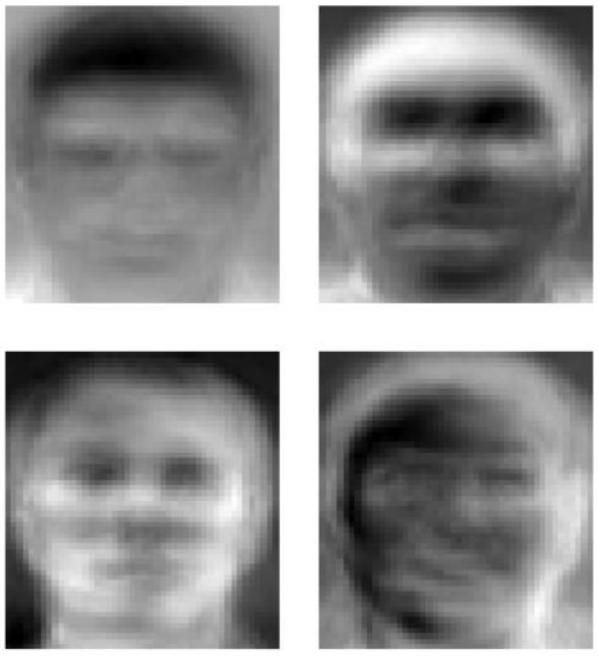


图 6：剑桥人脸数据特征向量

每个1024维特征向量可以映射到N个最大的特征向量，并可以表示为这些特征脸的线性组合。这些线性组合的权重确定人的身份。因为最大特征向量表示数据中的最大方差，所以这些特征脸描述信息量最大的图像区域（眼睛，鼻子，嘴等）。只考虑前N（例如，N=70）个特征向量，特征空间的维数大大减少了。

2.3 表情识别

我们选取经过处理后的 7 类表情，即高兴、难过、惊慌、害怕、愤怒、讨厌、平静做为标准，每种表情使用 20 个样本，试验中一共运用了 6×20 幅表情图像数据。采用 PCA 算法获取面部变化的主要因素，最后使用最近邻分类器对获得的特征进行识别与分析。

表情分类的依据如图 7 所示：

表情	额头、眉毛	眼睛	脸的下半部
惊奇	①眉毛抬起，变高变弯 ②眉毛下的皮肤被拉伸 ③皱纹可能横跨额头	①眼睛睁大，上眼皮抬高，下眼皮下落 ②眼白可能在瞳孔的上边和/或下边露出来	下颌下落，嘴张开，唇和齿分开， 但嘴部不紧张，也不拉伸
恐惧	①眉毛抬起并皱在一起 ②额头的皱纹只集中在中部， 而不横跨整个额头	上眼睑抬起，下眼皮拉紧	嘴张，嘴唇或轻微紧张，向后拉； 或拉长，同时向后拉
厌恶	眉毛压低，并压低上眼睑	在下眼皮下部出现横纹，脸颊 推动其向上，但并不紧张	①上唇抬起 ②下唇与上唇紧闭，推动上唇向上， 嘴角下拉，唇轻微凸起 ③鼻子皱起　④脸颊抬起
愤怒	①眉毛皱在一起，压低 ②在眉字间出现竖直皱纹	①下眼皮拉紧，抬起或不抬起 ②上眼皮拉紧，眉毛压低 ③眼睛瞪大，可能鼓起	①唇有两种基本的位置：紧闭， 唇角拉直或向下，张开，仿佛要喊 ②鼻孔可能张大
高兴	眉毛参考：稍微下弯	①下眼睑下边可能有皱纹，可能鼓起， 但并不紧张 ②鱼尾纹从外眼角向外扩张	①唇角向后拉并抬高 ②嘴可能被张大，牙齿可能露出 ③一道皱纹从鼻子一直延伸到嘴角外部 ④脸颊被抬起
悲伤	眉毛内角皱在一起，抬高， 带动眉毛下的皮肤	眼内角的上眼皮抬高	①嘴角下拉 ②嘴角可能颤抖

图 7：脸部运动特征具体表现

3. 研究结果：人脸表情分析软件的GUI演示

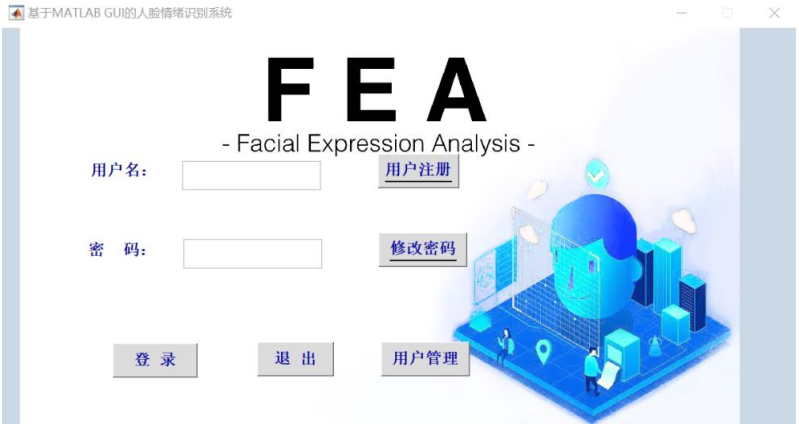
在完成了系统的整体调试之后，为了系统操作的完整性以及人们使用该系统时的保密性，我们利用 MATLAB 里面的 GUI 功能将整个系统设计成了完整的软件。这样，当我们想要进行人脸表情分析的时候，只要进入软件一步一步操作，就可以看到情绪分析的结果。同时，我们提取“人脸情绪分析”的英文首字母给软件设计了一个名字：FEA。即“Facial Expression Analysis”的首字母。

3.1 软件总体界面设计

一个完整的软件必定有自己的规则与操作规范，所以整个软件共有两个界面：登录界面和操作界面。在进入登录界面时，如果你是新用户，则需要注册一个账号，然后登录账号就可以进入操作系统。进入操作系统之后，我们首先录入一张静态图像，然后根据系统里的步骤，点击“人脸检测”按钮，就可以提取出该图像中人脸的位置，然后点击“表情分析”按钮就可以看到表达情绪的词，这个词就代表了该图像中人脸所反映的情绪。

3.2 GUI演示结果展示

(1)首先进入登录界面：



在这个界面中输入自己的账号即可登录系统。

（2）登录账号后就会进入操作界面：

此时，点击“进入系统”按钮就可以进入系统内部，根据界面中的按钮显示点击相应按钮即可进行人脸表情分析。在这里，我们以“笑脸”图像为例：在界面中点击“LoadImage”按钮，将自己想要分析的图像录入进系统，然后点击“人脸检测”按钮，就可以得到图像中待识别的人脸图像。整体效果如图 8 所示：

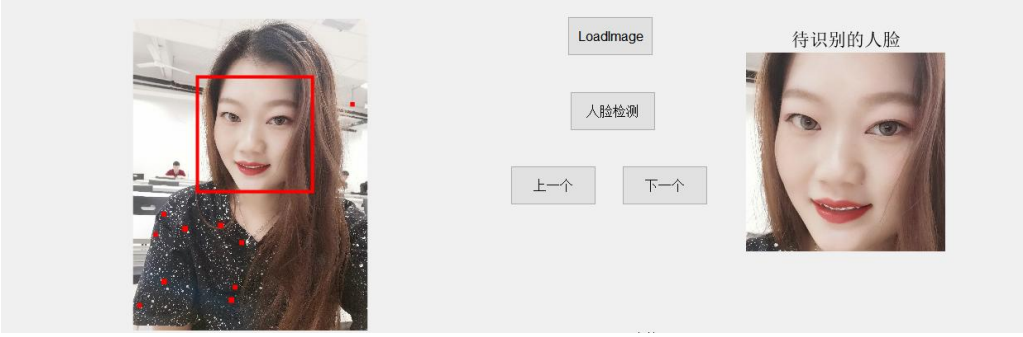


图 8：系统人脸检测结果展示

在图 8 中我们可以看出左边的图像为录入的静态图像，右边方框里就是图像中待识别的人脸图像，也就是进行表情识别的图像。

完成人脸检测后，就进入表情分析的环节了。此时我们只要点击“识别”按钮就可以进行表情识别与情绪分析了，具体效果如图 9 所示：

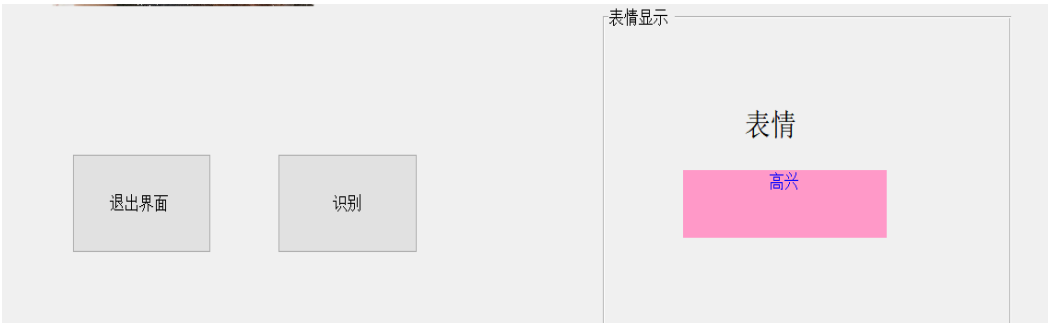


图 9：表情分析结果显示

从图 9 中我们可以看到表情显示一栏中，粉红色方框里出现了“高兴”一词。这就说明了对于刚刚录入的笑脸图像而言，图像中的人脸表情所表达的情绪为高兴。

最后点击“退出界面”按钮就可以退出整个系统软件了。

结论与讨论

本文将人脸识别技术应用于对人类的情绪管理中，不仅适应了当代智能时代的发展，更加促进了移动通信技术在各领域的商业化运用。研究人脸表情所反映的情绪，将机器赋予人的感情这一研究使得未来机器人与人类和平共处成为了可能。本文主要简述了MATLAB图像处理技术，讨论了人脸特征提取方法、表情分析的相关算法等，以及展示了GUI结果演示。其中重点探讨了以下内容：

1、图像预处理技术。我们通过将获得的图像进行灰度化处理和灰度化增强来获得图像中的人脸部位，从而避免因光照等因素带来的影响，使得特征提取时可以更加清晰方便。文章中也展示了灰度化增强的效果和算法。

2、特征提取算法：PCA算法。我们在比较讨论后，决定采用PCA算法来获得特征提取的数据。这种算法被称为模式分类中最经典算法之一，它主要是将高维数据转成低维数据，从而更加简洁地获得“主分量”，减少其余信息的干扰，特征提取的结果也会更加精确。文章对PCA算法的原理和应用进行了较为详细的介绍。

3、GUI结果演示。整个项目的创新点在于最后的GUI软件设计，在完成整个系统的设计之后，我们利用MATLAB中的GUI功能，将整个系统设计成一个软件，这样，当人们使用操作整个系统时，就可以方便简洁，一站到底，并且具有一定的隐私性。文章中以“笑脸”图像为例展示了GUI软件的显示结果。

系统的测试结果表明，本算法能够根据人脸静态图像识别出情绪状态。

工作展望

本次项目研究中也有很多需要改进的地方，比如特征提取算法模块，如果我们能够结合PCA算法与LBP算法共同进行特征提取，那将会使得提取的效果更加精确完善。同时在情绪分析这一部分，如果我们可以采用更加具体的方法，分析出每个表情所表达的情绪的程度，如高兴程度为98%，那么这样的系统则会对情绪分析地更加精确，也会更加适用于各行各业。今后，我们希望可以继续深入研究，不断完善这类系统的体系。

当然，对于人脸表情分析系统本身而言，我们的初衷就是希望与时俱进，不仅仅可以扩大智能时代的发展范畴，更重要的是希望这样的系统经过改善之后可以广泛应用于未来的车联网事业和远程医疗行业，这样就可以大大地促进5G的发展速度，也是对移动通信技术的一种支持，更是表达我们对自己通信工程专业的热爱与专注。

参考文献

刘浩,韩晶.(2016).MATLAB R2016a 完全自学一本通.北京:电子工业出版社.

Hesays.(2014）.基于 PCA 的人脸识别算法.CSDN 博客.https://github.com/bdseal.

自由的小干部.(2016).基于 MATLAB 的人脸识别系统设计与仿真.百度文库. https://wenku.baidu.com/view.

Okada Y, Fukuta K, &Nagashima T. (2011).Iterative forward on cross- validation approach and its application to infant cry classification. Proceedings of the International Multi Conference of Engineers and Computer Scientists.

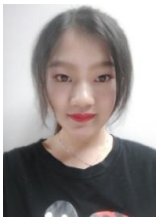
Patil HA.(2009).Infant identification from their cry. Proceedings of the 7th International Conference on Advances in Pattern Recognition,107-10.

Petroni M, Malowany AS, Johnston CC, &Stevens BJ.(1995). Classification of infant cry vocalization using artificial neural net- works (ANNs). Proceedings of the International Conference on Acoustics, Speech and Signal Processing, 3475-3478.

Kadipa Aung Myo Han.(2020).Black Gram Plant Nutrient Deficiency Classification in Combined Images Using Convolutional Neural Network.Ukrit Watchareeruetai 2020 8th International Electrical Engineering Congress(iEECON).

Petroni M, Malowany AS, Johnston CC, &Stevens BJ.(1995). Identification of pain from infant cry vocalizations using artificial neural networks (ANNs). Proceedings of the SPIE's Symposium on OE/Aerospace Sensing and Dual Use Photonics,729–738.

作者简介：



刘轩绮 女，南京工业大学浦江学院计算机与通信工程学院通信工程1702班本科生，班级文娱委员，曾担任校心理协会学术交流部部长。获得江苏省第十五届高等数学竞赛二等奖，校第四届高等数学竞赛三等奖，江宁大学城第二届手语操大赛三等奖，2017-2018年校“优秀学生会工作者”和“优秀共青团员”，2018-2019年校学业优秀奖学金二等奖，获得中小学教师资格考试合格证明，参加2019年赴泰国际暑校并获得“用于商业应用的机器学习”项目结业证书。

徐平平 女，南京工业大学浦江学院计算机与通信工程学院院长，东南大学移动通信国家重点实验室通信与信息系统学科教授博导，主要研究方向，5G通信及移动云计算，移动自组织网络，车联网通信，物联网应用等。

王树军 男，南京工业大学浦江学院计算机与通信工程学院助教，通信工程师，主要研究无线通信方向。

基金资助项目：

[1]南京工业大学浦江学院 2019 教改课题：基于 OBE 理念的“云大物智”应用型人才培养模式研究(No:2019JG002Z)
[2]南京工业大学浦江学院 2019 人才培养工程计划项目课题:ICT 人才科学思维能力培养体系研究与实践(njpj2019-01-03)
[3]南京工业大学浦江学院 2019 大学生创新创业训练计划项目课题：人脸识别算法在情绪管理中的应用研究（PJ20191390530）

应用型本科院校《通信原理》课程教学初探

RESEARCH ON TEACHING PRINCIPLES OF COMMUNICATIONS COURSE IN APPLICATION-ORIENTED COLLEGES

王树军 [1*]，徐平平 [2]，杨小琴 [3]

Shujun Wang[1*], Pingping Xu[2], Xiaoqin Yang[3]

南京工业大学浦江学院

Nanjing Tech University Pujiang Institute

[*]**Corresponding author, E-mail:** shujun.wang@qq.com

摘要

通信原理是电子信息类专业的专业核心课程，理论性强，抽象程度高，难学难教。根据应用型本科院校的特点和人才培养方案改革的要求，结合在教学过程中的体会，本文总结了一些教学方法，旨在提高学生学习的积极性和主动性，提高通信原理课堂教学的效果，更好地贯彻培养方案改革的目标和原则。。

关键词：通信原理，应用型本科，教学

Abstract

With the features of strong theory and abstract content, Principles of Communications is a specialized key course of electronic information specialties and hard to study and teach. Based on characters and requirements of application-oriented colleges and teaching experiences, this paper summarizes some teaching methods to mobilize the enthusiasm and initiative of students and to improve the effect of classroom teaching of principles of communications.

Keywords: Cooperative Communication, Relay selection, Power allocation, Energy efficiency

引言

通信原理是通信工程、电子信息工程、物联网工程等本科专业开设的专业核心课程，主要讲述通信系统的基本组成和信息传输的基本原理，对相关系统的性能进行理论分析。该课程是学习各种通信技术、各类通信网的前提和基础，具有承上启下的重要作用。如果对通信的基本原理没有一个清晰的认知，就很难对传输过程中出现的问题和故障进行分析和解决，在继续学习新的通信技术时就会遇到比较大的障碍，影响学生终生学习的能力和在本专业的长期发展。

通信学科本身的理论性非常强，通信原理课程的重要性和通信学科的这种特点分不开。自从克劳德.香农奠定了信息论基础以后，通信理论的发展日趋完善，形成了自身特有的理论框架和分析方法。通信理论的发展推动着新的通信技术不断产生，并应用到生产实践中去。同时通信理论也直接指导着工程实践，特别是随着软件无线电技术的发展，通信理论和物理设备开发实现的距离日趋缩短，可以说，通信学科是理论性和实践性结合得非常紧密的一门学科。通信原理渗透到当前所有的通信技术中，大至广播电视、卫星通信、移动通信、互联网、物联网，小至USB、SATA、DVI、HDMI等计算机设备的连接[1]。通信原理课程中基本理论、基本概念

的掌握对后续专业课程的学习，对学生毕业以后的工程实践都起着长期的、关键的支撑作用。通信原理的内容决定了其在电子信息类课程中的核心地位，研究通信原理课程教学过程中的问题具有重要的价值。

随着高等院校教学改革的深入，应用型本科院校在满足必要理论基础的条件下，逐渐强化应用性、实践性的特点。为了进一步突出学校特色，适应新的形势发展的需要，许多学校在新的人才培养方案的修订工作中，缩短了总的学时，压缩了理论课程的学时，提高了实习实践类课程的比例。理论课时的压缩，对通信原理这种理论性强、实用价值较高的课程带来了巨大的压力。如何在较短的学时内有效的让学生掌握必备的知识，形成自主学习的能力，成为该课程任课教师必须要思考和解决的问题。本文将针对应用型本科院校通信原理教学课程中遇到的一些问题，结合在教学过程中的体会，对在新形势下如何改进教学方法、提高教学效果进行探讨。

教学中遇到的问题和应用型本科院校的要求

应用型本科院校是随着我国工业化和教育产业的发展而提出来的一种新的教育类型，培养的应用型人才既要有足够的理论基础和专业素质，又能适应地方和行业发展的需要，满足我国产业升级和发展的人才需求。应用型本科院校不同于传统的、以科学为核心的精英教育，培养的人才要具备较强的实践能力，能够适应当地行业的当前需要。另一方面，又不同于以职业技术为唯一目标的职业学院，本科院校的培养要着眼于产业和人才的中长期发展，毕业生应该具备足够的理论基础和专业素质，有一定的创新能力，以适应技术快速进步和产业不断升级的需要。应用型本科人才的要求需要院校在人才的培养目标、培养过程、培养模式、管理体系的各个方面都要做出重大的变革，尤其是在课程体系方面，需要增强应用能力和实践能力方面的训练内容，调整理论课程和实践课程的比重[2]。

教学方法和教学内容的改变必须要结合学生的特点，才能实现教学目标的达成。根据一些对民办学校和独立学院学生的调查，此类院校学生基本素质较好，思维相对活跃，但也存在一些问题，如目标不够明确，自我管理能力较差，学习基础相对薄弱，学习的主动性和积极性不够，畏难情绪较为突出[3]。尤其是理工科学生，大学三年级的课业负担较重，学习倦怠的状况比较严重[4]，这些结论同作者在教学过程中的感受基本一致。新的教学方法离不开学生在课堂之上和课堂之外的积极配合，教学效率才能提高，教学效果才会改善。如何让学生主动学习、自主学习，如何调动学生学习的积极性，是应用型民办院校教学中遇到的一个较为严重问题。

通信原理以信息论为基础，理论性强，抽象程度高，对学生的数学分析能力要求较高。该课程要求的先修课程也比较多，包括微积分、复变函数、概率论、随机过程等数学课程，以及信号与系统、电路理论、高低频电子线路等内容，如果学生先修课程的基础不好，对本课内容的学习和理解影响非常大。尤其是随机过程，概念较难理解，该课程一般在研究生阶段开设，本科阶段不单独开课。在通信原理中，用到了大量随机过程的知识和结论，对学生的学习构成了较大的障碍，严重影响了学习的兴趣。作为一门专业课程，通信原理的实践性非常强，涉及到了大量通信系统和通信技术，如模拟通信中的标准调幅、双边带调制、单边带调制、调频，数字通信中的ASK、FSK、PSK、DPSK、QAM、OFDM技术及各种变种，内容非常丰富。如果内容组织的不好，就会显得非常庞杂，学生学起来找不到头绪，压力非常大，学习兴趣锐减。

在通信原理的教学过程中，学生在课程初期有一定的兴趣，但是随着内容难度的提高，尤其是作业的难度如果偏高，容易导致自信心受挫，学习积极性下降，学生的分化开始出现，并逐渐增强。在后续实践类课程的教学中也发现，部分学生基本概念和基本原理的掌握程度不足，严重影响了后续课程的学习和实践技能的培养。

改进通信原理课程教学的一些方法

抽象是通信理论的一个重要特点。通信学科的发展非常快，课程的内容非常丰富，课时在不断地压缩，学生的基础又不牢固，对通信原理的教学造成了巨大困难[5]。根据在通信原理课程教学中的一些经验，结合应用型本科院校教学的要求和学生特点，本文总结了一些作者在通信原理课程教学中的思考和看法。

1. 转变教育观念，提高教学理论水平，深化对应用型本科人才培养的认识

作为教师，应该掌握教学过程的一般规律，认识到教学活动的复杂性，才能根据学生的特点，调整教学目标，更新教育观念，完成应用型本科人才的培养要求。通信原理作为一门专业核心课程，要完成两个基本任务，一是要学生掌握信息传输的基本概念和基本理论，二是要培养学生对通信知识的应用和实践能力。没有通信理论做基础，应用能力就无从谈起。只具有理论知识，就达不到应用型人才的培养目标；不结合具体实践，对理论知识的理解也只会停留于表面，无法实现有意义的学习，难以长期保持巩固，对知识的灵活运用、学习迁移和能力转化都会造成困难。

在应用型本科院校中，要强化教学过程中学生主体性的认识[6]。应用型本科院校，特别是民办院校中，学生学习的主动性和积极性欠缺，教学过程容易陷入单向灌输的形式，导致教学效率低下，基本的教学目标都难以完成。因此，在教学过程中，要通过多种教学方法，加强教学过程的互动，激发学生的学习兴趣，调动学生的积极性，启发学生的创新思维。

教学要着眼于学生的发展，才能有效处理各种矛盾。应用型本科人才要达到理论基础和应用实践能力两个目标，这两个目标如何实现在理论界仍有许多不同的看法。这种目标的双重性为具体教学带来了困难，在课时有限的情况下，每个教师都要权衡理论和应用所占的具体比重。只要教师对知识掌握和能力发展之间的关系有了清楚的认识，根据自己在本学科的经验，从学生在行业中的发展出发，就能够根据具体情况，不断总结调整，实现两个目标的统一。

2. 降低内容难度，增强学生的自信心

麦克利兰和阿特金森的成就动机理论认为，对于避免失败的大学生，减少竞争、降低难度有利于维持学习的动力。由于民办本科院校中，尤其是在三年级阶段，学习动力严重不足，通信原理课程本身难度较高，学生基础不够牢固，学习有一定的障碍，应尽量帮助学生克服困难，调低课程目标，增强自信，保持学习的积极性[7]。

通信原理课程对高等数学、概率论、信号与系统、随机过程等先修课程和内容要求较高，部分学生对这些内容的掌握比较薄弱，又由于时间的因素遗忘较多。在用到这些内容时，如果不能理解，容易放弃对后面内容的学习。因此，在进行教学设计时，要注意对相关的内容进行补充和回顾。复习时，要处理好完整性和枯燥之间的矛盾，增加复习的频度，减少每次复习内容的数量。例如傅里叶变换，在第二章讲授确知信号分析时，以定义为主，集中回顾。在计算功率和能量时，回顾帕斯瓦尔定理；在分析带通白噪声时，回顾门函数的傅里叶变换形式；分析幅度调制时，回顾频移性质。课程经常用到三角变换公式、概率密度函数的计算，遇到相关内容时要注意回顾。

本课程使用的教材是樊昌信教授编著、国防工业出版社出版的《通信原理》，本教材多次再版，错误较少，内容丰富，为多所学校选用，也是学生考研的主要参考教材。本教材理论性较强，例题较少，学生学习起来有一定难度，特别是课后习题难度较高。教学时要对内容做适当精简，对例题适当补充。布置作业时要注意降低难度，结合课程目标，选择应用性较强的习题。

3. 合理设置课程目标，突出教学重点，实现分层次教学

在学时压缩以后，对教学内容必须有所选择，将有限的学时资源更多地分配到重点内容上，才能保证课程目标的实现。在课堂教学中，学生的注意力和有效听课时间有限。根据学生的基础和能力的差异，对内容的掌握程度要有所选择。在应用型本科院校的教学中，理论以应用为导向，为学生的发展提供必要的基础。在通信原理的教学中，应合理设置课程目标，实现分层次教学。学生首先要理解信息、传输等基本概念，在此基础上掌握传输的基本原理、基本方法和性能度量，进而理解实际系统中，为了提高性能而采用的复杂方法。所有学生都应该掌握通信中的基本概念和基本原理，在课堂教学中应反复强调，考试主要考查基础知识的掌握和基本应用的能力。在此基础上，讲解通信中的基本分析方法和重要定理的推导。

4. 突出应用，采用多种教学手段化解教学难点

枯燥的理论教学会让学生厌烦，丧失学习的兴趣，通信原理作为学生接触的第一门专业课，应该紧密联系实际应用，既能让学生意识到本门课的重要性，也对具体理论有更加直接的认识，加深印象和记忆。比如信道分类和模型，可以和实际系统联系起来。通过对单边带调制、残留边带调制应用的介绍，可以使学生更加明确具体技术的优点和缺点，理解在工程实践中如何各种技术的特点，对具体技术进行权衡和选择，增加学生的兴趣。教师也可以创设一个场景，例如火星和地球通信，提出对通信的要求，学生运用所学知识，选择使用具体的通信技术。在课堂中突出教学内容的应用性，在每门课程中贯彻实践应用性本科人才的培养目标，培养实践能力和创新能力。

在课程教学中，根据教学内容的不同，选择使用多种手段，使知识更易于被学生理解。在模拟调制和数字调制技术的分析比较时，可以使用图表方式展示结论，便于学生直观理解。在 PCM 编码的教学中，可以将编码过程以动画的显示展现出来，加深学生的理解和记忆。在进行定义推导和应用的演算中，板书与理论相结合，学生更容易跟上教师的思路。学生在对信息的讲解中，使用打电话、看报纸的例子说明信息的传输即是不确定性的减少这一本质概念，将抽象的概念实例化。在讲解复用时，可以用发言规则、车道划分等例说明什么是复用，怎么实现多用户通信。由于应用型本科院校中学生抽象思维能力不足，尽量采用一些较为直观的方法讲解抽象的概念，降低学习的门槛，也有利于调动学生的积极性。

5. 课程内容结构化，用核心概念和系统结构贯穿整个教学过程

教材第一章讲述了通信系统的基本结构，在讲授各章内容时，可以用图形的方式明确本部分在通信系统中的位置，明确其输入输出形式，学生对各部分内容的作用和目的的认识更为清晰，对各部分内容的整体性认识更加深入。同时，信号在各模块间传输时形式会有变化，通过对具体变化的了解，学生对各模块的工作原理，对一些基本概念的理解会更深刻，如多进制与调制，比特与码元，信息速率与符号速率，速率与带宽，噪声与性能等。

通信原理课程的内容较多，但有一些基本思想贯穿在各章内容之中。如带宽和功率、有效性和可靠性、性能与复杂度等关系和矛盾的折中与平衡多处可见，加强前后内容、概念的比较有利于学生建立各部分之间的联系，加深对概念和原理的理解。通过编制知识结构图，加速知识的内化，使知识虽多而不繁，虽散而不乱。

6. 加强与学生互动，活跃课堂气氛

课堂教学中单向的讲授容易降低学生思维的活跃性和兴奋度，不利于学生注意力的集中，可以通过提问、设问、启发诱导等多种方式与学生互动。一方面可以了解学生对知识的掌握程度，另一方面也可以促使学生积极思考，强化对知识的理解掌握。对一些基本概念，如比特与码元、编码和调制这些概念，可以经常性的提问，根据具体的章节内容从不同的角度去认知，提高教学的过程化程度，深化基本概念的掌握。增加与学生的互动，适当提高紧张度，有利于延长学生在课堂上的有效学习时间。

总结与讨论

通信原理课程理论性强，是一门教师不容易教、学生不容易学的课程，又面临课时压缩的问题。同时，应用型本科院校根据自身的特点提出了应用实践和理论基础并重等新的目标和要求，每个教师都在思考如何实现应用型本科人才的培养目标。本文以通信原理课程为例，结合作者的经验和应用型本科院校学生的特点，提出了一些课堂教学中的思考和意见，对学生积极性的提高有一定的改善。在今后的教学中，需要进一步平衡各种关系，加强对教学方法的改进和创新。

资助

本文受到以下项目资助，南京工业大学浦江学院2019教改项目：《通信原理》课程双对标的研究与实践（项目编号2019JG003Y），电路与电子技术开放实验设计与实践（项目编号2019JG001Y）；南京工业大学浦江学院2019人才培养工程计划项目课题：ICT人才科学思维能力培养体系研究与实践（项目编号njpj2019-01-03），特向南京工业大学浦江学院表示感谢。

参考文献

[1] 樊昌信，曹丽娜. 通信原理[M]，第7版. 北京：国防工业出版社，2015.
[2] 张蕾. 我国新建应用型本科院校课程结构困境研究[D]. 喀什：喀什大学，2013.
[3] 季艳萍. 民办与公办高校大学生学习适应性发展的比较性研究[D]. 南京：南京师范大学，2008.
[4] 张冬梅. 工科院校大学生学习倦怠与心理健康关系的研究[D]. 长春：东北师范大学，2010.
[5] 杨桂芹. “通信原理”教学研究[J].计算机教学，2008 (24): 101-105..
[6] 周川. 简明高等教育学[M]. 南京：河海大学出版社，2013.
[7] 谭顶良. 高等教育心理学[M]. 南京：河海大学出版社，2013.

基于 OpenCV 的车牌检测系统
LICENSE PLATE DETECTION SYSTEM BASED ON OPENCV

黄大胜 [1]，李 娟 [2]，林绅洋 [3]，张伟华 [4]，李奇 [5]
Huang Dasheng[1], Li Juan[2], Lin Shenyang[3], Zhang Weihua[4], Li Qi[5]

南京工业大学浦江学院
NANJING TECH UNIVERSITY PUJIANG INSTITUTE
*Corresponding author: 黄大胜，李 娟，林绅洋，张伟华，李奇
E-mail: 2590609710@qq.com

摘要

由于经济的快速发展以及人口基数逐渐加大，当下私家车的拥有量呈现上升趋势。庞大的车辆基数使得人工记录查验车牌难以保障时效性，庞锋（2012）认为我们需要开发出稳定高效的视频及图像车牌识别系统，解决交通智能管理的问题，而车辆牌照识别技术可以满足车辆在各种场景的应用需求。本文介绍使用计算机视觉技术从监控视频以及图像中检测出车辆牌照，并进行识别。

本文通过分析研究车辆图像牌照提取识别的相关技术技术，并解析现有算法流程，深入了解图像车牌检测的的流程及步骤，搭建了适合复杂背景的车牌图像检测识别系统。该系统车牌检测包括车牌区域检测、车辆牌照字符分割以及字符核对三大模块。采用能抗噪声的Canny边缘检测算法，使得图像边缘轮廓分割清晰，从而让车辆牌照的识别更加准确。

采用了C++程序语言在VS2015开发平台上利用Open CV跨平台开源视觉库进行该系统的编程设计，该系统实现了检测图像中车辆牌照的功能，通过算法识别牌照并提取显示出来，有着稳定性高的优势，应用场景种类众多。

关键词：Open CV, Canny算子，车牌检测

Abstract

Because of the rapid economic development, and now the population base is growing, private car ownership is on the rise. The huge vehicle base makes it difficult to guarantee the timeliness of ma nually recording and checking license plates. Pang feng (2012) believes that we need to develop a stab le and efficient video and image license plate recognition system to solve the problem of intelligent tr affic management,and vehicle license plate recognition technology can meet the application needs of vehicles in various scenarios. This paper introduces the use of computer vision technology to detect a nd identify vehicle license plates from surveillance video and images.

By analyzing and summarizing the research results of license plate recognition technology at home and abroad, this paper builds a license plate image detection and recognition system suitable for complex background. The license plate recognition system consists of three parts: license plate region detection, vehicle license plate character segmentation and character check. The Canny edge detection algorithm which can suppress noise is adopted to make the image edge contour segmentation clear, so as to make the recognition of vehicle license plate more accurate.

Using C++ programming language on the VS2015 development platform, OpenCV cross-platform open source visual library was used for the programming design of the system. The system realized the function of detecting the vehicle license plate in the image, recognized the license plate through the algorithm and extracted and displayed it, which has the advantage of high stability and a wide variety of application scenarios.

Keywords: Open CV, Canny operator, license plate detection

引言

车辆牌照号码是每辆车独有的特征，具有唯一的辨识性，曾泉（2016）认为车牌识别技术是交通管理系统中的重要成分。

本课题中车牌检测是利用OpenCV跨平台开源视觉库对图像进行相关处理，准确地对获取的车辆进行检测车牌区域，识别其车牌号码并从文字库中比对显示出来。其相关功能包括：降低车辆堵塞可能性、车辆牌照的识别并记录、不法车辆的识别、通过牌照查询车主的个人信息以及进行追踪监控。车牌识别系统主要有以下应用:

1.车流统计

2.超速违章拍照处罚

3.公共区域车辆出入库管理

4.停车场无人化管理

5.降低高速收费环节人工管理

物联网的迅速发展使得车牌识别技术趋于完善，逐步用于交通智能管理，在各大市场具有很强的应用性。

本课题通过将获取的车辆图像进行灰度转换、Canny算法、膨胀腐蚀处理等一系列操作识别并提取车牌号。

车牌检测流程如图1所示:

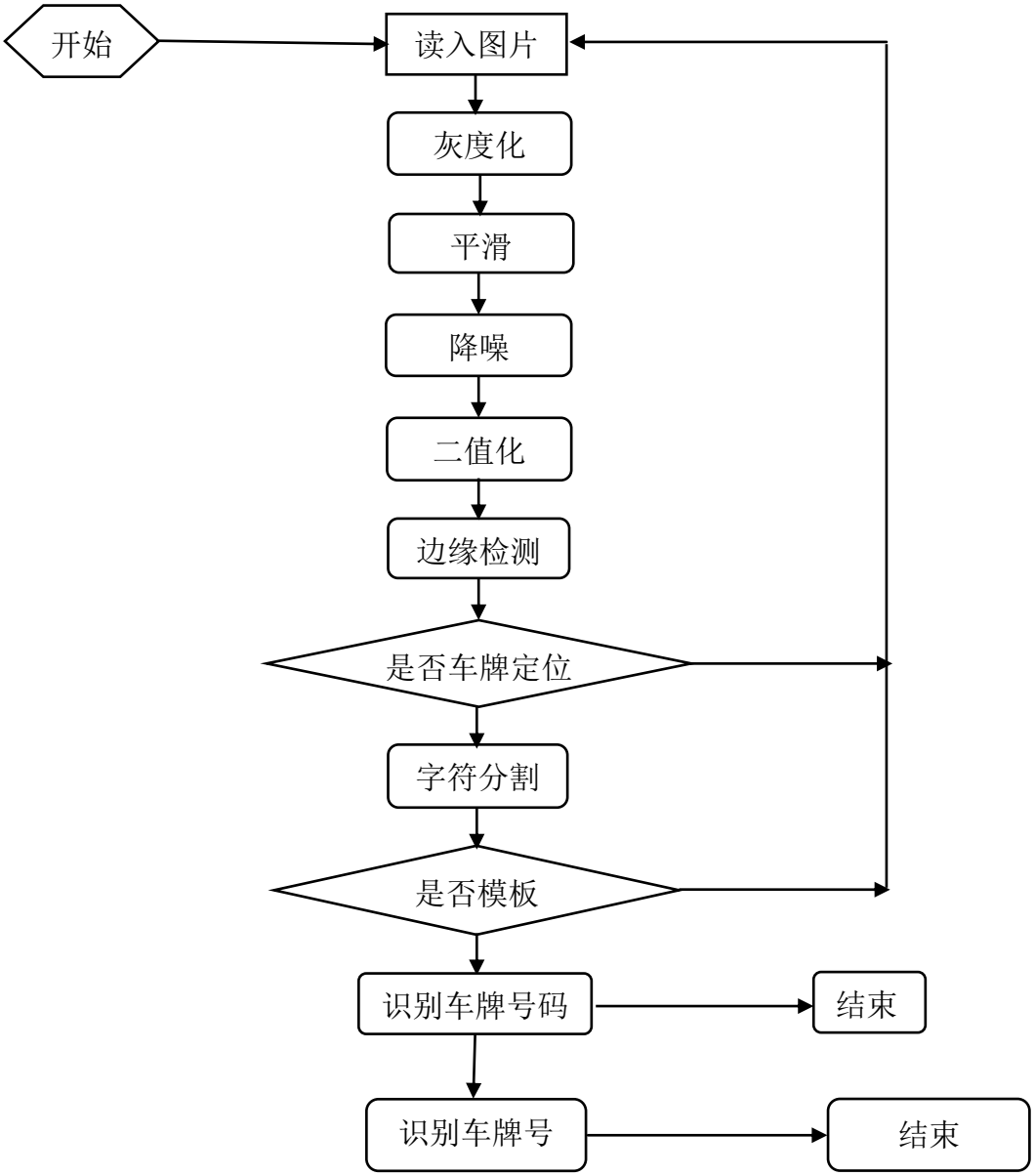


图1

1. 车牌预处理

利用OpenCV调用摄像头，进行视频捕捉，截取视频帧画面并对获取图像进行相关操作。

在实际图像处理中，需要采取灰度图片进行下一步操作，将各种格式的图像转变成灰度图像后能减少复杂背景对后续操作的影响，降低无关因素的干扰，以便于计算。陈扬（2014）提出在保留图像轮廓和特征的基础上，灰度图仍不足以完全准确地反映整幅图像的轮廓和纹理，通过中值滤波进行处理能实现更好的效果，最终除去图像中的小区域，提高图像准确性。

2. Canny算法

2.1 Canny算子简介

Canny边缘检测算子在1986年被人提出，能在背景复杂的图像中准确提取出有效的图像信息并极大程度减少计算量，目前在视觉图像处理领域被广泛使用。

2.2 Canny边缘检测

（1）消除噪声

由于车辆背景的复杂性和不可辨性，图像质量在处理过程中会大大降低。不同种类因素的影响，会降低图像的信息准确度，使用高斯平滑滤波器卷积降噪可以有效地降低这种影响。

郭瑞峰（2017）等人认为，为了消除图像中噪声便于下一步计算，可以使用高斯平滑滤波对图像进行降噪。其原理是对检测到的每个像素点在图像中进行遍历式扫描，根据算法选择确定适合的值，最终进行降噪处理，避免了噪声对该系统后续操作产生的不良影响，若噪声未能消除干净，将使之后的计算产生极大地误差。

一维高斯分布：

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

二维高斯分布：

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

（2）计算梯度幅值和方向

图像中的边缘有多个方向，本文检测图像的对角边缘以及图像中的水平和垂直方向，为达到这一目的，本系统利用四个Canny算子进行计算。设Gx代表水平，Gy代表垂直，这两方向的一阶导数值可用边缘检测算子进行相应计算，虽然无法完全覆盖所有方向，但是提高了效率，并将计算值返回，由此便可以确定像素点的梯度G。

$$M[x,y] = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}$$

$$\theta[x,y] = \arctan\left(\frac{G_x(x,y)}{G_y(x,y)}\right)$$

2.3 非极大值抑制

本系统为了达到消除边缘检测中无用信息的目的，采取对非极大值进行抑制的方法。该步骤是指沿着梯度方向上进行采取相关操作，其作用在于使图像中检测到的边缘减少模糊程度，更加精准的描述出边缘。对图像进行梯度计算后，仍然有边缘轮廓不清晰的问题存在。这是由于对图像处理不够全面引起的，非极大值抑制则可以解决这个问题，从而避免边缘过于模糊而影响后续操作，提高了图像处理的准确性。

在梯度图像中对每个像素进行非极大值抑制算法的操作时，需先记录当前像素梯度强度，并将其与正负梯度的方向上两个像素进行相互比较，当它们的像素相比差异达到最大时，则令该像素点为边缘点，否则该像素点不是边缘点并被抑制。

2.4 双阈值检测

非极大值抑制使经过上述处理的图像实际边缘显示的更加准确。但该结果并不完全准确，由于车辆背景的复杂性，以上步骤并不能完全表示实际边缘。为了降低这些因素对图像结果的不良影响，在这一操作中应避免高梯度值的使用，使用图像中较弱的梯度值对无用的边缘像素进行筛选。

3. 滞后阈值

滞后阈值是该算法中处理图像的最后操作，需要设定高低两个阈值，与双阈值检测有异曲同工之妙。对图像中某一像素位置的幅值与设定的高低阈值进行比较，若数值不一样，且幅值大于高阈值时，则保留该像素，称为边缘像素。反之该像素将不被称为边缘像素且不被保留。如果某一像素位置的幅值不属于上述两种情况，即在两个数值之间时，那么在图像中当遇到比该像素更高的像素并且两者进行连接时将会被保留。

4. 形态学（膨胀腐蚀）处理

腐蚀操作可以分割图像中单独的图像元素，使得核周围的最小值成为中心值，膨胀操作与腐蚀操作完全不同，是两个极端。膨胀操作使得彼此无间隔的元素建立一种密切连接的关系。通过腐蚀去除对提取字符造成影响的细小部分，便于将车牌字符准确的提取出来。该步骤中应避免多次重复膨胀腐蚀操作。

5. 自适应二值化处理

自适应阈值不限定一个常数，而是采取自适应方法，对检测到的图像进行相关二值化处理。周明辉（2013）认为二值化就是根据阈值将图像分成黑色和白色两部分。二值化的处理是对图像中的无用信息进一步筛选，降低无用信息对处理结果的影响，使得图像的边缘更加凸显。本项目使用了OpenCV开源视觉库中自带的二值化处理，处理过后使得车牌数字信息表现的更加明显。

6. 字符提取

本项目采取了两种方法进行像素级分割，分别是水平以及垂直映射像素。

分析中国汉字及阿拉伯数字的特点，对其特征进行提取，利用算法提取出车牌字符在图像中的轮廓，利用tesseract_ocr识别算法识别图像中的汉字及数字，遍历字符库后，计算其欧氏距离，最终提取车牌号码并显示出来。

字符提取过程中对每竖排的像素值进行检测，当像素值大于0的像素数大于等于六个时，便默认此竖排是有数字的，对是否有数字的列进行分析，将其转化成数组,并对数组的值进行相应的运算。如果存在大于等于三个像素周围像素值为1的像素，则默认该像素值与周围像素值相同，对该像素值进行同化。当像素周围像素为0的个数小于等于1时，则默认为该像素为0，对该像素进行上述相同操作。

本系统实际测验如图1所示：

图1

8. 总结

视频及图像车牌识别技术是城市交通管理系统的重要组成部分，在商场、公路以及一些公共区域具有非常多的应用，为车辆的管理带来了极大便利，促进了车辆管理的发展。

本文对图像牌照提取识别的相关技术技术进行深入分析研究，结合现有算法，利用跨平台开源视觉库OpenCV在vs2015平台开发出了图像车牌识别系统。

本课题来源于南京工业大学浦江学院与南京一科技企业的校企合作项目，车牌检测是其中一重要模块，发挥重要作用。

本课题在检测车牌过程中，由于国内车牌具体特征基本一致，利用了垂直边缘检测算法来识别车牌，采用高斯平滑滤波去噪来使车牌区域明显、使图像灰度化减少无关变量的影响，最终使用边缘检测算法识别了车牌的区域。并对图像车牌字符进行分割，最终识别字符轮廓并在文字库中进行比对，成功检测提取出字符。

该图像车牌识别系统实现了对车辆图像的车牌提取以及识别，具有高稳定性、高效率的特点。随着车牌检测技术不断地改进，该图像识别系统还有巨大的完善空间，可以结合其他新型技术产生新的应用。

参考文献

庞锋. (0). *车辆特征检测研究与实现*. (Doctoral dissertation, 电子科技大学).
曾泉. (2016). *复杂背景下基于OpenCV的车牌识别系统研究*. (Doctoral dissertation).
陈扬. (0). *基于OpenCV的机动车牌照识别系统*. (Doctoral dissertation).
郭瑞峰, 袁超峰, 杨柳, & 彭光宇. (2017). 基于opencv的机器视觉尺寸测量研究. *计算机工程与应用*, *053*(009), 253-257.
周明辉. (2013). *运动车辆车牌识别算法的研究与实现*. (Doctoral dissertation).

项目编号

[1]南京工业大学浦江学院2019教改课题：电路与电子技术开放实验设计与实践（2019JG001Y）
[2]南京工业大学浦江学院人才培养目标项目：立德树人视角下大学生担当精神的培育途径研究（njpj2019-1-05）