Journal of Information Technology and Education Science

University of Computer Studies (Taungoo)

JITES 2019

Journal of Information Technology and Education Science

Vol-1
No-1
Dec
2019

ISBN: 978-99971-0-759-6
Volume-1, No-1 December, 2019

JOURNAL OF INFORMATION TECHNOLOGY AND EDUCATION SCIENCE
VOL. 01, NO. 1, 2019

ISBN: 978-99971-0-759-6
Publisher: University of Computer Studies (Taungoo)

JITES2019

University of Computer Studies (Taungoo)
Yangon-Mandalay Road,
Taungoo Township,
Bago Region, Myanmar.

JOURNAL OF

INFORMATION TECHNOLOGY AND EDUCATION SCIENCE

JITES$_{2019}$ VOL. 01, NO.1, 2019

UNIVERSITY OF COMPUTER STUDIES (TAUNGOO)

Journal of Information Technology and Education Science
Vol. 01, No. 1, 2019

Contents

Information Technology

Quantitative Association Rule Based On Exam Report System

San San Yu, Kyi Zar Nyunt
University of Computer Studies (Kyaing Tong),
University of Computer Studies (Taungoo)
sanyumaw06@gmail.com, kyizar81@gmail.com

ABSTRACT: *At present, large quantities of data are generated. Data mining and automated knowledge extraction in this data belong to the major contemporary challenges in the decision making process. For this task association rule mining is one of the most often used methods. It discovers unexpected data dependencies in databases. The research area association rule mining has focused predominantly on databases categorical data only. However, many real-world databases contain quantitative attributes and current solutions for this case are so far inadequate. To reflect this issue and to support educational organizations, we present a quantitative association rule mining based exam report system, especially for University of Computer Studies. To be relevant to our application area, a quantitative association rules mining method is applied it in this exam report system. This system produces as accurate results in a reasonable time. It provides rules based on subject's marks of students and percentage of pass.*

Keywords: Data mining; association rules; quantitative association rule mining;

1. INTRODUCTION

There are several data sources and information systems available at the University of Computer Studies. The management and administrative processes often need to view the exam marks data as a whole for the purpose of producing of exam reports. With the use of this exam report system (i) the report of pass students together with percentage and (ii) the report of passed students for desired percentage can be viewed promptly. As the flexibility and performance of this approach make a good impression, an idea of a university-wide mark-analysis and exam reporting system has been arisen.

Pattern mining plays an essential role in many data mining tasks. Association rule mining is one of the most important fields of pattern mining. It is a fundamental problem of data mining tasks that aims to find frequently occurring subsets in a sequence of sets as an initial step and association rules from these frequently subsets as a result. It is a very young research field born in 1993 [2].

It has abroad applications not only in its natural form such as customers relationship management, e-commerce, bioinformatics, DNA analysis, protein analysis, inductive databases, query expansion, network intrusion detection, and so on; but also as a subroutine in various other problems such as correlations, classification, clustering, web-mining, and so on.

In this paper, we present a quantitative association rule mining based exam report system, especially for University of Computer Studies. Quantitative association rules are multidimensional association rules in which the numeric attributes are dynamically discretized during the mining process so as to satisfy some mining criteria, such as maximizing the confidence or compactness of the rules mined. To be relevant to our application area, a quantitative association rules mining method is applied it in this exam report system. Our system can be viewed as the electronically management information support system (EMISS) which can give users an ease analysis for the pass/fail exam result percentage, the rules which satisfy this percentage and the list of passed students as report.

The aim of our system is to facilitate the work of the department teachers and superiors to support exam report electronically. The expected benefit from this system is that the system converts the analysis task manual to electronic that exam marks analysis tasks can be performed in a reasonable time including creating the most efficient exam report system. Moreover, it is likely to be sure that evaluation is consistent and fair and human errors will not affect the process.

First we briefly describe the quantitative association rule mining. This will be followed by an introduction to our system. Finally, the empirical study is explained with a description of the dataset and the results of this system are presented.

2. ASSOCIATION RULE MINING

Data mining on transactional database focuses on the mining of association rules, finding the correlation between items in the transaction records. Association rules are widely used in various areas such as telecommunication networks, market and risk management, inventory control etc. The two basic parameters of Association Rule Mining (ARM) are: support and confidence. Support(s) of an association rule is defined as the percentage/fraction of records. Confidence of an association rule is defined as the percentage/fraction of the number of transaction. Association Rule mining is one of the most important and well researched techniques of data mining. It aims to extract interesting correlations, frequent patterns, associations or casual structures among sets of items in the transaction databases or other data repositories. Association Rule Mining is to find out association rules that satisfy the predefined minimum support and confidence from a given database.

Quantitative attributes are partitioned to different ranges according to the predefined hierarchies and attributes are replaced by those ranges prior to the mining process. After this process the relevant data can be stored in the relational database. The relevant data can also be stored the in the data cube, which is more suitable for multiple dimensional association rule since data cube itself is multidimensional. An example of association rule on transaction database is exam report system.

Association rule mining finds interesting association or correlation relationships among a large set of data items. The discovery of interesting association relationships among huge amounts of business transaction records can help in many business decision making processes, such as catalog design, cross-marketing, and loss-leader analysis.

The rules $A \Rightarrow B$ holds in the transaction set D with supports, where s is the percentage of transactions in D that contain $A \cup B$ (i.e., both A and B). This is taken to be the probability, $P(A \cup B)$.

$$\text{Support } (A \Rightarrow B) = P(A \cup B)$$

For association rules of the form "$A \Rightarrow B$" where A and B sets of items.

$$\text{Support } (A \Rightarrow B) = \frac{\text{\# tuples containing both A and B}}{\text{Total \# of tuples}}$$

A set of items is referred to as an itemset. The "itemste" is more commonly used than "item set". The occurrence frequency of an itemset is the number of transactions that contain the itemset. This is also known, as the frequency, support count, or count of the itemset. An itemset satisfies minimum support if the occurrence frequency of the itemset is greater than or equal to the product of min-sup and the total number of transactions in D. The number of transactions required for the itemset to satisfy minimum support is therefore referred to as the minimum support count. If an itemset satisfied minimum support, then it is a frequent itemset. Itemsets satisfying minimum support were referred to as large. This term, is somewhat confusing as it has connotations to the number of items in an itemset rather than the frequency of occurrence of the set. The more recent term frequent.

Association rule mining is a two-step process:
1. Find all frequent itemsets: By definition, each of these itemsets will occur at least as frequently as pre-determined minimum support count.
2. Generate strong association rules from the frequent itemsets: By definition, these rules must satisfy minimum support and minimum confidence [4].

2.1. Quantitative Association Rule Mining

Mining association rules that aims at finding strong relations between attributes has already long been studied [1]. However, mining association rules on both categorical and numeric attributes, also called quantitative association rules, has been less studied. Basically, this task involves several problems. Numeric attributes are usually defined on a wide range of different values. It is useless to work on all possible numeric value will not appear frequently. A classical way to deal domains into intervals. This is called discretize "correctly" numeric attributes with respect to Minsupp and MinConf [2]. Items found in relational tables have many different attributes. These attributes may be either quantitative or categorical. Any valued attribute will be treated as quantitative and will be used to derive the quantitative association rules [7].The quantitative measures to test hypothetical generalizations (Hoepfl, 1997), and they also emphasize the measurement and analysis of causal relationships between variables (Denzin and Lincoln, 1998). Quantitative Association Rule Mining the problem or concept to be studied, and perhaps generate hypotheses to be tested [6].

Quantitative association rules are multidimensional association rules in which the numeric attributes are dynamically discretized during the mining process so as to satisfy some mining criteria, such as maximizing the confidence or compactness of the rules mined. Quantitative association rules having two quantitative attributes on the left-hand side of the rule, and one categorical attribute on the right-hand side of the rule.

$$A_{quan1} \wedge A_{quan2} \Rightarrow A_{cat}$$

Where A_{quan1} and A_{quan2} are tests on quantitative ranges and A_{cat} tests a categorical attribute from the task-relevant data [4].

2.1.1. Mining Association Rules using Static Discretization of Quantitative Attributes

Place In this approach, quantitative attributes are discretized using predefined concept hierarchies. This discretization occurs prior to mining. For instance, a concept hierarchy for exam mark may be used to replace the original numeric values of this attribute by ranges, such "0....20", "21...50", "51....75", "76...100",

etc. Here discretizing is static and predetermined. The discretized numeric attributes, with their range values, can then be treated as categorical attributes (where each range is considered a category). We refer to this as mining multidimensional association rules using static discretization of quantitative attributes.

Quantitative attributes, are discretized prior to mining using predefined concept hierarchies, where numeric values are replaced by ranges. Categorical attributes may also be generalized to higher conceptual levels if desired. If the resulting task-relevant data are stored in a relational table, then the Apriori algorithm requires just a slight modification so as to find all frequent predicate sets rather than frequent itemsets (i.e., by searching through all of the relevant attributes, instead of searching only one attribute, like buys). Finding all frequent k-predicate sets will require k or k+1 scans of the table. Other strategies, such as hashing, partitioning, and sampling may be employed to improve the performance [4].

2.1.2. Mining Association Rules using Binning Method

In the second approach, the bins are further combined during the mining process. This discretization process is dynamic and established so as to satisfy some mining criteria, such as maximizing the confidence of the rules mined. Because this method treats the numeric attribute values as quantities rather than as predefined ranges or categories, association rules mined from this approach are referred to as quantitative association rules.

2.1.3. Mining Quantitative Association Rules using Distance-Based Approach

Distance-based approach discretizes the quantitative attributes so as to capture the semantic meaning of such interval data. This dynamic discretization procedure considers the distance between data point. Hence, such quantitative association rules are also referred to as distance-based association rules.

Quantitative association rules where quantitative attributes are discretized initially by binning methods and the resulting intervals are then combined.

A disadvantage of association rules is that they do not allow for approximations of attribute values. This motivates the mining of distance-base association rules, which capture the semantics of interval data while allowing for approximation in data values. A two-phase algorithm can be used to mind distance-base association rules. The first phase employs clustering to find the intervals or clusters, adapting to the amount of available memory. The second phase obtains distance-base association rules by searching for groups of clusters that occur frequently together.

3. OUR EXAM RESULT SYSTEM

In this section, we introduce our system gives exam result for a given passed students percentage as well as exam result together with passed percentage for a given rule to pass an exam. The result will show with figure.

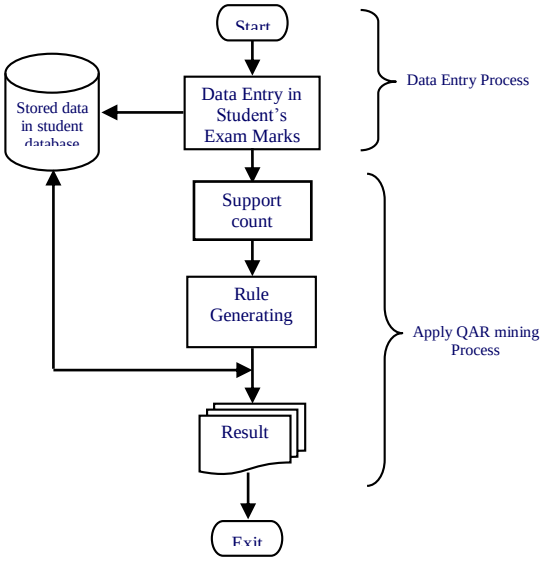


Figure 3.1. Overview of System Flow Diagram

This system, two main processes: Data entry process and apply quantitative association rule mining process. In the first part, there are stored students name, roll number and marks in the database. In second one, we firstly assume support count (percentage) to generate rule, and extract result from student database academic to this rule. It the part support count is the part that user defined support count (percentage). Generation is the rule that is evolved by giving percentage. The mark by rule is output the result, comparing marks of each student in the stored database. Based on Aprori algorithm in association rule mining, it is outputting the result of students. In this system, the results by using algorithm and formula are compared with graph.

When we use exam report for the results to be subsequently received. Examination data is store in the database system. Examination records, exam data (such as result), are provided together in University of Computer Studies' student record. Once the examination results have been imported, you can get a

number of useful exam result report, using quantitative association rule mining.

Evolving rules produces exactly percentage based on rules. It also produces results of students based on rules. The exam result of pass students for desired percentage can be viewed graph.

3.1. Experiment Results

Transactional data

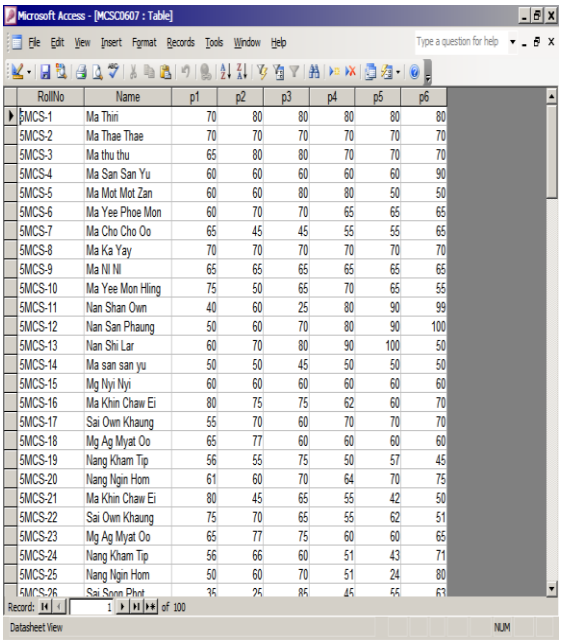
Academic	List of item-IDs
Module I	P1,P2,P3
Module II	P1,P2,P3
1st Year	P1,P2,P3,P4,P5,P6,P7
2nd Year	P1,P2,P3,P4,P5,P6,P7
3rd Year	P1,P2,P3,P4,P5,P6,P7
4th Year	P1,P2,P3,P4,P5,P6,P7
M.C.Sc	P1,P2,P3,P4,P5,P6
M.C.Tech	P1,P2,P3,P4,P5,P6
M.I.Sc	P1,P2,P3,P4,P5
D.C.Sc	P1,P2,P3,P4,P5,P6

Formula = (int) sup-count * record /100

Figure 3.2. Candidate itemsets where the minimum support count is defined

We present the results of our examination report system in this section. We use the sample M.C.Sc student datasets of the University of the Computer Studies, Kyaing Tong. The datasets contain the variables such as student ID, Roll No, Name, Paper1, Paper 2, Paper 3, Paper 4, Paper 5 and Paper 6. The example records are shown at Table 3.1.

Table 3.1. M.C.Sc 2015-16 to 2019-2020 Students Exam Result Records (100) in University of Computer Studies, Kyaing Tong



3.1.1. Resulted Rule

The desired percentage(%) of the passed student or the desired rule has been used in our system in apply. According to these, the result of the system can be one of the forms

(i) passed students result for the given percentage as in Figure 3.3 or

(ii) The exam result of pass students for desired percentage can be viewed graph Figure 3.4.

The student Dataset involve the variables such as 1st Year 2015-16 to 2019-2020, 2nd Year 2015-16 to 2019-2020,….., and so on. Firstly, we must select the dataset and we choose M.C.Sc 2015-16 to 2019-2020.

Example;

Input: Support count = 65% (passed percentage)

Students records =100

Formula = (int) support count * record/100

Formula= 65%*100 =65

Rules is;

Input : 65% (passed percentage)

This system tends to evaluate the students' result according to the Association Rule Mining. Thus we have to generate the Rule from the user preferred support count.

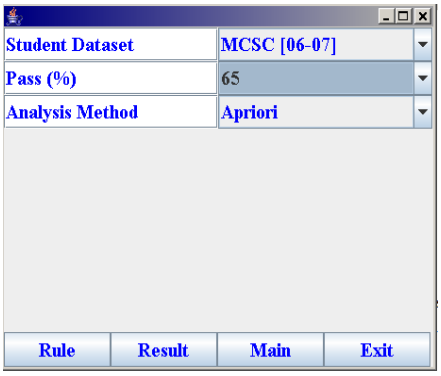


Figure 3.3. Support count Entry and Rule

These systems, thus firstly, gets the preferred input support count and generate the associated rule for that support count.

Output: P1≥40 ^ P2≥40 ^ P3≥25 ^ P4≥55 ^ P5≥47 ^ P6≥45

In rule generating phase, we mine minimum values of each paper on sorted order of counting value.

Finally, we mine corresponding the student result according to the associated rule.

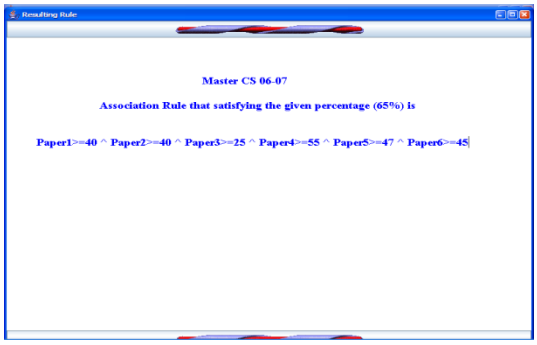
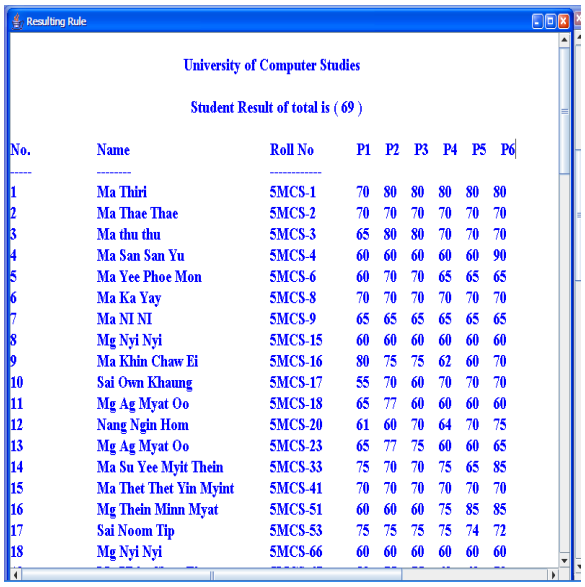


Figure 3.4. Resulting Rule



No.	Name	Roll No	P1	P2	P3	P4	P5	P6
1	Ma Thiri	5MCS-1	70	80	80	80	80	80
2	Ma Thae Thae	5MCS-2	70	70	70	70	70	70
3	Ma thu thu	5MCS-3	65	80	80	70	70	70
4	Ma San San Yu	5MCS-4	60	60	60	60	60	90
5	Ma Yee Phoe Mon	5MCS-6	60	70	70	65	65	65
6	Ma Ka Yay	5MCS-8	70	70	70	70	70	70
7	Ma NI NI	5MCS-9	65	65	65	65	65	65
8	Mg Nyi Nyi	5MCS-15	60	60	60	60	60	60
9	Ma Khin Chaw Ei	5MCS-16	80	75	75	62	60	70
10	Sai Own Khaung	5MCS-17	55	70	60	70	70	70
11	Mg Ag Myat Oo	5MCS-18	65	77	60	60	60	60
12	Nang Ngin Hom	5MCS-20	61	60	70	64	70	75
13	Mg Ag Myat Oo	5MCS-23	65	77	75	60	60	65
14	Ma Su Yee Myit Thein	5MCS-33	75	70	70	75	65	85
15	Ma Thet Thet Yin Myint	5MCS-41	70	70	70	70	70	70
16	Mg Thein Minn Myat	5MCS-51	60	60	60	75	85	85
17	Sai Noom Tip	5MCS-53	75	75	75	75	74	72
18	Mg Nyi Nyi	5MCS-66	60	60	60	60	60	60

Figure 3.5. Passed Student Result for 65%

- According to Quantitative Association Rule, Students Exam Result is 69.
- According to formula process, Student Exam Result is 65.

The passed student result together with the formula of percentage and using Quantitative Association Rule process generating exam result graph.

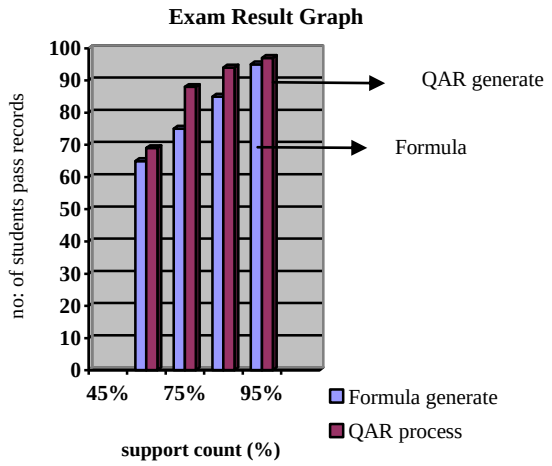


Figure 3.6. Exam Resulted Passed in Student List Graph

- values generated by formula of academics year
- result generated by Quantitative Association Rule process

4. CONCLUSIONS

The goal of our system is the effective analysis of students' marks to assist in students' marks. It can be used to monitor the progress of students e.g. exam result passed percentage and also to set attainment targets such as rules identification for exam result passed students. Thus we can conclude that our system can fulfill in some what extent for the administrative purpose by producing the quality exam report. It evolves marks identified for each subjects. Quantitative Association Based on the rules on exam report system, it shows the results of required class from database. It produces results of students who match with rules and formula.

5. FUTURE WORK

The system is not perfect, since it cannot produce that credit marks and distinction marks by quantitative association rule mining. The system can be extended to get these.

REFERENCES

[1] Association Rule Mining I, COMP 290-90 Seminar Fall 2004, The University of North Carolina at Chapel Hil.

[2] A.S. Aouissi, C. Vrain, and C. Nortet, "QuantMiner: A Genetic Algorithm for Mining Association Rules", *CCLS*, Columbia University, 475 RiverSide Drive, New York.

[3] F. Karel, and J. Kl´ema "Quantitative association rule mining in genomics using apriori knowledge", Department of cybernetics, Czech Technical University in Prague, Technick´a 2, Praha 6, 166 27 karelf1@fel.cvut.cz, klema@labe.felk.cvut.cz.

[4] J.Han, and M. Kamber "Data Mining: Concepts and Techniques".

[5] Johnson, B. R. (1997). "Examining the validity structure of qualitative research. Education", 118(3), 282-292.

[6] "Understanding Reliability and Validity in Qualitative Research" Nahid Golafshani University of Toronto, Toronto, Ontario,Canada.

[7] "Using Quantitative Information for EfficientAssociation Rule Generation"Bruno Pôssas, Wagner Meira Jr., Márcio Carvalho and Rodolfo Resende Department of Computer Science Federal University of Minas Gerais Belo Horizonte – MG – Brazil {bavep,meira,mlbc,rodolfo} @dcc.ufmg.br.

Audio Watermarking Based on MP3 Encoding Process

Myint Myint Than
Faculty of Computer Systems and Technologies, University of Computer Studies, Taungoo, Myanmar
m2t.myintthan@gmail.com

ABSTRACT: *Digital watermarking is the process that imperceptibly watermarks, the multimedia product with a specific watermark for the purpose of content authentication, data monitoring and tracking, and copyright protection. Amongst the applications, the most prominent usage of watermarking technique is helping in identifying the origins of different multimedia files and resolving ownership disputes. The aim of the study is to develop an audio watermarking technique based on MP3 encoding process. The hiding process takes place at the heart of the Layer III encoding process namely in the bit stream formatting.*

Keywords: audio watermarking; MP3 encoding

1. INTRODUCTION

Nowadays multimedia data (such as audio, image, and video) are normally stored in digital form, which can be replicated and modified by general users. The widespread use of Internet and wireless networking has made the distribution of multimedia data much easier than ever before. The tendency is further accelerated by a proliferation of smart phones and portable devices in recent years.

People around the world keep creating and spreading much multimedia data each day. Unfortunately, the illegal use of multimedia data is also rampant in digital age. Protection against intellectual property infringement increasingly becomes an important issue. Digital audio watermarking is an important and popular technique along the original musical audio content producers. Watermarks can be utilized as evidence for proving the ownership. This creates the problem of protecting the intellectual copyrights of multimedia data. Digital watermarking is the process of embedding a persistent digital identity into all forms of media content, providing the means for effective management and tracking of digital assets on the Web. Digital watermarks contain imperceptible digital data that can convey anything the owner chooses, including ownership information, contact details, and usage rights.

A performance analysis of DCT and DWT AW based on SVD is also presented by [1]. A novel audio watermarking method is also proposed to satisfy the IHC evaluation criteria by using different wavelet filters [8]. An aware extraction technique of audio watermarking using spread spectrum technique of audio watermarking using spread spectrum methodology has also been presented. This technique is different from previous spread spectrum AQT because here extraction process is blind and unique then other techniques. [3]. A mixture of DWT SVD and quantized indexed modulation is also a marvelous enhancement in the field of audio watermarking. This is also a blind technique and shown good results as compared to the previous techniques[4].

The aim of the study is to propose Audio watermarking system will hide information in MP3 files during the encoding process. The hiding process takes place at the heart of the Layer III encoding process namely in the bit stream formatting. The following section describes some techniques of audio watermarking and explains the related methodologies of proposed system.

2. SURVEY OF OTHER AUDIO WATERMARKING TECHNIQUES

2.1. Audio Watermarking Tools

Developer Alex Radzishevsky is offering a discount to Synthtopia readers on his app Audio Watermarking Tools (AWT) [1].

AWT is a set of software utilities for embedding (and retrieving) digital watermarks within (from) audio files. AWT lets you digitally and inaudibly sign audio content (music track).

AWT Features:

- cross-platform software (Windows, Linux, Mac OS, …)
- intuitive GUI front-end with drag-and-drop support
- console tools for easy automation and quick deployment on servers
- detailed documentation with examples and tips
- free evaluation package
- flexibly configurable algorithm
- very high watermarking data rate
- one license for any amount of machines (no binding to hardware or OS).

2.2. Secure Digital Music Initiative (SDMI)

SDMI plans to implement its digital rights architecture in two phases. In Phase I, portable music players will incorporate a watermarking and tracking system, allowing the players to continue to support both

secured and unsecured formats[7]. Then, when record labels are ready to begin distributing music with encryption security built-in, a screening technology will signal to the portable device that it is time to upgrade its software to play Phase II encoded music.

SDMI-compatible portable players will refuse to play pirated digital music files that are encoded with digital rights management information. Under the guidelines for the standard, a digital music file will be copy-able but future generations of the file cannot be reproduced, i.e. additional copies cannot be made from what is already a copy of the original music file.

3. PROPOSED FRAMEWORK OF AUDIO WATERMARKING SYSTEM

Audio watermarking system will hide information in MP3 files during the encoding process. Figure1 shows MP3 encoding process that is some parts of audio watermarking system

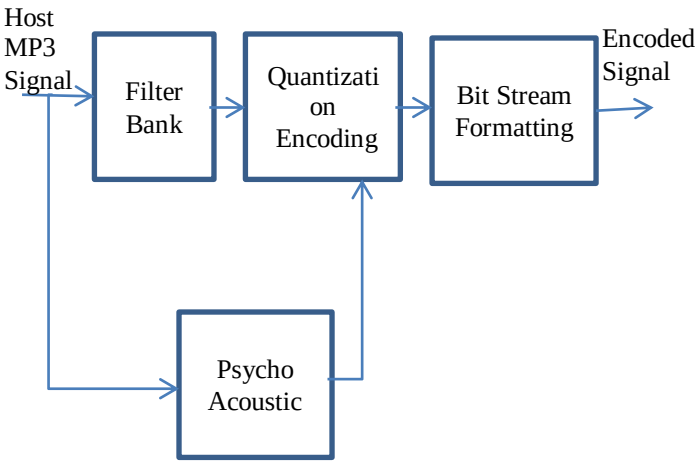


Figure 1. MP3 encoding

The hiding process takes place at the heart of the Layer III encoding process namely in the bit stream formatting. The part2_3_length variable contains the number of main_data bits in the MP3 bit stream. We encode the bits as its parity bit by changing the bit of watermark. Only randomly chosen part2_3_length value are modified; the selection is done using pseudo random bit generator. Figure 2 shows the process of audio watermarking system.

The documentation for this structure was generated from MP3 signal.

frame:: side info::granule::channel::part2_3_length

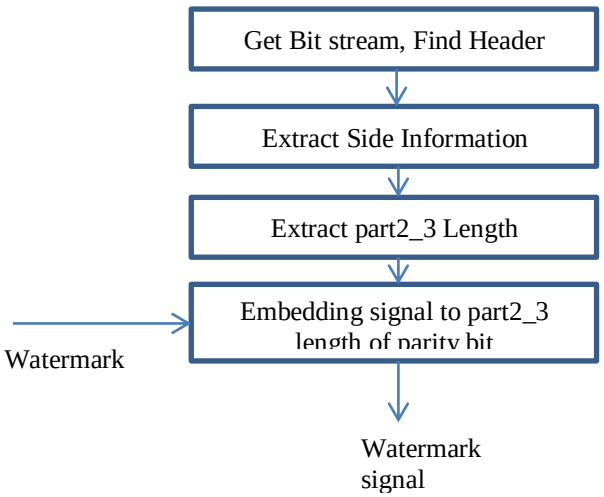


Figure 2. Block diagram of audio watermarking process

4. RELATED METHODOLOGIES

4.1. A Premier in MPEG Audio

MPEG-1 Audio Layer 3, or MP3 audio, is a pervasive digital format. Designed by the Moving Picture Experts Group (MPEG), MP3 uses a lossy form of audio compression to greatly reduce the size required to store audio information. MP3 files are created by using an encoding application to transform an audio track into an MP3. Using MP3 audio encoding, the size of an audio file can be reduced to about 1/10th of the size of the same file with no compression.

MP3 files are organized into sections designated “frames,” which can be abstractly thought of as a smaller version of the standard header/data format. As the audio data is encoded into an MP3 file, frames are created and the audio data is encapsulated into the frames. Each frame begins with a 32-bit header, which is used to synchronize the audio playback and specify characteristics of audio encapsulation; including bitrate, sampling rate information, and a flag bit that signals the presence of lack of a cyclic redundancy check within each MP3 frame for quality control.

The frame header of an MP3 file contains important technical information regarding the frame. Important information that can be found in the frame header includes:

• The frame sync block that enables the program that decodes the MP3 files to synchronize the bit stream with the MP3 file

• The MPEG Audio version ID and Layer descriptor that is used to encode the MP3 file

• Indicator bits to specify the presence of a cyclic redundancy check, padding bits, and copyright

• Audio information, including bitrate index, and stereo/mono channel mode

4.2. Bit Stream Formating

Below the figure3 describes the layout of a MP3 frame. Each block in the diagram indicates a size of 1 bit.

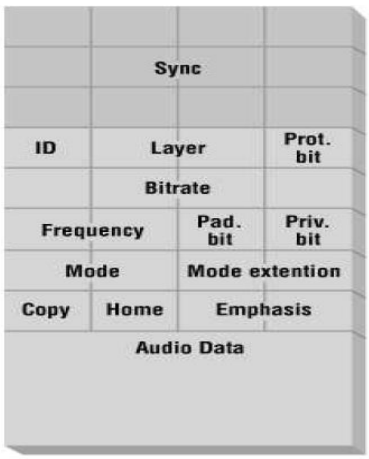


Figure 3. MP3 frame structure

Each frame in a MP3 bit stream is further split into 2 granules. The side information for each granule in a frame contains information needed to decode the main data. Table describes the component fields along with the size in bits for the side information. The size represents the requirements in single channel mode as well as the double that would be needed in dual channel mode.

Table 1. Side information for each granule

Name of MP3 bit stream	Number of bits
part2_3_length	12-24
big_values	9-18
global_gain	8-16
scalefac_compress	4-8
window_switching_flag	1-2
block_type	2-4
mixed_block_flag	1-2
table_slect[3]	10/20-15/30
subblock_gain[3]	9-18
region0_count	4-8
region1_count	3-6
preflag	1-2
scalefac_scale	1-2
count1table_select	1-2

5. RESULT ANALYSIS

The proposed steganalysis technique is implemented and tested on a set of 400 wav files. The audio samples include songs (pop, blue, rap, country, rock and r&b) nature noise etc. To assess the robustness against various attacks that are considered in this audio watermarking techniques are as follows:

(1) Resampling: conducting down sampling to 22,050 Hz and then up-sampling back to 44,100 Hz.

(2) Noise corruption: adding zero-mean white Gaussian noise to the watermarked audio signal with SNR=30 dB.

(3) MPEG compression: compressing and decompressing the watermarked audio signal with and MPEG layer III coder at a bit rate of 32 or 64 kbps.

When watermark are embedded into the MP3 signal during bit stream formatting process, which is later stage of MP3 compression, it will be difficult to distinguish whether the perceptual distortion is caused by the MP3 compression or watermarking. This technique can be adjustable to different degree of robustness, quality and different amount of data. For watermark security, parity based data hiding can be effective.

Part2_3_length and global_gain value from side information of auido signal. This process is done in MP3 encoding process. This section describes analyzing part2_3_length of MP3 file with compression rate 32kbps, 64kbpsfor highest amplitude rate of rock music file. Figure 4 presents the comparison for part2_3_length of MP3 file with 32kbps.

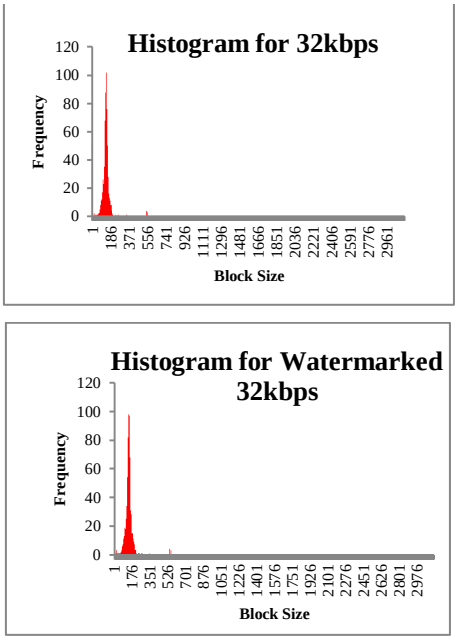


Figure 4. Comparison of part2_3_length over 32kbps compression rate (rock)

Figure 5 shows the comparison for part2_3_length of MP3 file with 64kbps. Block size (part2_3_length) are varied from the original MP3 files.

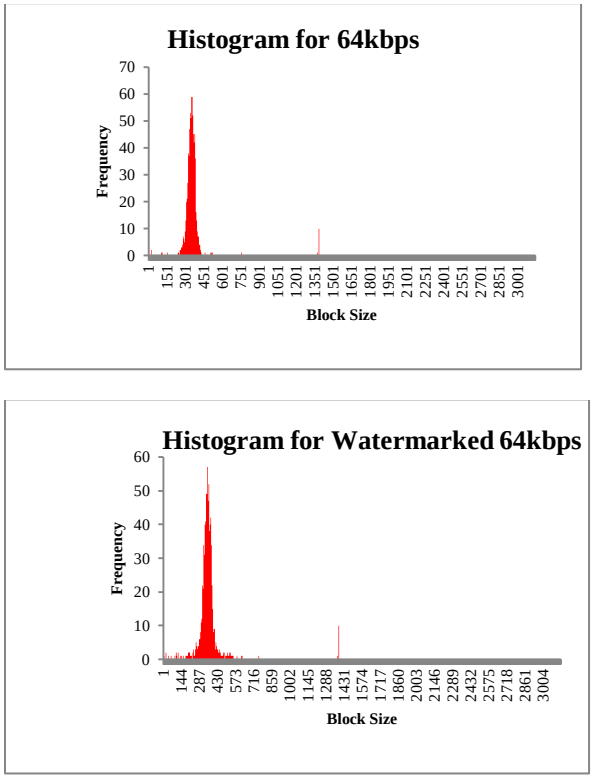


Figure 5. Comparison of part2_3_length over 64kbps compression rate (rock)

Figure 6 demonstrates that the steganography has a high robustness against various noise levels. Above 30% of noise level, adding zero-mean white Gaussian noise to the watermarked audio signal with SNR=30 dB.

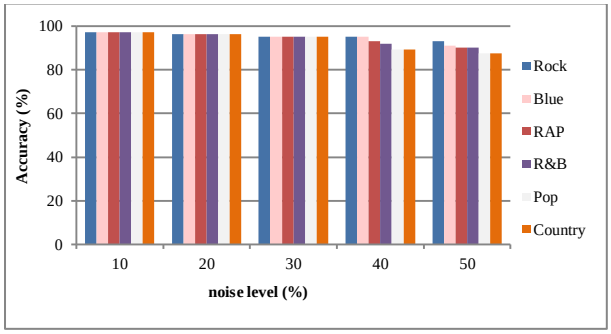


Figure 6: Testing accuracy with white gaussian noise

6. CONCLUSION

The proposed audio watermarking is a technique in which extra information is added to the data for securing the MP3 data from unauthorized use. Watermarks can be utilized as evidence for proving the ownership. This created the problem of protecting the intellectual copyrights of multimedia data. Evaluation analysis for this study showed that in above 30% of noise level, the performance accuracy still remained over 85%. The experimental result revealed that the proposed audio watermarking system is useful for identifying the origins of different multimedia files and resolving ownership disputes.

REFERENCES

[1] D.Yan, R.Wang "Steganalysis for MP3Stego using differential statistics of quantization step" Digital Signal Processing http:// dx.doi.org/10.1016/j.dsp.2013.02.013

[2] N.V.Lalitha, S. U. R. (2016). Performance Analysis of DCT and DWT Audio Watermarking based on SVD. Paper presented at the International Conference on Circuit, Power and Computing Technologies [ICCPCT].

[3] Rangkun Li, S. X., Huazhong Yang. (2016). Spread spectrum audio watermarking based on perceptual characteristic aware extraction. IET Signal Processing, 1-8.

[4] Sachin Balawant Takmare, R. K. G., Gajendra Singh Chandel. (2012). Voice Based Watermarking Technique for Relational Databases. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH, 1,(10).

[5] Singh, B. K. Data Hiding in Audio Using Matlab Software. International Journal of Advances in Engineering Science and Technology, 2(3).

[6] Subir. (2016). DWT-DCT based Blind Audio Watermarking using Arnold Scrambling and Cyclic Codes. Paper presented at the 3rd International Conference on Signal Processing and Integrated Networks (SPIN).

[7] Sujit M. Deokar, B. D. (2015). Blind Audio Watermarking Based On Discrete Wavelet and Cosine Transform. [Coference paper]. International Conference on Industrial Instrumentation and Control (ICIC)

[8] Toshiki Ito, H. K. (2016). Audio Watermarking Using Different Wavelet Filters. Springer International Publishing Switzerland(978-3-319-31960-5 26), 312–320.

[9] Yekta Said Can, F.A., and Melih Evren Burus. (2014). A Novel Spread Spectrum Digital Audio Watermarking Technique. Journal of Advances in Computer Networks, 2(1).

Teachers' Performance Appraisal by Using Fuzzy Expert System

Phyo Yatanar Lin[1], Nilar Thein[2]
[1]*Tutor, Faculty of Computer System and Technology, University of Computer Studies (Taungoo),*
[2]*Tutor, Faculty of Computer System and Technology, University of Computer Studies (Taungoo)*
phyoyatanarlin@gmail.com[1], nilarthein2016@gmail.com[2]

ABSTRACT: *The teachers' performances are played the most important role in teaching learning process. It is considered that faculty quality has direct bearing on developing and upholding quality in education. The more qualified teachers increase, the more outstanding students develop. It is difficult to decide the qualified or to evaluate the performance of teachers in real world. Using fuzzy expert system is the easiest way to solve these problems. Through fuzzy logic, linguistic variables, like Poor, Satisfy, Good and Very Good are defined by assigning weighted values to these qualitative facts. To compute the values of the final outputs, we apply all of the processes of fuzzy logic such as fuzzification, fuzzy inferencing and defuzzication based on the specified mathematical formulae. In this paper, a fuzzy expert system is designed to combine the knowledge and expertise of human experts with the reasoning capabilities. In this way, it is a great support for teachers' performance evaluation.*

Keywords: Fuzzy Logic; Expert System; Linguistic Variables; Crisp Values; Defuzzification; Membership Functions

1. INTRODUCTION

The performance of a teacher depends on various criteria such as teaching preparation, teaching process, personal ability, and professional ethics and so on. These criteria can be varied according to the decision makers and can be evaluated many various methods. Here, fuzzy expert system acts as an expert decision maker. An expert system which is the computer program using Artificial Intelligent (AI) [15] [16] techniques is a computer system that uses knowledge, facts and working memory and applies an appropriate reasoning technique to solve problems in a given field that is normally needed by the services of human experts.

An expert system can be designed based on a set of rules. As human beings have feelings and emotions, the expert system will be solving problems of the users without favoring the users. That's all, a fuzzy expert system is an expert system that uses fuzzy logic without using Boolean logic. A fuzzy expert system is a collection of membership functions [1] [2] [7] that is used to reason about data. Although the conversional expert systems are mainly symbolic reasoning engines, the fuzzy expert systems are intended to numerical processing.

In this paper, we discuss making the effective decision of the teachers' performances through AI technology. The fuzzy expert system [3] will be considering various inside for teachers such as research orientation, publication, work experience, promotion exam marks, job satisfaction, personal ability, etc. The eleven criteria are used in the system and the system will show the level of teacher quality as an output after calculation. The proposed system can be used to measure the performance of the teachers' quality in situation whenever and wherever required and is flexible enough that it could be enhanced with time.

2. RELATED WORKS FOR FUZZY EXPERT SYSTEM

Mir Anamul Hasan, Khaja Md. Sher-E-Alam and Ahsan Raja Chowdhury described a project work aiming to develop a web-based fuzzy expert system for diagnosing human diseases [4] [5]. Fuzzy systems use linguistic rules to describe systems. That research project focuses on the research and development of a web-based clinical tool designed to improve the quality of the exchange of health information between health care professionals and patients. Practitioners can also use that web-based tool to corroborate diagnosis. The proposed system is experimented on various scenarios in order to evaluate its performance. In all the cases, proposed system exhibited satisfactory results.

Lovi Raj Gupta and Avneet Kaur Dhawan implemented Diagnosis, Modeling and Prognosis of Learning System using Fuzzy Logic and Intelligent Decision Vectors [12]. In that paper fuzzy Expert Systems are used that are based on fuzzy logic and intelligent decision vectors to handle the quantitative as well as qualitative aspects in measuring the performance of an Educational Institution. The Academic performance of any institution is governed by various parameters that need to be studied in linguistic form. In the present work, a structured mathematical model is developed for individualistic and interdependent effects of these factors. Through fuzzification, they have converted the crisp values into linguistic variables [6] like very good, good, medium, low, high, very high. The two prime functions, each for

in-class and out-class activities are formulated then a control function for weaving the parameters within the function is crafted. A regulating function to encompass the dependencies of two prime functions is framed. A decision vector to engross both the prime and control function is originated to suggest the modifications on the present practices for enhancement of academia and overall performance of the institute [9] [13].

3. BACKGROUND THEORY

3.1. Fuzzy Logic

The central concept of the Fuzzy Logic [8] is the fuzzy sets. In real time situation number of times the boundaries of demarcation are not sharp enough leading to multiple outcome dependent upon context, person and ambient conditions. Boolean logic allows statements to be only 100% "TRUE" or "FALSE", in contrast the Fuzzy logic allows statements to be partially "TRUE" and partially "FALSE" at the same time. This peculiar nature of fuzzy logic has underlying strength that has made a powerful tool for problems solving in broad range of applications with evaluation being not an exception. The real world is an analogical in nature and not a discrete. Fuzzy logic though appears superficially to be an extension of a multi-value logic; its goal and application are different from multi-value logic. In general, approximation or fuzzy reasoning is the deduction of a precise conclusion emerging out of possible and imprecise initial sets.

3.2. Fuzzy Set

The basic idea of the fuzzy set theory is that an element belongs to a fuzzy set with a certain degree of membership. Thus, a proposition is not either true or false, but may be partly true or partly false to any degree. This degree is usually taken as a real number in the interval [0, 1]. A fuzzy set has three principal components: A degree of membership measure along the vertical Y-axis. The possible domain values for the set are along the horizontal X-axis. The set membership function is a continuous curve that connects a domain value to its degree of membership in the set.

3.3. Membership Function

The fuzzy membership function is a set of function values corresponding to input variables. The function that ties a number to each element x of the universe is called the membership function μ(x). An item's grade of membership is normally a real number between 0 and 1, often denoted by the Greek letter-μ. The membership functions can be represented by graph such as straight lines and triangular shape, etc. [14]

- Straight lines(increasing, decreasing)

$$l/(x_l, x_r, x) = \begin{cases} 0, & x < x_l \\ \dfrac{x - x_l}{x_r - x_l}, & x_l \le x \le x_r \\ 1, & x > x_r \end{cases} \quad (1)$$

$$l/(x_l, x_r, x) = \begin{cases} 1, & x < x_l \\ 1 - \dfrac{x_r - x}{x_r - x_l}, & x_l \le x \le x_r \\ 0, & x > x_r \end{cases} \quad (2)$$

- Triangular shape curves(4)

$$l/(x_l, x_c, x_r, x) = \begin{cases} 0, & x < x_l \\ \dfrac{x - x_l}{x_c - x_l}, & x_l \le x \le x_c \\ 1 - \dfrac{x - x_c}{x_r - x_c}, & x_c \le x \le x_r \\ 0, & x > x_r \end{cases} \quad (3)$$

3.4. Fuzzification

In the fuzzification process, the membership functions defined on the input variables are applied to their actual values, to determine the degree of truth for each rule premise. The degree of truth for a rule's premise is sometimes referred to as its alpha. If a rule's premise has a nonzero degree of truth then the rule is said to fire. Fuzzification means crisp value which is converted into Fuzzy input value with help of suitable membership function. Therefore, it is the process of converting the input crisp values into a set of membership values in the interval {0, 1} in the corresponding fuzzy sets. The membership function defined for each fuzzy set is applied on the input parameter to determine the degree of truth for each rule premise.

3.5. Fuzzy Inference Engine

In the inference sub processes, the truth value for the premise of each rule is computed, and applied to the conclusion part of each rule. This inference results in one fuzzy subset to be assigned to each output variable for each rule. The inference engine generates fuzzy outputs from the fuzzy rules ignited by the inputs. A fuzzy rule includes two parts: an **IF** part (called antecedent) and a **THEN** part (called consequent or

conclusion). Fuzzification and defuzzification need to access the corresponding membership function. Inferencing rules need to find the fired rules. All membership functions and rules are stored in knowledge base. The inference engine is the only one who accesses the knowledge base [17].

3.6. Defuzzification

There are more common defuzzification methods, two of them are the CENTROID and MAXIMUM methods. In the CENTROID method, the crisp value of the output variable is computed by finding the variable value of the center of gravity of the membership function for the fuzzy value.

In the MAXIMUM method, one of the variable values at which the fuzzy subset has its maximum truth value is chosen as the crisp value for the output variable. There are several variations of the MAXIMUM method that differ only in what they do when there is more than one variable value at which this maximum truth value occurs. One of these, the Weighted Average Method method, returns the average of the variable values at which the maximum truth value occurs. The formula of Weighted Average Method is shown in equation (4).

$$x^* = \frac{\sum \mu(x).x}{\sum \mu(x)} \qquad (4)$$

Where, $\sum$ denotes the algebraic summation and x is the element with maximum membership function.

4. PROPOSED SYSTEM

An expert system using fuzzy logic is used to handle uncertain and qualitative factors in the decision making. Qualitative factors can't be measured numerically and hardly to decide correctly. The proposed Teachers' Performance Appraisal System is also a qualitative factor, we apply the fuzzy approach to make the useful decision outputs. There are many criteria in real world to evaluate teachers' performance [10] [11]. In this paper, we uses eleven criteria and its sub-criteria which are really suited for universities, institutes and many private or public schools. These input criteria are got from the evaluation form which is collected from the teachers, colleagues, head of department and students. The system predefined the membership functions for attributes that are needed to calculate the fuzzy concept. When the inputs come to the system, the system will find each attribute value by using fuzzy concept according to the membership functions. The rules-based is defined by teacher experts. The output is the probability value between 0 and 1 and makes decision based on these values.

In this paper, the eleven input attributes such that research orientation, personal ability, supervision, professional ethics, administrative skills, teaching process, publication, work experience, promotion exam mark, job satisfaction, punctuality and their weight values are shown in Table 1. Among them, the first (9) attributes are defined as fuzzy term sets and (8) linguistic variables (Low, Medium, High, Very High) and (Poor, Satisfied, Good, Very Good) are denoted. The former linguistic variables Low, Medium, High and Very High are used for nine fuzzy term sets and the later linguistic variables Poor, Satisfied, Good, Very Good are used for the final fuzzy output which means Decision. These linguistic variables vary according to the weight values.

The first (6) attributes consists of corresponding sub attributes which are also defined weight values presented in the following Table 1. In each attribute, the total weight values for all sub-attributes must be one. Weighted Average Method is used in this proposed system. The weight values of last two attributes (job satisfaction and punctuality) are already setting up from the proposed system by predefined weight 0.05.

Figure 1 describes the detail design of the proposed system. Firstly, the system accepts the user's inputs and calculates the fuzzy values for the first nine fuzzy term sets by using membership function. The last two attributes values are using in rules matching. They are not needed to search with fuzzy calculation. Then, the system inferences rules from database and aggregates fuzzy values. To develop the final decision, the system defuzzifies by using defuzzification method. Finally, the system displays the final decision that is calculated from fuzzy values by using the corresponding membership function.

Table 1. Input Attributes and Weight Values of the System

No.	Attribute Names	Weight Values
1	**Research Orientation**	
	1.1. Research Potential	0.05
	1.2. Standard of Project	0.05
	1.3.Participations and Organization of workshop seminars and conferences	0.2
	1.4. Research Production	0.5
	1.5.Membership in Research Societies	0.2
2	**Personal Ability**	
	2.1.Intellectual Ability	0.1
	2.2.Self Confidence	0.1
	2.3.Problem Solving Skills	0.2
	2.4.Cooperative	0.2
	2.5. Intelligence	0.1
	2.6.Reliability	0.2
	2.7.Flexibility and Adaptability	0.1
3	**Supervision**	
	3.1.Controlling students in class	0.3
	3.2.Student Supervision	0.3
	3.3.Supervision activities other than teaching	0.3
	3.4.Interpersonal Relationships	0.1
4	**Professional Ethics**	
	4.1.Temperament and manners	1/6
	4.2.Interaction with students	1/6
	4.3.Interaction with colleagues	1/6
	4.4.Interaction with officers	1/6
	4.5.Interaction with lower staff	1/6
	4.6.Interaction with visitors/ guests	1/6
5	**Administrative Skills**	
	5.1.Leadership	0.2
	5.2.Decision making skills	0.2
	5.3.Care of rules and regulation	0.2
	5.4.Controlling crises situation and unce	0.2
	5.5.Listening suggestion of others	0.2
6	**Teaching Process**	
	6.1.Proficiency in teaching	0.2
	6.2.Personal Interest in teaching	0.1
	6.3.Presentation and communication skills	0.2
	6.4.Lecture preparation	0.2
	6.5.Question trackling	0.1
	6.6.Fairness in marking	0.1
	6.7.Content knowledge	0.1
7	**Publication**	Numerical Value
8	**Work experience**	Numerical Value
9	**Promotion exam marks**	Numerical Value
10	**Job satisfaction**	Yes (0.05)/ no (0)
11	**Punctuality**	Yes (0.05)/ no (0)

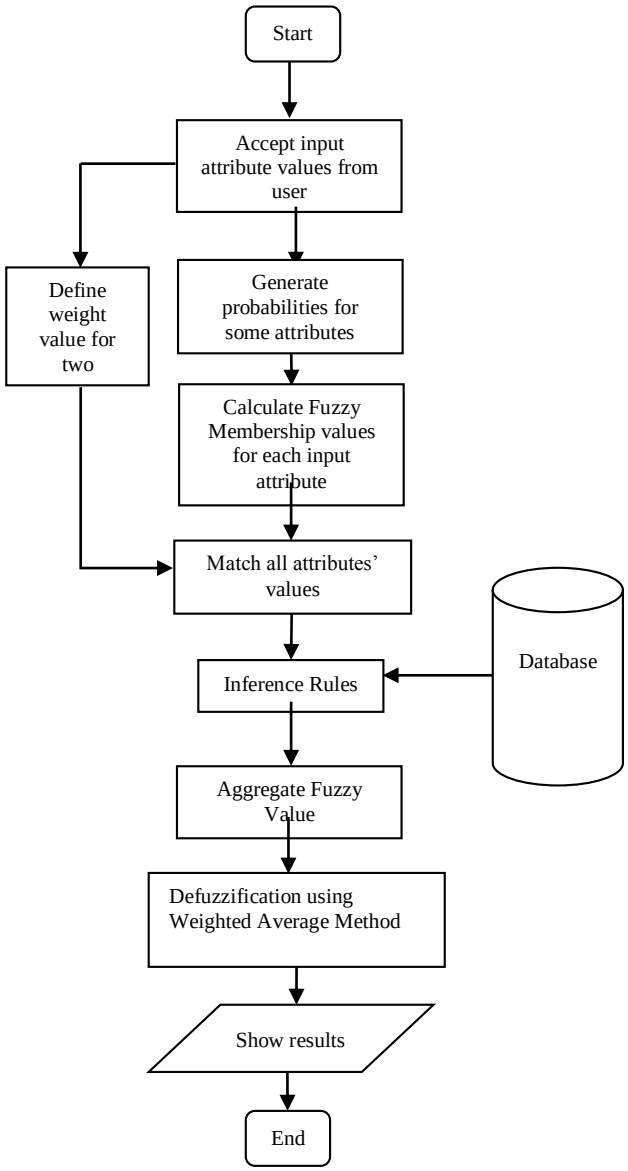


Figure1. Detail proposed system design

The fuzzy rules for the first (6) attributes are as the following form.

For Research orientation,
IF user select **Research Potential** THEN w=0.05
IF user select **Standard of Project** THEN w=0.05
IF user select **Participations and Organization of workshop seminars and conferences** THEN w=0.2
IF user select **Research Production** THEN w=0.5
IF user select **Membership in Research Societies** THEN w=0.2
For Personal Ability,
IF user select **Intellectual Ability** THEN w=0.1
IF user select **Self Confidence** THEN w=0.1
IF user select **Problem Solving Skills** THEN w=0.2
IF user select **Cooperative** THEN w=0.2
IF user select **Intelligence** THEN w=0.1
IF user select **Reliability** THEN w=0.2
IF user select **Flexibility and Adaptability** THEN w=0.1
For Supervision,
IF user select **Controlling students in class** THEN w=0.3
IF user select **Student Supervision** THEN w=0.3
IF user select **Supervision activities other than teaching** THEN w=0.3
IF user select **Interpersonal Relationships** THEN w=0.1
For Professional Ethics,
IF user select **Temperament and manners** THEN w=1/6
IF user select **Interaction with students** THEN w=1/6
IF user select **Interaction with colleagues** THEN w=1/6
IF user select **Interaction with officers** THEN w=1/6
IF user select **Interaction with lower staffs** THEN w=1/6
IF user select **Interaction with visitors/guests** THEN w=1/6
For Administrative Skill,
IF user select **Leadership** THEN w=0.2
IF user select **Decision making skill** THEN w=0.2
IF user select **Care of rules and regulations** THEN w=0.2
IF user select **Controlling crises situation and uncertainty** THEN w=0.2
IF user select **Listening suggestion of others** THEN w=0.2
For Teaching Process,
IF user select **Proficiency in Teaching** THEN w=0.2
IF user select **Personal Interest in teaching** THEN w=0.1
IF user select **Presentation and communication skills** THEN w=0.2
IF user select **Lecture preparation** THEN w=0.2
IF user select **Question trackling** THEN w=0.1
IF user select **Fairness in marking** THEN w=0.1
IF user select **Content knowledge** THEN w=0.1

To implement the proposed system for a teacher, the first step is to accept the input attributes according to the teacher performances as shown in the following table 2. Then, the system calculates the total weight of selected attributes. And then, the system finds the fuzzy sets such that {Medium, High}, {High}, {Low, Medium}, etc. by using the corresponding membership functions in the fuzzification stage as shown in Table 2.

Table 2. Input Data and Fuzzification proces

Attributes	Sub Attributes selected by user	Total Weight	Name of Group that received by calculating Membership Function
1.Research orientation	1.1.Research potential 1.3.Perticipation and organization of seminars 1.4.research production	0.7	{Low,Medium,High, Very High} ={0.0, 0.0, 1.0, 0.0} = {High}
2.Personal Ability	2.1.Intellectual Ability 2.3.Problem Solving Skill 2.4.Cooperative 2.5.Intelligence	0.8	{Low,Medium,High, Very High} ={0.0, 0.0, 0.5, 0.0} = {High}
3.Supervision	3.1.Controlling Students in Class 3.2.Students Supervision	0.6	{Low,Medium,High, Very High} ={0.0, 0.0, 0.5, 0.0} = {High}
4.Professional Ethics	4.1.Temperament and Manners 4.4.Interaction with Officers 4.5.Interaction with Lower Staff	0.51	{Low,Medium,High, Very High} ={0.0,0.4499999999 9999996,0.05000000 000000006, 0.0} = {Medium, High}
5.Administrative Skill	5.1. Leadership 5.2.Decision Making Skill 5.5. Listening Suggestions of other	0.6	{Low,Medium,High, Very High} ={0.0, 0.0, 0.5, 0.0} = {High}
6.Teaching Process	6.1. Proficiency in Teaching 6.4. Lecture Preparation 6.5. Question Tacking 6.6. Fairness in Marking 6.7. Content Knowledge	0.7	{Low,Medium,High, Very High} ={0.0, 0.0, 1.0, 0.0} = {High}
7.Numbers of Publication		6	{Low,Medium,High, Very High} ={0.0, 0.5, 0.0, 0.0} = {Medium}
8.Working Experience (Years)		8	{Low,Medium,High, Very High} ={0.4, 0.111111111111111 1, 0.0, 0.0} = {Low, Medium}
9.Promotion Exam Mark		77	{Low,Medium,High, Very High} ={0.0, 0.0, 0.30000000000004, 0.4} = {High, Very High}
10. Job Satisfaction		Yes	0.05
11. Punctuality		Yes	0.05

After calculating the membership function, identified the weight values for Rule matching of fuzzy sets. These are shown in the following Table 3 and Table 4 respectively.

Table 3. Weight Values for Rule Matching of Fuzzy set

Fuzzy Set	Weight Value
Low	0.01
Medium	0.05
High	0.075
Very High	0.1
Yes	0.05
No	0

Table 4. Range Values for Rule Matching

Range Value	Decision
$0.09 \leq x < 0.55$	Poor
$0.55 \leq x < 0.65$	Satisfy
$0.65 \leq x < 0.8$	Good
$0.8 \leq x \leq 1$	Very Good

And then, the number of rules for results fuzzy sets can be generated by using the following equation.

$$N_R = 2^i \quad (5)$$

Where, N_R = Number of Rules

i = Number of Fuzzy Sets that are generated from Table 3

After calculation the weight values for all rules, Defuzzification (Weighted Average Method) is implemented to make the final decision. The Defuzzification process of the system's calculation is as shown in the following:

x=(0.6*0.30000000000000004+0.6*0.2+0.6*0.1111111111111111+0.6*0.1111111111111111+0.6*0.05000000000000006*0.6*0.05000000000000006 06) / (0.4+0.2+0.1111111111111111+ 0.1111111111111111+0.05000000000000006+ 0.05000000000000006+ 0.05000000000000006+ 0.05000000000000006)

x=0.6

where, x=Defuzzification Value

After completion of the Defuzzification process, the implemented Fuzzy Expert System makes the final decision by using the classification range in the following Table 6.

Table 6. Classification of Decision

Input Field	Range	Fuzzy Set
Decision	0.0 – 0.4 0.2 - 0.6 0.4 – 0.8 0.6 – 1	Poor Satisfied Good Very Good

The defuzzified value x=0.6 is included in the above Fuzzy Set decision: Satisfied, Good and Very Good. Then the system calculates the corresponding membership functions for final decision as shown in the followings:

$$Satisfied(x) = \begin{cases} 0, & x < 0.2 \\ \frac{x - 0.2}{0.4 - 0.2}, & 0.2 \leq x < 0.4 \\ 1 - \frac{x - 0.4}{0.6 - 0.4}, & 0.4 \leq x \leq 0.6 \\ 0, & x > 0.6 \end{cases} \quad (6)$$

$$Good(x) = \begin{cases} 0, & x < 0.4 \\ \frac{x - 0.4}{0.6 - 0.4}, & 0.4 \leq x < 0.6 \\ 1 - \frac{x - 0.6}{0.8 - 0.6}, & 0.6 \leq x \leq 0.8 \\ 0, & x > 0.8 \end{cases} \quad (7)$$

$$VeryGood(x) = \begin{cases} 0, & x < 0.6 \\ \frac{x - 0.6}{0.8 - 0.6}, & 0.6 \leq x < 0.8 \\ 1 - \frac{x - 0.8}{1 - 0.8}, & 0.8 \leq x \leq 1 \\ 0, & x > 1 \end{cases} \quad (8)$$

The system calculates the final decision upon defuzzified value x=0.6 using equation (6), (7) and (8) that satisfied the value x.

Satisfied =1-[(0.6-0.4)/ (0.6-0.4)] =1-1=0

Good = [(0.6-0.4)/ (0.6-0.4)] =1

Very Good= [(0.6-0.6)/ (0.8-0.6)] =0

Decision = {Satisfied, Good, Very Good} = {0, 1, 0}

Above two fuzzy outputs, the maximum value is selected for the best decision. Thus, the teachers' performance is **'Good'** in this implemented example as shown in Figure 2. To display the result clearly and flexibly, each attribute and decision can be represented by the membership graph as shown in Figure 3.

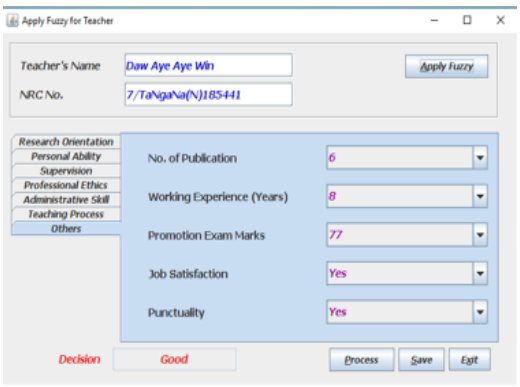


Figure 2. Calculation of Final Decision

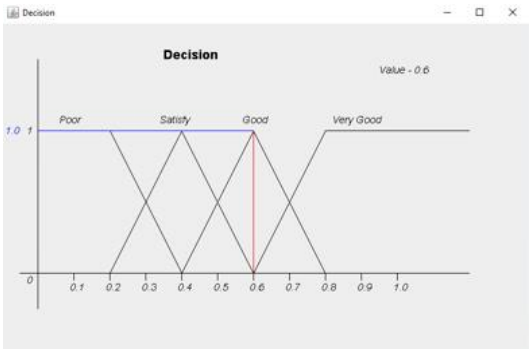


Figure 3. Final Decision Membership Graph

5. CONCLUSION

In this paper, the uncertain knowledge of the problem domain have been handled through the integration of expert system technology with fuzzy

logic concept. Moreover, it can be applied for the control systems and other applications to improve the efficiency and simplicity. This system provides valuable result for head or principle of the universities as an expert decider. By evaluating the teachers' performance, many qualified and outstanding teachers' can be decided in real world. This proposed system can help a decision maker for better decision. The fuzzy expert system's model may be used in other similar problems like courses selection, supervisor selection, courses designs and all types of employees' assessment in universities and other private organizations.

6. RESULTS AND DISCUSSION

In this proposed fuzzy approach for best teacher evaluation has been implemented in Java Programming Language. The proposed system has been tested with 50 teachers. For each teacher ranks of all criteria were fuzzified by means of the Straight Line and Triangular Shape Curve membership functions. The output (final decision) was calculated and then defuzzified by calculating the Average of Maximum method. The decision of proposed system is compared with the recorded interview data from the teachers, students and staffs of Computer University. The following Table 7 is the comparison of traditional decision and fuzzy decision.

Table 7- Comparison of Traditional and Fuzzy Decision

Decision	*Traditional Decision*	*Fuzzy Decision*
Poor	4%	2%
Satisfied	48%	46%
Good	40%	42%
Very Good	8%	10%

REFERENCES

[1] Bilgic, Taner & Turksen, "Measurement of Membership Function: Theoretical and Experimental Work", In D. Dubois and H. Prade (editors) Handbook of Fuzzy Systems, Vol. 1, Fundamentals of Fuzzy Sets, chapter 3, Kluwer Academic Publishers, Burhan, 1999, pp.195-202.

[2] Zhao, Jin and Bose, "Evaluation of Membership Function for Fuzzy Logic Controlled Induction Motor Drive", Proceeding of the IEEE 28th Annual Conference of the Industrial Electronics Society (IECON'02). Sevilla, Spain: 229-234, B.K. 2002.

[3] D. W. Patterson, "Introduction to Artificial Intelligence and Expert Systems", Prentice Hall of India Private Limited, New Delhi, pp.326-327, 2006.

[4] M. Anamul Hasan, K. Md. Sher-E-Alam and A. Raja Chowdhury, "Human Disease Diagnosis Using a Fuzzy Expert System", Journal of Computing, Volume 2, Issue 6, ISSN 2151-9617, June 2010.

[5] Nguyen H. P., and Kreinovich V., "Fuzzy Logic and Its Applications in Medicine", International Journal of Medical Informatics, 62, 2001, pp. 165–173.

[6] L.A.Zadeh, "The concept of a linguistic variable and its application to appropriate reasoning", Information Sciences, 1975, pp. 43-80.

[7] A.Voloshyn, G. Gnatienko and E. Drobot, "Fuzzy Membership Functions in a Fuzzy Decision Making Problem", International Journal "Information Theories & Applications", Vol.10.

[8] L.A. Zadeh, "Fuzzy sets," Information and Control," vol 8, pp. 338-353, 1965.

[9] R. Singh Yadav, "Modeling Academic Performance Evaluation Using Soft Computing Techniques: A Fuzzy Logic Approach", Research Scholar, Department of Computer Science and Engineering, Sharda University, Gr. Noida

[10] Mujis.D, "Measuring Teacher Effectiveness: Some methodological Reflections", Educational Research and Evaluation, 12(1), pp. 53-74.

[11] Kyriakides, L., Demetriou, D., &Charalambous, C., "Generating criteria for evaluation teachers through tracher effectiveness research", Educational Research, 48(1), 2006, pp.1-20.

[12] L. Raj Gupta and A. Kaur Dhawan, "Diagnosis, Modeling and Prognosis of Learning System using Fuzzy Logic and Intelligent Decision Vectors", International Journal of Computer Applications (0975 – 8887), Volume 37– No.6, January 2012.

[13] S. Patil1, A. Mulla and R.R. Mudholkar, "Best Student Award – A Fuzzy Evaluation Approach", International Journal of Computer Science and Communication, Vol. 3, No. 1, January-June 2012, pp. 9-12.

[14] T. Ozen and J. M. Garibaldi, "Effect of Type-2 Fuzzy Membership Function Shape on Modelling Variation in Human Decision Making", University of Nottingham, Nottingham, NG8 1BB, UK.

[15] D. W. Patterson, "Introduction to Artificial Intelligence and Expert Systems", Prentice Hall of India Private Limited, New Delhi, pp.326-327, 2006.

[16] N.P. Padhy, "Artificial intelligence and Intelligent System," Oxford University Press, p.p. 337, 2005.

[17] S. Singh, R.S. Yadav, D. Singh, "Fuzzy inference Mechanism and Rule Induction Algorithm for Performance Evaluation", proceedings of National Conference on Advances and Trends Web Based Computing (ATWBC-2010) organized by Department of Computer Science and Applications, M.G. Kashi Vidyapith, Varanasi, India, pp 86-93, 2010.

A Survey of Comparison in Decision Tree Algorithms for Classification

Kyi Zar Nyunt[1], Wint Aye Khaing[2], San San Yu[3]
[1, 2] *Faculty of Information Science, University of Computer Studies (Taungoo)*
[3] *Faculty of Computer Science, University of Computer Studies (Kyaing Tong)*
kyizar81@gmail.com, wintayekhaing5@gmail.com, sanyumaw06@gmail.com

ABSTRACT: *Data mining is the useful tool to discovering the knowledge from large data. Different methods and algorithms are available in data mining. Classification is most common method used for finding the mine rule from the large database. Decision tree method generally used for the Classification, because it is the simple hierarchical structure for the user understanding and decision making. Various data mining algorithms available for classification based on Artificial Neural Network, Nearest Neighbour Rule and Baysen classifiers but decision tree mining is simple one. In this paper, student qualitative data has been taken from Educational data mining and the performance analysis of the decision tree algorithm ID3, C4.5 and CART are compared.*

Keywords: Educational Data Mining; Decision tree; ID3; C4.5; CART;

1. INTRODUCTION

Classification is a form of data analysis that extracts models describing important data classes. Such models, called classifiers, predict categorical (discrete, unordered) class labels. Many classification methods have been proposed by researchers in machine learning, pattern recognition, and statistics. Most algorithms are memory resident, typically assuming a small data size. Recent data mining research has built on such work, developing scalable classification and prediction techniques capable of handling large amounts of disk-resident data. Classification has numerous applications, including fraud detection, target marketing, performance prediction, manufacturing, and medical diagnosis [6].

Decision tree (DT) has been widely used to build classification models, due to its simple representation that resembles the human reasoning [4]. There are various decision tree algorithms. Among them this paper includes comparison of famous decision tree algorithms (Iterative Dichotomiser 3 (ID3) algorithm, C4.5 algorithm and Classification And Regression Tree (CART) algorithm) and their advantages and disadvantages.

Educational Data Mining is an emerging field that can be applied to the field of education, it concerns with developing methods that discover knowledge from data originating from educational environments [3]. Decision tree algorithms can be used in educational field to understand performance of students [4].

2. DECISION TREE

Decision tree can be constructed relatively fast compared to other methods of classification. Trees can be easily converted into SQL statements that can be used to access databases efficiently. Decision tree classifiers obtain similar and sometimes better accuracy when compared with other classification methods. Decision tree algorithm can be implemented in a serial or parallel fashion based on the volume of data, memory space available on the computer resource and scalability of the algorithm [9].

The tree is a structure build from elements called nodes and branches. Three types of nodes can be distinguished: a root, internal and terminal (leaf). The root and the internal nodes denote tests on the attributes and each branch represents the outcome of a test. Each leaf node holds a class label i.e. the final decision. The general algorithm for building decision tree has two phases: growth and pruning.

In the first phase a decision tree is built by selecting the best test attribute as the root of the decision tree. Based on the test the learning set is split into two subsets (two children nodes). Then, repeat recursively the procedure on each branch to induce the remaining levels of the decision tree until all instances in a leaf belong to the same class. Different algorithms use different metrics to determine the best way to generate the test in the node and split the records. The most common impurity measures (metrics) are: Gini Index, Information Gain, and Gain Ratio.

The second, pruning phase may be done only on the fully grown tree. The goal is to reduce the size of decision trees by removing "insignificant" nodes or even sub-trees (sections of a node that may be based on noisy or erroneous data). A tree that is too large risks over-fitting the training data and poorly generalizing to new samples [4].

The benefits of having a decision tree are as follows

- It does not require any domain knowledge.
- It is easy to comprehend.
- The learning and classification steps of a decision tree are simple and fast [10].

3. DECISION TREE INDUCTION ALGORITHMS

3.1. ID3 (Iterative Dichotomiser 3)

In the decision tree method, information gain approach is generally used to determine suitable property for each node of a generated decision tree. Therefore, we can select the attribute with the highest information gain (entropy reduction in the level of maximum) as the test attribute of current node. In this way, the information needed to classify the training sample subset obtained from later on partitioning will be the smallest. So, the use of this property for partitioning the sample set contained in current node will make the mixture degree of different types for all generated sample subsets reduced to a minimum. Hence, the use of an information theory approach will effectively reduce the required dividing number of object classification [5].

For constructing a decision tree information gain is calculated for each and every attribute and attribute with the highest information gain becomes the root node. The rest possible values are denoted by arcs. After that, all the outcome instances that are possible are examined whether they belong to the same class or not. For the instances of the same class, a single name class is used to denote otherwise the instances are classified on the basis of splitting attribute [3].

ID3 is based on the Concept Learning System (CLS) algorithm that is in fact a recursive top-down divide-and-conquer algorithm. The ID3 family of decision tree induction algorithms uses information theory to decide which attribute shared by a collection of instances to split the data on next. Attributes are chosen repeatedly in this way until a complete decision tree that classifies every input is obtained. If the data is noisy, some of the original instances may be misclassified. It may be possible to prune the decision tree in order to reduce classification errors in the presence of noisy data. The speed of this learning algorithm is reasonably high, as is the speed of the resulting decision tree classification system [7].

Continuous attributes can be handled using the ID3 algorithm by discretizing or directly, by considering the values to find the best split point by taking a threshold on the attribute values [9].

ID3 is a supervised learning algorithm, [10] builds a decision tree from a fixed set of examples. The resulting tree is used to classify future samples. ID3 algorithm builds tree based on the information (information gain) obtained from the training instances and then uses the same to classify the test data. ID3 algorithm generally uses nominal attributes for classification with no missing values [2].

***Inputs**: R: a set of non- target attributes, C: the target attribute, S: training data.*

***Output**: returns a decision tree*

Start

Initialize to empty tree;

If** S is empty **then

***Return** a single node failure value*

End If

If** S is made only for the values of the same target **then

***Return** a single node of this value*

End if

If** R is empty **then

***Return** a single node with value as the most common value of the target attribute values found in S*

End if

D ← the attribute that has the largest Gain (D, S) among all the attributes of R

{d_j j = 1, 2, ..., m} ← Attribute values of D

{S_j with j = 1, 2, ..., m} ←The subsets of S respectively constituted of d_j records attribute value D

***Return** a tree whose root is D and the arcs are labeled by d1, d2, ..., d_m and going to sub-trees ID3 (R-{D}, C, S1), ID3 (R-{D} C, S2), .., ID3 (R-{D}, C, Sm)*

End

Figure 1. Pseudocode of ID3 algorithm

In Figure 1, the pseudo code of this algorithm is very simple. Given a set of attributes not target C_1, C_2, …,C_n, C the target attribute, and a set S of recoding learning. When start the algorithm, it is initialize to empty tree. If S is empty, the algorithm returns a single node failure value. If S is made only for the values of the same target, it returns a single node of this value. If R is empty, it returns a single node with value as the most common value of the target attribute values found in S. D is the attribute that has the largest Gain (D, S) among all the attributes of R. Attribute values of D are d_1, d_2 to d_m. S_1, S_2 to S_m are subsets of S respectively constituted of dj records attribute value D. Root is D and the arcs are labeled by d_1, d_2, …, d_m and going to sub-trees ID3(R-{D}, C, S1), ID3 (R-{D} C, S2), .., ID3 (R-{D}, C, Sm).

3.2. C4.5

C4.5 is an algorithm used to generate a decision tree developed by Ross Quinlan. C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for

classification, and for this reason C4.5 is often referred to as a statistical classifier [7]. C4.5 algorithm uses information gain as splitting criteria. It can accept data with categorical or numerical values. To handle continuous values it generates threshold and then divides attributes with values above the threshold and values equal to or below the threshold. C4.5 algorithm can easily handle missing values. As missing attribute values are not utilized in gain calculations by C4.5 [8].

It is also known as J48 algorithm. It is also based on Hunt's algorithm. It is serially implemented like ID3. C4.5 handles both categorical and continuous attributes to build a decision tree. In order to handle continuous attributes, C4.5 splits the attribute values into two partitions based on the selected threshold such that all the values above the threshold as one child and the remaining as another child. It removes the biasness of information gain when there are many outcome values of an attribute.

It is suitable for real world problems as it deals with numeric attributes and missing values. The algorithm can be used for building smaller or larger, more accurate decision trees and the algorithm is quite time efficient. This algorithm is used to handle continuous attributes e.g. temperature. C4.5 improves computational efficiency [9].

Pruning takes place in C4.5 by replacing the internal node with a leaf node thereby reducing the error rate. It accepts both continuous and categorical attributes in building the decision tree (Anju Rathee). It has an enhanced method of tree pruning that reduces misclassification errors due to noise and too many details in the training data set [1].

3.3. CART

It stands for Classification And Regression Trees. It was introduced by Breiman in 1984. It builds both classifications and regression trees. The classification tree construction by CART is based on binary splitting of the attributes. CART also based on Hunt‟s algorithm and can be implemented serially. Gini index is used as splitting measure in selecting the splitting attribute. CART is different from other Hunt‟s based algorithm because it is also use for regression analysis with the help of the regression trees. The regression analysis feature is used in forecasting a dependent variable given a set of predictor variables over a given period of time. CARTS supports continuous and nominal attribute data and have average speed of processing [5].

CART is a non-parametric decision tree learning technique that produces either classification or regression trees, depending on whether the dependent variable is categorical or numeric, respectively. The word binary implies that a node in a decision tree can only be split into two groups. CART uses gini index as impurity measure for selecting attribute. The attribute with the largest reduction in impurity is used for splitting the node's records. CART accepts data with numerical or categorical values and also handles missing attribute values. It uses cost-complexity pruning and also generate regression trees [8].

CART has enhanced features and capabilities that address the short-comings of CART giving rise to a modern decision tree classifier with high classification and prediction accuracy [1].

Table 1. Comparison between different Decision Tree Algorithm

ID3	**ID3**, or Iternative Dichotomizer, was the first of three Decision Tree implementations developed by Ross Quinlan. It builds a decision tree for the given data in a top-down fashion, starting from a set of objects and a specification of properties Resources and Information each node of the tree, one property is tested based on maximizing information gain and minimizing entropy, and the results are used to split the object set. This process is recursively done until the set in a given sub-tree is homogeneous (i.e. it contains objects belonging to the same category). The ID3 algorithm uses a greedy search. It selects a test using the information gain criterion, and then never explores the possibility of alternate choices.
C4.5	**C4.5**, Quinlan's next iteration. The new features (versus ID3) are: (i) accepts both continuous and discrete features; (ii) handles incomplete data points; (iii) solves over-fitting problem by (very clever) bottom-up technique usually known as "pruning"; and (iv) different weights can be applied the features that comprise the training data. Of these, the first *three* are very important--and I would suggest that any DT implementation you choose has all three. The fourth (differential weighting) is much less important.
CART	**CART**, or *Classification And Regression Trees* is often used as a generic acronym for the term Decision Tree, though it apparently has a more specific meaning. In sum, the CART implementation is very similar to C4.5; the one notable difference is that CART constructs the tree based on a numerical splitting criterion recursively applied to the data, whereas C4.5 includes the intermediate step of constructing *rule set*s.

4. ATTRIBUTE SELECTION MEASURES

The most popular attribute selection measures are – Entropy (Information Gain), Gain Ratio and Gini Index.

4.1. Entropy

Entropy is a measure of uncertainty associated with a random variable. The entropy increases with the increase in uncertainty or randomness and decreases with a decrease in uncertainty or randomness. The value of entropy ranges from 0-1.

$$Entropy(D) = \sum_{i=1}^{c} -p_i log_2(p_i) \quad (1)$$

where p_i is the non-zero probability that an arbitrary tuple in D belongs to class C and is estimated by $|C_{i,D}|/|D|$. A log function of base 2 is used because as stated above the entropy is encoded in bits 0 and 1.

4.2. Information Gain

ID3 uses information gain as its attribute selection measure. Claude Shannon studied the value or "information content" of messages and gave information gain as a measure in his Information Theory [3]. Information Gain is the difference between the original information gain requirement (i.e. based on just the proportion of classes) and the new requirement (i.e. obtained after the partitioning of A).

$$Gain(D,A) = Entropy(D) - \sum_{j=1}^{v} \frac{|Dj|}{|D|} Entropy(D_j) \quad (2)$$

Where,
D: A given data partition
A: Attribute
V: Suppose we partition the tuples in D on some attribute A having v distinct values
D is split into v partition or subsets, $\{D_1, D_2, .. D_j\}$ were D_j contains those tuples in D that have outcome a_j of A.
The attribute that has the highest information gain is chosen.

4.3. Gain Ratio

The information gain measure is biased towards tests with many outcomes. That is it prefers to select attributes having a large number of values. As each partition is pure, the information gain by partitioning is maximal. But such partitioning cannot be used for classification.

C4.5 (a successor of ID3) uses this attribute selection measure named Gain Ratio which is an extension to the information gain. Gain Ratio differs from information gain, which measures the information with respect to a classification that is acquired based on some partitioning. Gain Ratio applies kind of information gain using a "split information" value defined as:

$$SplitInfo_A = -\left|\sum_{j=1}^{v} \frac{|Dj|}{|D|} log_2\left(\frac{|Dj|}{|D|}\right)\right. \quad (3)$$

The Gain Ratio is then defined as:

$$Gain\ Ratio\ (A) = \frac{Gain\ (A)}{SplitInfoA(D)} \quad (4)$$

A splitting attribute is selected which is the attribute having the maximum Gain Ratio. The gain ratio becomes unstable if the split information tends to 0. A constraint is added to avoid such condition, whereby the information gain of test selected must be large- at least as great as the average gain over all tests examined.

4.4. Gini Index

Gini Index is an attribute selection measure used by the CART decision tree algorithm. The Gini Index measures the impurity D, a data partition or set of training tuples as:

$$Gini(D) = 1 - \sum_{i=1}^{m} p_i^2 \quad (5)$$

Where p_i is the probability that a tuple in D belongs to class C_i and is estimated by $|C_{i,D}|/|D|$. The sum is computed over m classes. The attribute that reduces the impurity to the maximum level (or has the minimum gini index) is selected as the splitting attribute [3].

5. STRONG AND WEAKENESS OF DECISION TREE ALGORITHMS

The following table shows the strong and weakness of ID3, C4.5 and CART.

	Strong	**Weakness**
ID3	-The training data is used to create understandable prediction rules. -It builds the fastest as well as a short tree. -ID3 searches the whole dataset to create the whole tree. -It finds the leaf nodes thus enabling the test data to be pruned and reducing the number of tests. -The calculation time of ID3 is the linear function of	-For a small sample, data may be over-fitted or over-classified. -For making a decision, only one attribute is tested at an instant thus consuming a lot of time. -Classifying the continuous data may prove to be expensive in terms of computation, as many trees have to be generated to see where to break the continuum. -One disadvantage of ID3 is that when given a large number of input

	the product of the characteristic number and node number	values, it is overly sensitive to features with a large number of values [2].
C4.5	-C4.5 is easy to implement. -C4.5 builds models that can be easily interpreted. -It can handle both categorical and continuous values. -It can deal with noise and deal with missing value attributes.	-A small variation in data can lead to different decision trees when using C4.5. -For a small training set, C4.5 does not work very well.
CART	-CART can handle missing values automatically using surrogate splits. -Uses any combination of continuous/discrete variables. -CART automatically performs variable selection. -CART can establish interactions among variables. -CART does not vary according to the monotonic transformation of predictive variable [7].	-CART may have unstable decision trees. -CART splits only by one variable. -Non-parametric.

6. COMPARISON

In this paper, three decision tree algorithms such as ID3, C4.5 and CART are compared using student's qualitative data. The domain values for some of the variables are defined for the present investigation as follows.

ParQua - Parent's Qualification are obtained. Here, The student's parent Qualification is specified whether they are educated or uneducated.

LivLoc - Living Location is obtained. Living Location is divided into two classes: Rural – for student's coming from rural areas, Urban – for student's coming from urban areas.

Eco – Economical background is obtained. The student's family income status is declared and it is divided into three classes. Low, Middle and High.

FRSup - Family and Relation Support is obtained (To find whether the student gets the moral support from family and relation for his studies). Family and Relation support is divided into three classes: Low – they did not get support from anyone, Middle – they only get support sometimes, High – they get full support from parent's and as well as from relation.

Res – Resource (Internet/Library access) are obtained (To check whether the students are able to access the internet and library). Resources are divided into three classes. Low – they have not accessed the internet and library, Middle - sometimes they have accessed the resource, High – they have accessed the both internet and library regularly.

Att – Attendance of student. Minimum 70% of attendance is compulsory to attend the semester examination; special cases are considered for any genuine reason. Attendance is divided into three classes: low – below 50%, Middle - > 79% and < 69%, High – >80% and <100%.

Result – Results are obtained and it is declared as response variables. It is divided into four classes: Fail – below 40%, Second - >60% and <69, Third - >59% and < 50% and First – above 70%.

Table 2. Student Qualitative Data and its Variables

Variable	Description	Possible Values
ParQua	Parent's Qualification	{Educated, Uneducated}
Livloc	Living Location	{Urban, Rural}
Eco	Economic Status	{High, Middle, Low}
FRSupp	Friends and Relative Support	{High, Middle, Low}
Res	Resource Accessibility	{High, Middle, Low}
Att	Attendance	{High, Middle, Low}
Result	Result	{First, Second, Third, Fail}

Table 2, describes the classification accuracy of ID3, C4.5 and CART algorithms when we applied on the collected student data sets using 10-fold cross validation is observed as follows,

Table 3. Decision Tree Classifier Accuracy

Decision Tree Algorithms	Correctly Classified Instances	Incorrectly Classified Instances	Unclassified Classified Instances
ID3	50%	47.5%	2.50%
C4.5	54.17%	45.83%	0%
CART	55.83%	44.17%	0%

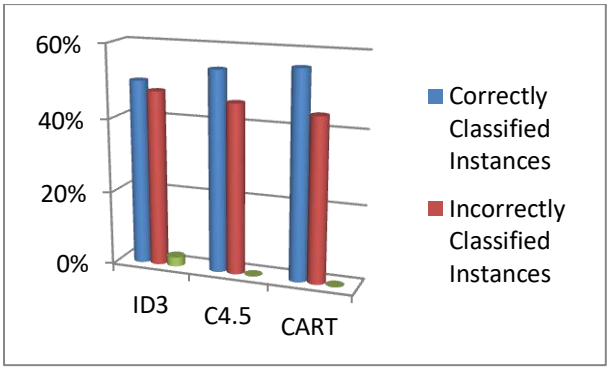


Figure 2. Comparisons of Classifiers with its classification instances

Table 4. Classifier Accuracy

Algorithm	Class	TP Rate	FP Rate
ID3	First	0.71	0.438
	Second	0.167	0.141
	Third	0.2	0.25
	Fail	0.333	0.278
C4.5	First	0.829	0.64
	Second	0.105	0.059
	Third	0	0.076
	Fail	0.313	0.087
CART	First	0.914	0.76
	Second	0.053	0.01
	Third	0.067	0.095
	Fail	0.063	0.038

Four classes are Fail –below 40%, Second - >60% and <69, Third - >59% and < 50% and First – above 70%. These three algorithms are compared based on the classification accuracy and TP rate. In this comparison the TP rate of CART is 0.914 and the class is 'FIRST' and it yields highest accuracy of 55.83% the other two decision tree algorithms. The TP rate of C4.5 algorithm is 0.829 and the class is 'FIRST' and it yields classification accuracy of 54.17% than the ID3 algorithm.

From the experimental results and the analysis of decision tree, the classification accuracy and class of accuracy of CART algorithm is considered to be best when it compared to other decision tree algorithms. This comparative analysis also identified the influence of qualitative parameters in student education performance.

7. CONCLUSION

Decision tree induction is a top-down recursive tree induction algorithm, which uses an attribute selection measure to select the attribute tested for each nonleaf node in the tree. This paper studied famous decision tree algorithms (ID3, C4.5 and CART) using different attribute selection measures. Each algorithm has got its own pros and cons as given in this paper. The efficiency of various decision tree algorithms can be analyzed based on their accuracy and time taken to derive the tree. The experimentation result shows that the CART has the best classification accuracy when compared to ID3 and C4.5. This experimentation significance also concludes that student's performance in examinations and other activities are affected by qualitative factors.

REFERENCES

[1] Anuja Priyam, Abhijeet, Rahul Gupta, Anju Rathee, and Saurabh Srivastava, "Comparative Analysis of Decision Tree Classification Algorithms", International Journal of Current Engineering and Technology, ISSN 2277-4106, Vol.3, No.2, June 2013.

[2] Badr HSSINA, Abdelkarim MERBOUHA, Hanane EZZIKOURI, Mohammed ERRITALI, "A comparative study of decision tree ID3 and C4.5", (IJACSA) International Journal of Advanced Computer Science and Application, Special Issue on Advances in Vehicular Ad Hoc Networking and Applications.

[3] Bhumika Gupta, Aditya Rawat, Akshay Jain, Arpit Arora and Naresh Dhami, "Analysis of Various Decision Tree Algorithms for Classification in Data Mining", International Journal of Computer Applications (0975-8887), Volume 163-No 8, April 2017.

[4] Dariusz Jankowski and Konrad Jackowski, "Evolutionary Algorithm for Decision Tree Induction", 13th IFIP International Conference on Computer Information Systems and Industrial Management (CISIM), Nov 2014, Ho Chi Min City, Vietnam. pp.23-32, 10.1007/978-3-662-45237-0_4.

[5] Himani Sharma and Sunil Kumar, "A Survey on Decision Tree Algorithms of Classification in Data Mining", International Journal of Science and Research (IJSR), April 2016.

[6] Jiawei Han, Micheline Kamber and Jian Pei, "Data Mining Concepts and Techniques", Third Edition.

[7] Kiran Singh and Raunak Sulekh, "The Comparison of Various Decision Tree Algorithms for Data Analysis", International Journal of Engineering And Computer Science ISSN: 2319-7242, Volume 6 Issue 6 June 2017, Page No.21557-21562.

[8] Mr. Brijain R Patel and Mr. Kushik K Rana, "A Survey on Decision Tree Algorithm for Classifition", 2014 IJEDR, Volume 2, Issue 1, ISSN:2321-9939.

[9] O..O. Adeyemo & T..O Adeyeye and D. Ogunbiyi, "Comparative Study of ID3/C4.5 Decision Tree and Mulilayer Percepton Algorithms for the Prediction of Typhoid Fever", African Journal of Computing & ICT, Vol 8 No. 1, March 2015.

[10] S. Venkata Krishna Kumar and P. Kiruthika, "An Overview of Classification Algorithm in Data Mining", International Journal of Advanced Research in Computer and Communication Engineering, Vol.4, Issue 12, December 2015.

Relaxed Web Development System with Advanced Encryption Standard (AES) Algorithm

Nay Nandar Linn
University of Computer Studies, Taungoo
snoopylinn@gmail.com, naynandarlinn@ucstaungoo.edu.mm

ABSTRACT: *The relaxed web development gadget is proposed as web-based totally application. Nowadays, global communications is very vital to make sure that messages are not modified by unauthorized individuals. On-line communications did not initially include aid for security. The contents of messages within the system ought to be stored mystery or the sender of the messages ought to be authenticated. For the Web security, Advanced Encryption Standard (AES) algorithm is used in login page. By incorporating access control features to the login forms, a site administrator can secure and restrict access with the Login Protection module. This paper must protect information from unauthorized changed. In this system, AES algorithm provides combining the best safety performance, flexibility, implementability and efficiency.*

Keywords: Cryptography, AES algorithm, Security, Secret key

1. INTRODUCTION

Nowadays, millions of users generate and exchange the information across a wide range of fields, such as financial and legal files, medical reports, and banking services via the internet. These examples of applications deserve special treatment from the point of view of protection, not only in the transmission of such data, but also in its processing. Protection threats take the forms of programs being eavesdropped, masqueraded, abused or refused. So, new technologies and strategies are increasingly more required for information security whilst sending statistics over the internet. The motivation of this paper is to defend sensitive facts with an excessive stage of security. This paper considers key established order protocols and related cryptographic techniques which gives shared secrets and techniques among two or extra events.

Due to electronic data security, the private and public sectors used different techniques and methods to protect sensitive data from intruders. Cryptography is one of the most important and common techniques used by two critical processes, encryption and decryption, to secure data from attackers. Encryption is the method of data encryption that prevents intruders from reading the original data easily. This stage can change the original data (Plaintext) into unreadable format known as Cipher text. Decryption is the next process to be performed by the authorized person.

Security of information systems could be used with many well-known security algorithms that can be modified with different settings for these algorithms. There are many considerations for security settings, including major factors such as the type of cipher that proves safety usability, the time consumption of the device, packet size, the data type used and the battery power consumption.

To achieve the confidentiality of information, all symmetric and asymmetric encryption techniques are used. A single key is used in symmetric encryption, and all people receiving the message must have that secret key to communicate safely. Asymmetric encryption consists of a pair of keys, a private key, and a public key. By using the digital signature, asymmetric encryption techniques can ensure authenticity and non-repudiation.

Cryptography is computer technology and security of communication. It is the science of secret codes, allowing communication confidentiality through an insecure channel. Through prohibiting unauthorized modification of use, it protects against unauthorized parties. Using a cryptographic system in general, it converts a plaintext into a ciphertext using a password[10].

Resource sharing is a major motivation to develop this program. Servers can handle resources and clients can access them. The Internet is addressed as an example of resource sharing and its key features are implemented Encryption and decryption is the most effective way to achieve data security, particularly in the resource sharing process.

Realizing the various types of security threats you may encounter is critical. The three main attack forms are: information disclosure, data theft, and service denial.Therefore secret keys were used in encryption in this system. By using these secret keys company sensitive data can carry out securely from online communication[13].

2. CRYPTOGRAPHY

Cryptography implemented in various ways and circumstances is important for protection and trust: privacy, authenticity, encryption and non-repudiation[3].

Privacy is the assurance the data is not made available or disclosed to unauthorized individuals, entities, or processes. Integrity is the guarantee that data by substitution, addition, or deletion is not intentionally changed in transit. Authentication is the guarantee that the source from which it appears to come. Non-repudiation is the assurance that this message is actually sent by a mechanism to prove that the sender[2].

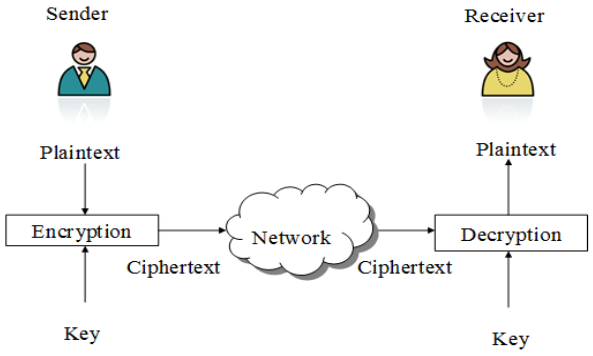


Figure1. Components of Cryptography

Figure 1 shows the components of cryptography. The initial intelligible message is called the plaintext and ciphertext is called the coded message or unintelligible message. On the plaintext, the encryption algorithm performs various replacements and transformations. The algorithm for decryption converts the ciphertext back into plaintext. The recipient uses an algorithm for encryption, and the receiver uses an algorithm for decryption.

2.1 Types of Cryptography

Cryptographic techniques are typically divided into two generic types:

1. Secret-key Cryptography
2. Public-key Cryptography

Symmetric encryption algorithm uses shared secret keys; copies of the secret key must be obtained by the sender and recipient and the key must be stable. Before the advent of public-key encryption, symmetric encryption was the only form of encryption used. Single-key, secret-key, one-key, private-key, or standard encryption is symmetric encryption[6].

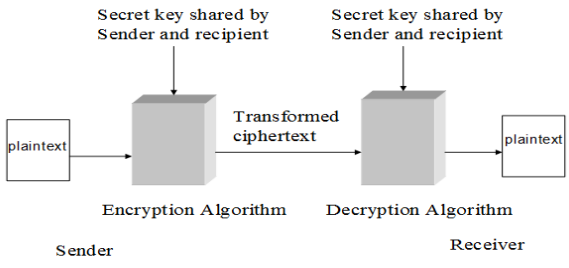


Figure2. Secret-key Cryptography

The asymmetric encryption algorithm uses pairs of public / private key, a message sender uses a public key to encrypt the message, and the recipient uses a private key to decrypt the text. Asymmetric encryption is called two keys or encryption with a public key.

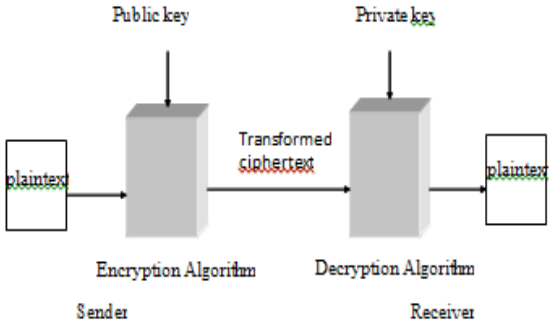


Figure3. Public-key Cryptography

2.2 Cryptographic goals

There are other common cryptographic problems that need to be solved and they can be just as important if not more important depending on who attacks you and what you are trying to secure against attackers. Confidentiality, integrity, authentication and non-repudiation are the cryptographic objectives protected in this text[14].

3. ADVANCED ENCRYPTION STANDARD (AES)

The Advanced Encryption Standard (AES) has been designed from the ground up to be fast, unbreakable and capable of supporting the smallest imaginable computing devices[4,5].AES is a substitution-permutation network, a collection of substitution and permutation mathematical operations, and its careful definition implies that each bit of output depends on each bit of input. The DES algorithm transforms 64-bits plaintext input encrypted output using a 56-bits secret key. Along with a lager key space, the AES had to be a 128-bits block cipher that is process 128-bits block of plaintext input at a time. AES also had to support 128-bit, 192-bit and 256-bit key setting and be more efficient than DES.

Advanced Encryption Standard (AES) is the block cipher encryption. A block cipher differs from a stream cipher as it encodes a symbol group in one step. Block cipher has either 64-bit or 128-bit block sizes, which means they process the plaintext in 64-bit or 128-bit blocks. For a given secret key, the changing from the original data to encrypted data is set. That is, mapping the same plaintext to the same ciphertext with the same secret key[10].

3.1 AES Data Structure

The AES block cipher receivesa 128-bit plaintext, and provides a 128-bit ciphertext under the control of a 128-, 192-, or 256-bit secret key. To map the plaintext to ciphertext, it is a Substitution-Permutation Network layout with a single set of steps called a round that are repeated 9, 11 or 13 times[4, 5].The standard round function of Advanced Encryption Standard (AES) contains four steps:

1. SubBytes
2. ShiftRows
3. MixColumns
4. AddRoundKey

Each round uses its own 128-bit *round key*, which is derived from the supplied secret key through a process known as a *key schedule*. This distributes the key entropy through each of the round keys. If that entropy is not adequately distributed, this causes trouble of all sorts such as identical keys, associated keys, and other similar distinctive attacks.AES treats the 128-bit input as a vector of 16 bytes organized in a column major 4x4 matrix called the *state*. That is, the first byte maps to $a_{0,0}$, the third byte to $a_{3,0}$,the fourth byte to $a_{0,1}$, and the 16th byte maps to $a_{3,3}$ in Figure 4.

$a_{0,0}$	$a_{0,1}$	$a_{0,2}$	$a_{0,3}$
$a_{1,0}$	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$
$a_{2,0}$	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$
$a_{3,0}$	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$

Figure4. The AES State Diagram

The entire forward AES cipher contains

1. AddRoundKey (round=0)
2. for round = 1 to Nr-1 (9, 11 or 13 depending on the key size) do
    1. SubBytes
    2. ShiftRows
    3. MixColumns
    4. AddRoundKey(round)
3. SubBytes
4. ShiftRows
5. AddRoundKey(Nr)

Figure 5 is the AES general structure of encryption and decryption.

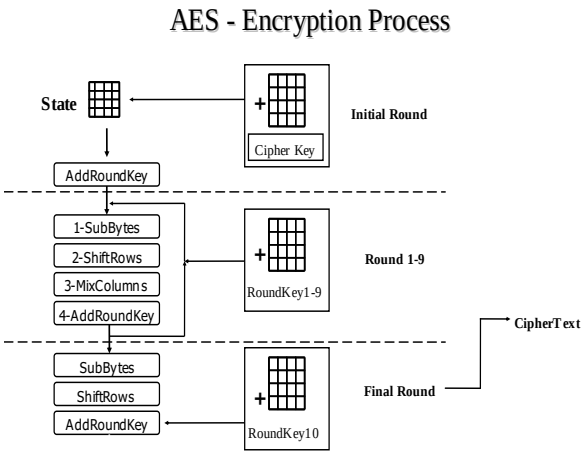


Figure 5. AES Algorithm Overview

3.2 Key Schedule

The key schedule is responsible for turning the input key into the Nr+1 required 128-bit round keys. The algorithm in Figure 6 will compute the round keys.

Input:

NkNumber of 32-bit words in the key (4, 6 or 8)

w Array of 4*(Nk+1) 32-bit words

Output:

w Array setup with key

1. Preload the secret key into the first Nk words of w in big endian fashion.
2. i = Nk
3. while (i < 4*(Nr+1)) do
    1. temp = w[i – 1]
    2. if (i mod Nk = 0)
        i. temp = SubWord(RotWord(temp)) XOR Rcon[i/Nk]
    3. else if (Nk> 6 and i mod Nk = 4)
i. temp = SubWord(temp)
4. w[i] = w[i-Nk] xor temp
5. i = i + 1

Figure6. The AES Key Schedule

The key schedule provides two additional functions. SubWord() takes the input of 32-bit and sends each byte in parallel through the substitution table of AES SubBytes. RotWord() cyclically rotates the word eight bits to the left. The Rcon table is an array of

the first 10 powers of the polynomial g(x)=x modulo that the AES polynomial only stored in the 32-bit word's most important byte[12].

4. IMPLEMENTATION OF SECURE WEB DEVELOPMENT SYSTEM

A relaxed web development machine is used wherein components placed at networked computer systems speak and coordinate their moves best by passing messages. This system is implemented for web in any local or international networked corporations. This device can be used to secure information that want to securely save and retrieve in. It's far crucial to defend the communication channels and the interfaces of any machine that handles data that might be the challenge of attacks e.g. Personal mail, electronic commerce and different financial transactions.

4.1 System Design

The detailed process is

- User A write projects in server computer.
- User A must protect information from unauthorized changed.
- Therefore secret keys were used in encryption in this process.
- User A sent secret key to User B and D that these users can modify this project.
- User C is unauthorized person who read this project but he cannot replace, insert or delete.

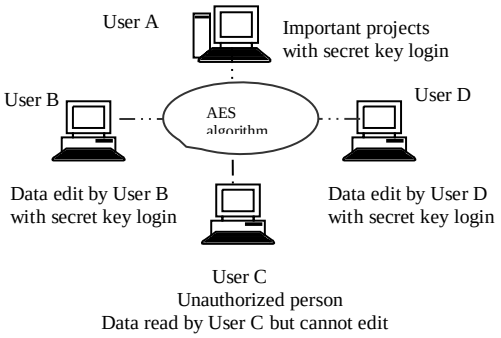


Figure7. Design of the System

The design of the system is shown in Figure 7. The system has two user types: administrator and user. First, both of them can be viewed projects on home page if they are interested in. Administrator has admin login to enter the edit project page. Advanced Encryption Standard algorithm used in this place. Administrator can modify every parts of the project. Admin shares secret key to authorized persons. So these persons can also be modified this project if necessary.

Everyone who read this task but he/she cannot substitute, insertion or deletion to it. If person is wanted to be modified this mission, person wishes his/her username and password to login the gadget. Then, this selects cipher key and decrypt it after which test it with administrator' password. If efficaciously login, user can reach edit undertaking page. For this reason this user can exchange the entirety. Un-authorized user has now not actual password and key who read this venture however he cannot regulate every elements of it. He time and again tried getting this password but he cannot get. Because of this login is carried out via the usage of advanced encryption standard (AES) set of rules.

4.2 System Flow Diagram

Figure8 describes the step flow for the administrator encryption process. Administrator write important project in server computer and modified every part of the project. Admin must protect information from unauthorized changed and creates admin login page. To log in to the system, Admin needs its username and password. Advanced Standard Encryption (AES) algorithm is used here.So that admin choose secret key for password encryption process. To get cipher text, the password will be encrypted with this key after the key has been got. Secure password has been got because of encryption. So, no one understands this password until decryption. Users cannot reach edit project page without secret key. If a user wanted to be modified this project, the user must know secret key. Admin shares secret key to authorized uses so that they can be modified this project if they are necessary to use.

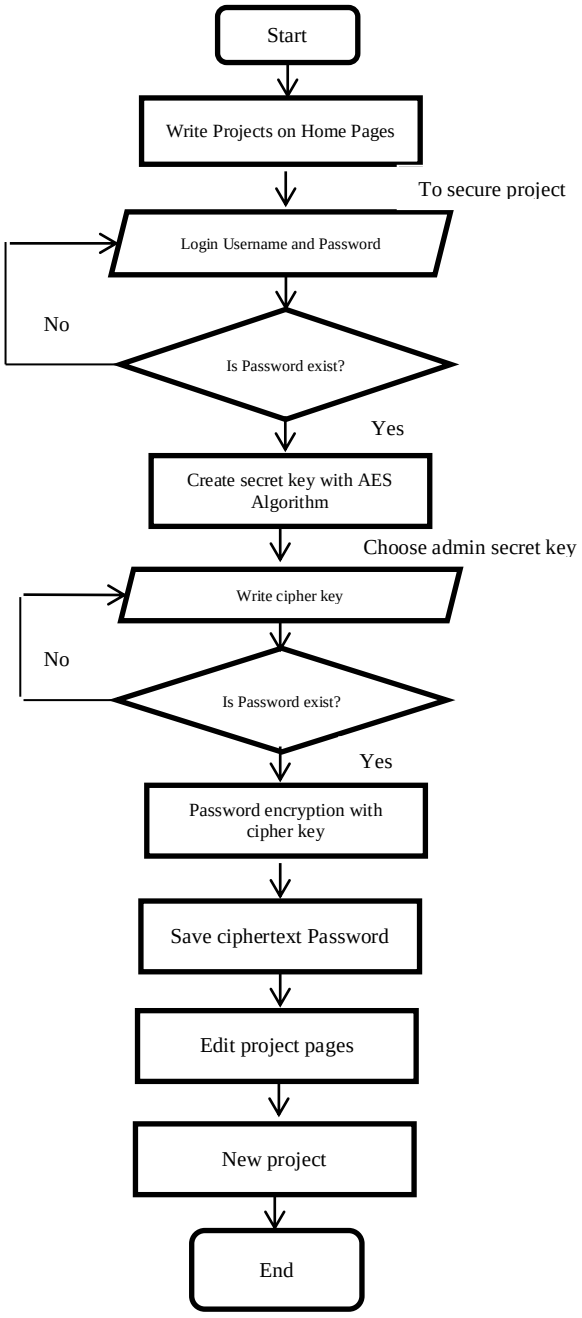


Figure8. Encryption for the System (administrator)

Figure 8 describes the flowchart for user encryption process. If a user is interested in, he/she can be viewed projects on home page. If a user is wanted to go edit project page, user pays his/her username and password to login the system. The administrator creates the secret key with Advanced Encryption Standard (AES) algorithm. Then administrator chooses the cipher key and encrypt it with password. The administrator save the ciphertext password. If the login is successful, the user can reach edit project page.

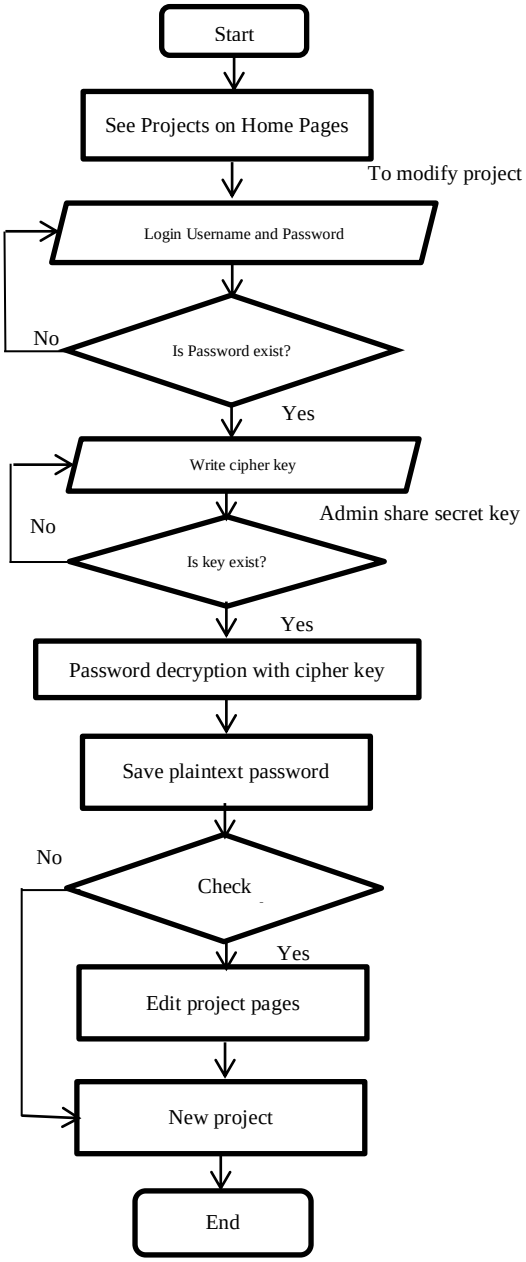


Figure9. Decryption for the System (user)

Figure 9 is the flowchart for user decryption process. User can be viewed projects on home page if interested in. If user is wanted to go edit project page, user pays his/her username and password to login the system. Then, choose cipher key and decrypt it and then check it with administrator password. If the login is successful, user can reach edit project page. Thus this user can change everything. Unauthorized user has not true password and key who read this project but he cannot modify every parts of it. He repeatedly tried getting this password however he cannot get. Because of this login is implemented by using Advanced Encryption Standard (AES) algorithm.

5. USABILITY EVALUATION FOR AES SECURITY

Rijndael is designed to withstand all known attacks, such as linear or differential attacks, so AES is certainly safer than DES and 3DES. Rijndael, however, is a very algebraic cipher and it wasn't long before the first algebraic attacks on AES were written. They use the fact that AES can be written as a multivariate quadratic equations system that is over-defined[7]. It would be an alternative to break AES to solve such a system in less time than to search the entire key space. But no one has yet been able to prove or disprove the effectiveness of this scheme of attack. Therefore AES will share the fate of all computationally secure ciphers, and time can give better than brute force logical attacks.

Therefore, this proposed system protects the confidentiality of digital data stored on computer system or transmitted via the internet or other computer network. This encryption mechanism plays a vital role in the security assurance of IT systems and communication. This system prevents third parties from recovering any of the original data or even any information about the data, from encrypted data.

5.1 Performance Testing

The result of the testing performance with this combined system have been documented and shown in Figure (10), Figure (11) and Figure (12). Figure (10) shows the encryption time of four different files. Figure (11) compares the performance result based on two different key lengths, 128-bit and 256-bit. In Figure (12), the results compare between an Intel Pentium 4 Celeron Processor, 3.66 GHz, 1 G RAM, running Microsoft Window XP SP (2) and an Intel Pentium 4 Celeron Processor, 2.67 GHz, 256 RAM, running Microsoft Window XP SP (2). By examining these Figures, we concluded that performance depends on three factors: file size, key length and type of processor. The more the file size is large, the more the encryption time takes. Similarly, when the key length bit increased, the encryption time also takes long. And the fast performance can gain by using high speed processor.

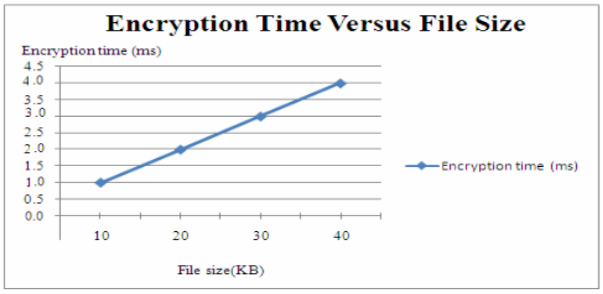


Figure10. Encryption Time Versus File Size

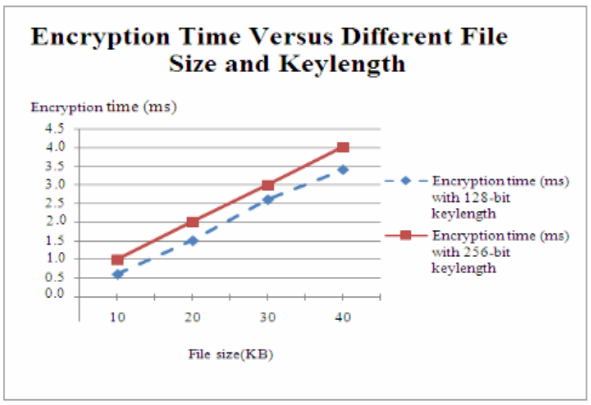


Figure11. Encryption Time Versus Different File Size and Key length

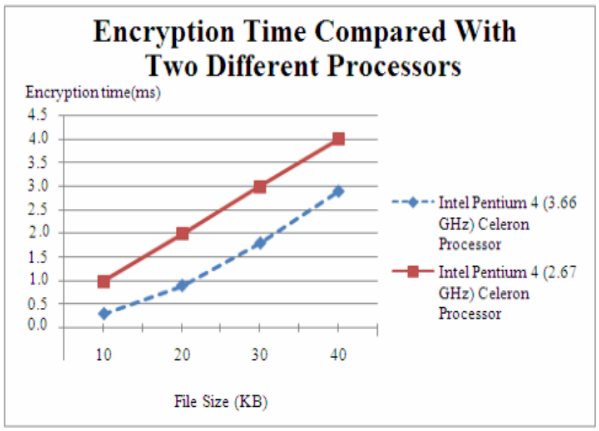


Figure12. Encryption Time Compared With Two Different Processors

It can also reduce the encryption time and can increase the whole system performance. By using this system, it does not need to keep the algorithm secret; just need to keep only secret key. Moreover, with the use of AES, it is relatively easy to implement and requires little memory. Therefore, this system provides not only authentication but also confidentiality.

6. CONCLUSION

Since the rise of the Internet, the message is transferred from one place to another in everywhere. In this case, information security has been one of the most important factors in information technology and communication. Cryptography is important for protection and trust in different ways and circumstances: privacy, authenticity, encryption, and non-repudiation. The login pages of the AES will serve functions above. This paper shows AES ' immediate impact on private organizations, businesses, state, and financial institutions ' practices. This system is important for securing any system's communication channels and interfaces that manage data that may be subject to attacks. There is a need to secure original information from falling into the wrong

hands in this era of ubiquitous digital communication, malware and hackers. By using this system, our project cannot change necessary. This system is provided secure and effective way to manage access to an organization's information.

REFERENCES

[1] A. Dandalis, V. K. Prasanna, D. P.Rolim, "A Comparative Study of Performance of AES Final CandidatesUsing FPGAs", the third AES conference, New York, April 2010.

[2] A. Lee, "Guideline for Implementing Cryptography", NIST Special Publication 800-21.

[3] B. Schneier, "Applied Cryptography: Protocols, Algorithms, and Source Code in C", John Wiley & Sons, 2016.

[4] J. Daemen and V. Rijmen, "AES Proposal:Rijndael, the Rijndael Block Cipher", 2009, pp. 1–45.

[5] J. Daemen and V. Rijmen, "AES Proposal:Rijndael, AES Algorithm Submission".

[6] J. Daemen, L.R. Knudsen, V. Rijmen, "Linear frameworks for block ciphers", to appear in Design, Codes and Cryptography.

[7] J. Nechvatal, et. al., "Report on the Development of the Advanced Encryption Standard (AES)", National Institute of Standards and Technology, October 2, 2016.

[8] M.Mylonas, D. J. Holding., K. J. Blow,"DES Developed In Handel-C", 2012.

[9] P. Bora, T. Czajka, "Implementation of the Serpent Algorithm Using Alter a FPGA Devices".

[10] S. Johnson, T. St Denis, "Cryptography for Developers", Elliptic Semiconductor Inc.,and Author of the LibTom Project, ISBN-1597491047.

[11] V. FISCHER, M. DRUTAROVSKY: "Two Methods of Rijndael Implementation in Reconfigurable Hardware", Proc. of 3rdInt. Workshop on Cryptographic Hardware and Embedded Sys-Teems, May 2011, pp. 77–92.

[12] "Advanced Encryption Standard (AES)", FIPS PUB 197, National Institute of Standards and Technology, U.S. Department of Commerce, November 2001.

[13] ISO/IEC 9797, "Information technology - security techniques - data integrity mechanism using a cryptographic check function employing a block cipher algorithm", International Organization for Standardization, Geneva, 2014 (second edition).

[14] AES page available via http://www.nist.gov/CryptoToolkit.

Study on Audio Watermarking Based on Arnold Transformation with DWT-DCT

Khin Myo Kyi
University of Computer Studies, Taungoo, Myanmar
khinmyokyi1984@gmail.com

ABSTRACT: *Digital information management system has played an increasingly important role in modern music production industry. Digital watermark is an efficient technology for protecting copyright information. In this study, Arnold transformation with DCT-DWT is applied in audio watermarking. In this system, firstly, the audio convert signal was divided evenly, the audio segments were selected from the divided audio, then by doing the DWT and DCT embedded the watermark to the low frequency coefficients of those audio segments which were selected. To evaluate the performance of the proposed audio watermarking, Signal to noise ratio (SNR) and Normalized correlation (NC) are calculated.*

Keywords: audio watermarking; Arnold transformation; DWT; DCT;

1. INTRODUCTION

Digital watermarking is one of the possible solutions of the multimedia data ownership problem. Embedding of watermark in audio signals is to be made in such a way that it does not degrade the audibility of the signal. Some of the popular audio watermarking method are least significant bit (LSB) coding [4], echo hiding scheme [1] and spread spectrum watermarking [2]. An adaptive audio watermarking based on support vector regression is proposed in [6].

The proposed watermark embedding scheme accomplishes perceptual transparency after watermark embedding by exploiting the masking effect of the human auditory system. Chen et al. in 2008 proposed a novel multipurpose audio watermarking scheme to make audio watermarking accomplish both copyright protection and content authentication with localization. In [3], which proposed that in order to provide authentication the watermark image is scrambled with Arnold transformation and then embedded into the Discrete cosine transform (DCT) original audio signal. Yassine in 2012 [7] proposed an innovative watermarking scheme for audio signal based on double insertion of the watermark in DWT-DST domain of the host signal by using a gray scale log image.

This paper aim to propose audio watermarking system based on Arnold transformation with DWT-DCT. Watermark is a binary image which was transformed using Arnold transformation. The rest of the paper is shown in below; section 2 describes the proposed framework, and some related methodologies describes in section 3.

2. PROPOSED SYSTEM

The audio convert signal was divided evenly, in audio watermarking embedding. From the divided audio, the selected audio file are extracted, moreover by doing the DWT and DCT embedded the watermark to the low frequency coefficients of those audio segments which were selected. A random private key is used during watermark embedding for more security. The following figure 1 describes the step by step procedure of the audio watermarking system.

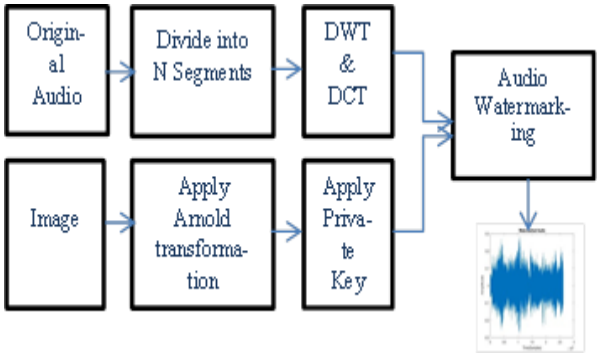


Figure1. Proposed framework

3. METHODOLOGIES

3.1. Arnold Scrambling Transform

Scrambling transformation as a means of encrypted technology is applied in the pretreatment stage of the watermarking, after scrambling transformation, one meaningful watermarking will become a meaningless [5]. The K x K binary watermark image W is transformed into W' by Arnold transformation to lower the autocorrelation coefficient of image and then confidentiality of watermark is strengthened. The Arnold transformation is given by

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} (mod\ N_1) \qquad (1)$$

By the above formula the corresponding inverse transform formula can be obtained:

$$\begin{bmatrix} a_1' \\ b_1' \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} (mod\ N_1) \qquad (2)$$

3.2. DWT Transform

The principle objective of the wavelet transform is to hierarchically decompose an input signal into a series of successively lower frequency approximation sub band and their associated detail sub bands. For the dyadic wavelet decomposition, at each level, the low frequency approximation sub band and detail sub band contain the information needed to reconstruct the low frequency approximation signal at the next higher resolution level. Wavelet techniques provide excellent space and frequency energy compaction, in which energy tends to cluster spatially in each sub band. For DWT [5u], the link between the spatial/temporal domain signals, f(t), and the DWT of f(t), d(k,l), is

$$f(t) = \sum_{k=-\alpha}^{\alpha} \sum_{l=-\alpha}^{\alpha} d(k,l) 2^{-\frac{k}{2}} \psi(2^{-k}t - l) \qquad (3)$$

Where $\psi(\bullet)$ denotes the mother wavelet.

3.3. DCT transform

The discrete cosine transform is a technique for converting a signal into elementary frequency components. The most common DCT definition of 1-D sequence of length N is

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) cos\left[\frac{\pi(2x+1)u}{2N}\right] \qquad (4)$$

For u=0, 1, 2,…, N-1. Similarly, the inverse transform is defined as,

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) cos\left[\frac{\pi(2x+1)u}{2N}\right] \qquad (5)$$

For x=0, 1, 2, …., N-1. In both equations (1) and (2) $\alpha(u)$ is defined as,

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} for\ u = 0 \\ \sqrt{\frac{2}{N}} for\ u \neq 0 \end{cases} \qquad (6)$$

It is clear from eq. that for u=0,

$$C(0) = \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} f(x) \qquad (7)$$

The first transform coefficient is, referred to as DC coefficient, the average value of the sample sequence. The other transform coefficients are called the AC coefficients.

4. RESULTS AND DISCUSSION

To verify the effectiveness of proposed algorithm, we carried out the simulation in Matlab2014a platform. In this study, different types of audio signal are used for cover audio such as (pop, rock, r&b, country, jazz). Various types of images are tested in this system for embedded signals. According to the analysis, we can see the implementation results are shown as below. In the experiment of embedded process, the parameter values: Arnold scrambling period is 24 and the numbers of transform times are 10. Convert the transformed watermark to a vector which are divided into audio segments. Figure 2 represents the results of tested host signal and embedded image.

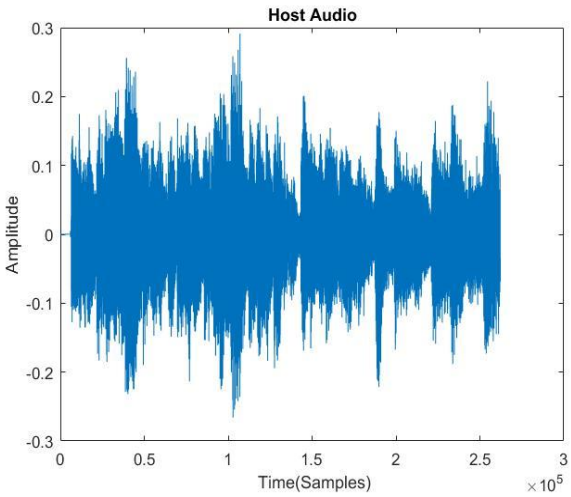


Figure 2 (a). Host audio signal

Figure 2 (b). Original embedded image

Figure 3 shows that image are extracted without any attacks by this algorithm are very similar to the original image. Figure 3(b) shows the waveform comparison of the carrier audio and watermarked audio (only show a short fragment lasting about three seconds) without performing any attack. In this figure, it can be seen that two waveform figures have no obvious changes before and after the watermark was embedded into carrier audio.

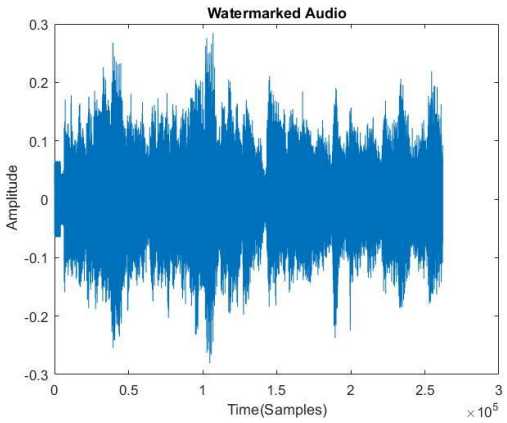


Figure 3 (a). watermarked Audio

Figure 3 (b). Extracted watermark image

4.1. Performance Analysis

In this section the experiment results for listening evaluations and robustness against signal processing attacks separately for Method 1 and Method 2 have been tested. Signal to noise ratio (SNR) and Normalized correlation (NC) are the parameters used for performance analysis. “Sample” referred to the different type of audio signal which are embedded with various images.

In this testing, the value of alpha is taken as 2.5*10-2 that gives a decent SNR for various thresholds chosen in determining the embedding regions. The original and the watermarked signal are shown in previous section. Signal to noise ratio (SNR) is calculated for each value of threshold and is shown in table1. The analysis of absolute difference of amplitude of host and watermarked audio signal are shown in figure 4.

Table 1. Variation of SNR for different values of thresholds

Threshold	SNR (db)				
	Sample 1	Sample 2	Sample 3	Sample 4	Sample5
0.95	67.50	56.2	61.9	58.4	56.4
0.90	66.25	55.1	60.2	56.3	55.1
0.85	61.86	53.2	57.4	54.3	52.2
0.80	60.81	50.0	56.1	52.9	49.0
0.75	58.47	46.7	54.2	51.7	46.1
0.70	56.71	45.3	52.1	48.1	45.2

The structural similarity index (SSIM) is a perceptual metric that quantifies image quality degradation caused by processing such as data compression or by losses in data transmission. It is one of performance analysis for audio watermarking with image embedded signal. It is full reference metric that requires two images from the same image. The measured compression ratio in bits/pixel for the processed image is shown in figure 4.

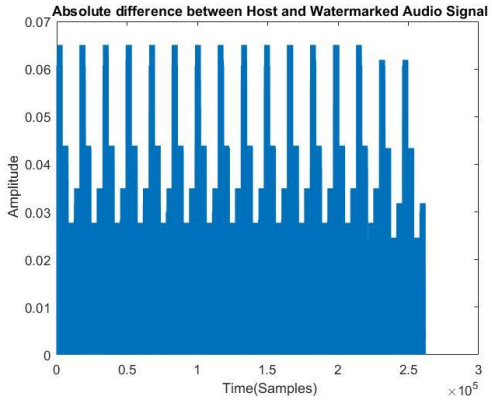


Figure 4(a). Absolute difference between host and watermarked audio signal

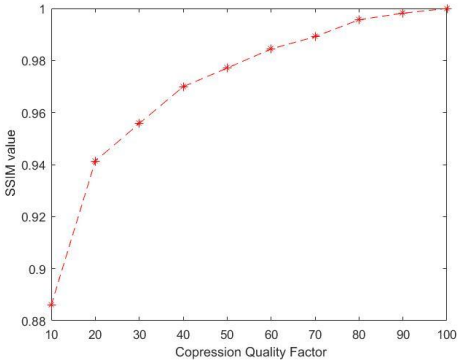


Figure 4 (b). Structural similarity index

5. CONCLUSIONS

Audio watermarking is an active research area that has been driven by the need to solve the copyright protection problem of digital audio products. In this scheme was scrambling the watermark with Arnold transformation and embed quantitatively it into the audio signal. The experimental results have illustrated the robust nature our watermarking scheme. The result revealed that simulation results demonstrated the audibility and robustness of the proposed audio watermarking algorithm.

REFERENCES

[1] Bender, Gruhl, Morimoto and Lu, "Techniques for Data Hiding" IBM Systems Journal 35(3), 1996, pp.- 313–336.

[2] Kirovski D and Malvar H, "Spread-spectrum Watermarking of Audio Signals" IEEE Transactions on Signal Processing 51(4), 2003, pp. 1020–1033.

[3] Lalitha, N.V. , Suresh, G. and Talagarupa, P. (2012), "Audio authentication using Arnold and Discrete cosine transform", International conference on computing Electronic and electrical technologies [ICCEET].

[4] Miroslav Goljan, Jessica J. Fridrich and Rui Du, "Distortion-free Data Embedding for Images" LNCS 2137, 2001 pp. 27–41.

[5] Sharma, S., Rajpurohit, J. and Dhankar, S. (2012), "Survey on different level of audio watermarking techniques", International journal of computer application (0975-8887), Vol.49, No. 10.

[6] Xiangyang Wang, Wei Qi, and Panpan Niu, "A New Adaptive Audio Watermarking based on Support Vector Regression" IEEE Transactions on Audio, Speech, and Language Processing, 15 (8), 2007, pp. 2270-2277.

[7] Yassine, H. , Bachir, B. and Aziz, K. (2012), "A secure and high robust audio watermarking system for copyright protection", International Journal of Computer application (0975-8887) Vol. 53, No. 17

Tuning of PID Controller using Ziegler-Nichols Method for Room Temperature Control System

Myat Thu Zar [1], Nang Mwe Seng [2], Su Mon Thwin [3]
[1, 2, 3] *Department of Mechatronic Engineering, Pyay Technological University*
mthuzar360@gmail.com, nangmweseng@gmail.com,
sumonthwinsulay@gmail.com

ABSTRACT: *In modern lifestyle the energy demand for room heating is growing rapidly around the world and tendency for indoor comfort. These energy leads to design of control strategies to make the system energy efficient. The heating and ventilation is a technology that provides thermal comfort and air quality. Heating system is the most important part of this technology and several methods have been proposed for controlling it. In this research, an efficient room heating system has been designed using controller. The design has a temperature controller with the purpose to sense the temperature of a room and maintain it to a fixed desired value. The proposed design uses PID controller as a control method that maintains the temperature of simulated heater to the desired point. PID gain parameters are tuned with Ziegler-Nichols method. All design and simulations have been done in Simulink of Matlab platform. The simulation result shows the rise time is* 22 sec, settling time is 144 sec and percent overshoot of 3.86% with PID controller. *It has been concluded that the proposed control strategy is very advantages due to small overshoot, fast response time*.

Keywords: Temperature control; PID; Ziegler-Nichols tuning; Thermal modelling; Matlab/Simulink

1. INTRODUCTION

Maintaining a desired room temperature has increasing problem for home and offices as the weather is taking an unpredictable. Several studies have shown that comfort, health and productivity in offices improved by individual thermal control [1]. In order to solve this problem, stronger controller is needed to maintain the desired temperature range in home and officies. Room temperature control is a procedure where variation of temperature is measured, and by manipulating the flow of heat energy, desired room temperature is maintained [2]. Temperature control system is based on closed loop feedback control system. Mostly; home use thermostat that constantly asses the current room temperature and controls a heater or air conditioner. Temperature is an essential physical quantity which is found in most application of home appliances, laboratory, and industrial processes. Temperature becomes a significant leading parameter in thermodynamic and the study of heat dynamic of a room is very important in this paper.

Temperature control system is the most important part of technology and several methods have been proposed for controlling. PID controller is a conventional controller applied in most industrial applications, although the development of modern control system which used of algorithms theoretical or technology in control systems. PID controller has simple structure, robust performance and becomes a fundamental controller for today modern control algorithms [3]. It has a simple control rules and less parameter adjustments and obtaining the suitable PID parameters.

It includes tuning of PID controllers using Ziegler-Nichols method, which control the air flow rate entering the room. Ziegler-Nichols presents a tuning formula, based on time response and experiences. Although it lacks selection of parameters and has an excessive overshoot in time response, it still opens the way of tuning parameters. Second method of Ziegler-Nichols tuning based on PID tuning formula for set point regulation accommodate the response speed and overshoot [4]. PID controller for air flow rate control is developed using the second method of Ziegler-Nichols.

Matlab/Simulink is an interactive tool for modelling, simulating, and analyzing dynamic system [5]. It has been assisting engineers to design and test system models for different fields of engineering and enables engineers to build graphical block diagrams, evaluate system performance, and refine their design.

This paper presents PID parameters and software modules developed in Matlab/Simulink for tuning PID controllers. This paper is structured as follow: section II is proposed model of room heating temperature control system. Section III describes the methodology and in section IV, the simulation and result are presented for PID control. Section V is the conclusion of this paper.

2. PROPOSED MODEL

In simulation of a controller using Simulink MATLAB, the first thing to make is a mathematical model of the room. In a room heating system, there is three main components, specifically are a heater, temperature sensor and room. The relation between all three of them can be seen in Figure1.

Figure1 shows workflow of the heater when heat up the room temperature. Air from the inside of room being absorbed into pipeline. After that, the air will flow through the heater. The air that has been heated, will flow back to the room.

There are serval things that are important to know in creating mathematical model for room heating system. They are a room thermal characteristic, a heater thermal characteristic, temperature control and inside/ outside temperature of the room. From five points that affect a room heating system, there are four variables which corresponds.

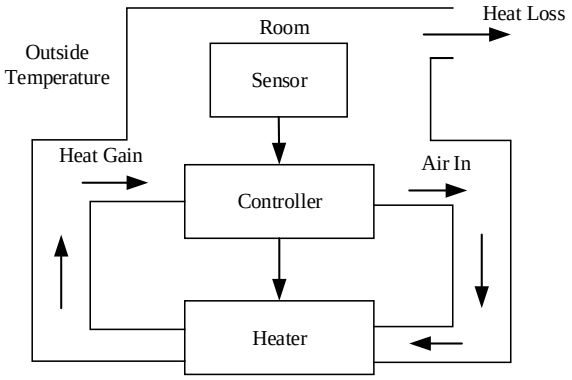


Figure 1. Proposed model of room heating temperature control system

They are thermal energy which comes from heater to the room (Q_{gain}), thermal energy which comes from room to the surroundings (Q_{loss}), room temperature (T_{room}), and temperature outside of the room (T_{out}). And then T_{heater} indicates the air temperature of heater. Then the increase of thermal energy that flow by convection from the heater to the room (Q_{gain}) can be given as

$$Q_{gain} = m_{heaterair} c_{air} (T_{heater} - T_{room}) \quad (1)$$

Where, $m_{heaterair}$ is the mass of air entering in the room and c_{air} is specific heat capacity of air. The main actuating block of the control system includes a fan to pass the heated air inside the room. The rate which the room gains thermal energy from the heater is

$$\frac{dQ_{gain}}{dt} = \frac{dm_{heaterair}}{dt} c_{air} (T_{heater} - T_{room}) \quad (2)$$

Controller is designed to actuate the air flow rate that in turn helps manipulating the room temperature. Room thermal energy which has obtained from the heater will reduce as big as (Q_{loss}). This is because heat transfer is by conduction through the room wall or window. The rate of thermal energy loss from the room is given as:

$$\frac{dQ_{loss}}{dt} = \frac{T_{room} - T_{out}}{R} \quad (3)$$

Where, R indicates equivalent thermal resistance of room and T_{out} is ambient temperature. To measure the rapidity of room temperature change, a heater to room thermal energy that loss from inside the room to surroundings. The rate at which the room temperature changes is defined by equation given below:

$$\frac{dT_{room}}{dt} = \frac{1}{m_{roomair} c_{air}} \left(\frac{dQ_{gain}}{dt} - \frac{dQ_{loss}}{dt} \right) \quad (4)$$

By integrating the rate, actual temperature can be measured. Difference between this actual temperature and desired temperature is the error signal. Controller takes this signal as it's input. The country normal minimum temperature for the period 1981-2010 over Myanmar is 19.7˚C [6]. So, room temperature (T_{room}) is choose as 20˚C and outdoor temperature (T_{out}) is choose a range from 10˚C to 20˚C.

Table 1. Parameters value for thermal model of room

Parameters	**Values**
Equivalent thermal resistance (R)	3.86 x 10^{-7} Hour Degree/Joule
Specific heat capacity (c_{air})	1000Joule/Kilogram Degree
Air temperature from the heater (T_{heater})	70˚C
Room temperature (T_{room})	20˚C
Outdoor temperature (T_{out})	10˚C to 20˚C

3. METHODOLOGY

A PID controller refers to the first letters of the individual terms that make up the standard three term controller. These are P for the proportional term, I for the integral term and D for derivate term.

3.1. PID Controller Parameters

The "three-term" functionalities are heighted by the following.

- The proportional term reduces the rise time and will reduce, but never eliminate, the steady state error. Its output is proportional to input error signal which provides an overall control action proportional to the error signal. While the steady-state errors will highlight if there is only proportional term in the control system.
- The integral term eliminates steady state errors through low frequency compensation by an integrator but it may transient response.
- The derivative term increases the stability of the system, reduces the overshoot, and improves the transient response. Its output proportional to the derivative of input error signal and improves transient response through high frequency compensation by a differentiator.
- The individual effects of these three terms on closed loop performance are summarized in Table 2 [4].

Table 2. Effects of independent P, I, and D Tuning

Closed Loop Response	Rise Time	Overshoot	Settling Time	Steady State Error
Increasing K_p	Decrease	Increase	Small Increase	Decrease
Increasing K_i	Small Decrease	Increase	Increase	Large Decrease
Increasing K_d	Small Decrease	Decrease	Decrease	Minor Change

3.2. PID Control

A proportional integral derivative controller (PID) is widely used in industrial control systems. They provide control signals that are proportional to the error between the reference signal and the actual output (proportional action), to the integral of the error (integral action), and to the derivative of the error (derivative action), namely

$$U(t) = K_p\left[e(t) + \frac{1}{T_i}\int_0^t e(t)dt + T_d\frac{de(t)}{dt}\right] \quad (5)$$

Where u(t) and e(t) denote the control and the error signals respectively, and K_p, T_i and T_d are parameters to be tuned. The corresponding transfer function is given as;

$$K(s) = K_p\left[1 + \frac{1}{T_i(s)} + T_d(s)\right] \quad (6)$$

The structure of PID controller is simple, it is the most extensive control method to be used in industry. The PID controller is mainly to adjust an appropriate proportional gain (K_p), integral to a gain (K_i), and differential gain (K_d) to achieve the optimal control performance. The PID controller system block diagram of this paper is shown in Figure 2.

Transfer function can also be expressed as;

$$K(s) = \frac{U(s)}{E(s)} = K_p + \frac{K_i}{s} + K_d s \quad (7)$$

The main features of PID controllers are the capacity to eliminate steady state error of the response to a reference signal (because of integral action) and the ability to output change (when derivative action employed).

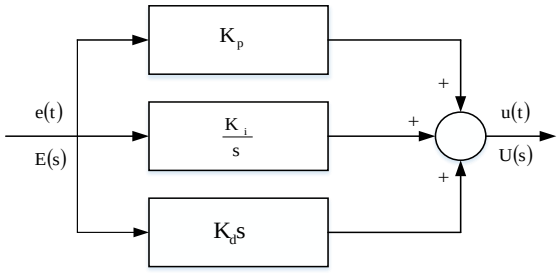


Figure 2. PID controller block diagram

3.3. Ziegler-Nichols Tuning Method (Z-N)

The Ziegler-Nichols method is a simple and clear PID controller tuning method. It is implemented by setting the I and D gains to zero and "P" gain is increased until the output of the control loop oscillates with constant amplitude. Then the oscillation periods are used to set P, I, and D gains. The set taken to simulate PID parameters tuning with Ziegler-Nichols method are as following:

- Set the controller to P-only mode, that is say assign the integral gain K_i=0; the derivative gain K_d=0 and the controller gain K_p=1;
- Set the controller gain, K_p up slowly and observe the output response. When K_p results in a sustained periodic oscillation in the output, mark this critical value of K_p as K_u, the ultimate gain. Also, measure the period of oscillation P_u, referred to as the ultimate period.
- Using K_u and P_u, Ziegler-Nichols tuning method prescribes the following values for K_p, T_i, and T_d as in Table 3 [5].

Table 3. Ziegler-Nichols tuning method

Type of controller	K_p	T_i	T_d
P	$0.5K_u$	∞	0
PI	$0.45K_u$	$1/1.2P_u$	0
PID	$0.6K_u$	$0.5P_u$	$0.125P_u$

The Ziegler-Nichols tuning rule gives PID loops best disturbance rejection performance.

4. SIMULATION AND RESULT

A room dimension 6m x 9m x 4m and the supplied electrical heater power to the heat zone is 2kW. A test model as considered is taken for study of room temperature control PID with Z-N tuning algorithm. "Ziegler-Nichols method" has been used for the purpose tuning. As in the above method, the K_i and K_d gain are first set to zero. The K_p gain is increased from 0 to ultimate gain K_u at the output of the starts to oscillate K_u and the oscillation period P_u are used to set gains.

From Figure 3, the ultimate gain "K_u" is choosed as 2 and the ultimate period "P_u" is choosed as 0.015. The tuning method gives the values as "K_p" = 1.2, "T_i" = 0.0075, and "T_d" = 0.0019. Figure 4 shows the block diagram of room temperature control system.

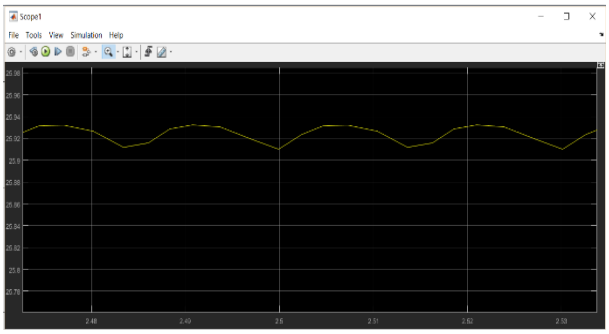


Figure 3. The oscillation period P_u at the ultimate gain K_u=2

Mathematical model of the whole system is constructed with the environment MATLAB/Simulink. Figure 5 (a) shows the block diagram of room temperature control system. The model of the room temperature control in Simulink is shown in Figure 5 (b). Here, room temperature is measured and compared with the desired value. The difference is considered to be error signal. Controller is to make this error signal to zero. Outside room temperature is varying in between 10˚Cto20˚C.

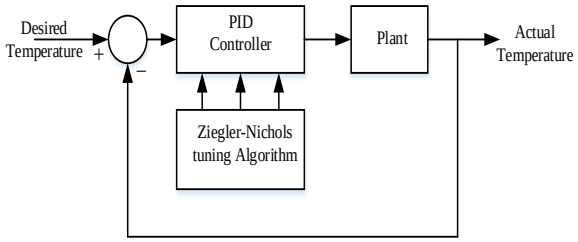


Figure 4. Block diagram of room temperature control system

In figure 6, it can be found that traditional on/off control temperature regulation are fluctuated around the desired temperature. Figure 7 shows output signal of the Ziegler-Nichols tuning with PID controller. The room temperature starts increasing from the initial room temperature of 20˚C and gradually settles to the desired value 26˚C. From the simulation results, the value of rise time is 22 sec, settling time is 144 sec and percent overshoot of 3.86 % with PID controller. According the result, the steady state error is zero and the overshoot is within the limit of applicable.

Table 4. Parameters values

Parameters	PID Controller
Rise Time	22 sec
Settling Time	144 sec
Percent Overshoot	3.86 %

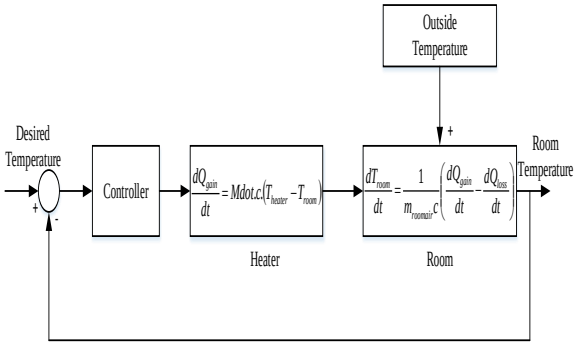


Figure 5(a). Block diagram of room temperature control system

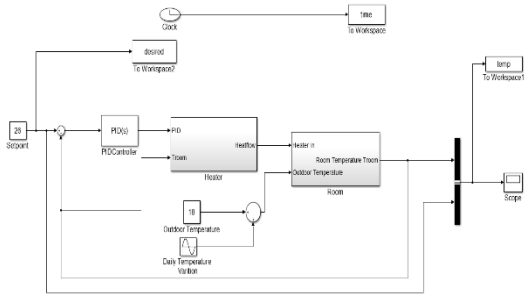


Figure 5(b). Simulink block diagram of room temperature control system

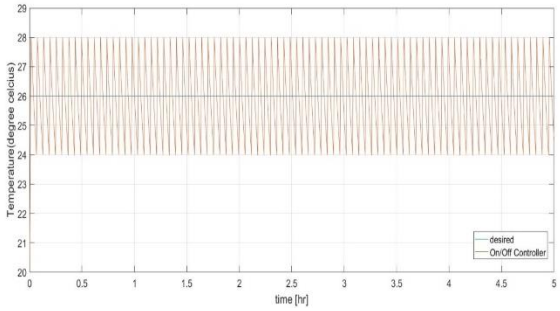


Figure 6. Response of On/Off control with set point temperature 26˚C

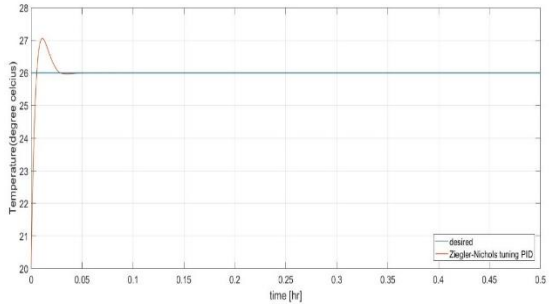


Figure 7. Response of PID Control with set point temperature 26˚C

5. CONCLUSIONS

In this paper, a thermal modeling of room and air heater is carried out by energy conversion equation. The model of room temperature is created in the SIMULINK environment of MATLAB. The performance of on/ off controller and PID controller are shown in figure 6 and 7 for the transient response of room temperature control. For on/off controller, the output temperature is oscillated with high frequency as shown in figure 6. In PID control, the system transient response observed that the set point is reached faster as well as to the steady state position. Moreover, the rise time and the settling time are short and makes the response to a bit oscillation around the desired input. It can be concluded that the Ziegler-Nichols tuning method is the best for gain tuning of PID controller which presents satisfactory performances for minimum rise time, minimum settling time, and minimum overshoot.

ACKNOLEDGMENTS

The author wishes to express her gratitude to Dr. Nyaunt Soe, Rector of Pyay Technological University, for his encouragement. The author is deeply grateful to Dr. Khin Thu Zar Win, Professor and Head, Department of Mechatronic Engineering, Pyay Technological University, for her willingness to share her ideas and helpful suggestions. The author greatly expresses her do not number thanks to all persons will contribute in different ways to support in preparing this paper.

REFERENCES

[1] E.G.Ochieng, N.Jones, " Integration of Energy Efficient Technologies in UK Supermarkets," April 2014, Vol. 67, pp. 388-393

[2] D.W.U.Perea, CF.Pfeiffer and N.O Skeie "Modelling the Heat Dynamic of a Residential Building Unit : Application to Norwegian Building," Modelling Idenfication and Control, ISSN 1890-1328, 2014,Vol.35,No.1, pp-43-57.

[3] J.C.Basilio and S.R Matos, "Design of PI and PID Controllers with Transient Performance Specification," IEEE Trans. Education, 2002, Vol.45, Issue No-4,pp.364.

[4] Kiam Heong Ang, Gregory Chong, and Yun Li. "PID System Analysis," Design and Technology, IEEE Transaction on control system technology, 2005, pp. 13-17

[5] Suping Li, Quansheng Jiang. "Study on PID Parameters Tuning Method Based on MATLAB/SIMULINK," IEEE, 2011, pp. 408-411.

[6] Dr.Lai Lai Aung, Mrs.Ei Ei Zin, Ms Pwint Theingi, Ms Naw Elvera, Ms Phyu Phyu Aung, Ms Thu Thu Han, Ms Yamin Oo, and Dr Reidun Gangstop Skaland, "Myanmar Climate Report", No.9, 2017.ISSN 2387-4201, ME Treport Norwegin IV.

Certificate Index	*Author*
VOL01IS01001	*A NweSoe, Paing Thwe Soe, Theint Theint* *"Impact of Written Exam on Students' Project Performance"*
VOL01IS01002	*Aung San Min, San San Mon, Hla Myat Thandar, Min Min Aye* *"Microcontroller Based Automatic Water Level Controller"*
VOL01IS01003	*Aye Aye Htun* လွွဲပါဘူးသမုဆွေ တေးထပ်ကဗျာလေ့လာချက်
VOL01IS01004	*Aye Aye Khaing, Yu Ya Win* *"Local University-Industry Linkage and Job Perception"*
VOL01IS01005	*Aye Aye Lwin,Thin Thin Soe* *"Assessment of Toe River Water Quality around Maubin Township from Ayeyarwady Region in Myanmar"*
VOL01IS01006	*Aye Myint Sein, Myint Myint Khine, Aye Thida* *"Isolation of Lupeol Compound and Pharmacological Actions of Crataeva nurvala Ham. Bark (Kadet)"*
VOL01IS01007	*Hnin Wuit Yee, Hla Ngwe* *"Physicochemical Characterization of Water Soluble Chitosan and Its Acute Toxicity and Weight Loss Activity"*
VOL01IS01008	*Htay Htay Khaing* *"Preparation of Biodiesel From Waste Fried Palm Oil Using Zeolite Catalyst"*
VOL01IS01009	*Khaing Khin Aye, Thi Thi Swe* မြန်မာ့လုပ်ငန်းသုံးဆိုင်ရာ စကားပုံများ လေ့လာချက်
VOL01IS01010	*Khin Cho Latt, Chaw Ei Su* *"A Study On Some Noticeable Factors In Communicative English for Non-Native Speaker"*
VOL01IS01011	*Khin Cho Thant, Tun Tun Win, Thet Hnin Lwin* *"Effect of Gamma Radiation on Agronomic Characteristics of Maize (Zea Mays L.)"*
VOL01IS01012	*Khin Ma Lay* မဏိကုဏ္ဍလဝတ္ထုမှ ထူးခြားသောစာဟန်
VOL01IS01013	*Khin Moh Moh Thin, Hla Yin Moe, Linn Linn Aye* *" Calculate the Electrical Network by Using Matrices "*
VOL01IS01014	*Khin Myo Kyi* *"Study on Audio Watermarking Based on Arnold Transformation with DWT-DCT"*
VOL01IS01015	*Khin San Win, Ni Ni Aung, Thu Zar Myint, Hlaing Hlaing Myat)* *"Comparative Study on Nutritional Values, Antimicrobial Activities and Antioxidant Activities of Papaya Leaves and Tea Leaves"*
VOL01IS01016	*Khin San Win, Swe Swe Lwin* *"Determination of Antimicrobial and Antioxidant Activities of the Fruits of Momordica charantia L. (Bitter Gourd)"*
VOL01IS01017	*Kyi Kyi Hla* ထားဝယ်ဒေသ(ကျောက်ဆည်)ဒေသိယစကားလေ့လာချက်
VOL01IS01018	*Kyi Kyi Thwin, Nyi Nyi Htway* *"Innovative Strategies to Develop Speaking Performance: A Study of Teachers and Learners at UCS (Taungoo)"*
VOL01IS01019	*Kyi Pyar Moe, Myo Su Su Hlaing* *" Solution and Application of Heat Equation Using Method of Separation of Variables"*

VOL01IS01020 *Kyi Zar Nyunt, Wint Aye Khaing, San San Yu*
"A Survey of Comparison in Decision Tree Algorithms for Classification"

VOL01IS01021 *Lae Lae Khine, Nwe Ni Soe*
"Influence of Dopant Calcium Concentration on Electrical Behavior of $PbTiO_3$ Ceramic Capacitor"

VOL01IS01022 *Lae Lae Oo*
"An Integrated Approach to Teaching Verb (Tenses) for All Students"

VOL01IS01023 *Lwin Ko Oo, Nan Thidar Chit Swe, Than Zaw Oo, Ye Chan*
"Influence of Reducing Agents on the Formation of Reduced Graphene Oxide"

VOL01IS01024 *May Myo Swe,Yi Yi Maw, Su Hlaing Win*
ဦးသုခ၏ အကျော်ဒေးယျဗုဒ္ဓသမီးတော်များ ဝတ္ထုတိုများမှ သရုပ်ဖော်အဖွဲ့များ

VOL01IS01025 *Moe Moe Yee, Kyaw Naing*
"Mineral Elements, Essential Nutritive Elements and Toxic Elements in Some Myanmar Indigenous Medicines Containing Kyauk-thway"

VOL01IS01026 *Moe Sandar Naing, Ju Ju Khin, Kyi Kyi Wai*
"Assessment of Soil Quality in Ywar Thar Gyi Village, Kyaiklat Township, Ayeyarwady Region"

VOL01IS01027 *Mya Mya Sainn*
"Isolation and Characterization of Essential Oils from Houttuyania cordata Thunb.(Fish Mint)"

VOL01IS01028 *Myat Thu Zar, Nang Mwe Seng, Su Mon Thwin*
"Tuning of PID Controller using Ziegler-Nichols Method for Room Temperature Control System"

VOL01IS01029 *Myint Myint Swe, San San Htwe, Sandar Win*
"Design And Construction of Mosfet Inverter"

VOL01IS01030 *Myint Myint Than*
"Audio Watermarking Based on MP3 Encoding Process"

VOL01IS01031 *Nay Nandar Linn*
"Relaxed Web Development System with Advanced Encryption Standard (AES) Algorithm"

VOL01IS01032 *Nilar Tint*
မင်းသုဝဏ်၏ သမိုင်းနောက်ခံ ဇာတ်လမ်းကဗျာ လေ့လာချက်

VOL01IS01033 *Nwe Nwe Aung, Khin Phyo Win*
"Determination of Dyeing Process and Colour Fastness Properties of Natural Dye Extracted from Mango Bark (Mangifera indica L.) on Cotton Cloth"

VOL01IS01034 *Ohnmar Khine, Myo Kay Thwe*
"Appropriateness of Eclectic Approach to English Language Teaching"

VOL01IS01035 *Pa Pa San, Tin Myo Latt, Thandar Khaing, Nyein Nyein Aye, Yi Yi Lwin*
"Assessment of Soil Quality by Supporting With Pe-Bizat, Peanut Husk, Cowdung from TaungThaman Village, Amarapura Township, Mandalay Region"

VOL01IS01036 *Paing Thwe Soe, Nilar Soe, Theint Theint*
"Impact of Teaching Method on Academic Performance: An Experimental Research Study"

VOL01IS01037 *Paing Thwe Soe, Nay Nandar Linn, Thin Thin Htwe*
"The Role of Capacity Assessment in Research Capacity Building Strategy: An Experimental Study"

VOL01IS01038 *Phyo Yatanar Lin, Nilar Thein*
" Teachers' Performance Appraisal by Using Fuzzy Expert System"

VOL01IS01039 *Pyae Sone Aung, Zaw Naing*
"Removal of Organic Dye (Methylene Blue) by Activated Bentonite Clay"

VOL01IS01040 *San San Htwe*
"Numerical Solution to the Laplace Equation for a Magnetostatic Potential"

VOL01IS01041 *San San Yu, Kyi Zar Nyunt*
"Quantitative Association Rule Based on Exam Report System"

VOL01IS01042 *Soe Kyaw Kyaw, Yu Yu Hlaing*
"Evaluation of Quality Control Parameters of Four Different Brands of Paracetamol Tablets"

VOL01IS01043 *Thanda Aung, Phyo Phyo Wai, Wint War Htet Naing, Theint Theint Phyo, Than Than Myint, Ye Myint Aung, Ni Ni Than*
"Phytochemical Investigation and Antioxidant Activities of Three Selected Medicinal Plants from Ayeyarwady Region, Myanmar"

VOL01IS01044 *Thida Tin, Yee Lai Win, Myat Myat Thu*
"Synthesis and Characterization of Silver Nanoparticles Using Three Different Leaves"

VOL01IS01045 *Thinzar Htun*
"Enhancing the Different Types of Writing Skill of the Third Year Students at the University of Computer Studies, Taungoo"

VOL01IS01046 *Thwe Thwe Maw, Yin Yin San*
"Promoting the Students' Speaking Skills Through the Use of Communicative Language Teaching Approach"

VOL01IS01047 *Thwet Hnin Moe, Ye Win*
အင်းဝခေတ်ကျောက်စာ(၁၀)ချပ်လာဝါကျရိုးဖွဲ့ပုံလေ့လာချက်

VOL01IS01048 *Tin Tin Pyone, Wai Lwin Oo*
"Analysis of Elemental Concentrations on the Vegetables"

VOL01IS01049 *Win Cho, Kyi Kyi Khaing*
သော်တာဆွေ၏"ကြိုးခွဲရသည်"ဝတ္ထုတို၌ ဖလေ့သုံးစကားများ ဆန်းစစ်ချက်

VOL01IS01050 *Yi Mon Khine*
"Effective Motivations and Activities to Improve English Skills for University Students"

VOL01IS01051 *Yi Yi Thant*
"The Impact of Integrating Reading and Writing Skills at the Technological University's Students"

VOL01IS01052 *Yin Yin San, Thwe Thwe Maw*
"Using Prewriting Strategies for Promoting EFL Students' Writing Skills and Attitudes"

VOL01IS01053 *Zarchi San, Pan Wint Hmone Htwe, K Khaing Wint*
"Design and Construction of Dual Axis Solar Tracker in Pyay"

VOL01IS01054 *Zaw Min Tun, Aye Myat Mon*
"A Study of Quality Assurance in Universities"

VOL01IS01055 *Zin Nwe Khaing, Hla Yin Moe, San San Nwe*
"Optimization Problem Solving with Graphical and Simplex Method"

ISBN 978-99971-0-759-6