ISSN : 2165-4069(Online)
ISSN : 2165-4050(Print)

IJARAI

International Journal of
Advanced Research in Artificial Intelligence

Volume 4 Issue 3

www.ijarai.thesai.org

A Publication of
The Science and Information Organization

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN ARTIFICIAL INTELLIGENCE

THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org

Editorial Preface

From the Desk of Managing Editor…

Artificial Intelligence is hardly a new idea. Human likenesses, with the ability to act as human, dates back to Geek mythology with Pygmalion's ivory statue or the bronze robot of Hephaestus. However, with innovations in the technological world, AI is undergoing a renaissance that is giving way to new channels of creativity.

The study and pursuit of creating artificial intelligence is more than designing a system that can beat grand masters at chess or win endless rounds of Jeopardy!. Instead, the journey of discovery has more real-life applications than could be expected. While it may seem like it is out of a science fiction novel, work in the field of AI can be used to perfect face recognition software or be used to design a fully functioning neural network.

At the International Journal of Advanced Research in Artificial Intelligence, we strive to disseminate proposals for new ways of looking at problems related to AI. This includes being able to provide demonstrations of effectiveness in this field. We also look for papers that have real-life applications complete with descriptions of scenarios, solutions, and in-depth evaluations of the techniques being utilized.

Our mission is to be one of the most respected publications in the field and engage in the ubiquitous spread of knowledge with effectiveness to a wide audience. It is why all of articles are open access and available view at any time.

IJARAI strives to include articles of both research and innovative applications of AI from all over the world. It is our goal to bring together researchers, professors, and students to share ideas, problems, and solution relating to artificial intelligence and application with its convergence strategies. We would like to express our gratitude to all authors, whose research results have been published in our journal, as well as our referees for their in-depth evaluations.

We hope that this journal will inspire and educate. For those who may be enticed to submit papers, thank you for sharing your wisdom.

Editor-in-Chief
IJARAI
Volume 4 Issue 3 March 2015
ISSN: 2165-4069(Online)
ISSN: 2165-4050(Print)

Editorial Board

Reviewer Board Members

CONTENTS

Digital Library of Expert System Based at Indonesia Technology University

Dewa Gede Hendra Divayana[1]
Chair of Information Technology Department
Indonesia Technology University
Bali, Indonesia

I Made Sugiarta[3]
Lecture of Mathematics Education
Ganesha University of Education
Bali, Indonesia

I Putu Wisna Ariawan [2]
Lecture of Mathematics Education
Ganesha University of Education
Bali, Indonesia

I Wayan Artanayasa[4]
Chair of Sport & Health Education Department
Ganesha University of Education
Bali, Indonesia

***Abstract*—Digital library is a very interesting phenomenon in the world of libraries. In this era of globalization, the digital library is needed by students, faculty, and the community in the search for quick reference through internet access, so that students, faculty, and the community does not have to come directly to the library. Accessing collections of digital libraries can also be done anytime and anywhere. Digital Library development also occurred at Indonesia Technology University. That University offers a digital library based of expert system. The concept of digital library is utilizing science expert system in the process of cataloging and searching digital collections. By using this digital library based of expert system, users can search the collection, reading collection, and download the desired collection by online system. The digital library based of expert system at Indonesia Technology University is built using the PHP programming language, MySQL database as a data base management system, and developed the method of forward chaining and backward chaining as inference engine.**

Keywords—Digital Library; Expert System; Forward Chaining; Backward Chaining

I. INTRODUCTION

In the current era of globalization, information technology has a very important role in supporting the activities carried out by the community. The development of information technology very rapidly leads us toward the use of documents in digital form.

Library as one of the sources of knowledge needs to be organized and presented for the system services can be accessed from anywhere and anytime with the involvement of information technology in the form of an integrated system, so the user does not have to come directly to the library. This phenomenon is supported by the use of the Internet that facilitate processing, dissemination, and accessing information in digital form from anywhere and anytime. The development of increasingly sophisticated technology also affects the formation of a library in a digital form.

Digital library is a very interesting phenomenon in the world of libraries. In this era of globalization, the digital library is needed by students, faculty, and the community in the search for quick reference through internet access, so that students, faculty, and the community do not have to come directly to the library. Accessing collections of digital libraries can also be done anytime and anywhere. Digital Library development also occurred at Indonesia Technology University. That University offers a digital library based of expert system. The concept of digital library is utilizing science expert system in the process of cataloging and searching digital collections. By using this digital library based of expert system, users can search the collection, reading collection, and download the desired collection by online system.

II. LITERATURE REVIEW

A. Digital Library

In [1], The library has become digital: processes such as mass digitization, web archiving, and to a smaller extent digital preservation, are no longer isolated but disseminated among relevant production teams within the library.

In [2], A digital library (DL) is a library in which collections are stored in digital formats (as opposed to print, microform, or other media) and accessible by computers.

In [3], A digital library is a particular kind of information system which consists of a set of components, typically a collection (or collections) of computer system offering diverse services on a technical infrastructure, people, and the environment or usage.

In [4], Digital libraries are set of library activities and services which facilitate by electronic means the processing, transmission and display of information.

B. Expert System

In [5], an expert system is a computer program designed to simulate the problem-solving behaviour of a human who is an expert in a narrow domain or discipline. An expert system is normally composed of a knowledge base (information, heuristics, etc.), inference engine (analyses the knowledge base), and the end user interface (accepting inputs, generating outputs). The concepts for expert system

development come from the subject domain of artificial intelligence (AI), and require a departure from conventional computing practices and programming techniques.

In [6], an Expert system is software that simulates the performance of human experts in a specific field. Today's expert systems have been used in many areas where require decision making or predicting with expertise.

In [7], an expert system is a set of programs that manipulate encoded knowledge to solve problems in a specialized domain that normally requires human expertise.

In [8], Expert System is a branch of Artificial Intelligence that makes extensive use of specialized knowledge to solve problems at the human expert level.

In [9], the Expert System (ES) is one of the well-known reasoning techniques that is utilized in diagnosis applications domain. In ES, human knowledge about a particular expertise to accomplish a particular task is represented as facts and rules in its knowledge base.

In [10], an expert system is the computer system that emulates the behaviour of human experts in a well-specified manner, and narrowly defines the domain of knowledge. It captures the knowledge and heuristics that an expert employs in a specific task. An overview of current technologies applied with an expert system that is developed for Database Management System, Decision Support System, and the other Intelligent Systems such as Neural Networks System, Genetic Algorithm, etc.

In [11], expert system is an artificial intelligence system that combines knowledge base with inference engine so that it can adopt the ability of the experts into a computer, so the computer can solve problems such as the often performed by experts.

C. Forward Chaining

The inference engine contains the methodology used to perform reasoning on the information in the knowledge base and used to formulate conclusions. Inference engine is the part that contains the mechanism and function of thought patterns of reasoning systems that are used by an expert. The mechanism will analyze a specific problem and will seek answers, conclusions or decisions are best. Because the inference engine is the most important part of an expert system, that plays a role in determining the effectiveness and efficiency of the system.

There are several ways that can be done in performing inference, including the Forward Chaining. In [12], an inference engine using forward chaining searches the inference rules until it finds one where the IF clause is known to be true.

When found it can conclude, or infer, the THEN clause, resulting in the addition of new information to its dataset. In other words, it starts with some facts and applies rules to find all possible conclusions. Therefore, it is also known as Data Driven Approach.

In [13], forward chaining is matching facts or statements starting from the left (first IF).

D. Backward Chaining

Also in [13], backward chaining is matching facts or statements starting from the right (first THEN). In other words, the reasoning starts from the first hypothesis, and to test the truth of this hypothesis to look for the facts that exist in the knowledge base.

In [14], an inference engine using backward chaining would search the inference rules until it finds one which has a THEN clause that matches a desired goal. If the IF clause of that inference rule is not known to be true, then it is added to the list of goals (in order for goal to be confirmed it must also provide data that confirms this new rule). In other words, this approach starts with the desired conclusion and works backward to find supporting facts. Therefore, it is also known as Goal-Driven Approach.

In [15], backward chaining systems are good for diagnostic and classification tasks, but they are not good for planning, design, process monitoring, and quite a few other tasks. In backward chaining, the search is goal directed, so rules can be applied that are necessary to achieve the goal.

E. Digital Library Based of Expert System

In [16], Digital Library Based of Expert System is a digital library that implements the basic concept of expert system includes a knowledge base and the inference engine in helping service mechanism. The knowledge base is used for storage and cataloging process of making collections that exist in digital library, while the inference engine is used to search the detail collection that available in digital libraries such as the ability to work like an expert.

III. METHODOLOGY

A. Object dan Research Site

1) Research Object is Digital Library Based of Expert System

2) Research Site at Indonesia Technology University.

B. Data Type

In this research, the authors use primary data, secondary data, quantitative data and qualitative data.

C. Data Collection Techniques

In this research, the authors use data collection techniques such as interviews, observation, and documentation.

D. Analysis Techniques

Analysis techniques used in this research is descriptive statistical.

IV. RESULT AND DISCUSSION

A. Result

1) Early Trial

At this early trial, the authors conducted a limited scale testing of the digital library based of expert system that have been made previously by involving five staff at Indonesia Technology University to perform *white box* and *black box* testing. This test can be done by giving 10 questionnaires

early trials digital library based of expert system to staff at Indonesia Technology University.

Diagram form of answers score percentage given by the respondents in early trial can be described as follows:

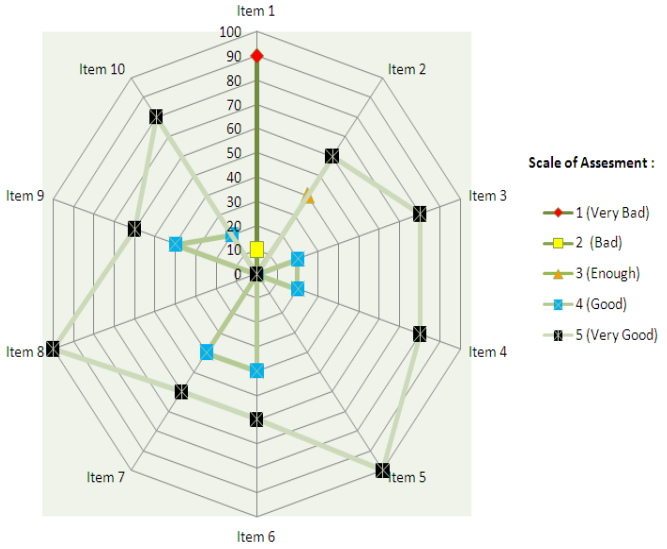


Fig. 1. Percentage Diagram of Respondents Answer Score In Early Trial

Based on the diagram above, it can be seen that the results of early trials of the digital library based of expert system, find a constraint that is the answer to a very bad score by 90% of the questions on the questionnaire 1st initial trials. This is due to the unavailability of the form for the create of a new username and password for administrator in the future if there is a mutation of the staff who operate the digital library based of expert system. Given these constraints, then the system needs to be revised again.

2) Field Trial

At this field trial, the authors tested in a larger scale, involving an expert is understood about the digital library and ten staff at Indonesia Technology University. This test can be done by giving 15 questionnaires field trials for digital library based of expert system to the librarian and staff at Indonesia Technology University.

Diagram form of answers score percentage given by the respondents in field trial can be described as follows:

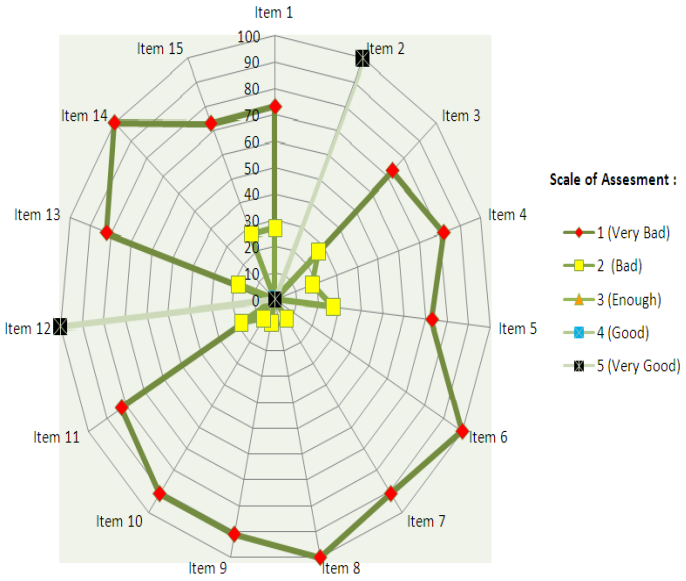


Fig. 2. Percentage Diagram of Respondents Answer Score In Field Trial

Based on the diagram above, it can be seen that the results of a field trial of the digital library based of expert system, the presence of obstacles that scores are very bad answer the score of 73% to the question 1st, 3rd, 5th, and 15th, 82% of the questions 4th, 11th and 13th, at 91% of the questions 7th, 9th, 10th, and at 100% of the questions 6th, 8th and 14th on field trial questionnaire.

This is due to the unavailability of the collection input or edits form if in the future there is a new journal and article. Of the constraints are found, then the system needs to be revised to obtain collection more interactive and dynamic.

3) Usage Test

At this usage test, the authors conducted a trial involving with the use of 50 people (users). The test is performed to test the operation of the overall form available on digital library based of expert system that has undergone revisions to field trials. This test can be done by giving the user satisfaction questionnaire to users who visited Indonesia Technology University.

Diagram form of answers score percentage given by the respondents in usage test can be described as follows:

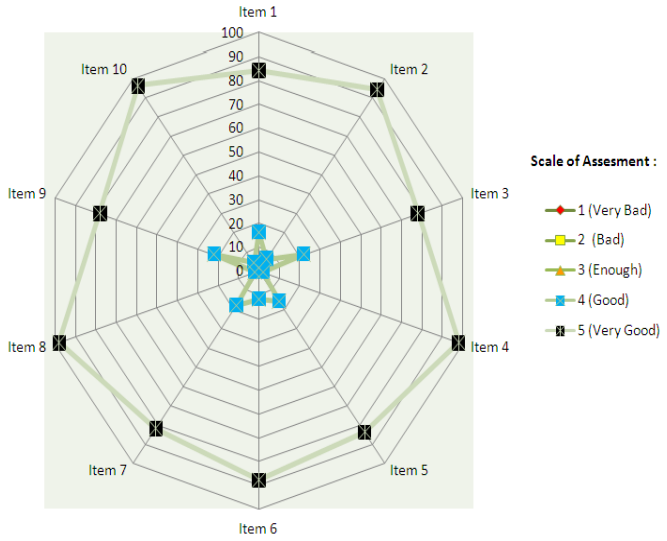


Fig. 3. Percentage Diagram of Respondents Answer Score In Usage Test

Based on the diagram above, it can be seen that the results of testing the use of the digital library based of expert system outline already looks very good and not found again the constraints in terms of technical operation (inputing and editing a new collection) as well as the principle method of expertise (forward chaining and bacward chaining method). This is evidenced by the percentage scoring very good response by 78% of questions 3rd, and 9th. Percentage scoring very good response by 82% of questions 7th. Percentage scoring very good response by 84% of questions 1st and 5th. Percentage scoring very good response by 88% of statement 6th. Percentage scoring very good response by 94% of question 2nd. Percentage scoring very good response by 96% of questions 10th.

As well as scoring 98% of the questions 4th and 8th trial usage. And it would be even better if the digital library based of expert system added help fasility form for written in accordance with the suggestions of the respondents to the improvement of the system, so as to explain the performance of the expert system and the function of the buttons in the design of digital library based of expert system overall with easy to understand and simple language.

B. Discussion

1) Knowledge Base

Knowledge base is used to build the expert system obtained from multiple sources of knowledge, containing data on digital library based of expert system at Indonesia Technology University. The knowledge base contained in a digital library based of expert system at Indonesia Technology University can be described by the following table.

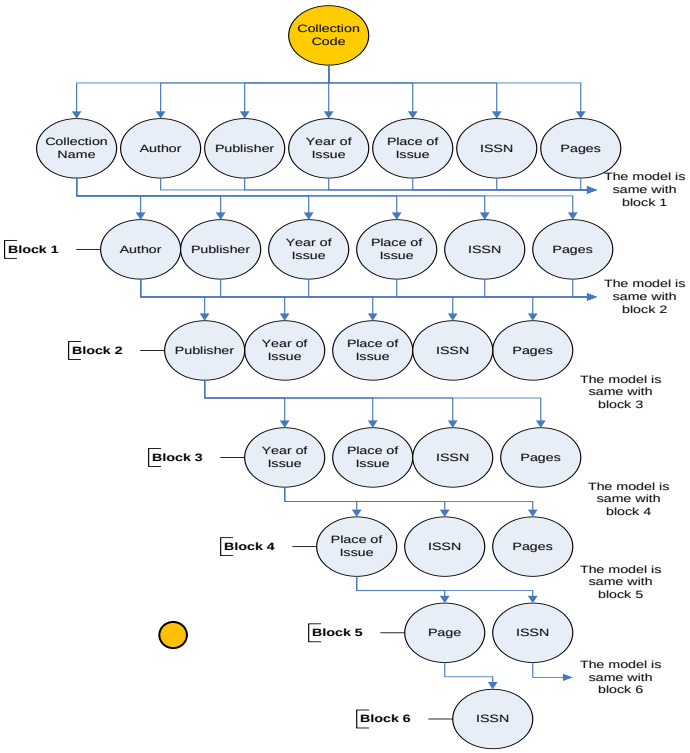
TABLE I. KNOWLEDGE BASE ON DIGITAL LIBRARY BASED OF EXPERT SYSTEM AT INDONESIA TECHNOLOGY UNIVERSITY

No	Properties/Identity	Journal	Book	Magazine	Article
1.	Collection Code	√	√	√	√
2.	Collection Name	√	√	√	√
3.	Author	√	√	√	√
4.	Publisher	√	√	√	√
5.	Year of Issue	√	√	√	√
6.	Pages	√	√	√	√
7.	Place of Issue	√	√	√	√
8.	ISSN	√	√	√	√

2) Shows forward chaining concept in digital library based of expert system

Application of forward chaining method in a digital library can be explained by the following chart:

For example : search by collection code



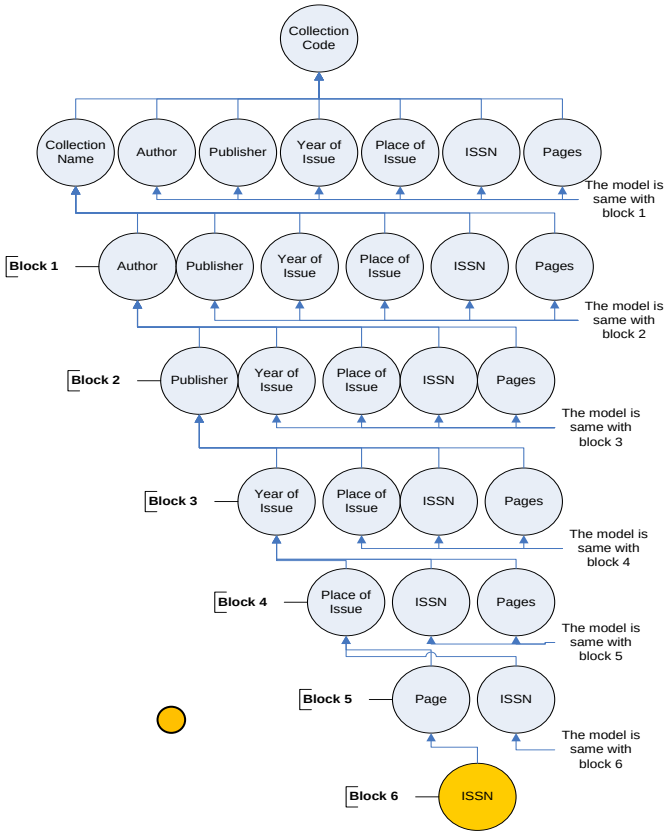
Explanation : = Search by

Fig. 4. Forward Chaining Concept in Digital Library Based of Expert System

3) Shows backward chaining concept in digital library based of expert system

Application of backward chaining method in a digital library can be explained by the following chart:

For example : search by ISSN



Explanation : = Search by

Fig. 5. Backward Chaining Concept in Digital Library Based of Expert System

4) Trials forward chaining and backward chaining performed by respondents

Respondents who did this trial was a expert and 19 staff at Indonesia Technology University conducted the field trials. The trial results are shown in table II.

Based on the table results of *trials forward chaining and backward chaining performed by respondents* mentioned above, it can be analyzed that the forward chaining and backward chaining method has been run in accordance with the rule of expert system inference engine.

To view the *forward chaining and backward chaining* method has been run in accordance with the rules can be seen in the percentage diagram of rules conformance testing.

TABLE II. TRIALS FORWARD CHAINING AND BACKWARD CHAINING METHOD

Respondent	Method	Properties/Identity								%
		CC	CN	AT	PL	YI	PG	PI	IS	
RS.01	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.02	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.03	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.04	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.05	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.06	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.07	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.08	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.10	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.11	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.12	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.13	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.14	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.15	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.16	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.17	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.18	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.19	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100
RS.20	FC	√	√	√	√	√	√	√	√	100
	BC	√	√	√	√	√	√	√	√	100

Explanation :

FC	: Forward Chaining	BC	: Backward Chaining
CC	: Collection Code	CN	: Collection Name
AT	: Author	PL	: Publisher
YI	: Year of Issue	PG	: Pages
PI	: Place of Issue	IS	: ISSN

As for the form of percentage diagram of rules conformance testing given by the respondents can be described as follows:

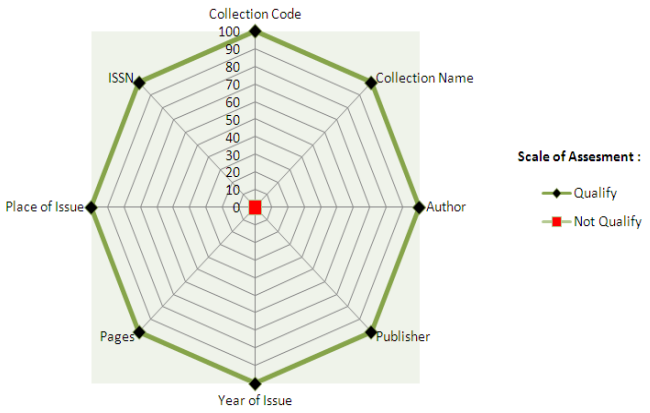


Fig. 6. Answer Percentage Diagram of Rules Conformance Testing

Based on the diagram above, it can be seen that the results of testing the suitability of digital library based of expert system rules is an outline already looks qualify. This is evidenced by the answer percentage of collection code, collection name, author, publisher, year of issue, pages, place of issue and ISSN according to the rules of backward chaining and forward chaining in the field of testing and each get a percentage of 100%.

5) Implementation of Digital Library Based of Expert System

a) Main Menu Page

Fig. 7. Main Page

This main menu page used as link to home menu, about us menu, book case menu, and contact menu.

b) Membership Registration Form

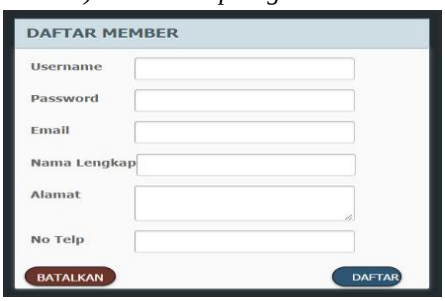


Fig. 8. Membership Registration Form

This form is used by users who will register to become a member, that all fields must be filled in accordance with the user's identity, and then click the register button to register and click cancel button to cancel the registration.

c) Member Login Form

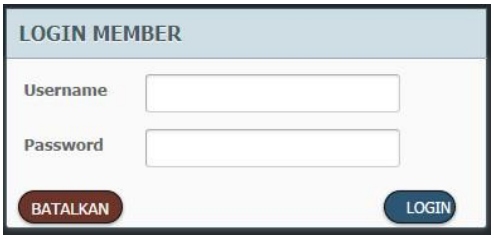


Fig. 9. Member Login Form

This form contains the username and password for the members, so that members can search and download a collection of all the collections that exist in digital library based of expert system at Indonesia Technology University.

d) Latest Collection Menu Page

After login, the user will go to the latest collection menu, the menu contains the latest collections of Digital Library based of expert system. The following figure is a latest collection menu display.

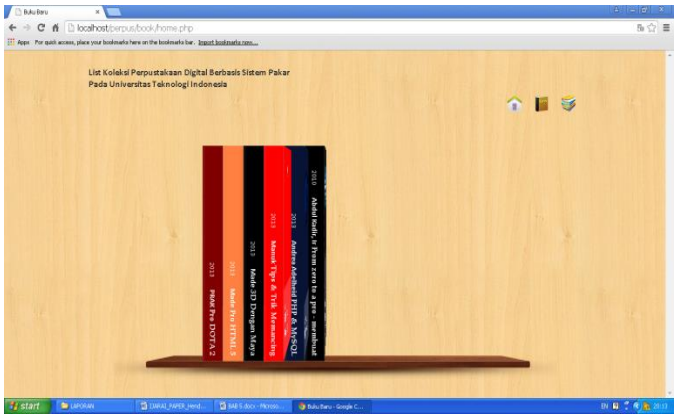


Fig. 10. Latest Collection Menu Page

e) Collection List Page

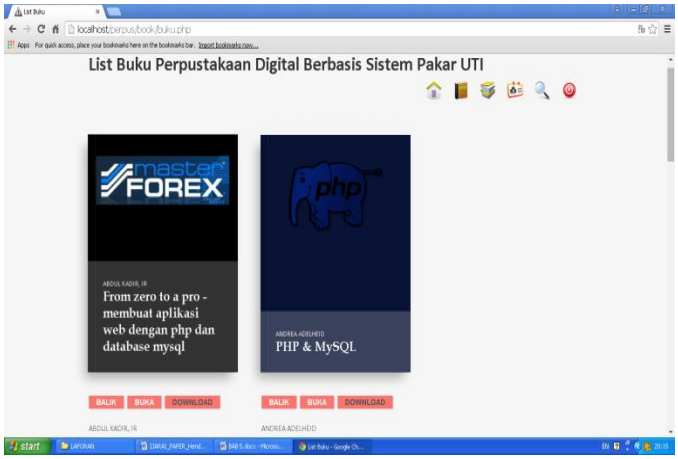


Fig. 11. Collection List Page

This collection list page is a page that is used to display a list of collections that exist in the digital library based of expert system at Indonesia Technology University and displays the complete details of the collection identity.

f) Collection Search Page

This form contains the search facilities of data on existing collections in digital library based of expert system using a keyword based input the desired category.

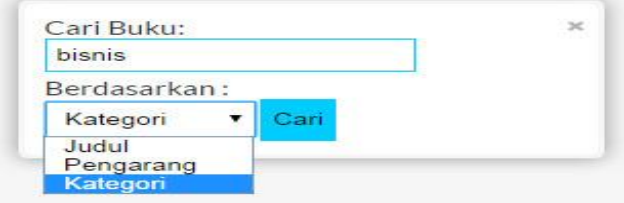


Fig. 12. Collection Search Page

g) Collection Search Result Page

Collection Search Result Page is the page serves to display the collection search result page. The following figure is a collection search results page display.

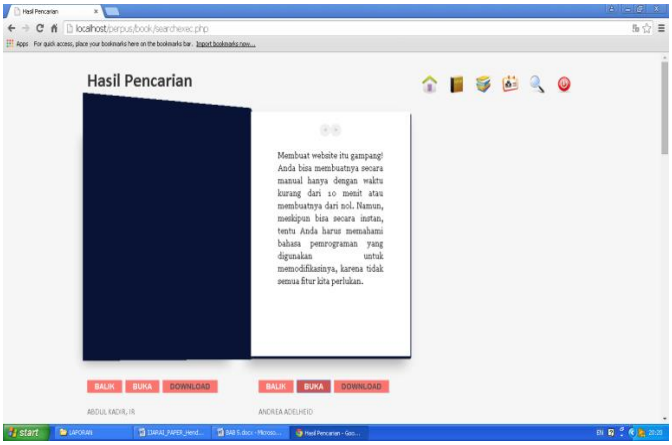


Fig. 13. Collection Search Result Page

h) Administrator Login Form

Fig. 14. Administrator Login Form

This form contains the username and password for the Admin, click on login button if you want to login and click cancel button if you want to cancel.

i) Administrator Page

Administrator page is a page that is used by administrators to system perform processing, such as processing of knowledge base and put the rule into an expert system inference engine. The following figure is a administrator page display.

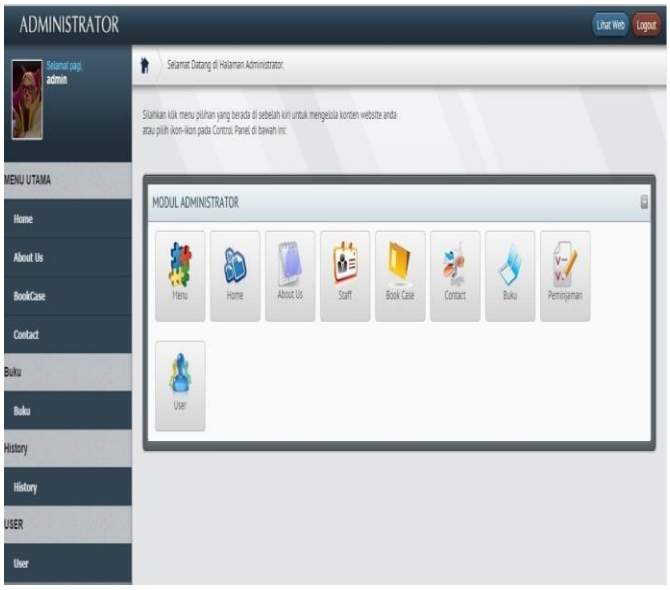


Fig. 15. Administrator Page

V. CONCLUSIONS

Based on the analysis that has been made and the results of the discussion in the previous section, then some conclusions can be drawn as follows:

a) With this digital library based of expert system can help the performance of existing conventional systems towards a computerized system, so as to speed up service.

b) With digital library based of expert system makes it easy for users to search the collection (books, journals, magazines, and articles) in digital form through the facilities of the internet without having to come directly to the campus library.

c) With the digital library digital library based of expert system, the concept of a real expert system can already be applied and useful in solving the existing problems in human life, especially in the field of library.

d) By looking at the optimal use of knowledge base and also the use of forward chaining and backward chaining inference engine implemented as in this digital library based of expert system, it can also be the accuracy result of the system is working at 100%.

ACKNOWLEDGMENT

The authors express their gratefulness to staff at Indonesia Technology University for inspiring words and allowing them to use the examination data. They generously thank Mr. Dayung, President of Indonesia Technology University, and Mr. Ketut Semadi, Dean of Computer Faculty, Indonesia Technology University.

REFERENCES

[1] E.Bermès, and L.Fauduet, "The Human Face of Digital Preservation: Organizational and Staff Challenges, and Initiatives at the Bibliothèque

nationale de France," in The International Journal of Digital Curation vol.6, 2011,pp.226-237.

[2] D.I.Greenstein, Thorin, and S. Elizabeth. The Digital Library: A Biography. Washington: Digital Library Federation, 2002.

[3] N. Fuhr, G. Tsakonas, T. Aalberg, M.Agosti, P. Hansen, and S. Kapidakis,"Evaluation of digital libraries," in International Journal on Digital Libraries vol. 8, 2007, pp.21–38.

[4] K. Towolawi, and Oluwakemi, "School Library Media Specialist's Awareness and Perception of Digital Library Services: A Survey," in Ozean Journal of Social Sciences vol. 6, 2013, pp.77-89.

[5] Y.A. Nada, "Construction of Powerful Online Search Expert System Based on Semantic Web," in International Journal of Advanced Computer Science and Applications vol.4, 2013, pp.181-187.

[6] Y. Qu, F. Tao, and H. Qui, "A Fuzzy Expert System Framework Using Object Oriented Techniques," in IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 2008, pp. 474-477.

[7] Y. Erdani, "Developing Recursive Forward Chaining Method in Ternary Grid Expert Systems," in International Journal of Computer Science and Network Security, vol.11, No.8, 2011, pp.126-130.

[8] J.C. Giarratano, and G. Riley, Expert Systems : Principles and Programming 4th Edition. USA : PWS Publishing Co, 2004.

[9] A. A. Hopgood, Intelligent Systems for Engineers and Scientists (2nd Edition). USA : CRC Press, 2001.

[10] E. Turban, and J. E. Aronson, Decision Support Systems and Intelligent System. NJ, USA: Prentice-Hall Inc, 2001.

[11] H. Divayana, "Development of Duck Diseases Expert System with Applying Alliance Method at Bali Provincial Livestock Office," in International Journal of Advanced Computer Science and Applications, Vol. 5, No. 8, 2014,pp.48-54.

[12] RC Chakraborty, 2010. Knowledge Representation: AI Course Lecture 15-22.Retrieved December, 2012, from www.myreaders.info/ html/aritificial_intelligence.html

[13] S.Kusumadewi, Artificial Intelligence (Technique and Application) 1st Edition. Yogyakarta : Graha Ilmu, 2003.

[14] K. Donna, A Comparison of Forward and Backward Chaining Algorithms For use in a Technical Support Expert System Used for Diagnosing Computer Virus Issues. Chico : Computer Science Department California State University, 2009.

[15] T. Sharma, *et.al.*, "Study of Difference Between Forward and Backward Reasoning," in International Journal of Emerging Technology and Advanced Engineering, vol. 2 No. 10, 2012, pp.271-273.

[16] H. Divayana, *et.al.*, Digital Library Based of Expert System at Indonesia Technology University. Bali : Indonesia Technology University, 2015.

AUTHORS PROFILE

Dewa Gede Hendra Divayana, S.Kom., M.Kom., M.M., Ph.D. was born in Denpasar, Bali, in 1984. He received his Ph.D. in Information Technology from Corllins University, USA. He worked as Lecturer of Expert System, Chair of Information Technology Department, Faculty of Computer, Indonesia Technology University, Bali, Indonesia.

Drs. I Putu Wisna Ariawan, M.Si. was born in Ulakan, May 19th 1968. He worked as Lecturer of Mathematics Education at Ganesha University of Education. And also He worked as Guest Lecturer of Graph Theory at Faculty of Computer, Indonesia Technology University, Bali, Indonesia.

Drs. I Made Sugiarta, M.Si. was born in Badung, in 1967. He worked as Lecturer of Mathematics Education at Ganesha University of Education. And also He worked as Guest Lecturer of Dicrete Mathematic at Faculty of Computer, Indonesia Technology University, Bali, Indonesia.

I Wayan Artanayasa, S.Pd., M.Pd. was born in Manikliyu, in 1973. He worked as Lecturer of Sport & Health Education at Ganesha University of Education. And also He worked as Guest Lecturer of Human-Computer Interaction at Faculty of Computer, Indonesia Technology University, Bali, Indonesia.

Application of Machine Learning Approaches in Intrusion Detection System: A Survey

Nutan Farah Haq
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh

Musharrat Rafni
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh

Abdur Rahman Onik
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh

Faisal Muhammad Shah
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh

Md. Avishek Khan Hridoy
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh

Dewan Md. Farid
Department of Computer Science and Engineering
United International University
Dhaka, Bangladesh

***Abstract*—Network security is one of the major concerns of the modern era. With the rapid development and massive usage of internet over the past decade, the vulnerabilities of network security have become an important issue. Intrusion detection system is used to identify unauthorized access and unusual attacks over the secured networks. Over the past years, many studies have been conducted on the intrusion detection system. However, in order to understand the current status of implementation of machine learning techniques for solving the intrusion detection problems this survey paper enlisted the 49 related studies in the time frame between 2009 and 2014 focusing on the architecture of the single, hybrid and ensemble classifier design. This survey paper also includes a statistical comparison of classifier algorithms, datasets being used and some other experimental setups as well as consideration of feature selection step.**

***Keywords*—Intrusion detection; Survey; Classifiers; Hybrid; Ensemble; Dataset; Feature Selection**

I. INTRODUCTION

The Internet has become the most essential tool and one of the best sources of information about the current world. Internet can be considered as one of the major components of education and business purpose. Therefore, the data across the Internet must be secure. Internet security is one of the major concerns now-a-days. As Internet is threatened by various attacks it is very essential to design a system to protect those data, as well as the users using those data. Intrusion detection system (IDS) is therefore an invention to fulfill that requirement. Network administrators adapt intrusion detection system in order to prevent malicious attacks. Therefore, intrusion detection system became an essential part of the security management. Intrusion detection system detects and reports any intrusion attempts or misuse on the network. IDS can detect and block malicious attacks on the network, retain the performance normal during any malicious outbreak, perform an experienced security analysis.

Intrusion detection system approaches can be classified in 2 different categories. One of them is anomaly detection and the other one is signature based detection, also known as misuse detection based detection approach [4, 41]. The misuse detection is used to identify attacks in a form of signature or pattern. As misuse detection uses the known pattern to detect attacks the main disadvantage is that it will fail to identify any unknown attacks to the network or system. On the other hand, anomaly detection is used to detect unknown attacks. There are different ways to find out the anomalies. Different machine learning techniques are introduced in order to identify the anomalies.

Over the years, many researchers and scholars have done some significant work on the development of intrusion detection system. This paper reviewed the related studies in intrusion detection system over the past six years. This paper enlisted 49 papers in total from the year 2009 to 2014.This paper enlisted the proposed architecture of the classification techniques, algorithms being used. A Statistical comparison has been added to show classifier design, chosen algorithms, used datasets as well as the consideration of feature selection step.

This paper is organized as follows: Section 2 provides the research topic overview where a number of techniques for intrusion detection have been described. Section 3 represents a statistical overview of articles over the years on the algorithms that were frequently used, the datasets for each experiment and the consideration of feature selection step. Section 4 includes the discussion and conclusion as well as some issues which have been highlighted for future research in intrusion detection system using machine learning approaches.

II. RESEARCH PAPER OVERVIEW

A. Machine Learning Approach

Machine learning is a special branch of artificial intelligence that acquires knowledge from training data based on known facts. Machine learning is defined as a study that allows computers to learn knowledge without being programmed mentioned by Arthur Samuel in 1959.Machine learning mainly focuses on prediction. Machine learning techniques are classified into three broad categories such as – supervised learning, unsupervised learning, and reinforcement learning.

1) Supervised Learning

Supervised learning is also known as classification. In supervised learning data, instances are labeled in the training phase. There are several supervised learning algorithms. Artificial Neural Network, Bayesian Statistics, Gaussian Process Regression, Lazy learning, Nearest Neighbor algorithm, Support Vector Machine, Hidden Markov Model, Bayesian Networks, Decision Trees(C4.5,ID3, CART, Random Forrest), K-nearest neighbor, Boosting, Ensembles classifiers (Bagging, Boosting), Linear Classifiers (Logistic regression, Fisher Linear discriminant, Naive Bayes classifier, Perceptron, SVM), Quadratic classifiers are some of the most popular supervised learning algorithms.

2) Unsupervised Learning

In unsupervised learning data instances are unlabeled. A prominent way for this learning technique is clustering.

Some of the common unsupervised learners are Cluster analysis (K-means clustering, Fuzzy clustering), Hierarchical clustering, Self-organizing map, Apriori algorithm, Eclat algorithm and Outlier detection (Local outlier factor).

3) Reinforcement Learning

Reinforcement learning means computer interacting with an environment to achieve a certain goal. A reinforcement approach can ask a user (e.g., a domain expert) to label an instance, which may be from a set of unlabeled instances.

B. Single Classifiers

One machine learning algorithm or technique for developing an intrusion detection system can be used as a standalone classifier or single classifier. Some of the machine learning techniques have been discussed in this study which have been found as frequently used single classifiers in our studied 49 research papers.

1) Decision Tree

Creating a classifier for predicting the value of a target class for an unseen test instance, based on several already known instances is the task of Decision tree (DT). Through a sequence of decisions, an unseen test instance is being classified by a Decision tree [11]. Decision tree is very much popular as a single classifier because of its simplicity and easier implementation [14]. Decision tree can be expanded in 2 types: (i) Classification tree, with a range of symbolic class labels and (ii) Regression tree, with a range of numerically valued class labels [11].

2) Naive Bayes

On the basis of the class label given Naive Bayes assumes that the attributes are conditionally independent and thus tries to estimate the class-conditional probability[15]. Naive Bayes often produces good results in the classification where there exist simpler relations. Naive Bayes requires only one scan of the training data and thus it eases the task of classification a lot.

3) K-nearest neighbor

Various distance measure techniques are being used in K-nearest neighbor. K-nearest neighbor finds out k number of samples in training data that are nearest to the test sample and then it assigns the most frequent class label among the considered training samples to the test sample. For classifying samples, K-nearest neighbor is known as an approach which is the most simple and nonparametric[8]. K-nearest neighbor can be mentioned as an instance-based learner, not an inductive based [35].

4) Artificial Neural Network

Artificial Neural Network (ANN) is a processing unit for information which was inspired by the functionality of human brains [23]. Typically neural networks are organized in layers which are made up of a number of interconnected nodes which contain a function of activation. Patterns are presented to the network via the input layer, which communicates to one or more hidden layers where via a system of weighted connections the actual processing is done. The hidden layers then link to an output layer for producing the detection result as output.

5) Support Vector Machines

Support vector machine (SVM) was introduced in mid-1990's [5]. The concept behind SVM for intrusion detection basically is to use the training data as a description of only the normal class of objects or which is known as non-attack in intrusion detection system, and thus assuming the rest as anomalies [51]. The classifier constructed by support vector machines methodology discriminates the input space in a finite region where the normal objects are contained and all the rest of the space is assumed to contain the anomalies [9].

6) Fuzzy Logic

For reasoning purpose, dual logic's truth values can be either absolutely false (0) or absolutely true (1), but in Fuzzy logic these kinds of restrictions are being relaxed [60]. That means in Fuzzy logic the range of the degree of truth of a statement can hold the value between 0 and 1 along with '0' and '1'[11].

C. Hybrid Classifiers

A hybrid classifier offers combination of more than one machine learning algorithms or techniques for improving the intrusion detection system's performance vastly. Using some clustering-based techniques for preprocessing samples in training data for eliminating non-representative training samples and then, the results of the clustering are used as training samples for pattern recognition in order to design a classifier. Thus, either supervised or unsupervised learning approaches can be the first level of a hybrid classifier [11].

D. Ensemble Classifiers

The classifiers performing slightly better than a random classifier are known as weak learners. When multiple weak learners are combined for the greater purpose of improving the performance of a classifier significantly is known as Ensemble classifier [11].Majority vote, bagging and boosting are some common strategies for combining weak learners [15].Though it is known that the disadvantages of the component classifiers get accumulated in the ensemble classifier, but it has been producing a very efficient performance in some combination. So researchers are becoming more interested in ensemble classifiers day by day.

III. Statistical Comparisons of Related Work

A. Distribution of Papers by Year of Publication

The survey comprises 49 research papers in the time frame between 2009 and 2014. It discussed 8 papers from each of the year 2009, 2010 and 2012.The highest number of papers are studied from the year 2011.The number of papers from that year is 11. 10 papers are enlisted for the year 2013 and 4 papers from 2014.Fig.1 depicts the percentage of distribution of papers by year of publication.

B. Classifier design

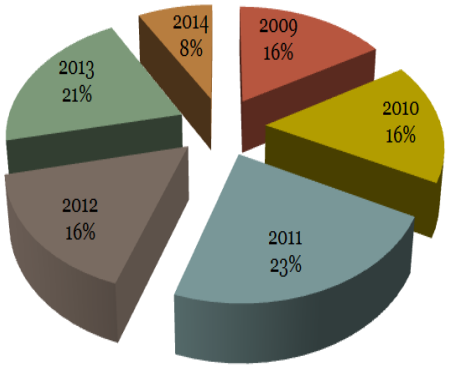


Fig. 1. Year-wise distribution of papers

Intrusion detection method can be categorized in 3 categories namely single, hybrid and ensemble[11] .Fig.2 depicts the number of research papers in terms of single, hybrid and ensemble classifiers used in each year. According

TABLE I. Total Numbers of Research Papers for the Types Of Classifier Design

Classifier design type	No. of research paper	References
Single	20	(D. Sa´nchez, 2009)[12], (Su-Yun Wua, 2009)[50], (Jun Ma, 2009)[27], (Mao Ye, 2009)[31], (Feng Jiang, 2009)[16], (Yung-Tsung Hou, 2010)[58], (Min Seok Mok, 2010)[34], (Han-Ching Wu, 2010)[22], (Chengpo Mua, 2010)[10], (Wang Dawei, 2011)[53], (G. Davanzo, 2011)[17], (Levent Koc, 2012)[29], (Carlos A. Catania, 2012)[9], (Inho Kang, 2012)[26], (Prabhjeet Kaur, 2012)[38], (Yusuf Sahin, 2013)[59], (S. Devaraju, 2013)[42], (Guillermo L. Grinblat, 2013)[21], (Mario Poggiolini, 2013)[32], (Adel Sabry Eesa, 2014)[2].
Hybrid	22	(Kamran Shafi, 2009)[28], (M. Bahrololum, 2009)[30], (Gang Wang, 2010)[18], (Woochul Shim, 2010)[55], (Muna Mhammad T. Jawhar, 2010)[37], (Ilhan Aydin, 2010)[25], (Seung Kim, 2011)[45], (I.T. Christou, 2011)[24], (Mohammad Saniee Abadeh, 2011)[36], (Shun-Sheng Wang, 2011)[47], (Su, 2011)[49], (Seungmin Lee, 2011)[46], (Yinhui Li, 2012)[57], (Bose, 2012)[6], (Prof. D.P. Gaikwad, 2012)[39], (A.M.Chandrashekhar, 2013)[1], (Mazyar Mohammadi Lisehroodi, 2013)[33], (Dahlia Asyiqin Ahmad Zainaddin, 2013)[13], (Seongjun Shin, 2013)[44], (Gisung Kim, A novel hybrid intrusion detection method integrating anomaly detection with misuse detection, 2013)[19], (Wenying Feng, 2014)[54], (Ravi Ranjan, 2014)[40].
Ensemble	7	(Tich Phuoc Tran, 2009)[52], (C.A. Laurentys, 2011)[7], (Dewan Md. Farid M. Z., 2011)[15], (Yang Yi, 2011)[56], (Siva S. Sivatha Sindhu, 2012)[48], (Dewan Md. Farid L. Z., 2013)[14], (Akhilesh Kumar Shrivas, 2014)[3]

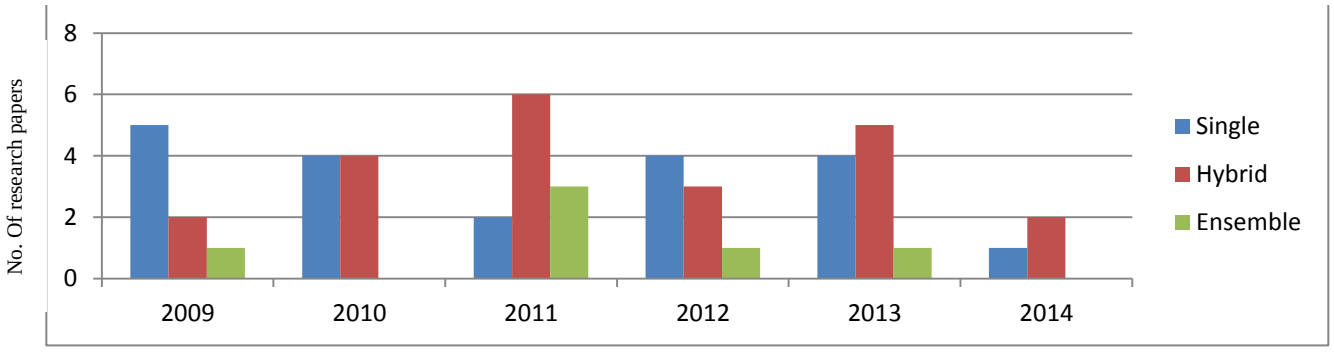


Fig. 2. Year wise distribution of research papers for the types of classifier design

to the statistical comparison between the enlisted papers, hybrid classifiers have the highest number of literatures in the time frame mentioned earlier with a total number of 22. What comes later in terms of study is single classifiers which have been studied in 20 papers.

C. Single classifiers

Fig. 3 depicts the number of single learning algorithms used as classifiers. The number of research papers in the single classifier architecture using different classification techniques, e.g. Bayesian, SVM, DT, ANN, KNN, Fuzzy Logic enlisted in this survey paper is twenty. Table II enlists the proposed algorithms used in all the articles reviewed in this paper. Table IV shows Year wise distribution of single classifiers regarding results and citation of each article.

Support vector machine and Artificial neural network are the most popular approaches for single learning algorithm classifiers. Though we have taken 49 related papers and number of comparative samples is less but the comparison result implies that Support Vector machine is by far the most common and considered single classification technique. On the contrary, Fuzzy logic seems to be less considerable among the single classifiers over the enlisted literatures.

D. Ensemble classifiers

Multiple weak learners are combined in Ensemble classifiers. Table III depicts the articles using ensemble classifiers in intrusion detection system. Statistics shows AdaBoost is the most commonly used learning algorithm along with majority voting. Table III also enlists the detection rate of each of the classifier and the citation of each article throughout the time period.

E. Hybrid classifiers

Table V depicts Year wise distribution of Hybrid classifiers regarding results and citation of each article. Hybrid classifiers in intrusion detection have established in the mainstream study due to the performance accuracy in recent times Statistics shows hybrid classifiers have the highest number of articles in the Year of 2011. The table also shows the used algorithms in each article and their performance in intrusion detection system.

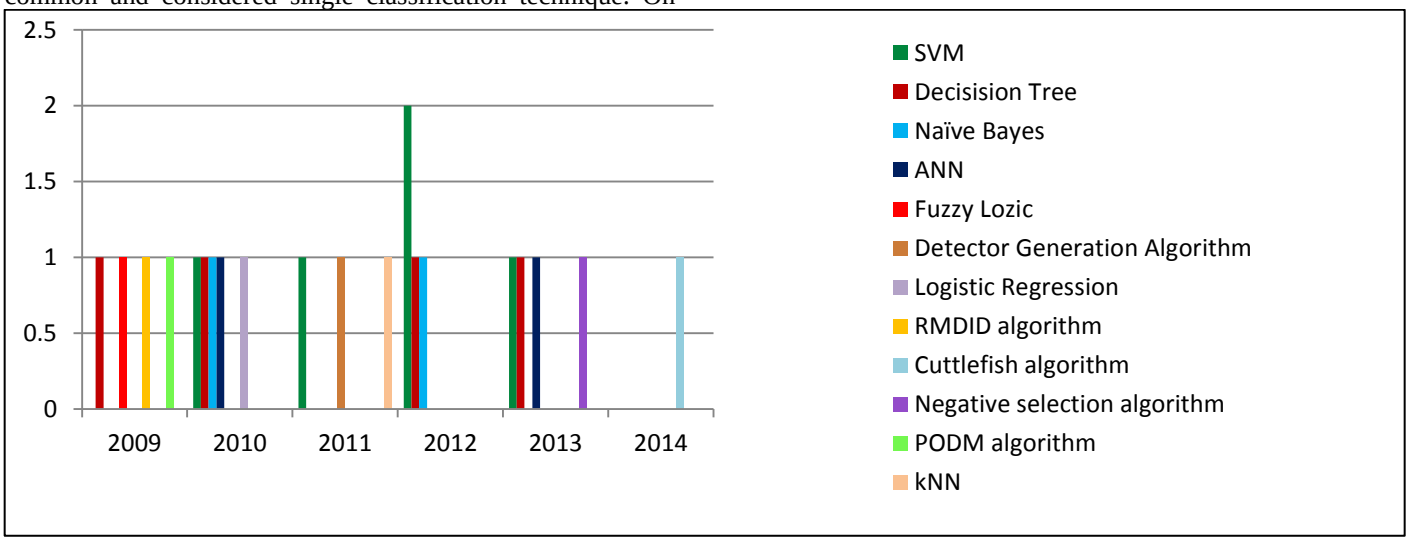


Fig. 3. Distribution of Single classifiers over the Years

TABLE II. ALGORITHMS USED IN SINGLE TYPE OF CLASSIFIER DESIGNED BASED RESEARCH PAPERS

Algorithm	Research paper Title	Reference
Naive Bayes	• A network intrusion detection system based on a hidden naive bayes multiclass classifier. • Malicious web content detection by machine learning.	(Levent Koc, 2012)[29] ; (Yung-Tsung Hou, 2010)[58]
Support Vector Machine	• An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection. • A differentiated one-class classification method with applications to intrusion detection. • Abrupt change detection with One-Class Time Adaptive Support Vector Machines. • Malicious web content detection by machine learning. • Anomaly detection techniques for a web defacement monitoring service.	(Carlos A. Catania, 2012)[9] ; (Inho Kang, 2012)[26] ; (Guillermo L. Grinblat, 2013)[21] ; (Yung-Tsung Hou, 2010)[58]; (G. Davanzo, 2011)[17].
Decision Tree	• Madam id for intrusion detection using data mining. • A cost-sensitive decision tree approach for fraud detection. • Data mining-based intrusion detectors. • Malicious web content detection by machine learning.	(Prabhjeet Kaur, 2012)[38]; (Yusuf Sahin, 2013)[59] ; (Su-Yun Wua, 2009)[50] ; (Yung-Tsung Hou, 2010)[58].
Artificial Neural Network	• Detection of accuracy for intrusion detection system using neural network classifier. • Neural networks-based detection of stepping-stone intrusion.	(S. Devaraju, 2013)[42] ; (Han-Ching Wu, 2010)[22].

Fuzzy Logic	• Data mining-based intrusion detectors.	(Su-Yun Wua, 2009)[50].
Detector Generation Algorithm	• Evolving boundary detector for anomaly detection	(Wang Dawei, 2011)[53].
Negative Selection algorithm	• Application of the feature-detection rule to the Negative Selection Algorithm	(Mario Poggiolini, 2013)[32].
Logistic regression	• Random effects logistic regression model for anomaly detection	(Min Seok Mok, 2010)[34].
RMDID	• Projected outlier detection in high-dimensional mixed-attributes data set.	(Mao Ye, 2009)[31].
PODM	• Information inconsistencies detection using a rule-map technique	(Jun Ma, 2009)[27]
Cuttlefish algorithm	• A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems.	(Adel Sabry Eesa, 2014)[2].
Sequence-based Outlier Detection algorithm	• Some issues about outlier detection in rough set theory.	(Feng Jiang, 2009)[16].
K-nearest neighbour (KNN)	• Anomaly detection techniques for a web defacement monitoring service.	(G. Davanzo, 2011)[17].

TABLE III. YEAR WISE DISTRIBUTION OF ENSEMBLE CLASSIFIERS REGARDING RESULTS AND CITATION OF EACH ARTICLE

Year	Research Paper Title	Reference	Algorithm used	Result (%)	Citation
2009	Novel intrusion detection using probabilistic neural network & adaptive boosting	(Tich Phuoc Tran, 2009)[52]	• NN • AdaBoost • BSPNN	DR : 94.31	14
2011	A novel artificial immune system for fault behavior detection	(C.A. Laurentys, 2011)[7]	• GA • Majority Vote	DR : 97.85	17
	Adaptive intrusion detection based on boosting & naive Bayesian classifier	(Dewan Md. Farid M. Z., 2011)[15]	• NB • AdaBoost	DR : 99.75	14
	Incremental SVM based on reversed set for network intrusion detection	(Yang Yi, 2011)[56]	• SVM • ISVM	DR : 81.377	30
2012	Decision tree based light weight intrusion detection using a wrapper approach	(Siva S. Sivatha Sindhu, 2012)[48]	• Neural ensemble decision tree	DR : 98.38	44
2013	An adaptive ensemble classifier for mining concept drifting data streams	(Dewan Md. Farid L. Z., 2013)[14]	• NB • C4.5 • AdaBoost	DR : 92.65	13
2014	An ensemble model for classification of attacks with feature selection based on KDD-99 & NSL-KDD data set	(Akhilesh Kumar Shrivas, 2014)[3]	• ANN • Bayesian Network • Gain ratio FS	DR : 97.53 (using NSL-KDD) DR: 99.41 (using KDD-99)	[a]

[a]Not cited yet.

TABLE IV. YEAR WISE DISTRIBUTION OF SINGLE CLASSIFIERS REGARDING RESULTS AND CITATION OF EACH ARTICLE

Year	Research Paper Title	Reference	Algorithm used	Result (%)	Citation
2009	Association rules applied to credit card fraud detection	(D. Sa´nchez, 2009)[12]	• Association rule methodology	Certainty factor : 80.08	64
	Data Mining based intrusion detectors	(Su-Yun Wua, 2009)[50]	• C4.5	DR : 70.62 FAR: 1.44	67
	Some issues about outlier detection in rough set theory	(Feng Jiang, 2009)[16]	• Outlier Detection algorithm	DR: SEQ based : 90 DIS based : 92	30
	Projected outlier detection in high dimensional mixed attributes data set	(Mao Ye, 2009)[31]	• PODM algorithm	DR: Credit approval data : 70 Breast Cancer Data : 80 Mushroom Data : 96 Synthetic Data : 97	24
	Information inconsistencies detection using a rule map technique	(Jun Ma, 2009)[27]	• RMDID algorithm	Error scales = 5.0% Inconsistent entries in Train Set = 5, Test Set = 4	1
2010	Malicious web content detection by machine learning	(Yung-Tsung Hou, 2010)[58]	• Naïve Bayes • DT • SVM • AdaBoost	Accuracy : NB : 58.28 DT : 94.74 SVM: 93.51 Boosted DT: 96.14	39
	Random effect logistic regression model for anomaly detection	(Min Seok Mok, 2010)[34]	• Logistic regression model.	Classification accuracy : Training dataset : 79.43 (Normal) 20.57(Attack) Validation dataset: 79.17 (Normal) 20.83 (Attack)	8
	An intrusion response decision making model based on hierarchical task network planning	(Chengpo Mua, 2010)[10]	• Hierarchical task network planning	Roc curve : excellent	20
	Neural Networks based detection of stepping stone intrusion	(Han-Ching Wu, 2010)[22]	• Neural Network	Accuracy : 99.0	13

2011	Evolving boundary detectors for anomaly detection	(Wang Dawei, 2011)[53]	• Detector Generation algorithm	DR : Iris Dataset : 99.28 considering Self radius = 0.08 Boundary threshold = 0.04 KDD dataset : DOS : 94.5 Probing : 93.64 U2R: 78.85 R2L: 50.69 considering Self radius = 0.05 Boundary threshold = 0.025	6
	Anomaly detection techniques for a web defacement monitoring service	(G. Davanzo, 2011)[17]	• K nearest neighbor • Support Vector machine	FPR: K nearest neighbor : 19.43 SVM :6.45	3
2012	A network intrusion detection system based on Hidden Naïve bayes multiclass classifier	(Levent Koc, 2012)[29]	• Hidden Naïve Bayes	Accuracy : 93.73 Error rate: 6.28	45
	An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection	(Carlos A. Catania, 2012)[9]	• Support Vector machine	DR : 88.64 (80% attack) 98.37 (1% attack)	11
	A differentiated one-class classification method with applications to intrusion detection	(Inho Kang, 2012)[26]	• Support Vector machine	DR : M=200* Targeted attack : 96.9 (4.7 % more than ordinary detection)	17
	Madam id for intrusion detection using Data mining	(Prabhjeet Kaur, 2012)[38]	• Decision Tree (J48)	FP rate :75.00 Precession : 1.7 Recall: 66.7	7
2013	A cost sensitive Decision tree approach for fraud detection	(Yusuf Sahin, 2013)[59]	• Decision Tree	TPR: Direct cost : 74.6 Class Probability : 92.1 CS-Gini : 92.8 Cs-IG: 92.6	9
	Detection of accuracy for intrusion detection system using neural network classifier	(S. Devaraju, 2013)[42]	• Neural Network	Accuracy : FFNN : 79.49 ENN: 78.1 GRNN: 58.74 PNN:85.50 RBNN: 83.51	4
	Abrupt change detection with one class time adaptive Support Vector Machine	(Guillermo L. Grinblat, 2013)[21]	• Support Vector Machine	495.9 sequences correctly classified within 500 sequences.	3
	Application of feature –detection rule to the negative selection algorithm	(Mario Poggiolini, 2013)[32]	• Negative Selection algorithm	Feature Detection rule : 0.9375 HD rule : 0.7686 RCHK(No MHC rule):0.8258 RCHK(Global MHC rule) : 0.5155 RCHK(MHC) rule : 0.9482	3
2014	A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection system	(Adel Sabry Eesa, 2014)[2]	• Cuttlefish algorithm	AR : 73.267 DR: 71.067 FPR: 17.685	[b]

[b]Not cited yet.

TABLE V. A DETAILED INFORMATION ON RESEARCH PAPERS DESIGNED WITH HYBRID CLASSIFIER

Year	Research Paper Title	Reference	Algorithm(s) used	Result (%)	Citation
2009	Anomaly intrusion detection design using hybrid of unsupervised and supervised neural network	(M. Bahrololum, 2009)[30]	• NN	TP rate : 97.00(Dos) 71.65(Probe) 26.69(R2L)	11
	An adaptive genetic-based signature learning system for intrusion detection	(Kamran Shafi, 2009)[28]	• GA	Accuracy : 92 FA rate : 0.84	31
2010	A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering	(Gang Wang, 2010)[18]	• ANN. • Fuzzy clustering.	Accuracy : 96.71 Precision : 99.91(Dos) 48.12(Probe) 93.18(R2L) 83.33(U2R)	114
	A distributed sinkhole detection method using cluster analysis	(Woochul Shim, 2010)[55]	• Hierarchical cluster analysis.	DR : 96.61	7
	Design Network Intrusion Detection System using hybrid Fuzzy-Neural Network	(Muna Mhammad T. Jawhar, 2010)[37]	• Fuzzy C-means clustering. • NN	Accuracy : 100(Dos) 100(U2R) 99.8(Probe) 40(R2L) 68.6(Unknown)	21
	Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection	(Ilhan Aydin, 2010)[25]	• Negative selection. • Clonal selection. • KNN.	Accuracy : 97.65	51
2011	Detecting fraud in online games of chance and lotteries	(I.T. Christou, 2011)[24]	• LOF. • K-means clustering. • EXAMCE.	DR : 98	3
	Fast outlier detection for very large log data	(Seung Kim, 2011)[45]	• Kd-tree indexing. • Approximated KNN. • LOF.	Gained time efficiency : 293-8727	11
	Design and analysis of genetic fuzzy systems for intrusion detection in computer networks	(Mohammad Saniee Abadeh, 2011)[36]	• Fuzzy genetic based machine learning methods: (i)Michigan,(ii)Pitsburg,(iii)IRL.	DR : 88.13 (Mitchigan) 99.53 (Pitsburg) 93.2 (IRL)	21

	An Integrated Intrusion Detection System for Cluster-based Wireless Sensor Networks	(Shun-Sheng Wang, 2011)[47]	• BPN. • ART. • Rule based method.	Accuracy: 95.13	24
	Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbor classifiers	(Su, 2011)[49]	• GA. • KNN.	Accuracy : 97.42 (with known attack) Accuracy : 78 (with unknown attack)	16
	Self-adaptive and dynamic clustering for online anomaly detection	(Seungmin Lee, 2011)[46]	• SOM. • K-means clustering	DR : 83.4 (offline) 86.4 (online)	14
2012	An efficient intrusion detection system based on support vector machines and gradually feature removal method	(Yinhui Li, 2012)[57]	• K-means clustering. • SVM. • Ant colony.	DR : 98.6249	40
	The combined approach for anomaly detection using neural networks & clustering techniques	(Bose, 2012)[6]	• SOM. • K-means clustering.	DR : 98.5 (Dos)	2
	Anomaly Based Intrusion Detection System Using Artificial Neural Network and Fuzzy Clustering	(Prof. D.P. Gaikwad, 2012)[39]	• ANN. • Fuzzy clustering.	*	6
2013	Fortification of hybrid intrusion detection system using variants of neural networks & support vector machines	(A.M.Chandrashekhar, 2013)[1]	• Fuzzy C-means clustering. • Fuzzy neural network. • SVM.	Accuracy : 98.94 (Dos) 97.11 (Probe) 97.80 (U2R) 97.78 (R2L)	2
	Hybrid of fuzzy clustering Neural network over NSL data set for intrusion detection system	(Dahlia Asyiqin Ahmad Zainaddin, 2013)[13]	• Fuzzy clustering.	Recall : 99.1 (Dos) 94.1 (Prob) 78 (U2R) 89 (R2L)	4
	A hybrid framework based on neural network MLP & K-means clustering for intrusion detection system	(Mazyar Mohammadi Lisehroodi, 2013)[33]	• K-means clustering. • MLP	DR : 99.99 (Dos) 99.97 (Probe) 99.99 (U2R) 99.98 (R2L)	[c]
	Advanced probabilistic approach for network intrusion forecasting and detection	(Seongjun Shin, 2013)[44]	• Markov chain. • K-means clustering. • APAN.	DR : 90	9
	A novel hybrid intrusion detection method integrating anomaly detection with misuse detection	(Gisung Kim, A novel hybrid intrusion detection method integrating anomaly detection with misuse detection, 2013)[19]	• C4.5. • 1-class SVM.	DR : 99.98 (with known attack) 97.4 (with unknown attack) Training time : 21.375 sec Testing time : 10.13 sec	9
2014	Mining network data for intrusion detection through combining SVMs with ant colony networks	(Wenying Feng, 2014)[54]	• CSOACN (self organized ant colony network) • SVM • CSVAC(combining support vectors with ant colony)	DR : 94.86 FP : 6.01 FN : 1.00	10
	A new clustering approach for anomaly intrusion detection	(Ravi Ranjan, 2014)[40]	• C4.5. • SVM. • K-means clustering.	DR : 96.12 (Dos) 90.10 (R2L) 70.51 (U2R) 70.13 (R2L). Accuracy : 96.38 False alarm rate : 3.2	4

[c]Not cited yet

F. Used Dataset in Researches

Datasets are assigned for default tasks e.g., Classification, Regression, Function learning, Clustering. Datasets reviewed by this paper is for classification purpose. As Fig.4 depicts, by far the most common dataset being used is KDD cup 1999 dataset. This dataset contains 4,000,000 instances and 42 attributes. The number of papers using KDD cup 1999 data set yields a peak in 2011 and in total 20 research papers has mentioned KDD Cup 1999 as their dataset.

Car evolution dataset [32] contains 1,728 instances with 6 attributes, attribute types are categorical. Wisconsin Breast cancer [16] has multivariate data types, all 10 attributes are integer types and it has 699 instances. Glass [32] dataset with multivariate data types and 214 instances It has 10 real attributes. Mushroom dataset [32] contains 22 categorical attributes and 8,124 instances. Lympography dataset [16] contains 18 categorical attributes and 148 instances. Yeast dataset [24] have 8 real attributes with 1,484 instances. Fisher-Iris dataset [25] contains 4 real attributes with 150 instances.Bicup2006 dataset and CO2 dataset [27] have 1,323 and 296 instances respectively. Public datasets like DARPA 1998, DARPA 2000, Fisher-Iris dataset, NSL KDD datasets are used in many related studies. Study also shows that few private or non-public datasets used over the time frame. Although the study briefly highlights public datasets like KDD cup 99, DARPA 1998, DARPA 2000 being considered as standard datasets for intrusion detection system. DARPA dataset contains around 1.5 million traffic instances [36]. NSL-KDD dataset was proposed by removing all redundant instances from KDD'99. Thus, NSL-KDD dataset is more efficient than KDD'99 in getting more accurate evaluation of different learning techniques [19]. Some of the datasets were randomly used by the researchers. Table VI shows the year-wise distribution of randomly used dataset.

TABLE VI. YEAR-WISE DISTRIBUTION OF RANDOMLY USED DATASET

Data Set	2009	2010	2011	2012	2013	2014	Total
Car Evaluation					1		1
Glass					1		1
DAMADICS			1				1
Yeast			1				1
Ionosphere			1				1
Musk			1				1
Malicious Web pages		1	1				2
Bicup2006	1						1
CO2	1						1
Lympography	1						1

G. Feature Selection

Feature Selection is an important step for the improvement of the system performance. Feature selection is considered before the training phase. Feature selection points out the best features and eliminates the redundant and irrelevant features. Table VII shows the year-wise distribution of feature selection step consideration. Table VII implies that out of 49 studies, 21 used feature selection step for their proposed architecture. It also shows that the number of papers using feature selection yields a peak in the year 2012, where out of 8 papers in that year 7 used feature selection step. On the contrary, in 2009 the scenario was completely opposite. Though we have taken 49 related papers and number of differences in those papers are trivial but the comparison result implies that 21 experiments used feature selection where 28 experiments did not. It implies that feature selection is not a popular procedure in intrusion detection. Table VII and VIII overview the year-wise distribution of feature selection considered in related studies and the count of paper.

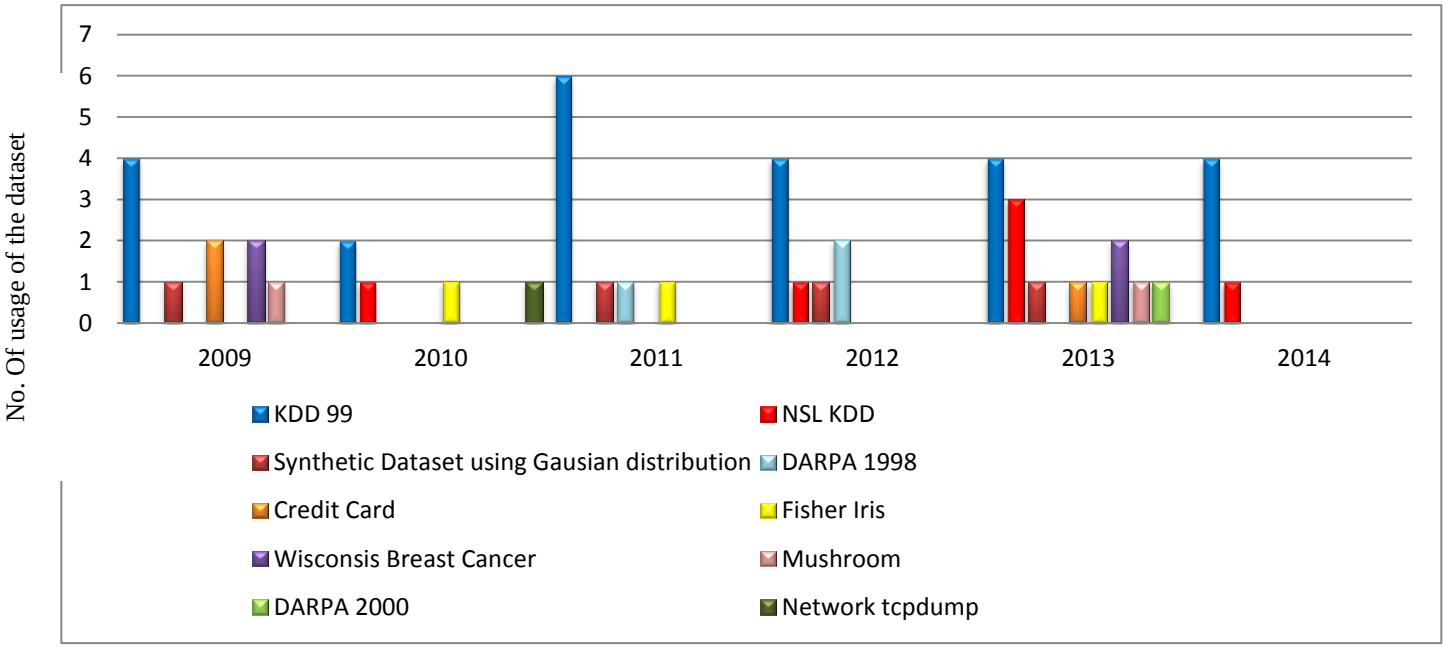


Fig. 4. Distribution of popular datasets over the years

TABLE VII. YEAR-WISE DISTRIBUTION OF FEATURE SELECTION CONSIDERED

Feature Selection Considered	2009	2010	2011	2012	2013	2014	Total
YES	1	3	4	7	4	2	21
NO	7	5	7	1	6	2	28

TABLE VIII. DISTRIBUTION OF RESEARCH PAPERS CONSIDERING THE FEATURE SELECTION STEP

Feature Selection	No. of research papers	Research papers
YES	21	A.m.chandrashekhar, k. (2013)[1]. adel sabryeesa, z. o. (2014)[2]. Akhilesh Kumar Shrivas, A. K. (2014)[3]. Bose, A. A. (2012)[6] Carlos A. Catania, F. B. (2012)[9]. Inho Kang, M. K. (2012)[26]. Levent Koc, T. A. (2012)[29].M. Bahrololum, E. S. (2009)[30]. Mario Poggiolini, A. E. (2013)[32]. Min Seok Mok, S. Y. (2010)[34]. Prabhjeet Kaur, A. K. (2012)[38]. S. Devaraju, S. R. (2013)[42]. Seongjun Shin, S. L. (2013)[44]. Shun-Sheng Wang, K.-Q. Y.-C.-W. (2011)[47]. Siva S. Sivatha Sindhu, S. G. (2012)[48]. Su, M.-Y. (2011)[49]. Woochul Shim, G. K. (2010)[55]. Yang Yi, J. W. (2011)[56]. Yinhui Li, J. X. (2012)[57]. Yung-Tsung Hou, Y. C.-S.-M. (2010)[58]. Yusuf Sahin, S. B. (2013)[59].
NO	28	C.A. Laurentys, R. P. (2011)[7] Chengpo Mua, Y. L. (2010)[10] D. Sa´nchez, M. V. (2009)[12] Dahlia Asyiqin Ahmad Zainaddin, Z. M. (2013)[13]. Dewan Md. Farid, L. Z. (2013)[14] Dewan Md. Farid, M. Z. (2011)[15] Feng Jiang, Y. S. (2009)[16] G. Davanzo, E. M. (2011)[17] Gang Wang, J. H. (2010)[18] Gisung Kim, S. L. (2013)[19](Ravi Ranjan, 2014)[40] Guillermo L. Grinblat, L. C. (2013)[21] Han-Ching Wu, S.-H. S. (2010)[22] I.T. Christou, M. B. (2011)[24] Ilhan Aydin, M. K. (2010)[25]. Jun Ma, J. L. (2009)[27] Kamran Shafi, H. A. (2009)[28]Mao Ye, X. L. (2009)[31]. Mazyar Mohammadi Lisehroodi, Z. M. (2013)[33]. Mohammad Saniee Abadeh, H. M. (2011)[36]. Muna Mhammad T. Jawhar, M. M. (2010)[37]. Prof. D.P. Gaikwad, S. J. (2012)[39] Seung Kim, N. W.-H. (2011)[45]. Seungmin Lee, G. K. (2011)[46]. Su-Yun Wua, E. Y. (2009)[50]. Tich Phuoc Tran, L. C. (2009)[52]. Wang Dawei, Z. F. (2011)[53]. Wenying Feng, Q. Z. (2014)[54].

IV. Discussion, Future Work and Conclusion

Uses of different classifier techniques in intrusion detection system is an emerging study in machine learning and artificial intelligence. It has been the attention of researchers for a long period of time. This paper has identified 49 research papers related to application of using different classifiers for intrusion detection published between 2009 and 2014. Though this survey paper cannot claim to be an in-depth study of those studies, but it presents a reasonable perspective and shows a valid comparison of works in this field over those years. The following issues could be useful for future research:

- Removal of redundant and irrelevant features for the training phase is a key factor for system performance. Consideration of feature selection will play a vital role in the classification techniques in future work.
- Feature selection has many algorithms to work with. Using different feature selection algorithms and working with the best possible one will be helpful for the classification techniques and also increase the consideration of feature selection step in intrusion detection.
- Uses of single classifiers or baseline classifiers in performance measurement can be replaced by hybrid or ensemble classifiers.

References

[1] A.M.Chandrashekhar, K. (2013). Fortification of hybrid intrusion detection system using variants of neural networks & support vector machines. International Journal of Network Security & Its Applications (IJNSA) .

[2] Adel Sabry Eesa, Z. O. (2014). A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. Expert Systems with Applications,ELSEVIER .

[3] Akhilesh Kumar Shrivas, A. K. (2014). An Ensemble Model for Classification of Attacks with Feature Selection based on KDD99 and NSL-KDD Data Set. International Journal of Computer Applications .

[4] Anderson, J. (1995). An introduction to neural networks. Cambridge: MIT Press.

[5] Bernhard E Boser, I. M. (1992). A Training Algorithm for Optimal Margin Classiers. Proceedings of the 5th Annual ACM Workshop on Computational , 144-152.

[6] Bose, A. A. (2012). THE COMBINED APPROACH FOR ANOMALY detection using neural networks & clustering techniques. Computer Science & Engineering: An International Journal (CSEIJ) .

[7] C.A. Laurentys, R. P. (2011). A novel Artificial Immune System for fault behavior detection. Expert Systems with Applications,ELSEVIER .

[8] C.M.Bishop. (1995). Neural networks for pattern recognition. England: Oxford University.

[9] Carlos A. Catania, F. B. (2012). An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection. Expert Systems with Applications,ELSEVIER .

[10] Chengpo Mua, Y. L. (2010). An intrusion response decision-making model based on hierarchical. Expert Systems with Applications,ELSEVIER .

[11] Chih-Fong Tsai, Y.-F. H.-Y.-Y. (2009). Intrusion detection by machine learning: A review. expert systems with applications,ELSEVIER .

[12] D. Sa´nchez, M. V. (2009). Association rules applied to credit card fraud detection. Expert Systems with Applications,ELSEVIER .

[13] Dahlia Asyiqin Ahmad Zainaddin, Z. M. (2013). HYBRID OF FUZZY CLUSTERING NEURAL NETWORK OVER NSL DATASET FOR INTRUSION DETECTION SYSTEM. Journal of Computer Science.

[14] Dewan Md. Farid, L. Z. (2013). An Adaptive Ensemble Classifier for Mining Concept-Drifting Data Streams. Expert systems with Applications,ELSEVIER .

[15] Dewan Md. Farid, M. Z. (2011). Adaptive Intrusion Detection based on Boosting and. International Journal of Computer Applications .

[16] Feng Jiang, Y. S. (2009). Some issues about outlier detection in rough set theory. expert systems with application,ELSEVIER .

[17] G. Davanzo, E. M. (2011). Anomaly detection techniques for a web defacement monitoring service. Expert Systems with Applications,ELSEVIER .

[18] Gang Wang, J. H. (2010). A new approach to intrusion detection using Artificial Neural Networks and. Expert Systems with Applications,ELSEVIER .

[19] Gisung Kim, S. L. (2013). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications,ELSEVIER .

[20] Gisung Kim,J.C,S.K. (2012). A congestion-aware IDS node selection method for wireless sensor networks. IJDSN.

[21] Guillermo L. Grinblat, L. C. (2013). Abrupt change detection with One-Class Time-Adaptive Support Vector Machines. Expert Systems with Applications,ELSEVIER .

[22] Han-Ching Wu, S.-H. S. (2010). Neural networks-based detection of stepping-stone intrusion. Expert Systems with Applications,ELSEVIER .

[23] Haykin, S. (1999). Neural networks: A comprehensive foundation (2nd Edition). New Jersey: Prentice Hall.

[24] I.T. Christou, M. B. (2011). Detecting fraud in online games of chance and lotteries. Expert Systems with Applications,ELSEVIER .

[25] Ilhan Aydin, M. K. (2010). Chaotic-based hybrid negative selection algorithm and its applications in fault. expert systems with applications,ELSEVIER .

[26] Inho Kang, M. K. (2012). A differentiated one-class classification method with applications to intrusion detection. Expert Systems with Applications,ELSEVIER .

[27] Jun Ma, J. L. (2009). Information inconsistencies detection using a rule-map technique. Expert systems with applications,ELSEVIER .

[28] Kamran Shafi, H. A. (2009). An adaptive genetic-based signature learning system for intrusion detection. Expert Systems with Applications, ELSEVIER .

[29] Levent Koc, T. A. (2012). A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier. Expert Systems with Applications,ELSEVIER .

[30] M. Bahrololum, E. S. (2009). ANOMALY INTRUSION DETECTION DESIGN USING. International Journal of Computer Networks & Communications (IJCNC) .

[31] Mao Ye, X. L. (2009). Projected outlier detection in high-dimensional mixed-attributes data set. Expert systems with applications,ELSEVIER .

[32] Mario Poggiolini, A. E. (2013). Application of the feature-detection rule to the Negative Selection Algorithm. Expert Systems with Applications,ELSEVIER .

[33] Mazyar Mohammadi Lisehroodi, Z. M. (2013). A HYBRID FRAMEWORK BASED ON NEURAL NETWORK MLP AND K-MEANS CLUSTERING FOR INTRUSION DETECTION SYSTEM. Proceedings of the 4th International Conference on Computing and Informatics, ICOCI 2013 (p. Paper No. 020). Sarawak, Malaysia: Universiti Utara Malaysia.

[34] Min Seok Mok, S. Y. (2010). Random effects logistic regression model for anomaly detection. Expert Systems with Applications,ELSEVIER .

[35] Mitchell, T. (1997). Machine learning. New york: MacHraw Hill.

[36] Mohammad Saniee Abadeh, H. M. (2011). Design and analysis of genetic fuzzy systems for intrusion detection in. Expert Systems with Applications,ELSEVIER .

[37] Muna Mhammad T. Jawhar, M. M. (2010). Design Network Intrusion Detection System using hybrid. International Journal of Computer Science and Security.

[38] Prabhjeet Kaur, A. K. (2012). MADAM ID FOR INTRUSION DETECTION USING DATA MINING. International Journal of Research in IT & Management,IJRIM .

[39] Prof. D.P. Gaikwad, S. J. (2012). Anomaly Based Intrusion Detection System Using Artificial Neural Network & Fuzzy clustering. International Journal of Engineering Research & Technology (IJERT) .

[40] Ravi Ranjan, G. S. (2014). A NEW CLUSTERING APPROACH FOR ANOMALY INTRUSION DETECTION . International Journal of Data Mining & Knowledge Management Process (IJDKP) .

[41] Rhodes, B. M. (2000). Multiple self-organizing maps for intrusion detection. Baltimore, MD.

[42] S. Devaraju, S. R. (2013). DETECTION OF ACCURACY FOR INTRUSION DETECTION SYSTEM USING NEURAL NETWORK CLASSIFIER. International Journal of Emerging Technology and Advanced Engineering .

[43] Sahoo, R. R. (2014). A NEW CLUSTERING APPROACH FOR ANOMALY INTRUSION DETECTION. International Journal of Data Mining & Knowledge Management Process (IJDKP) .

[44] Seongjun Shin, S. L. (2013). Advanced probabilistic approach for network intrusion forecasting and detection. Expert Systems with Applications,ELSEVIER .

[45] Seung Kim, N. W.-H. (2011). Fast outlier detection for very large log data. Expert Systems with Applications,ELSEVIER .

[46] Seungmin Lee, G. K. (2011). Self-adaptive and dynamic clustering for online anomaly detection. Expert Systems with Applications,ELSEVIER.

[47] Shun-Sheng Wang, K.-Q. Y.-C.-W. (2011). An Integrated Intrusion Detection System for Cluster-based Wireless. Expert Systems with Applications.

[48] Siva S. Sivatha Sindhu, S. G. (2012). Decision tree based light weight intrusion detection using a wrapper approach. Expert Systems with Applications,ELSEVIER .

[49] Su, M.-Y. (2011). Real-time anomaly detection systems for Denial-of-Service attacks by weighted. Expert Systems with Applications,ELSEVIER .

[50] Su-Yun Wua, E. Y. (2009). Data mining-based intrusion detectors. Expert Systems with Applications,ELSEVIER .

[51] Tax, D. &. (1999). Data domain description using support vectors. Proceedings of the european symposium on artificial neural networks, 251-256.

[52] Tich Phuoc Tran, L. C. (2009). Novel Intrusion Detection using Probabilistic Neural. (IJCSIS) International Journal of Computer Science and Information Security.

[53] Wang Dawei, Z. F. (2011). Evolving boundary detector for anomaly detection. Expert Systems with Applications.

[54] Wenying Feng, Q. Z. (2014). Mining network data for intrusion detection through combining SVMs with ant colony networks. Future Generation Computer Systems,ELSEVIER .

[55] Woochul Shim, G. K. (2010). A distributed sinkhole detection method using cluster analysis. Expert Systems with Applications,ELSEVIER .

[56] Yang Yi, J. W. (2011). Incremental SVM based on reserved set for network intrusion detection. Expert Systems with Applications .

[57] Yinhui Li, J. X. (2012). An efficient intrusion detection system based on support vector machines and gradually feature removal method. Expert Systems with Applications,ELSEVIER .

[58] Yung-Tsung Hou, Y. C.-S.-M. (2010). Malicious web content detection by machine learning. expert systems with applications,ELSEVIER .

[59] Yusuf Sahin, S. B. (2013). A cost-sensitive decision tree approach for fraud detection. Expert Systems with Applications,ELSEVIER.

[60] Zimmermann, H.-J. (2010). Fuzzy set theory. Advanced Review John Wiley & Sons, Inc

Attribute Reduction for Generalized Decision Systems*

Bi-Jun REN, Yan-Ling FU
Department of Information Engineering
Henan College of Finance and Taxation
Zhengzhou, Henan 451464, China

Ke-Yun QIN
College of Mathematics
Southwest Jiaotong University
Chengdu, Sichuan 610031, China

***Abstract*—Attribute reduction of information system is one of the most important applications of rough set theory. This paper focuses on generalized decision system and aims at studying positive region reduction and distribution reduction based on generalized indiscernibility relation. The judgment theorems for attribute reductions and attribute reduction approaches are presented. Our approaches improved the existed discernibility matrix and discernibility conditions. Furthermore, the reduction algorithms based on discernible degree are proposed.**

Keywords—Rough set; generalized indiscernibility relation; positive region reduction; distribution reduction

I. INTRODUCTION

The theory of rough sets, proposed by Pawlak[6], is an extension of the set theory. Rough set theory has been conceived as a tool to conceptualize, organize, and analyze various types of data, in particular, to deal with inexact, uncertain or vague knowledge in applications related to artificial intelligence.

Information systems (sometimes called data tables, attribute-value systems, decision system etc.) are used for representing knowledge. A basic problem related to many practical applications of information systems is whether the whole set of attributes is always necessary to define a given partition of a universe. This problem is referred to as knowledge reduction, i.e., removing superfluous attributes from the information systems in such a way that the remaining attributes are the most informative. A large variety of approaches have been proposed in the literatures for effective and efficient reduction of knowledge. Of all paradigms, rough set theory is perhaps the most recent one making significant contribution to the field. Based on this theory and discernibility functions, some approaches for attribute reduction in complete and discrete decision systems are proposed[5,9,11,14,16].

In many practical situations, it may happen that the precise values of some of the attributes in an information system are not known, i.e. are missing or known partially. Such a system is called an incomplete information system. In order to deal with incomplete information systems, classical rough sets have been extended to several general models by using other binary relations or covers on the universe[1,2,7,8,10,15,18,19]. Based on these extended rough set models, the researchers have put forward several meaningful indiscernibility relations in incomplete information system to characterize the similarity of objects. For instance, Kryszkiewicz[3,4] introduced a kind of indiscernibility relation, called tolerance relation, to handle incomplete information tables. Stefanowski[12] introduced two generalizations of the rough sets theory to handle the missing value. The first generalization introduces the use of a non symmetric similarity relation in order to formalize the idea of absent value semantics. The second proposal is based on the use of valued tolerance relations. The tolerance relation has also been generalized to constrained similarity relation and constrained dissymmetrical similarity relation[2,13,17]. Accordingly, some attribute reduction approached for incomplete decision systems have been proposed. In this paper, an approach to attribute reduction for incomplete decision systems based on generalized indiscernibility relation is presented. Specifically, this study is not limited to a particular indiscernibility relation, but focus on the indiscernibility relation that satisfies reflexivity and symmetry. A general theory frame of attribute reduction for incomplete decision system will be presented. The paper is organized as follows: In Section 2, we recall some notions and properties of rough sets and decision systems. In Section 3, we propose an approach for positive region reduction. The reduction algorithm based on discernible degree is also presented. Section 4 is devoted to distribution reduction. The paper is completed with some concluding remarks.

II. GENERALIZED DECISION SYSTEMS

An information system is a triplet (U, A, F), where U is a nonempty finite set of objects called the universe of discourse, $A=\{a_1,\cdots,a_m\}$ is a nonempty finite set of attributes, $F=\{f_j; j\le m\}$ is a set of information functions such that $f_j(x)\in V_j$ for all $x\in U$, where V_j is the domain of attribute a_j. A decision system $(U, C\cup\{d\}, F)$ is a special case of an information system, where d is a special attribute called decision. The elements of C are called conditional attributes.

In a generalized decision system, we do not care about the information function, but focus on the indiscernibility relations generated by attributes. Concretely, a generalized decision system is a triple $S=(U, A, d)$, where U is a nonempty universe of objects, A is a set of conditional attributes, and d is a distinguished decision attribute. Each conditional attribute a determines an indiscernibility relation which is denoted by R_a.

This work has been supported by the National Natural Science Foundation of China (Grant No. 61473239), The Key Scientific and Technological Funds (Grant No. 142102310096) of Henan, China and Soft-scientific Item (Grant No. 142400410671) of Henan, China.

In what follows we suppose that R_a is reflexive. Additionally, the decision attribute d determine a partition $U/d=\{D_1,\cdots,D_r\}$ of U. If $x\in D_i$, then we take i as the decision value of x and denoted by $d(x)=i$.

Let $S=(U,A,d)$ be a generalized decision system. For any $B\subseteq A$, the indiscernibility relation generated by B is defined as $R_B=\bigcap_{a\in B}R_a$. For $x\in U$, the neighborhood of x related to R_B is denoted as $R_B(x)=\{y\in U;(x,y)\in R_B\}$. Obviously, $R_B(x)=\bigcap_{a\in B}R_a(x)$. Additionally, because of the reflexivity of R_B, $\{R_B(x);x\in U\}$ forms a cover of U.

Definition 2.1[2,15] Let $S=(U,A,d)$ be a generalized decision system. For any $B\subseteq A$, $X\subseteq U$, the lower approximation and upper approximation of X with respect to R_B are defined as

$$\underline{R_B}(X)=\{x\in U;R_B(x)\subseteq X\} \quad (1)$$

$$\overline{R_B}(X)=\{x\in U;R_B(x)\cap X\neq\varnothing\} \quad (2)$$

Theorem 2.1[2,15] Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$, $X,Y\subseteq U$. Then

(1) $\underline{R_B}(X)\subseteq X\subseteq\overline{R_B}(X)$.

(2)If $X\subseteq Y$, then $\underline{R_B}(X)\subseteq\underline{R_B}(Y)$, $\overline{R_B}(X)\subseteq\overline{R_B}(Y)$.

(3) $\underline{R_B}(X\cap Y)=\underline{R_B}(X)\cap\underline{R_B}(Y)$, $\overline{R_B}(X\cup Y)\subseteq\overline{R_B}(X)\cup\overline{R_B}(Y)$.

(4) $\overline{R_B}(X)=\sim\underline{R_B}(\sim X)$, $\underline{R_B}(X)=\sim\overline{R_B}(\sim X)$.

III. Attribute Reduction Based on Positive Region

The section is devoted to the discussion of positive region reduction of generalized decision systems.

Definition 3.1[11] Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$, $U/d=\{D_1,\cdots,D_r\}$. The positive region of d with respect to B is defined as

$$Pos_B(d)=\bigcup_{X\in U/d}\underline{R_B}(X)=\bigcup_{i\le r}\underline{R_B}(D_i) \quad (3)$$

The above definition shows that $x\in Pos_B(d)$ if and only if the objects in $R_B(x)$ have the same decision values. Thus, $Pos_B(d)$ is the set of all elements of U that can be uniquely classified to blocks of the partition U/d by means of B. If we take B as the set of conditional attributes, then $x\in Pos_B(d)$ means the decision rule with respect to x is definite.

Theorem 3.1[9] Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$, $U/d=\{D_1,\cdots,D_r\}$. Then

(1) $Pos_B(d)\subseteq Pos_A(d)$.

(2) $Pos_B(d)=Pos_A(d)$ if and only if $\underline{R_B}(D_i)=\underline{R_A}(D_i)$ for each $i\le r$.

(3) $x\in Pos_B(d)$ if and only if $x\in\underline{R_B}([x]_d)$.

Definition 3.2 Let $S=(U,A,d)$ be a generalized decision system. If $B\subseteq A$ such that $Pos_B(d)=Pos_A(d)$, then B is called a positive region consistent set of S. The minimal positive region consistent set of S (with respect to set inclusion relation) is called as positive region reduction of S.

Let $S=(U,A,d)$ be a generalized decision system, $x,y\in U$. We consider the following condition $\omega(x,y)$:

$$\omega(x,y):x\in Pos_A(d)\wedge d(x)\neq d(y).$$

We note that $\omega(x,y)$ is not symmetric to x and y.

Theorem 3.2[9] Let $S=(U,A,d)$ be a generalized decision system. If $x,y\in U$ satisfy $\omega(x,y)$, then $\alpha_A(x,y)\neq\varnothing$, where $\alpha_A(x,y)=\{a\in A;(x,y)\notin R_a\}$.

Theorem 3.3 Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$. B is a positive region consistent set of S if and only if $B\cap\alpha_A(x,y)\neq\varnothing$ for $x,y\in U$ satisfy $\omega(x,y)$.

Proof: Suppose that B is a positive region consistent set of S and $x,y\in U$ satisfy $\omega(x,y)$. Then $d(x)\neq d(y)$ and $x\in Pos_A(d)$. By $x\in Pos_A(d)=Pos_B(d)$, we have $R_B(x)\subseteq[x]_d$. Because of $[x]_d\cap[y]_d=\varnothing$, it follows that $R_B(x)\cap[y]_d=\varnothing$, and consequently $y\notin R_B(x)$. Thus there exists $a\in B$ such that $(x,y)\notin R_a$, namely $a\in\alpha_A(x,y)$, and thus $B\cap\alpha_A(x,y)\neq\varnothing$.

Conversely, we suppose that $B\cap\alpha_A(x,y)\neq\varnothing$ for $x,y\in U$ satisfy $\omega(x,y)$. It only need to prove $Pos_A(d)\subseteq Pos_B(d)$. For any $x\in U$, if $x\notin Pos_B(d)$, then $R_B(x)\subset[x]_d$. Thus there exists $y\in R_B(x)$ such that $y\notin[x]_d$. By $B\cap\alpha_A(x,y)=\varnothing$, we know that x,y do not satisfy $\omega(x,y)$. It follows that $x\notin Pos_A(d)$ by $d(x)\neq d(y)$. Thus $Pos_A(d)\subseteq Pos_B(d)$ as required.

This theorem shows that, with respect to positive region reduction, x and y need to be discerned if x,y satisfy $\omega(x,y)$. In this case, we let $\vee\alpha_A(x,y)=\bigvee_{a\in\alpha_A(x,y)}a$ denote the disjunction of all attributes in $\alpha_A(x,y)$, where each attribute is looked upon as a Boolean variable. In what follows, $\Delta^*=\bigwedge_{(x,y)\in D^*}\vee\alpha_A(x,y)$ is called the positive discernibility function of S, where $D^*=\{(x,y)\in U\times U;\omega(x,y)\}$. It is noted that R_A is reflexive, therefore, D^* need not to be symmetry in general.

Theorem 3.4 Let $S=(U,A,d)$ be a generalized decision system and Δ^* be the positive discernibility function of S. If

$$\Delta^{*\prime}=(a_{11}\wedge\cdots\wedge a_{1m_1})\vee\cdots\vee(a_{k1}\wedge\cdots\wedge a_{km_k})$$

is the reduced disjunctive form of Δ^*, then $\mathrm{Re}d=\{T_1,\cdots,T_k\}$ is the set of all positive region reductions of S, where $T_i=\{a_{i1},\cdots,a_{im_i}\}$ for each $i\le k$.

Proof: (1) For any $i \le k$, T_i is a positive region reduction of S. In fact, if there exist $x, y \in U$ such that x, y satisfy $\omega(x,y)$ and $T_i \cap \alpha_A(x,y) = \varnothing$, then we let all Boolean variable in T_i be assigned 1 and the other Boolean variables be assigned 0. It follows that $\Delta^* = 0$ because $\vee \alpha_A(x,y) = 0$ and $\Delta^{*\prime} = 1$ because $a_{i1} \wedge \cdots \wedge a_{im_i} = 1$. This contradicts the fact that $\Delta^{*\prime}$ is the disjunctive form of Δ^*. Thus T_i is a positive region consistent set.

We suppose that there exists a proper subset $T \subset T_i$ such that $T \cap \alpha_A(x,y) \neq \varnothing$ for any $(x,y) \in D^*$. By the property of Boolean function, there exist $j \le k$ such that $T_j \subseteq T$. It follows that $T_j \subset T_i$. This contradicts the fact that $\Delta^{*\prime}$ is the reduced disjunctive form of Δ^*. Thus T_i is a positive region reduction of S.

(2) We suppose that B is a reduction of S. It follows that $B \cap \alpha_A(x,y) \neq \varnothing$ for $(x,y) \in D^*$. It follows that there exist $i \le k$ such that $T_i \subseteq B$. Because T_i is a positive region consistent set, we have $T_i \subseteq B$. Thus, $\{T_1, \cdots, T_k\}$ is just the set of all positive region reductions of S.

If R_A is reflexive and symmetric, then $\alpha_A(x,y) = \alpha_A(y,x)$ for any $x, y \in U$. Hence we have the following corollary.

Corollary 3.1 Let $S = (U, A, d)$ be a generalized decision system and $U = \{x_1, x_2, \cdots, x_n\}$. If R_a is reflexive and symmetric for any $a \in A$, then the positive discernibility function of S is

$$\Delta^* = \bigwedge_{(x,y) \in D_1^*} \vee \alpha_A(x,y)$$

where $D_1^* = \{(x_i, x_j); 1 \le j < i \le n, \omega_1(x_i, x_j)\}$, $\omega_1(x,y)$ represents the condition: $(x \in Pos_A(d) \vee y \in Pos_A(d)) \wedge d(x) \neq d(y)$.

Theorem 3.5 Let $S = (U, A, d)$ be a generalized decision system and R_a an equivalence relation for any $a \in A$. If $x \in Pos_A(d)$, $y \notin Pos_A(d)$ and $d(x) = d(y)$, then there exists $z \notin Pos_A(d)$ such that $d(x) \neq d(z)$ and $\alpha_A(x,y) = \alpha_A(x,z)$.

Proof: It is trivial that R_A is an equivalence relation on U. We use $[y]_A$ to denote $R_A(y)$. By $x \in Pos_A(d)$, $y \notin Pos_A(d)$ we have $[x]_A \subseteq [x]_d$, $[y]_A \not\subset [y]_d$. It follows that there exists $z \in [y]_A$ such that $z \notin [y]_d$. Thus $d(y) \neq d(z)$, and hence $d(x) \neq d(z)$. By $z \in [y]_A$, we have $(y,z) \in R_a$ for any $a \in A$. In consequence,

$$\alpha_A(x,y) = \{a \in A; a(x) \neq a(y)\} = \{a \in A; a(x) \neq a(z)\} = \alpha_A(x,z).$$

Furthermore, by $z \in [y]_A$, it follows that $[z]_A = [y]_A$. Thus we have $z \notin Pos_A(d)$ by $y \notin Pos_A(d)$.

Remark: Let $S = (U, A, d)$ be a decision system and R_a an equivalence relation for any $a \in A$.

Skowron[11] proposed the discernibility conditions for object pairs that need to discern with respect to positive region reduction. The discernibility conditions are

$$\omega_S(x,y): x \in Pos_A(d) \wedge y \notin Pos_A(d);$$

$$\text{or } x \notin Pos_A(d) \wedge y \in Pos_A(d);$$

$$\text{or } x \in Pos_A(d) \wedge y \in Pos_A(d) \wedge d(x) \neq d(y).$$

According to above theorem, the object pair (x,y) that satisfies $d(x) = d(y)$ do not need to discern in the criterion of positive region reduction. To be specific, Skowrons' discernibility conditions can be simplified as following:

$$\omega_1(x,y): (x \in Pos_A(d) \vee y \in Pos_A(d)) \wedge d(x) \neq d(y).$$

In essence, based on Corollary 3.1, the discernibility condition is $\omega_1(x,y)$ when the indiscernibility relation satisfies reflexivity and symmetry.

Theorem 3.4 presents an approach to calculate the positive region reductions by discernibility function. Similarly as pointed out in [11], the approach is NP hard. In the following of this section, we present a heuristic algorithm based on discernibility matrix to calculate positive region reduction.

Let $S = (U, A, d)$ be a generalized decision system, $B \subseteq A$. By Theorem 3.3, B is a positive region consistent set of S if and only if $B \cap \alpha_A(x,y) \neq \varnothing$ for $x, y \in U$ satisfy $\omega(x,y)$. It follows that D^* is the set of element pairs that needs to be discerned with respect to positive region reduction. For an attribute $a \in A$, $\{(x,y) \in D^*; a \in \alpha_A(x,y)\}$ is the set of object pairs that a can discern. Thus, the bigger the set $\{(x,y) \in D^*; a \in \alpha_A(x,y)\}$, the more possible that a is an element of a reduction. Based on this observation, we propose the notion of discernible degree.

Definition 3.3 Let $S = (U, A, d)$ be a generalized decision system, $E = \bigcup_{(x,y) \in D^*} \alpha_A(x,y)$. For any $a \in E$, the positive region discernible degree $\lambda(a)$ of a is defined as

$$\lambda(a) = \frac{\left|\{(x,y) \in D^*; a \in \alpha_A(x,y)\}\right|}{\left|D^*\right|},$$

where $\left|\{(x,y) \in D^*; a \in \alpha_A(x,y)\}\right|$ and $\left|D^*\right|$ are cardinalities of $\{(x,y) \in D^*; a \in \alpha_A(x,y)\}$ and D^* respectively.

Intuitively speaking, the bigger the $\lambda(a)$, the more important the attribute a. We propose the following algorithm.

Algorithm 1

1) Input the generalized decision system $S = (U, A, d)$.
2) Compute the positive region $Pos_A(d)$ of d and $\alpha_A(x,y)$ for every $(x,y) \in D^*$.
3) Place $\alpha_A(x,y)$ in discernibility matrix DM_1.
4) Compute the positive region discernible degree $\lambda(a)$ for each $a \in \bigcup_{\alpha_A(x,y) \in DM_1} \alpha_A(x,y)$, where

$$\lambda(a)=\frac{\left|\{\alpha_A(x,y)\in DM_1; a\in\alpha_A(x,y)\}\right|}{\left|DM_1\right|}.$$

5) Choose a_1 *such that* $\lambda(a_1)=\max_{b\in E}\lambda(b)$ *(If there are more than one attributes with this property, then any one of the attribute may be chosen), delete* $\alpha_A(x,y)$ *which contain a from discernibility matrix* DM_1 *to obtain* DM_2.

6) Go back to step 3 till $DM_{i+1}=\varnothing$. *Then* $T=\{a_1,\cdots,a_i\}$ *is a positive region reduction.*

Example 3.1 We consider the generalized decision system $S=(U,A,d)$, where $U=\{x_1,x_2,x_3,x_4\}$, $A=\{a,b,c\}$, the neighborhoods are given by:

$R_a(x_1)=\{x_1,x_2\}$, $R_a(x_2)=\{x_2,x_3,x_4\}$, $R_a(x_3)=\{x_2,x_3\}$,

$R_a(x_4)=\{x_3,x_4\}$, $R_b(x_1)=\{x_1,x_2,x_4\}$, $R_b(x_2)=\{x_2,x_3\}$,

$R_b(x_3)=\{x_1,x_3,x_4\}$, $R_b(x_4)=\{x_4\}$, $R_c(x_1)=\{x_1,x_2\}$,

$R_c(x_2)=\{x_2,x_3,x_4\}$, $R_c(x_3)=\{x_2,x_3,x_4\}$, $R_c(x_4)=\{x_4\}$.

Furthermore, $U/d=\{D_1,D_2\}$, $D_1=\{x_1,x_2\}$, $D_2=\{x_3,x_4\}$. It follows that $R_A(x_1)=\{x_1,x_2\}$, $R_A(x_2)=\{x_2,x_3\}$, $R_A(x_3)=\{x_3\}$, $R_A(x_4)=\{x_4\}$. We note that R_A is reflexive, but not symmetric and transitive. By routine computation, $Pos_A(d)=\{x_1,x_3,x_4\}$,

$$DM_1=\begin{pmatrix} & x_1 & x_2 & x_3 & x_4\\ x_1 & & & \{a,b,c\} & \{a,c\}\\ x_2 & & & & \\ x_3 & \{a,c\} & \{b\} & & \\ x_4 & \{a,b,c\} & \{a,b,c\} & & \end{pmatrix}.$$

Thus $\lambda(a)=\frac{5}{6}$, $\lambda(b)=\frac{4}{6}$, $\lambda(c)=\frac{5}{6}$. Choose a, then

$$DM_2=\begin{pmatrix} & x_1 & x_2 & x_3 & x_4\\ x_1 & & & & \\ x_2 & & & & \\ x_3 & & \{b\} & & \\ x_4 & & & & \end{pmatrix},$$

and choose b, then $DM_3=\varnothing$. Thus $T=\{a,b\}$ is a positive region reduction.

Note: If we firstly choose c, then we obtain another positive region reduction $T=\{b,c\}$.

IV. Distribution Reductions for Generalized Decision Systems

Kryszkiewicz[3] proposed an rough set approach to incomplete information systems where the indiscernibility relation is a tolerance relation (reflexive and symmetric relation). In this section, we generalized the approach to generalized decision systems.

Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$. We define $d_B:U\to P(V_d)$ as

$$d_B(x)=d(R_B(x))=\{d(y); y\in R_B(x)\}.$$

Namely, $d_B(x)$ is the set of d attribute values of objects in $R_B(x)$. The mapping d_B is called decision function determined by B.

Definition 4.1 Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$. If $d_B=d_A$, then B is called a distribution consistent set of S, and the minimal distribution consistent set of S (with respect to set inclusion relation) is called a distribution reduction of S.

Theorem 4.1[11] Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$, $U/d=\{D_1,\cdots,D_r\}$. Then B is a distribution consistent set if and only if $\overline{R_B}(D_i)=\overline{R_A}(D_i)$ for each $i\le r$.

Theorem 4.2 Let $S=(U,A,d)$ be a generalized decision system, $B\subseteq A$. Then B is a distribution consistent set if and only if $B\cap\alpha_A(x,y)\neq\varnothing$ for any $(x,y)\in D_2^*$, where $D_2^*=\{(x,y); d(y)\notin d_A(x)\}$.

Proof: Necessity: Notice that when $(x,y)\in D_2^*$, we have $d(y)\notin d_A(x)$ and hence $y\notin R_A(x)$, $\alpha_A(x,y)\neq\varnothing$. Let $d_B=d_A$ and $(x,y)\in D_2^*$. By $d(y)\notin d_A(x)$ it follows that $d(y)\notin d_B(x)$. Thus $R_B(x)\cap[y]_d=\varnothing$, and $y\notin R_B(x)$. It follows that there exists $b\in B$ such that $(x,y)\notin R_b$, namely, $B\cap\alpha_A(x,y)\neq\varnothing$.

Sufficiency: For any $x\in U$, we have $d_A(x)\subseteq d_B(x)$. Suppose that u is a decision value of d, $u\notin d_A(x)$ and $d(y)=u$. For any $z\in[y]_d$, it follows that $d(z)=d(y)\notin d_A(x)$, and hence $(x,z)\in D_2^*$. Consequently, we have $B\cap\alpha_A(x,z)\neq\varnothing$, namely, there exists $b\in B$ such that $(x,z)\notin R_b$, and thus $z\notin R_B(x)$. It follows that $R_B(x)\cap[y]_d=\varnothing$, and in consequence $u=d(y)\notin d(R_B(x))=d_B(x)$. Thus $d_B(x)\subseteq d_A(x)$. It follows that $d_B=d_A$ and B is a distribution consistent set of S as required.

Let $S=(U,A,d)$ be a generalized decision system. In what follows, $\Delta^0=\bigwedge_{(x,y)\in D_2^*}\vee\alpha_A(x,y)$ is called the distribution discernibility function of S.

Corollary 4.1 Let $S=(U,A,d)$ be a generalized decision system. If $\Delta^{0*}=(a_{11}\wedge\cdots\wedge a_{1m_1})\vee\cdots\vee(a_{k1}\wedge\cdots\wedge a_{km_k})$ is the reduced disjunctive form of Δ^0, then $\mathrm{Re}d=\{T_1,\cdots,T_k\}$ is the set of all distribution reductions of S, where $T_i=\{a_{i1},\cdots,a_{im_i}\}$ for each $i\le k$.

Theorem 4.2 and Corollary 4.1 show the method of distribution reduction based on generalized indiscernibility relation, which only satisfies reflexivity. Obviously, the methods improve the conclusion of literature. Similarly, We propose the following algorithm to compute distribution reduction.

Algorithm 2

1) Input the generalized decision system $S=(U,A,d)$.
2) Compute D_2^* *and* $\alpha_A(x,y)$ *for every* $(x,y)\in D_2^*$.
3) Place $\alpha_A(x,y)$ *in discernibility matrix* DM_1^*.

4) Compute the distribution discernible degree $\lambda(a)$ *for each* $a \in \bigcup_{\alpha_A(x,y)\in DM_1^*} \alpha_A(x,y)$, *where*

$$\lambda(a) = \frac{\left|\{\alpha_A(x,y) \in DM_1^*; a \in \alpha_A(x,y)\}\right|}{\left|DM_1^*\right|}.$$

5) Choose a_1 *such that* $\lambda(a_1) = \max_{b\in E} \lambda(b)$ *(If there are more than one attributes with this property, then any one of the attribute may be chosen), delete* $\alpha_A(x,y)$ *which contain* a *from discernibility matrix* DM_1^* *to obtain* DM_2^*.

6) Go back to step 3 till $DM_{i+1}^* = \varnothing$. *Then* $T = \{a_1, \cdots, a_i\}$ *is a distribution reduction.*

The following theorem shows the connection between the concepts of distribution reduction and positive region reduction.

Theorem 4.3 Let $S = (U, A, d)$ be a generalized decision system, $B \subseteq A$, $U/d = \{D_1, \cdots, D_r\}$. If B is a distribution consistent set, then B is a positive region consistent set.

Proof: We suppose that B is a distribution consistent set. It follows that $\overline{R_B}(D_i) = \overline{R_A}(D_i)$ for each $i \le r$. Thus

$$\underline{R_B}(D_i) = \sim \overline{R_B}(\sim D_i) = \sim \overline{R_B}(\bigcup_{j\neq i} D_j) = \sim \bigcup_{j\neq i} \overline{R_B}(D_j)$$

$$= \sim \bigcup_{j\neq i} \overline{R_A}(D_j) = \sim \overline{R_A}(\bigcup_{j\neq i} D_j) = \sim \overline{R_A}(\sim D_i) = \underline{R_A}(D_i).$$

Consequently, B is a positive region consistent set.

V. CONCLUSIONS

Rough set under incomplete information has been extensively studied. Researchers have put forward several similarity relations on objects and some attribute reduction approaches for incomplete information systems. This paper is devoted to the study of positive region reduction and distribution reduction based on generalized indiscernibility relation.

The judgment theorems for positive region reduction and distribution reduction of generalized decision systems and attribute reduction approaches are presented. Furthermore, the reduction algorithms based on discernible degree are proposed. Based on this work, we can further probe the rough set model under incomplete information and its application in knowledge discovery.

REFERENCES

[1] Z.Bonikowski, E.Bryniarski, U.Wybraniec, "Extensions and intentions in the rough set theory," Information Sciences, vol. 107, pp. 149-167, 1998.

[2] L.H.Guan, G.Y.Wang, Generalized approximations defined by non-equivalence relations, Information Sciences, vol. 193, pp. 163-179, 2012.

[3] M.Kryszkiewicz, Rough set approach to incomplete information system, Information Sciences, vol. 112, pp. 39-49, 1998.

[4] M.Kryszkiewicz, Properties of incomplete information systems in the framework of rough sets, Rough Sets in Data Mining and Knowledge Discovery, Physica-Verlag, 1998, pp. 422-450.

[5] K.Marzena K, Comparative study of alternative types of knowledge reduction in inconsistent systems, International Journal of Intelligent Systems, vol. 16, pp. 105-120, 2001.

[6] Z.Pawlak, Rough sets, Int. J. Computer and Information Sci., vol. 11, pp. 341-356, 1982.

[7] Z.Pawlak, A.Skowron, Rough sets: Some extensions, Information Sciences, vol. 177, pp. 28-40, 2007.

[8] K.Qin, Z.Pei, J.Yang, Y.Xu, Approximation operators on complete completely distributive lattices, Information Sciences, vol. 247, pp. 123-130, 2013

[9] K.Qin, H.Zhao, Z.Pei, The reduction of decision table based on generalized indiscernibility relation, Journal of Xihua University, vol.32(4), pp.1-4, 2013

[10] A.M.Radzikowska, E.E.Kerre, A comparative study of fuzzy rough sets, Fuzzy Sets and Systems, vol. 126, pp.137-155, 2002

[11] A.Skowron, C.Rauszer, The discernibility matrices and functions in information systems, In: R. Slowinski (Ed.), Intelligent Decision Support-Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, London, pp.331-362, 1992.

[12] J.Stefanowski, A.Tsoukias, Incomplete information tables and rough classification, Computational Intelligence, 17(3)2001, pp. 545-566

[13] G.Y.Wang, Extension of rough set under incomplete information systems, Journal of Computer Research and Development (in Chinese), vol. 39, pp. 1238-1243, 2002.

[14] G.Y.Wang, H.Yu, D.Yang, Decision table reduction based on conditional information entropy, Chinese Journal of Computers(in Chinese), vol. 25, pp. 759-766, 2002.

[15] Y.Y.Yao, Relational interpretation of neighborhood operators and rough set approximation operator, Information Sciences, 111, pp.239-259, 1998.

[16] Y.Y.Yao, Y.Zhao, Discernibility matrix simplification for constructing attribute reducts, Information Sciences, 179, pp.867-882, 2009.

[17] X.Yin, X.Jia, L.Shang, A new extension model of rough sets under incomplete information, Lecture Notes in Artificial Intelligence, 4062, pp. 141-146, 2006.

[18] X.H.Zhang, B.Zhou, P.Li, A general frame for intuitionistic fuzzy rough sets, Information Sciences, 216, pp.34-49, 2012.

[19] X.H.Zhang, J.H.Dai, Y.C.Yu, On the union and intersection operations of rough sets based on various approximation spaces, Information Sciences, 292, pp.214-229, 2015.

For a Better Coordination Between Students Learning Styles and Instructors Teaching Styles

Sylvia Encheva
Stord/Haugesund University College
Bjørnsonsg. 45,
5528 Haugesund,
Norway

***Abstract*—While learning has been in the main focus of a number of educators and researches, instructors' teaching styles have received considerably less attention. When it comes to dependencies between learning styles and teaching styles the available knowledge is even less. There is a definite need for a systematic approach while looking for such dependencies. We propose application of refinement orders and relational concept analysis for pursuing further investigations on the matter.**

***Keywords*—Refinement orders; Relational concept analysis; Learning**

I. Introduction

Learning styles [7] in general refer to how people learn. In [18] they are described as - visual, aural, verbal, physical, logical, social and solitary. Students' learning styles in particular have been wildly discussed and structured in a number of models, [18]. According to Felder and Soloman's model [4] learners can be: active or reflective, depending on their tendencies to retain and understand information; sensing or intuitive, depending on whether they prefer learning facts or discover possibilities and relationships; visual or verbal, depending on their preferences to information been presented visually and verbally; sequential or global, depending on whether they are more comfortable gaining understanding in linear steps or in large jumps. Instructional methods for coping with different learning styles are also included.

The importance of addressing most common learning styles is emphasized in [5]. Students, whose learning styles are compatible with the teaching style of a course instructor tend to retain information longer, apply it more effectively, and have more positive post-course attitudes toward the subject than do their counterparts who experience learning/teaching style mismatches, [5]. It is pointed that students also differ in their preferences to the way presented information is organized: inductive - where facts and observations are given, and underlying principles are inferred, or deductive - where principles are given, consequences and applications are deduced. Strengths and weakness of different learning styles are further discussed and a multi style approach is recommended.

"If professors teach exclusively in a manner that favours their students' less preferred learning style modes, the students' discomfort level may be great enough to interfere with their learning. On the other hand, if professors teach exclusively in their students' preferred modes, the students may not develop the mental dexterity they need to reach their potential for achievement in school and as professionals.", [6].

The Myers-Briggs Type Indicator model [14] is mainly concerned with students' preferences from psychological point of view. They can be extraverts or introverts, sensors or intuitors, thinkers or feelers, judgers or perceivers. This leads to sixteen different learning style types, [14].

The Kolb's model is based on students' preferences for how to take information in and how to internalize information, [15]. The model consists of four types of learners concrete, reflective; abstract, reflective; abstract, active or concrete, active.

The Herrmann Brain Dominance Instrument model is concerned with preferences based on some brain functions, [12]. The four modes are analysis; methods and procedures; teamwork and communications; creative problem solving, systems thinking, synthesis, and design.

Learner-centered and teacher-centered teaching styles are discussed in [1].

Five teaching styles are described in [9] - expert, formal authority, personal model, facilitator, and delegator. Their advantages and disadvantages are also clearly formulated. Four teaching styles are identified in [19] - formal authority, demonstrator, facilitator, and delegator.

Our goal is to find a systematic way for detecting between students learning styles and lecturers teaching styles applying permutographs, [2] and relational concept analysis (RCA), [13], [10].

II. Preliminaries

Let P be a non-empty ordered set. If $sup\{x, y\}$ and $inf\{x, y\}$ exist for all $x, y \in P$, then P is called a *lattice* [3]. In a lattice, illustrating partial ordering of knowledge values, the logical conjunction is identified with the meet operation and the logical disjunction with the join operation, [8].

Definition 1: [2] The permutograph on a set X is the graph, denoted by Σ_X, whose set of vertices is the set $\mathcal{L}_X$ of linear orders on X and whose edges are defined by the following adjacency relation, denoted Adj, between two linear orders: for $L, L' \in \mathcal{L}_X$, $LAdjL'$ if $|L \cap L'^d| = 1$.

For two linear orders L and L' on X representing preferences, $d(L, L') = |L \cap L'^d| = |L \backslash L'|$ is the number of disagreements on preferences between these two orders, [2].

The geodesic distance $\delta(L, L')$ between two linear orders L and L' in a permutograph is the minimum number of commutations to carry out in order to go from one to the other, [2].

In RCA, input data is organized as a pair made of a set of objects-to-attributes contexts $\mathbf{K} = \{\mathcal{K}_i\}_{i=1,...,n}$ and a set of objects-to-objects binary relations $\mathbf{R} = \{r_k\}_{k=1,...,m}$. Here, a relation $r \in R$ links two object sets from two contexts, i.e., there are $i_1, i_2 \in \{1, ..., n\}$ (possibly $i_1 = i_2$) such that $r \subseteq O_{i_1} \times O_{i_2}$. Both contexts from $\mathbf{K}$ and relations from $\mathbf{R}$ are introduced as cross-tables,[11].

Definition 2: [11] (Relational Context Family (RCF)) An RCF is a pair $(\mathbf{K}, \mathbf{R})$ where:

- $\mathbf{K} = \{\mathcal{K}_i\}_{i=1,...,n}$ is a set of contexts $\mathcal{K}_i = (O_i, A_i, I_i)$

and

- $\mathbf{R} = \{r_k\}_{k=1,...,m}$ is a set of relations r_k where

$r_k \subseteq O_{i_1} \times O_{i_2}$ for some $i_1, i_2 \in 1, ..., n$.

Cluster analysis partition data into sets (clusters) sharing common properties, [2]. A frequently used tool in cluster analysis is a dissimilarity function d on a set of objects E, measuring the degree of dissemblance between the elements in E, [2].

III. Selections

Students are first properly introduced to the meaning of learning styles and are afterwards suggested to express their preferences via web based questionnaires.

We consider four groups of students formed according to gender and work experience. Criteria used under refinement order are based on learning styles models as in [4] and [5].

These four groups of students can be placed in seven sets due to application of one of the fore-mentioned criteria. Each of these seven sets has two subsets with 1 + 3 or 2 + 2 items in a subset.

After employing two of those criteria we obtain six sets following the refinement order. These six sets contain three subsets each with one or two items in a subset. All of them are placed in the 3rd row of the lattice in Fig. 1. In any of the six sets there is a couple of indiscernible elements (groups). Splitting these couples requires enforcement of yet another criterion. Set-valued functions developed in RCA are well suited for extracting knowledge from sets of students formed at different time periods.

Whether all criteria are to be applied or just some of them is up to a system modeling team. At the same time lattices as in Fig. 1 obtained from disjoint sets of students can be connected via RCA for extracting additional knowledge. The technical side of such processes is well explained in [11].

Four teaching styles described in [19] are ranked according to students' preferences as in permutograph in Fig. 2. The numbers in 1, 2, 3, 4 in Fig. 2 represent the four commonly understood teaching styles, i. e. formal authority (1), demonstrator (2), facilitator (3), and delegator (4). Teaching styles vary from topic to topic and students feedback can be followed by studying their responses, delivered via web

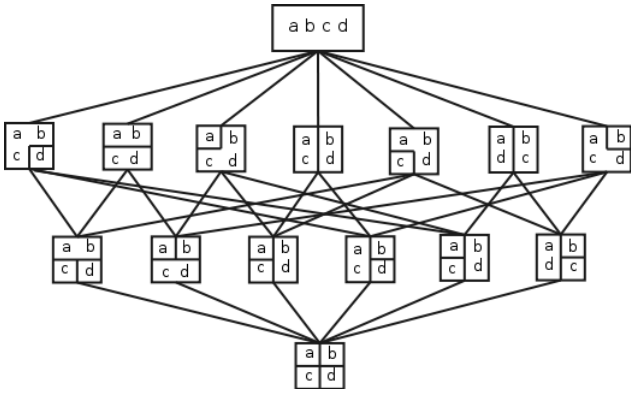


Fig. 1: Lattice of partitions

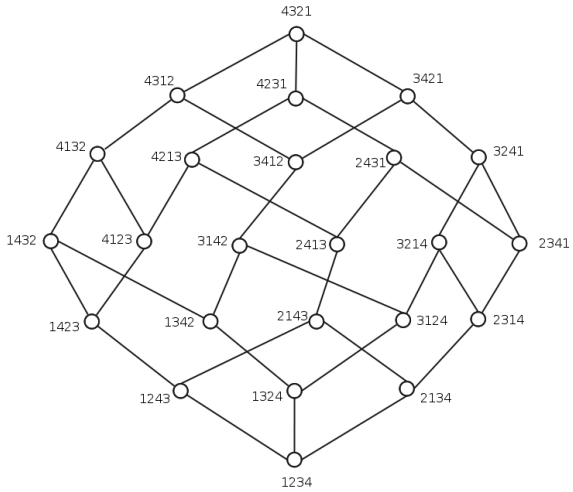


Fig. 2: Permutograph

based questionnaires. All possible orderings are summarized in Fig. 2, where two orderings connected by a strait line differ in positioning of two neighbor elements. This can be used while adjusting current teaching of new topics as well as performing further tuning of teaching the same course in the future.

Distances between vertices are applied while considering which group of users is effecting the order of preferences. Gender and age f. ex. are factors with a significant implication on preference orderings. This is to be incorporated in corresponding recommending processes. If users supply that type of information they would receive recommendations based on data from users with similar initial characteristics. If users do not provide such information they would receive recommendations based on the total collected data.

As an example about students' preferences one can look at the usual dilemma about orders in which problems are delivered by a course instructor: "learning how to apply a skill benefits more from blocked problem orders" while "learning when to apply a skill benefits more from interleaved problem orders", [16].

Another example is related to finding the degree to which students' preferences effect their progress in two consecutive

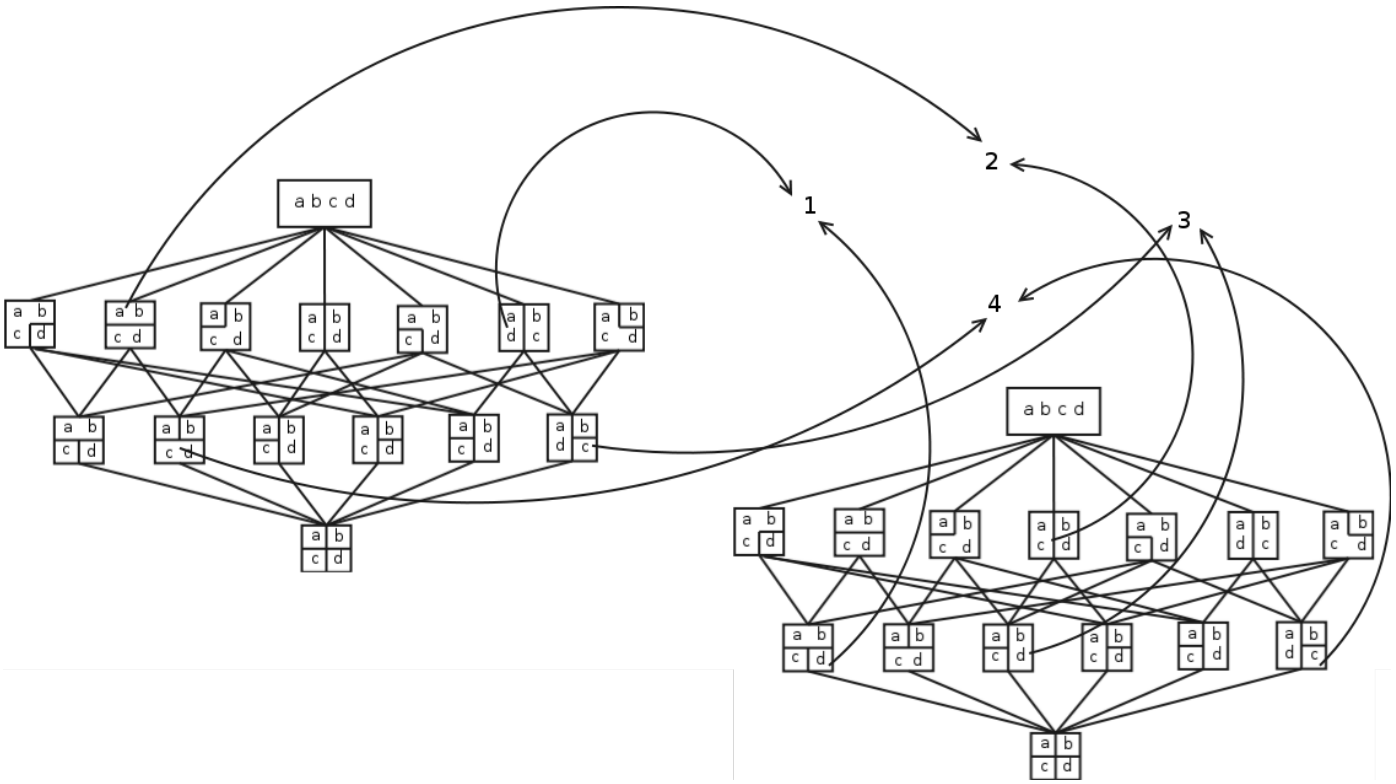


Fig. 3: Learning styles and preferences to teaching styles of two student groups

subjects when one of them is a prerequisite to the other. In a similar fashion one can work with related topics within a subject.

Distances between preferences can also be used to form clusters by joining existing clusters if they are within a predefined geodesic distance from a particular node, or creating new clusters with elements not within the geodesic distance. Once the clusters are formed an analysis of the reasons for their formations is to be performed.

Relational concept analysis is to be further applied for drawing conclusions about teaching different groups of students. Students groups, their learning styles and lectures teaching styles are to be collected in an information table as in [11]. The derived Hasse diagrams show correlations between learning styles and teaching styles, Fig. 3.

IV. Conclusion

There is no doubt about the existence of dependencies between students learning styles and lecturers teaching styles. Additional research has to be carried out in order to come up with meaningful recommendations to future instructors. Once a need for further tuning of a lecturer's teaching is established, additional efforts have to made for finding out what exactly has to be done. Both permutographs and relational concept analysis lend themselves very well to exploring compatibility between learning and teaching styles.

References

[1] K. R. Barrett, B. L. Bower and N. C. Donovan, *Teaching Styles of Community College Instructors*, American Journal of Distance Education, vol. 21(1), pp. 37–49, 2007.

[2] N. Caspard, B. Leclerc, and B. Monjardet, *Finite Ordered Sets Concepts, Results and Uses*, Cambridge University Press, 2012.

[3] B. A. Davey and H. A. Priestley, *Introduction to lattices and order*, Cambridge University Press, Cambridge, 2005.

[4] R. M. Felder and L. K. Silverman, *Learning Styles and Teaching Styles in Engineering Education*, Engineering Education, vol. 78 (7), pp. 674–681, 1988.

[5] R. M. Felder, *Reaching the Second Tier: Learning and Teaching Styles in College Science Education*, Journal of College Science Teaching, vol. 23(5), pp. 286–290, 1993.

[6] R. M. Felder, *Matters of style*, ASEE Prism, vol. 6(4), pp. 18–23, 1996.

[7] R. M. Felder and J. Spurlin, *Applications, reliability, and validity of the index of learning styles*, International Journal of Engineering Education, vol. 21(1), pp. 103–112, 2005.

[8] B. Ganter and R. Wille, *Formal Concept Analysis*, Springer, 1999.

[9] A. Grasha, *Teaching with Style*, Pittsburgh, PA: Alliance Publishers, 1996.

[10] M. R. Hacene, M. Huchard, A. Napoli, and P. Valtchev, *A proposal for combining formal concept analysis and description logics for mining relational data*. In: Kuznetsov, S., Schmidt, S. (eds.) Proc. of the 5th Intl. Conf. on Formal Concept Analysis (ICFCA07). LNCS, vol. 4390, pp. 51–65, 2007.

[11] M. R. Hacene, M. Huchard, A. Napoli and P. Valtchev, *Relational concept analysis: mining concept lattices from multi-relational data*, Ann. Math. Artif. Intell., vol. 67, pp. 81–108, 2013.

[12] N. Herrmann, *The Creative Brain*, Lake Lure, NC, Brain Books, 1990.

[13] M. Huchard, M. R. Hacene, C. Roume, and P. Valtchev, *Relational concept discovery in structured datasets*, Ann. Math. Artif. Intell, vol. 49 (1-4), pp. 39–76, 2007.

[14] M. H. McCaulley, *The MBTI and Individual Pathways in Engineering Design*, Engineering Education, vol. 80, pp. 537–542, 1990.

[15] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development* Englewood Cliffs, NJ, Prentice-Hall, 1984.

[16] N. Li, W. W. Cohen, and K. R. Koedinger, *Problem Order Implications for Learning Transfer*, Lecture Notes in Computer Science, vol. 7315, pp. 185–194, 2012.

[17] T. A. Litzinger, H. L. Sang, J .C. Wise, and R. M. Felder, *A psychometric study of the index of learning styles*, Journal of Engineering Education, vol. 96(4), pp. 309–319, 2007.

[18] M. Lumsdaine and M. Lumsdaine, *Thinking Preferences of Engineering Students: Implications for Curriculum Restructuring*, Journal of Engineering Education, vol. 84(2), pp. 193–204, 1995.

[19] R. Wittmann-Price, M. Godshall, and L. Wilson, *Certified Nurse Educator (CNE) Review Manual*, Springer Publishing Company, 2 edition, 2013.

Fuzzy Soft Sets Supporting Multi-Criteria Decision Processes

Sylvia Encheva
Stord/Haugesund University College
Bjørnsonsg. 45,
5528 Haugesund,
Norway

***Abstract*—Students experience various types of difficulties when it comes to examinations, where some of them are subject related while others are more of a psychological character. A number of factors influencing academic success or failure of undergraduate students are identified in various research studies. One of the many important questions related to that is how to select individuals endangered to be unable to complete a particular study program or a subject. The intention of this work is to develop an approach for early discovery of students who could face serious difficulties through their studies.**

***Keywords*—Soft sets; Uncertainties; Decision making**

I. Introduction

Exam failure is a serious problem for both students and the respective educational institutions where these students are enrolled inn. One of the important questions arising in such cases is related to early identification of students who are potentially in danger of exam failure.

Students experience various types of difficulties when it comes to examinations, where some of them are subject related while others are more of a psychological character. The former are usually more specific while the latter are more general. Some examples of the latter include anxiety, low level of concentration, increased stress level and sleep disorders, [14], [15]. A large number of factors influencing academic success or failure of university students is listed in [5]. Our intention is to identify students who might be in danger of not being able to complete a particular course at a very early stage of their enrollment and consequently provide them with individual recommendations. Since such processes are usually described in uncertain and unprecised ways handling them with methods from fuzzy soft set theory is proposed in this work. In the soft set theory [12], the initial description of the object has an approximate nature, [13]. Very useful group decision making methods based on intuitionistic fuzzy soft matrices are presented in [10]. In this paper one of their approaches is expended in a way that allows obtaining a set of interesting items that differ from the one with the highest score.

The rest of this work goes as follows. Definitions and statements are placed in Section II. The main results are presented in Section III, and a conclusion can be found in Section IV.

II. Soft Sets

Let U be an initial universe set and E_U be the set of all possible parameters under consideration with respect to U. The power set of U (i.e., the set of all subsets of U) is denoted by $P(U)$ and $A \subseteq E$, [1]. A soft set is defined in the following way:

Definition 1: [12] A pair (F, A) is called a soft set over U, where F is a mapping given by

$$F : A \to P(U).$$

Definition 2: [3] Let U be an initial universe, $P(U)$ be the power set of U, E be the set of all parameters and X be a fuzzy set over E. An FP-soft set F_X on the universe U is defined by the set of ordered pairs

$$F_X = (\mu_X(x)/x, f_X(x)) :$$
$$x \in E, f_X(x) \in P(U), \mu_X(x) \in [0,1]\},$$

where the function $f_X : E \to P(U)$ is called approximate function such that $f_X(x) = \emptyset$ if $\mu_X(x) = 0$, and the function $\mu_X : E \to [0,1]$ is called membership function of FP-soft set F_X. The value of $\mu_X(x)$ is the degree of importance of the parameter x, and depends on the decision makers requirements.

Definition 3: [3] Let $F_X \in FPS(U)$, where $FPS(U)$ stands for the sets of all FP-soft sets over U. Then a fuzzy decision set of F_X, denoted by F_X^d, is defined by

$$F_X^d = \mu_{F_X^d}(u)/u : u \in U$$

which is a fuzzy set over U, its membership function $\mu_{F_X^d}$ is defined by $\mu_{F_X^d} : U \to [0,1]$,

$$\mu_{F_X^d}(u) = \frac{1}{|supp(X)|} \sum_{x \in supp(X)} \mu_X(x) \chi_{f_X(x)}(u)$$

where $supp(X)$ is the support set of X, $f_X(x)$ is the crisp subset determined by the parameter x and

$$\chi_{f_X(x)}(u) = \begin{cases} 1, & u \in f_X(x), \\ 0, & u \notin f_X(x). \end{cases}$$

Definition 4: [9] The union of two soft sets (F, A) and (G, B) over a common universe U is the soft set (H, C), where $C = A \cup B$, and $\forall\ e \in C$,

$$H(e) = \begin{cases} F(e), & if \quad e \in A - B, \\ G(e), & if \quad e \in B - A, \\ F(e) \cup G(e), & if \quad e \in A \cap B. \end{cases}$$

It is denoted as $(F, A)\tilde{\cup}(G, B) = (H, C)$.

Definition 5: [9] The intersection of two soft sets (F, A) and (G, B) over a common universe U is the soft set (H, C), where $C = A \cap B$, and $\forall\ e \in C$, $H(e) = F(e)$ or $G(e)$ (as both are same set). It is denoted as $(F, A)\tilde{\cap}(G, B) = (H, C)$.

Definition 6: [9] Let (F, A) and (G, B) be soft sets over a common universe set U. Then

(a) $(F, A) \wedge (G, B)$ is a soft set defined by

$$(F, A) \wedge (G, B) = (H, A \times B),$$

where $H(\alpha, \beta) = F(\alpha) \cap G(\beta)$, $\forall(\alpha, \beta) \in A \times B$, and $\cap$ is the intersection operation of sets.

(b) $(F, A) \vee (G, B)$ is a soft set defined by

$$(F, A) \vee (G, B) = (K, A \times B),$$

where $K(\alpha, \beta) = F(\alpha) \cup G(\beta)$, $\forall(\alpha, \beta) \in A \times B$, and $\cup$ is the union operation of sets.

Soft set relations and functions are well presented in [2]. An intuitionsitic fuzzy soft sets based decision making is discussed in [8].

III. Attribute Selection

Suppose three advisors are forming a committee that has to select attributes indicating potential exam failure. Advisors' opinions are to be taken with weights 0.5, 0.3 and 0.2 respectively, as in [10]. Weight distributions can also be determined by a decision making body that is in charge of that project. It is worth mentioning that in case the three advisors are assumed to have different influence, then there are not that many weight combinations that can actually effect attribute choice. Thus, if the lowest weight is 0.1 then the highest has to be at least 0.5 and if the lowest weight is 0.2 then the highest can be 0.4 (this implies two advisers with equal weight 0.4), 0.5 or 0.6 (this implies two advisers with equal weight 0.2).

In our case the set of attributes to be considered contains the following elements

A 1 - health related issues,

A 2 - last education relevant to this study has been obtained at least five years ago,

A 3 - time consuming obligations outside of the study,

A 4 - preliminary test results,

A 5 - amount of time a student can devote to study that subject weekly,

A 6 - student absences from classes, tutorials, etc.,

A 7 - insufficient preliminary knowledge,

TABLE I: Attributes significance

	O 1	O 2	O 3
A 1	(0.83, 0.1)	(0.6, 0.2)	(0.6, 0.1)
A 2	(0.3, 0.51)	(0.58, 0.4)	(0.8, 0.1)
A 3	$(0.6, 0.18)$	$(0.71, 0.24)$	(0.31, 0.5)
A 4	(0.9, 0.05)	(0.8, 0.13)	(0.4, 0.52)
A 5	(0.55, 0.3)	(0.9, 0.01)	(0.5, 0.36)
A 6	(0.47, 0.21)	(0.66, 0.3)	(0.7, 0.22)
A 7	(0.8, 0.08)	(0.8, 0.15)	(0.83, 0.1)
A 8	(0.58, 0.12)	(0.3, 0.64)	$(0.69, 0.3)$
λ_{med} (E)	(0.6, 0.18)	(0.71, 0.24)	(0.69, 0.3)

TABLE II: Values for all attributes

	O 1	O 2	O 3	Attributes values
A 1	1	0	0	1
A 2	0	0	1	1
A 3	1	1	0	2
A 4	1	1	0	2
A 5	0	1	0	1
A 6	0	0	1	1
A 7	1	1	1	3
A 8	0	0	1	1

A 8 - opportunities to work together with other students.

Each attribute is rated applying values from the set $0.1, ..., 1.0$, where 1.0 is the most important. Notations in Table I are as follows: O 1, O 2, O 3 represent opinions of first, second and third advisor with respect to attributes A 1, A 2, ..., A 8. A number in the first position of each couple describes a degree to which that attribute is important and the second number describes a degree to which that attribute is not important. A threshold vector λ_{med} (E) is based on median. Values for threshold vectors in Table I are emphasized.

The paper continues following mainly the work presented in [3]. Due to the specific nature of this investigation it is needed to tune a bit their approach. Thus instead of applying predefined degrees of attributes' importance values from Table II are taken. In [4], [8], and [10] they were referred to as choice values. For our study this seems more reasonable. Otherwise it will be necessary to ask every student to supply such degrees of importance. Most students would find such requests difficult and very few would be able to provide meaningful responses.

Next important difference is that the goal here is to identify all students in danger to fail their exam while following [3] one would find one student who seems to have more problems than the rest of his classmates. To achieve this the fuzzy decision set F_X^d is calculated and all students within the last quartile are selected.

To avoid confusions the terms 'classical set' and 'fuzzy soft set' are used in every single case without assuming that one or the other is understood by convention.

The classical set of proposed attributes is $\{A1, A2, ..., A8\}$ as described above. Students who answered a Web based inquiry are denoted by $\{St1, ..., St20\}$. Their responses are assumed to be binary in this case, see Table III. Nonbinary scale can be used if a finer grading is found to be more beneficial. Once again, the idea is to keep it simple. Students should not be overload with too many questions and too many

TABLE III: Responses from students

	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 8	F^d_{St}
St 1	×	×		×	×			×	0.75
St 2	×		×		×	×		×	0.75
St 3		×	×		×	×	×		1
St 4	×			×		×	×	×	1.25
St 5		×	×		×			×	0.625
St 6		×	×	×		×	×		1.25
St 7	×	×		×	×	×		×	1
St 8	×		×	×	×		×		1.25
St 9		×	×		×			×	0.625
St 10	×	×	×		×	×		×	0.875
St 11		×			×	×	×	×	0.875
St 12	×		×	×		×			0.75
St 13	×	×	×		×			×	0.75
St 14		×	×	×	×	×			0.875
St 15	×	×					×	×	0.75
St 16	×	×		×	×		×		1
St 17	×			×		×			0.5
St 18		×	×		×	×		×	0.875

options to choose from.

Members of the fuzzy decision set F^d_{St} are shown in Table III. After working on 4-quantiles of F^d_{St} we believe that students belonging to the forth quartile should be the ones to begin with. In other words, students $St3, St4, St6, St7, St8, St16$ should receive personal advises on what ought to be done in order to avoid exam failure.

Initially experience from previous courses is used. The intention is to tune the system after some time when new data has been collected. When it comes to handling situations requiring aggregation methods the approach in [10] is suggested. In case different datasets are to be used for drawing conclusions, applying statements presented in [9] seems to be quite approporiate.

A. Discussion

Another method that can be used involves formal concept analysis, [6], [7]. This is a method supporting data analysis among many other things. Once the sets of attributes and objects, and their relations are well presented in an information table, a corresponding concept lattice can be depicted. Each node in that lattice contains all the students that share the same attributes. A fuzzy function indicating degrees to which each node contents reflect danger in exam failure has to be build.

It seems that fuzzy soft sets are well equipped to handle problems presented in this work because every student is treated individually. The outcomes of formal concept analysis studies are beneficial for group of students and as a result some details concerning individuals might be omitted. Formal concept analysis based methods can be very helpful while dealing with new students and/or new advisors.

IV. Conclusion

Exam failure is most of the time a result of internal and external factors. Among the internal ones are lack of commitment and motivation, fear of exams, personal or financial problems, etc. The external ones are related to overloaded study programmes, supervision quality, inadequate requirements and so on. To determine which factors are of the highest importance one should study particular educational institutions and students groups. More research has to be done in order to determine the correct value of a proper q-quantile, as well.

References

[1] M. I. Ali, F. Feng, X. Liu, W. K. Min, and M. Shabir, *On some new operations in soft set theory*, Computers and Mathematics with Applications, vol. 57, pp. 1547–1553, 2009.

[2] K. V. Babitha and J.J. Sunil, *Soft set relations and functions*, Computers and Mathematics with Applications, vol. 60, pp. 1840–1849, 2010.

[3] N. Çağman N., F. Çitak, and S. Enginoglu, *FP-soft set theory and its applications*, Annals of Fuzzy Mathematics and Informatics, vol. 2(2), pp. 219–226, 2011.

[4] F. Feng F. and Y. B. Jun, *An adjustable approach to fuzzy soft set based decision making*, Journal of Computer Applied Mathematics, vol. 234, pp. 10–20, 2010.

[5] W. J. Fraser and R. Killen, *Factors influencing academic success or failure of first-year and senior university students: do education students and lecturers perceive things differently?*, South African Journal of Education, vol. 23(4), pp. 254 – 260, 2003.

[6] B. Ganter and R. Wille, *Formal Concept Analysis*, Springer, 1999.

[7] B. Ganter, G. Stumme, and R. Wille, *Formal Concept Analysis: Foundations and Applications*, Lecture Notes in Artificial Intelligence, vol. 3626, Springer-Verlag, 2005.

[8] Y. Jiang, Y. Tang, and Q. Chen, *An adjustable approach to intuitionsitic fuzzy soft sets based decision making*, Applied Mathematical Modeling, vol. 35, pp. 824–836, 2011.

[9] P. K. Maji, R. Biswas, and A. R. Roy, *Soft set theory*, Computers and Mathematics with Applications, vol. 45, no. 4–5, pp. 555–562, 2003.

[10] J. Mao J., D. Yao D., and C. Wang, *Group decision making methods based on intuitionistic fuzzy soft matrices*, Applied Mathematical Modelling, vol. 37, pp. 6425–6436, 2013.

[11] N. Moha, J. Rezgui, Y-G. Gueheneuc, P. Valtchev, G. Boussaidi, *Using FCA to Suggest Refactorings to Correct Design Defects*, Concept Lattices and Their Applications, Lecture Notes in Computer Science, vol. 4923, pp. 269–275, 2008.

[12] D. Molodtsov, *Soft set theoryfirst results*, Computers and Mathematics with Applications, vol. 37, no. 4–5, pp. 19–31, 1999.

[13] M. M. Mushrif, S. Sengupta, and A. K. Ray, *Texture Classification Using a Novel, Soft-Set Theory Based Classification Algorithm*, ACCV 2006, Lecture Notes in Computer Science, vol. 3851, Springer-Verlag Berlin Heidelberg, pp. 246–254, 2006.

[14] S. S. Sazhin, *Teaching Mathematic to Engineering Students*, International Journal Engineering Education, vol. 14 (2), pp. 145–152, 1997.

[15] P. Vitasan, T. Herawanb, M. N. A. Wahabc, A. Othmana, and S. K. Sinnaduraic, *Exploring mathematics anxiety among engineering students*, Procedia - Social and Behavioral Sciences, vol. 8, pp. 482–489, 2010.