School of Computer Science
Universiti Sains Malaysia
Pulau Pinang

Data Mining in Network Traffic using Fuzzy Clustering

by

Shamsul Bin Mohamad

2003/2004 Academic Session

Submitted as partial fulfillment towards graduation requirements for
Masters of Science in Computer Science, Universiti Sains Malaysia

"This report is prepared as a partial fulfillment towards graduation requirements for Masters of Science in Computer Science, Universiti Sains Malaysia, Penang. This report and all the products of the project (source codes, system/application, user manual etc.) are the copyright of Universiti Sains Malaysia, Penang. No part of this report and project shall be reproduced, stored in a retrieval system, or transmitted by any means, electronic, mechanical, photocopying, recording, or otherwise, without permission from Universiti Sains Malaysia, Penang."

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deepest and sincere gratitude to my project supervisor, Dr. Rahmat Budiarto for his constant source of inspiration, encouragement and guidance.

I also wish to record my sincere appreciation to the staff members of School of Computer Science.

Finally, I would like to thank my beloved wife, daughter, parents and friends for giving the moral supports, guidance, advice and blessing.

ABSTRACT

Nowadays, in network traffic, we have various application such as HTTP, Telnet, SMTP, FTP and NetBIOS. These various application make it difficult for the network administrator to model certain network traffic. The network traffic model is very important to know whether that particular network traffic is normal or abnormal. In this project, I have developed a program to capture and filter the packets based on the application. The fuzzy clustering process are made using three algorithms : Fuzzy C-Means (FCM), Gustafsof-Kessel (GK) and Gath-Geva (GG) algorithm. The production of clustering are used to build rules.

TABLE OF CONTENTS

ABSTRACT

CHAPTER 1 INTRODUCTION

1.1 Background 1

1.2 Objectives 2

1.3 Scope of Project 3

CHAPTER 2 RELATED WORK

2.1 Introduction to Data Mining 4

2.2 Fuzzy Clustering 6

2.3 Takagi-Sugeno Model 15

2.4 Rules Extraction from Clusters 16

2.5 Building Fuzzy Models 17

CHAPTER 3 METHODOLGY

3.1 Capturing and Filtering the Packet 21

3.2 Clustering the Data 24

3.3 Analyzing the Data 28

CHAPTER 4 IMPLEMENTATION & RESULT

4.1 Capturing and Filtering the Data 29

4.2 Clustering the Data 31

4.3 Output 33

4.4 Analysis 39

CHAPTER 5 SUMMARY & FUTURE WORK **49**

BIBILIOGRAPHY **50**

DIAGRAM LIST

Figure		Page
2.1	Fuzzy Clusters	9
2.2	Identification by Fuzzy Clustering	19
3.1	PCs with a different MAC connected to a hub	21
4.1	Screen for capturing and Filtering the data	29
4.2	Screen for Clustering and Analyzing the Data	32
4.3	Clustering using Fuzzy C-Means Algorithm, 4 clusters and gaussian membership function type	33
4.4	Function that has been built after the FCM clustering process	34
4.5	Clustering using Fuzzy C-Means Algorithm, 4 clusters and gaussian membership function type	34
4.6	Clustering using GK Algorithm, 4 clusters and gaussian membership function type	35
4.7	Function that has been built after the GK clustering process	36
4.8	Function that has been built after the GG clustering process	36
4.9	Clustering using GG Algorithm, 4 clusters and gaussian membership function type	37
4.10	Clustering using GG Algorithm, 4 clusters and triangle membership function type	38
4.11	Clustering using GG Algorithm, 4 clusters and sigmoid membership function type	38
4.12	Clustering using Fuzzy C-Means Algorithm	42
4.13	Clustering using Gustaffson Kessel Clustering Algorithm	44
4.14	Clustering using Gath-Geva Clustering Algorithm	46

CHAPTER 1

INTRODUCTION

1.1 Background

Nowadays, Information Technology (IT) has widely expanded due to the growth of communication until it is known as Information Communication Technology (ICT). Starting from the simple network traffic for example PSTN that is used in telephone linkage until now, we have various new applications such as HTTP, Telnet, SMTP, DNS, and FTP. There are also high-speed network linkage like frame relay, ATM and high-speed LAN. Network traffic has become more complex. A complex traffic flows in ways that can be applied in practice. When a network performs poorly, its users often complain to the folks running it, demanding for improvements. To improve the performance, the operators must first determine exactly what is going on. To find out what is really happening, the operators must make some measurement.

Fuzzy clustering is one of the fuzzy logic branches. Fuzzy clustering algorithm partition the data set into overlapping groups such as that the clusters describe an underlying structure within the data. In order to obtain a good performance from a fuzzy clustering algorithm, a number of issues must be considered. These concern the shape and the volume of the clusters, the initialization of the clustering algorithm, the distribution of the data patterns and the number of clusters in the data. Fuzzy clustering has been used in various fields such as group-positioning [1], medical [2], financial [3], management [4] and network traffic [5].

1.2 Objectives

Network technologies have outpaced our abilities to effectively manage and engineer them. A major challenge for the networks is to develop models that can realistically capture the behavior of the network traffic. The performance of the networks depends crucially on the traffic assessment. Traffic models play a significant role in the analysis and characterization of network traffic and network performance. Accurate models enhance our understanding of the complicated network characteristics and behaviors by allowing us to study the effect of various model parameters on the network performance through simulation. Good traffic engineering methods relay on parameterized traffic models where the parameters are estimated from network measurements. Such models may serve as the basis for the development of methods and tools for quality assessment, providing more efficient control and management of information flow in the Internet.

Network applications such as FTP, WWW, mirroring etc. are presently operated with little or no knowledge about the characteristics of the underlying network. These applications could operate more efficiently if the characteristics of the network are known and/or are made available to the concerned application. For example, a network client may choose from amongst a host of servers serving the same document, depending on the estimated cost or speed of retrieval at that point of time. On the other hand a periodic activity like mirroring may be scheduled for that hour of the day, when the network is likely to be least loaded.

In this project I have developed a program that can capture and filter the packets based on the application. Next, the data gathered are clustered using Fuzzy C–Means (FCM), Gustafson Kessel (GK) Clustering and Gath-Geva Clustering (GG). The existing cluster are analyzed to produce a fuzzy model of network traffic pattern and rules that is based on the different application such as HTTP, NetBios, Telnet, SMTP, FTP and etc. This pattern is used to show the usage of the application based on time whether it is high, moderate or low.

1.3 Scope of Project

i. Developed a program to capture and filter the packets based on the application.

ii. The data are taken from a small network that involved only a few personal computers and a hub.

iii. The data used is not a real time data. Meaning that the collection of data must be done first before being process.

iv. Developed a program that is based on FCM, GK and GG to analyze the data that has been collected.

v. Produced rules to represent network traffic pattern using the three algorithms mentioned before.

CHAPTER 2

RELATED WORK

Basically, this chapter is divided into a few parts. The first part discussed on the data mining. This is followed by the introduction of fuzzy clustering, whereby the concentration is on the use of fuzzy clustering in network traffic. The last part will discussed on how the relationship between clustering algorithm with the production of fuzzy model and also fuzzy rules.

2.1 Introduction to Data Mining

In recent years we have seen increasing volumes of collected data of all sorts. With so much data available, it is necessary to develop algorithms, which can extract meaningful information from the vast stores. Searching for useful nuggets of information among huge amounts of data has become known as the field of data mining.

Data mining can be applied to relational, transaction, and spatial databases, as well as large stores of unstructured data such as World Wide Web. There are many data mining systems in use today, and applications include the U.S. Treasury detecting money laundering, National Basketball Association coaches detecting trends and patterns of individual players and teams, and categorizing patterns of children in the foster care system.

Data mining is the application of specific algorithms for extracting knowledge from data (Fayyad et. al., 1996). Typical kinds of knowledge extracted include association rules,

characteristic rules, classification rules, discriminate rules, clustering, and surveyed data mining techniques developed in several research communities according to the kinds of knowledge to be mined.

2.1.1 Data Mining Approaches

Data mining, like clustering, is an exploratory activity; so clustering methods are well suited for data mining. Clustering is often an important initial step of several in the data mining process. Some of the data mining approaches, which use clustering, are database segmentation, predictive modeling, and visualization of large database.

Segmentation, clustering methods are used in data mining to segment database into homogenous groups. This can serve purposes of data compression (working with the clusters rather than individual items), or to identify characteristics of subpopulations, which can be targeted for specific purposes (e.g., marketing aimed at senior citizens).

Predictive modeling, statistical methods of data analysis usually involve hypothesis testing of a model the analyst already has in mind. Data mining can aid the user in discovering potential hypothesis prior to using statistical tools. Predictive modeling uses clustering to group items and then infers rules to characterize the groups and suggest models. For example, magazine subscribers can be clustered based on a number of factors (age, sex, income, etc.) then the resulting groups characterized in an attempt to find a model which will distinguish those subscribes that will renew their subscriptions from those that will not.

Clusters in large databases can be used for visualization, in order to aid human analysts in identifying groups and subgroups that have similar characteristics. WinViz is a

data mining visualization tool in which derived clusters can be exported as new attributes, which can then be characterized by the system.

Nowadays, there is a lot of software that can be used to implement fuzzy clustering. For example, Matlab and DataEngine. In Matlab there are a few functions that are related with fuzzy clustering such as FCM. FCM function needs for a few parameters such as the number of cluster, maximum iteration, minimum improvement and also weighting exponent. Suppose we do not have a clear idea on how many clusters there should be for a given set of data, we can use subtractive clustering. The cluster estimates obtained from the subclust function can be used to initialize iterative optimization-based clustering methods (fcm) and model identification methods (like anfis).

DataEngine is the software tool for intelligent data analysis, which unites statistical methods with neural networks and fuzzy technologies. DataEngine can be used to determine and describe the characteristics of the customer groups within the data or to develop a classifier that is able to classify each customer according to these customer groups. DataEngine performs both simultaneously: the clustering process defining the classifier. DataEngine also specifies the correlation coefficient, r, as the measure of dependency between two selected features to examine the interdependencies between two different features. DataEngine also can be used to normalize the data, check for validity measure and plotting a graph.

2.2 Fuzzy Clustering

Clustering is a division of data into groups of similar objects. Each group, called cluster, consists of objects that are similar between themselves and dissimilar to objects of other groups. Representing data by fewer clusters necessarily loses certain fine details, but

achieves simplification. It represents many data objects by few clusters, and hence, it models data by its clusters. Data modeling puts clustering in a historical perspective rooted in mathematics, statistics, and numerical analysis. From a machine learning perspective clusters correspond to hidden patterns, the search for clusters is unsupervised learning, and the resulting system represents a data concept. Therefore, clustering is unsupervised learning of a hidden data concept.

There is a close relationship between clustering techniques and many other disciplines. Clustering has always been used in statistics and science. Typical applications include speech and character recognition. Machine learning clustering algorithms were applied to image segmentation and computer vision. Clustering can be viewed as a density estimation problem. This is the subject of traditional multivariate statistical estimation. Clustering is also widely used for data compression in image processing, which is also known as vector quantization.

Clustering algorithms, in general, are divided into two categories:

i. Hierarchical Methods (agglomerative algorithms, divisive algorithms)

ii. Partitioning Methods (probabilistic clustering, *k*-medoids methods, *k*-means methods)

Hierarchical clustering builds a cluster hierarchy. Every cluster node contains child clusters; sibling clusters partition the points covered by their common parent. Such an approach allows exploring data on different levels of granularity. Hierarchical clustering methods are categorized into agglomerative (bottom-up) and divisive (top-down). An agglomerative clustering starts with one-point (singleton) clusters and recursively merges two or more most appropriate clusters. A divisive clustering starts with one cluster of all

data points and recursively splits the most appropriate cluster. The process continues until a stopping criterion (frequently, the requested number k of clusters) is achieved.

Data partitioning algorithms divide data into several subsets. Because checking all possible subset possibilities may be computationally very consumptive, certain heuristics are used in the form of iterative optimization. Unlike hierarchical methods, in which clusters are not revisited after being constructed, relocation algorithms gradually improve clusters.

Clustering of numerical data forms the basis of many classification and system modeling algorithms. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior.

In the literature, many clustering algorithms have been introduced. There are two methods of clustering: hard clustering and fuzzy clustering. However, fuzzy clustering methods allow the objects to belong to a several clusters simultaneously. Objects on the boundaries between several classes are not forced to fully belong to one class, but rather are assigned membership degrees between 0 and 1 indicating their membership.

2.2.1 Fuzzy Clustering Example

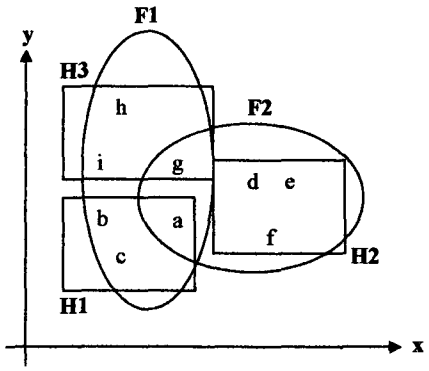

Figure 2.1: Fuzzy Clusters

In fuzzy clustering, each cluster is a fuzzy set of all the patterns. Figure 2.1 illustrates these ideas. The rectangles enclose three "hard" clusters in the data:

$$H_1 = \{a, b, c\}$$
$$H_2 = \{d, e, f\} \text{and}$$
$$H_3 = \{g, h, i\}$$

A fuzzy clustering algorithm might produce the two fuzzy clusters F_1 and F_2 depicted by ellipses. The patterns will have membership values in [0,1] for each cluster. For example, fuzzy cluster F_1 could be compactly described as

F_1={(a,0.7),(b,1.0),(c,0.9),(d,0.5),(e,0.0),(f,0.0),(g,0.2),(h,0.8),(i,0.9)}

and F_2 could be described as

F_2={(a,0.5),(b,0.0),(c,0.0),(d,0.5),(e,0.7),(f,0.8),(g,0.4),(h,0.0),(i,0.0)}

The ordered pairs (i, μ_i) in each cluster represent the i^{th} pattern and its membership value to the cluster μ_i. Larger membership values indicate higher confidence in the assignment of

the pattern to the cluster. A hard clustering can be obtained from a fuzzy partition by threshold the membership value.

2.2.2 Fuzzy Clustering Algorithms

In classical cluster analysis each data must be assigned to exactly one cluster. Fuzzy cluster analysis relaxes this requirement by allowing gradual memberships, thus offering the opportunity to deal with data that belong to more than one cluster at the same time. Most fuzzy clustering algorithms are objective function based. They determine an optimal classification by minimizing an objective function. In objective function based clustering usually each cluster is represented by a cluster prototype. This prototype consists of a cluster centre and maybe some additional information about the size and the shape of the cluster. The size and shape parameters determine the extension of the cluster in different directions of the underlying domain.

The degrees of membership to which a given data point belongs to the different clusters are computed from the distances of the data point to the cluster centers with regard to the size and the shape of the cluster as stated by the additional prototype information. The closer a data point lies to the centre of a cluster, the higher is its degree of membership to this cluster. Hence the problem to divide a dataset into c clusters can be stated as the task to minimize the distances of the data points to the cluster centers, since, of course, we want to maximize the degrees of membership. Most analytical fuzzy clustering algorithms are based on optimization of the basic c-means objective function, or some modification of it.

Hard clustering assigns each data point to one and only one cluster. The degree of membership for a point belonging to a cluster is 1 and often, real life data is not so easily partitioned. The boundaries between the clusters may be fuzzy.

Different from Hard Clustering, Fuzzy Clustering is a data point that belongs to multiple clusters with different degrees of membership. The degree of membership = "Belongingness" of a point to a cluster. The sum of the degrees of memberships adds up to 1. It can adapt to noisy data and classes that are not well separated.

Gustaffson-Kessel algorithm (GK) is an extension of FCM (Gustaffson and Kessel, 1979). Different distributions and sizes of clusters lead to sub-optimal results with FCM. GK used the cluster covariance matrix to capture the ellipsoidal properties of the clusters. GK extended the distance measure norm to include a positive definite matrix Mi. Mi is calculated using the cluster covariance matrix, which is adapted according to the actual shape of the individual clusters.

Gath-Geva Algorithm (Gath and Geva, 1989) combines Fuzzy C-Means with an adaptive distance measure known as Fuzzy Maximum Likelihood Estimation (FMLE). It gives optimal partition and accounts for:

- Variability of cluster shapes
- Variability of cluster densities
- Variability of number of data points in each cluster

Initial cluster centroids are generated through unsupervised learning. The optimal numbers of clusters are determined using cluster validity measures.

2.2.3 Fuzzy Clustering and Network Traffic

There are many researches made on network traffic and fuzzy clustering. Most of them focused on Intrusion Detection. Lee et al., [6] performed experiments on *sendmail* system call data and network *tcpdump* data. They used RIPPER [7] to generate classifiers for these datasets.

It is proven that one can construct concise and accurate classifiers to detect anomalies. An overview made on two general data mining algorithms has been implemented: the association rules algorithm and the frequent episodes algorithm. These algorithms can be used to compute the intra- and inter- audit record patterns, which are essential in describing program or user behavior. The discovered patterns can guide the audit data gathering process and facilitate feature selection. To meet the challenges of both efficient learning (mining) and real-time detection, an agent-based architecture for intrusion detection systems where the learning agents continuously compute and provide the updated (detection) models to the detection agents has been proposed.

In another paper Lee et al., [8] described how to use association rules and frequent episode algorithms to guide the process of audit data gathering and selection of useful features to build the classifiers.

Dokas et al., [9] has developed classification algorithms for intrusion detection. These algorithms are designed especially for learning from datasets in which the class of interest (i.e. the intrusion class) is significantly smaller than the class representing normal behavior. In this body of work, the authors discuss various outlier detection schemes for detecting network intrusions.

Dickerson et al., [10] developed the Fuzzy Intrusion Recognition Engine (FIRE) using fuzzy sets and fuzzy rules. FIRE uses the *Fuzzy C-Means Algorithm* developed by Bezdek to generate fuzzy sets for every observed feature. The fuzzy sets are then used to define fuzzy rules to detect individual attacks. FIRE does not establish any sort of model representing the quiescent state of the system, but instead relies on attack specific rules for detection.

FIRE is a network intrusion detection system that uses fuzzy systems to access malicious activity against computer networks. The system uses an agent-based approach to separate monitoring tasks. The individual agents perform their own fuzzification of input data sources. All agents communicate with a fuzzy evaluation engine that combines the results to individual agents using fuzzy rules to produce alerts that are true to a degree. The fuzzy systems are then tested using the data obtained from networks under simulated attacks. The results show that fuzzy systems can easily identify port scanning and denial of services attacks. The system can be effective at detecting some types of backdoor and Trojan horse attacks.

Hiren Shah et al. [11], has made a research on Fuzzy Clustering for Intrusion Detection. They have explored how fuzzy data mining and concepts introduced by the semantic web can operate in synergy to perform *Distributed Intrusion Detection*. The underlying premise of the intrusion detection model is to describe attacks as instances of an ontology using a semantically rich language, reason over them and subsequently classifies them as instances of an attack of a specific type. However, before an abnormality can be specified as an instance of the ontology, it first needs to be detected. Hence, in their intrusion detection model there are two phases, where the first phase uses data mining techniques to analyze low level data streams that capture process, system and network states and to detect anomalous behavior. The second phase reasons over instances of anomalous behavior specified according to their ontology.

During their experiment, they have employed a four-step approach to construct and test their model of outlier detection. First, they exercised the target system in an attack free environment, collecting data at the process, system and network levels. Then they used Principal Component Analysis (PCA) to reduce the dimensionality of the collected data.

Fuzzy clustering is then performed on the data to obtain clusters that model the quiescent state of the system. Finally, the system is placed under attack and data containing anomalous behavior is collected and compared to the clusters in order to test the efficacy of the initial phase of their intrusion detection model.

Marchette, D. [12], used two clustering methods: k-means and Approximate Distance Clustering (ADC) to described and applied to network data. These allow the clustering of machines into "activity groups", which consists of machines, which tend to have similar activity profiles. In addition, these methods allow the user to determine whether current activity matches these profiles, and hence to determine when there is "abnormal" activity on the network. A method for visualizing the clusters is described, and the approaches are applied to a data set consisting of a months worth of data from 993 machines.

From the research made to the use of fuzzy clustering in network traffic, basically we can say that fuzzy has been widely used to recognize network traffic whether there are normal or abnormal. The most important aspect is to build rules to recognize this network traffic. The explanation on Takagi-Sugeno model will be discussed in the next part of the report whereby the relationship between fuzzy clustering algorithm and Takagi-Sugeno model are discussed and how the rules are generated.

2.3 Takagi-Sugeno Model

The Takagi–Sugeno (TS) fuzzy model [13], uses crisp functions in the consequents. Hence, it can be seen as a combination of linguistic and mathematical regression modeling in the sense that the antecedents describe fuzzy regions in the input space in which consequent functions are valid. The TS rules have the following form:

R_i : If x is A_i then $y_i = f_i(x)$, i = 1,2, ... , K

Contrary to the linguistic model, the input x is a crisp variable (linguistic inputs are in principle possible, but would require the use of the extension principle to compute the fuzzy value of y_i). The functions f_i are typically of the same structure, only the parameters in each rule are different. Generally, f_i is a vector-valued function, but for the ease of notation it considers a scalar fi in the sequel. A simple and practically useful parameterization is the affine (linear in parameters) form, yielding the rules:

R_i: If x is A_i then $y_i = a_i^T x + b_i$, i = 1,2,...,K,

where a_i is a parameter vector and b_i is a scalar offset. This model is called an *affine TS model*.

2.4 Rules Extraction from Clusters

The task of fuzzy model construction is to determine both the nonlinear parameters of the membership functions and the linear parameters of the local models [14]. In general, there are two ways to obtain this information. Possibly, human experts are able to formulate their process knowledge in fuzzy rules. Unfortunately, this usually delivers only a rough idea of the plant behavior, as humans cannot sense all the details and might not be able to quantitatively express the observations. Therefore, numerous approaches have been

proposed which compute nonlinear dynamic fuzzy models from input–output measurement data [15].

- *Grid partitioning:* The number of input MSFs per input are typically chosen by prior knowledge. This approach severely suffers from the curse of dimensionality. To weaken its sensitivity to the input space dimensionality, the grid can be reduced to the regions where enough data is available or a multi-resolution grid can be used [16]. All grid-based approaches are restricted to a very low-dimensional problems and do not exploit the local complexity of the process.
- *Input space clustering:* The validity functions are placed according to the input data distribution [17]. Since the local process complexity (nonlinearity) is ignored this simple approach usually does not perform well.
- *Nonlinear local optimization:* Originally, the input MSFs and the rule consequent parameters have been optimized simultaneously. The current state-of-the-art method, however, is to optimize the rule premise parameters by nonlinear local optimization and the rule consequent parameters by global least squares in a nested or staggered approach as in ANFIS (adaptive neuro-fuzzy inference system) [18]. This approach is computationally expensive but typically yields very accurate results. However, a large number of parameters is optimized and over fitting often becomes a serious problem.
- *Genetic algorithms:* In order to circumvent the difficulties connected to the OLS algorithm, genetic algorithms can be applied for structure search [19]. Evolutionary algorithms offer a wide spectrum of different approaches. All of them, however, suffer from relatively slow convergence.

- *Product space clustering:* One of the most popular approaches applies the Gustafson-Kessel clustering algorithm to find hyper planes in the product space. It is (initially) assumed that the rule premise and consequent spaces are equivalent (**x** = **z**) and hyper planes are sought in the space spanned by $[x_1 x_2 ... x_{nx} y]$[2]

2.5 Building Fuzzy Models

Two common sources of information for building fuzzy models are the prior knowledge and data (process measurements). The prior knowledge can be of a rather approximate nature (qualitative knowledge, heuristics), which usually originates from "experts", i.e., process designers, operators, etc. In this sense, fuzzy models can be regarded as simple *fuzzy expert systems* [20].

For many processes, data are available as records of the process operation or special identification experiments can be designed to obtain the relevant data. Building fuzzy models from data involves methods based on fuzzy logic and approximate reasoning, but also ideas originating from the field of neural networks, data analysis and conventional systems identification. The acquisition or tuning of fuzzy models by means of data is usually termed *fuzzy identification.*

Two main approaches to the integration of knowledge and data in a fuzzy model can be distinguished:

1. The expert knowledge expressed in a verbal form is translated into a collection of if–then rules. In this way, a certain model structure is created. Parameters in this structure (membership functions, consequent singletons or parameters) can be fine-tuned using input output data. The particular tuning

BIBLIOGRAPHY

[1] T H Hsu, *An Application of Fuzzy Clustering in Group Positioning Analysis*. Proc. National Science Council. Vol 10, No 2, pp 157-167, Taiwan, 2000.

[2] G. Berks, D. Graf v. Keyslingk, J. Jantzen, M Dotoli, H. Axer, *Fuzzy Clustering – A Versatile Mean to Explore Medical Database*. ESIT 2000, pp. 453-457, Aachen, Germany, 2000.

[3] M. Michalopoulus, G. D. Dounias, N. Thomaidis, G. tselentis. *Decision Making using Fuzzy C-Means and Inductive Machine Learning for Managing Bank Branches Performances.*

[4] P. Alam, D. Booth, K. Lee. T. Thordarson. *The use of fuzzy clustering algorithm and self-organizing neural networks for idedntifying potentially failing banks : an experimental study*. Expert Systems with Applications 18 pp. 185-199, 2000.

[5] Susan M. Bridges, Rayford B. Vaughn. *Fuzzy Data Mining And Genetic Algorithms Applied To Intrusion Detection*. Presented at the National Information Systems Security Conference (NISSC), 2000.

[6] Wenkee Lee and Salvator J, Stolfo. *Data Mining Approaches for Intrusion Detection*. Proc. 1998. 7th USENIX Security Symposium, 1998.

[7] William W. Cohen. *Fast Effective Rule Indication*. In proceedings of the 12th International Conference on Machine Learning, 1995.

[8] Wenkee Lee, Salvator J. Stolfo and K. Mok. *Mining Audit Data to Build Intrusion Detection Models*. Proc. KDD-98, 1998.

[9] Paul Dokas, Levent Ertoz, Vipin Kumar, Aleksandar Lazarevic, Jaideep Srivastava and Pang-Nig Tan. *Data Mining for Network Intrusion.* Next Generation Data Mining, 2002.

[10] John E. Dickerson, Jukka Juslin, Ourania Loulousoula, and Julie A. Dickerson. *Fuzzy Intrusion Detection.* IFSA worls Congress and 20th North American Fuzzy Information Processing Society (NAFIPS) International Conference, 2001.

[11] Hiren Shah, Jeffrey Undercoffer and Anupam Joshi. *Fuzzy Clustering for Intrusion Detection.*

[12] David Marchette, *A Statistical Method for Profiling Network Traffic.*

[13] Takagi, T. and M. Sugeno (1985). *Fuzzy Identification Of Systems And Its Application To Modeling And Control. IEEE Trans. Systems, Man and Cybernetics 15*(1), 116–132.

[14] T.A. Johansen and B.A. Foss. *Constructing NARMAX Models using ARMAX Models. International Journal of Control*, 58(5):1125–1153, 1993.

[15] R. Babu¡ska and H.B. Verbruggen. *An Overview of Fuzzy Modeling for Control. Control Engineering Practice*, 4(11):1593–1606, 1996.

[16] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka. *Construction Of Fuzzy Classification Systems With Rectangular Fuzzy Rules Using Genetic Algorithms.* Fuzzy Sets & Systems, 65:237–253, 1994.

[17] K. Strokbro, D.K. Umberger, and J.A. Hertz. *Exploiting Neurons with Localized Receptive Fields to Learn Chaos. Journal of Complex Systems, 4(3):603–622, 1990.*

[18] J.S.R. Jang. *ANFIS: Adaptive-Network-Based Fuzzy Inference Systems.* IEEE Transactions on Systems, Man, and Cybernetics, 23(3):665–685, 1993.

[19] M. Tanaka, J. Ye, and T. Tanino. *Identi.Cation Of Nonlinear Systems Using Fuzzy Logic And Genetic Algorithms.* In IFAC Symposium on System Identi.cation, pages 301-306, Copenhagen, Denmark, 1994.

[20] Zimmermann, H.-J. *Fuzzy Sets, Decision Making and Expert Systems.* Boston: Kluwer Academic Publishers, 1987.

[21] Jang, J.-S.R. *ANFIS: Adaptive-Network-Based Fuzzy Inference Systems.* IEEE Transactions on Systems, Man & Cybernetics 23(3), 665–685, 1993.

[22] Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function.* Plenum Press, New York, 1981.

[23] Gustafson, D.E. and W.C. Kessel. *Fuzzy Clustering With A Fuzzy Covariance Matrix.* In Proc. IEEE CDC, San Diego, CA, USA, pp. 761–766, 1979.

[24] Babuska, R. and H.B. Verbruggen. *Identification Of Composite Linear Models Via Fuzzy Clustering.* In Proceedings European Control Conference, Rome, Italy, pp. 1207–1212, 1995.

[25] Gobithasan Rudrusamy, Azrudin Ahmad, Rahmat Budiarto, Azman Samsudin, Sureswaran Ramadass. *Fuzzy Based Diagnostics System for Identifying Network Traffic Flow Anomalies.* Proceedings of the International Conference on Robotics, Vision, Information and Signal Processing (ROVISP Jan 2003), Penang, Malaysia