© 2014,
Science line Publication
www.science-line.com

ISSN: 2322-5114
Journal of World's Electrical Engineering and Technology
J. World. Elect. Eng. Tech. 3(2): 53-58, 2014

JWEET

Comparison of Fuzzy Clustering Algorithms in Intrusion Detection System

Neda Jabbari[1]*, Jamshid Bagherzadeh[2]

[1] *Computer Engineering department, Science and Research Branch, Islamic Azad University, West Azerbaijan, Iran*
[2] *Computer Engineering department, Urmia University, Urmia, Iran*

*Corresponding author's Email: nedajabbari.nj@gmail.com

ORIGINAL ARTICLE
Received 13 April. 2014
Accepted 24 April. 2014

Abstract *– According to the growth of the Internet technology, there is a need to develop strategies in order to maintain security of system. One of the most effective techniques is Intrusion Detection System (IDS). Clustering which is commonly used to detect possible attacks is one of the branches of unsupervised learning. Fuzzy clustering algorithms play an important role to reduce spurious alarms and Intrusion detection, which have uncertain quality. This paper Compare and Review fuzzy c-means and Gath-Geva and Gustafson-Kessel algorithms in order to Intrusion detection in system.*

Keywords*: FCM, Gath-Geva, Gustafson-Kessel, IDS, Security*

INTRODUCTION

With increasing development of Internet in human society and more human use of online and network resources, the need for security in computer networks is felt more than ever. For a successful attack, first of all striker would require collecting data using various tools such as *whose* and *nslookup* to obtain IP address, domain name server, etc. The attacker then began to Probe and scan vulnerabilities to find existing vulnerabilities in the system to reach its goal. Now by creating some remote to local (R2L) attacks like guessing the password or overflow the buffer, try to create initial access. After that, the attacker doing user to root (U2R) attacks are caused Escalate privilege; U2R attack is actually such whole software that allows an attacker to run commands that are rated only by a root or of user-specific score. After performing these steps the attacker attempts to launch planned attack and with stealing and or modified confidential valuable data or changed web pages, etc. has caused disruption and creates intrusion systems in computer networks [1].

An intrusion detection system in computer networks is one of the security methods are used to detect various types of attacks and intrusions [2]. These systems are responsible for monitoring computer network activities to detect intruders on the management policy violations, security and malicious activities doing [3]. An Intrusion can be carried out by a legitimate user of the system or by illegal users systems [4]. Today with the increasing variety of attacks, detection and prevention of intrusion by antivirus and firewalls only would not be possible and IDS either hardware or software are used as one of the main mechanisms for securing of networks and computer systems which is generally firewall, Security complementary. In terms of speed and accuracy hardware systems based IDS are superior to software systems. But on the other hand software systems have the capability of more compatible with different operating systems, hence are more common and are usually a better choice [5].

James P. Anderson as the first person to evaluate the need to automatically log in the system in order to raise the security objectives is known. In 1980, Anderson released a report which is referred to as the basic activities in the field of intrusion detection [6-7]. Overall for coping with an attack on three main steps ahead: monitoring and evaluation, detection and reaction [8]. In the discovery phase, there are different ways according to [9] involves the misuse detection and anomaly detection that is more explained. One type of anomaly detection is learning are divided into two types self-learning on unsupervised learning and supervised learning; Clustering is one branch of unsupervised learning [10]. From a variety of clustering methods, clustering fuzzy algorithms can be cited.

Fuzzy clustering is another type that the probability of data is [0, 1] which belongs to these categories; one of the most important and applicable algorithms of fuzzy clustering is *C*-Mean fuzzy algorithm [11]. There are several criteria for clustering in this algorithm and the main one is the distance of any point from the center of cluster [12]. The other fuzzy clustering algorithm can be referred to Gath-Geva and Gustafson-Kessel.

This paper is organized as follows. Firstly, In the Section 2 and sub sections of them, we introduce types of

computer network's attacks and different types of attacks and the methods of intrusion detection in systems and then we have an overview of the types of fuzzy clustering and finally KDDCUP99 dataset is described. In Section 3, we review and evaluate the clustering methods such as FCM, Gath-Geva and Gustafson-Kessel Algorithms. Section 4 is simulation and evaluation of these methods. Finally, in the Section 5, we will offer conclusion of this paper.

Background

This section gives a brief description of the types of computer attacks, intrusion detection systems and as well as data set kddcup99.

A. Type of attacks

According to intrusion detection in systems, there are a variety of computer networks attack methods that can be divided into four general categories, DOS, Probes, U2R, and R2L.

- **DOS**: attacks to network or host sources. Attacker sends TCP packets with high traffic through the services. As a result this causes disorder in network normal data services. These sources include network bandwidth, data packets routing, server information, and memory and ability of calculation in servers. Victims of DOS attacks are powerful servers with fast network connections. Distributed Denial of Service (DDOS) attacks are other types of DOS attacks which are in distributed networks [6]. DOS attacks are typically divided into 6 groups.
- **Probes:** Network probes are usually attacks scanning computer networks to gather information or find known vulnerabilities, which are exploited for further or future attacks. The goal of this information gathering is to find out about computer and services that are present in a network as well as to detect the possibility of attack based on known vulnerabilities [13].
- **User to Root (U2R):** In U2R attack, the attacker starts with availability to normal user account, and in this way it can access the root [8]. These types of attacks are performed in victim's machine successfully and control the root [6]. There are several U2R attacks that the most important one is Buffer over Flow. This attack happens when a copy of program is copied with more data in static buffer without checking its capacity. Programmers solve these problems by exact techniques [14].
- **Remote to local (R2L):** A remote unwanted intrusion abuses user's legal account, and sends packet on the network [6]. In fact, this attack is caused when the attacker has the ability to send information packets through the victim machine and abuses of users' local availability vulnerable by sending packets in network. There are different ways to unallowable access to local account. Some of them are as follows: Warez master, Warez client, Spy, Phf, Multi-hop, Imap, Guess_paswd, and Ftp_write [15].

TABLE 1
FOUR TYPES OF ATTACKS FOR KDDCUP99 DATASET

Feature name	*Category*
Normal	Normal
Smurf	Dos
Neptune	Dos
Back	Dos
Treadrop	Dos
Pod	Dos
Land	Dos
Satan	Probe
Ipsweep	Probe
Portsweep	Probe
Nmap	Probe
Warezclient	R2L
Guess_passwd	R2L
Warezmaster	R2L
Imap	R2L
ftp_write	R2L
Multihop	R2L
Phf	R2L
Spy	R2L
Buffer_overflow	U2R
Rootkit	U2R
Loadmodule	U2R
Perl	U2R

B. Intrusion detection systems

An Intrusion Detection System (IDS) is a security technique attempting to detect various attacks. The basic principle of intrusion detection is based on the assumption that intrusive activities are noticeably different from normal ones and thus are detectable [16]. Many intrusion detection approaches have been suggested in the literature since Anderson's seminal report [17].

It has been identified mainly two techniques, namely misuse detection and anomaly detection [18]. The first approach, commonly known as misuse detection, is a rule-based approach that uses stored signatures of known intrusion events to detect known attacks. This approach has been highly successful in detecting occurrences of previously known attacks. However, it fails to detect new attack types and variants of known attacks whose signatures are not stored. When new attacks occur, the signature database has to be manually modified for future use (dynamic clustering). The second approach is commonly known as an anomaly detection approach. Any events which deviates the normal usage patterns are considered to be suspicious. It constructs the profile of user behavior or status of network traffic and compares

the observed behavior with the stored profile to determine whether an attack action occurs. The anomaly detection approach may have the advantage of detecting previously unknown attacks over the misuse detection approach. However, it may suffer from false alarm problem and radically changed user behaviors [19].

C. KDDCUP99 data set

KDDCUP99 data are collected based on DARPA innovation in 1998 for Intrusion detection system (IDS) designers that are used in several investigations to find the attacks and intrusion [20]. These data are simulated in seven weeks to intrusion detection, KDDCUP99 data have 41 properties which are divided to 4 parts [21]:

- 9 basic and SCD header features for each connection (similar to netflow)
- 9 time-based MCD header features constructed over a 2 window.
- 10 host-based MCD header features constructed over a 100 connection window to detect slow probes.
- 13 content-based features were constructed from the traffic payloads using domain knowledge.

FUZZY CLUSTERING ALGORITHMS

Clustering is an unsupervised classification that the classes have not been predefined [22]. In clustering process, the samples are divided into categories which the members are alike and called cluster [10]. In classic clustering, each input sample belongs to one cluster and cannot be a member of several clusters, so if a sample is like more than one clusters, it will be difficult for us to recognize that the sample belongs to which cluster, and this is the main difference between classic and fuzzy clustering. It shows that in fuzzy clustering a sample can belong to more than one cluster, and in fuzzy logic, belonging function of clusters doesn't have two values and can have any value between 0 and 1 [23]. Fuzzy clustering is an important problem which is the subject of active research in several real-world applications. Next, we introduce fuzzy clustering algorithm called fuzzy C-means (FCM), Gath-Geva, and Gustafson-Kessel.

A. FCM

A special case of the FCM algorithm was first reported by Dunn [24] in 1972. Dunn's algorithm was subsequently generalized by Bezdek [25], Gustafson and Kessel [26], and Bezdek et al. [27]. The FCM algorithm is based on the minimization of an objective function called *C-means functional*. It is defined by Dunn as:

$$J(X;U,V)=\sum_{i=1}^{c}\sum_{k=1}^{N}(m_{ik})^{m}\left\|X_k-V_i\right\|_A^2 \qquad (1)$$

where

$$V=[v_1,v_2,...,v_c] \quad v_i\in R^n \qquad (2)$$

Is a vector of cluster prototypes (centers), which have to be determined, and

$$D_{ikA}^2=\left\|X_k-V_i\right\|_A^2=(X_k-V_i)^T A(X_k-V_i) \qquad (3)$$

is a squared inner-product distance norm. Statistically, (1) can be seen as a measure of the total variance of X_k from V_i. The minimization of the c-means functional (1) represents a nonlinear optimization problem that can be solved by using a variety of available methods, ranging from grouped coordinate minimization, over simulated annealing to genetic algorithms. The most popular method, however, is a simple Picard iteration through the first-order conditions for stationary points of (1), known as the fuzzy c-means (FCM) algorithm. The stationary points of the objective function (1) can be found by adjoining the constraint (5) to J by means of Lagrange multipliers:

$$\bar{J}(X;U,V,l)=\sum_{i=1}^{c}\sum_{k=1}^{N}(m_{ik})^{m}D_{ikA}^2+\sum_{k=1}^{N}l_k\left(\sum_{i=1}^{c}m_{ik}-1\right) \qquad (4)$$

$$\sum_{k=1}^{c}m_{ik}=1, \quad 1\le i\le N \qquad (5)$$

and by setting the gradients of J with respect to U; V and λ to zero. If $D_{ikA}^2>0$; $m>1$, then $(U,V)\in M_{fc}\times R^{n}$ may minimize (1) only if

$$m_{ik}=\frac{1}{\sum_{j=1}^{c}\left(D_{ikA}/D_{jkA}\right)^{2/(m-1)}} \quad 1\le i\le c;\ 1\le k\le N \qquad (6)$$

and

$$V_i=\frac{\sum_{k=1}^{N}m_{ik}^{m}X_k}{\sum_{k=1}^{N}m_{ik}^{m}} \quad 1\le i\le c \qquad (7)$$

This solution also satisfies the remaining constraints (8) and (9). Note that equation (6) gives v_i as the weighted mean of the data items that belong to a cluster, where the weights are the membership degrees. That is why the algorithm is called c-means. One can see that the FCM algorithm is a simple iteration through (6) and (7).

$$m_{ij}\in[0,1], \quad 1\le i\le N, \quad 1\le k\le c \qquad (8)$$

$$0<\sum_{i=1}^{N}m_{ik}<N, \quad 1\le k\le c \qquad (9)$$

The FCM algorithm computes with the standard Euclidean distance norm, which induces hyper spherical clusters. Hence it can only detect clusters with the same shape and orientation, because the common choice of norm inducing matrix is $A = \boldsymbol{I}$ or it can be chosen as an $n\times n$ diagonal matrix that accounts for different variances in the directions in the directions of the coordinate axes of X:

$$A_D = \begin{bmatrix} (1/s_1)^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (1/s_n)^2 \end{bmatrix} \quad (10)$$

or A can be defined as the inverse of the $n \times n$ covariance matrix $A = F^{-1}$, with

$$F = \frac{1}{N} \sum_{k=1}^{N} (X_k - \bar{X})(X_k - \bar{X})^T \quad (11)$$

Here x denotes the sample mean of the data. Steps of fuzzy c-mean algorithm [28]:

- For the first clusters initial value for *k*, *m*, and *U* should be estimated.
- The center of clusters should be calculated by second formula.
- The dependence matrix should be calculated by in second step.
- If $\|U_1+1-U_1\| \leq \varepsilon$ the algorithm is finished, vice versa go to second step.

B. Gath-Geva

Many algorithms for fuzzy clustering depend on initial guesses of cluster prototypes, and on assumptions made as to the number of subgroups present in the data. Gath-Geva algorithm is derived from a combination of the fuzzy K-means algorithm and the fuzzy maximum likelihood estimation [29]. The Gath-Geva algorithm is an extension of Gustafson-Kessel algorithm that takes the size and density of the clusters into account. Gath-Geva clustering algorithm uses a distance norm based on the fuzzy maximum likelihood estimates [30]. Gath and Geva described an initialization strategy of unsupervised tracking of cluster prototypes in their2-layer clustering scheme, in which FCM and fuzzy ML estimation are effectively combined [29].

The fuzzy maximum likelihood estimates (FMLE) clustering algorithm employs a distance norm based on the fuzzy maximum likelihood estimates, proposed by Bezdek et al. [31].

$$D_{ik}(X_k, V_i) = \frac{\sqrt{\det(F_{wi})}}{a_i} \exp\left(0.5(X_k - V_i^{(l)})^T F_{wi}^{-1}(X_k - V_i^{(l)})\right) \quad (12)$$

Note that, contrary to the GK algorithm, this distance norm involves an exponential term and thus decreases faster than the inner-product norm. F_{wi} denotes the fuzzy covariance matrix of the *i*th cluster, given by:

$$F_{wi} = \frac{\sum_{k=1}^{N} (m_{ik})^w (X_k - V_i)(X_k - V_i)^T}{\sum_{k=1}^{N} (m_{ik})^w} \quad 1 \leq i \leq c \quad (13)$$

where $\omega = 1$ in the original FMLE algorithm, but we use the $\omega = 2$ weighting exponent, so that the partition becomes more fuzzy to compensate the exponential term of the distance norm. The diffrerence between the matrix *Fi* in GK algoritm and the F_{wi} define above is that the latter does not involve the weighting exponent *m*, instead of this it consists of $w = 1$. (The reason for using this *w* exponent is to enable to generalize this expression.) This is because the two weighted covariance matrices arise as generalizations of the classical covariance from two different concepts. The a_i is the prior probability of selecting cluster *i*, given by:

$$a_i = \frac{1}{N} \sum_{k=1}^{N} m_{ik} \quad (14)$$

The membership degrees *ik* are interpreted as the posterior probabilities of selecting the *i*-th cluster given the data point *xk*. Gath and Geva [8] reported that the fuzzy maximum likelihood estimates clustering algorithm is able to detect clusters of varying shapes, sizes and densities. The cluster covariance matrix is used in conjunction with an "exponential" distance, and the clusters are not constrained in volume. However, this algorithm is less robust in the sense that needs a good initialization, since due to the exponential distance norm, it converges to a near local optimum.

C. Gostafson-Kessel

Gustafson and Kessel [32] extended the standard fuzzy c-means algorithm by employing an adaptive distance norm, in order to detect clusters of different geometrical shapes in one data set [33]. Each cluster has its own norm-inducing matrix A_i, which yields the following inner-product norm:

$$D_{ikA}^2 = (x_k - v_i)^2 A_i (x_k - v_i) \quad 1 \leq i \leq c, 1 \leq k \leq N \quad (15)$$

The matrices A_i are used as optimization variables in the c-means functional, thus allowing each cluster to adapt the distance norm to the local topological structure of the data. Let *A* denote a *c*-tuple of the norm-inducing matrices: $A = [A_1; A_2; \ldots A_c]$. The objective functional of the *GK* algorithm is defined by:

$$J(X; U, V, A) = \sum_{i=1}^{c} \sum_{k=1}^{N} (m_{ik})^m D_{ikA_i}^2 \quad (16)$$

For a fixed *A*, conditions (8), (5) and (9) can be directly applied. However, the objective function (16) cannot be directly minimized with respect to *Ai*, since it is linear in *Ai*. This means that *J* can be made as small as desired by simply making *Ai* less positive definite. To obtain a feasible solution, *Ai* must be constrained in some way. The usual way of accomplishing this is to constrain the determinant of *Ai*. Allowing the matrix *Ai* to vary with its determinant fixed corresponds to optimizing the cluster's shape while its volume remains constant:

$$\|A_i\| = r_i, \quad r > 0 \quad (17)$$

Using the Lagrange multiplier method, the following expression for A_i is obtained:

$$A_i = [r_i \det(F_i)]^{1/n} F_i^{-1} \quad (18)$$

where F_i is the fuzzy covariance matrix of the ith cluster defined by:

$$F_i = \frac{\sum_{k=1}^{N} (m_{ik})^m (X_k - V_i)(X_k - V_i)^T}{\sum_{k=1}^{N} (m_{ik})^m} \quad (19)$$

Note that the substitution of (18) and (19) into (15) gives a generalized squared Mahalanobis distance norm between X_k and the cluster mean V_i where the covariance is weighted by the membership degrees in U. The numerically robust GK algorithm described by R. Babuska et al. [34] is used in this toolbox.

SIMULATION RESULTS

In this Section, we provide the simulation results that were obtained during our experiments. In this paper KDD CUP99 dataset is used. This dataset was conducted by MIT Lincoln Laboratory. There are two types of traffic in this dataset, normal and abnormal (attack). We consider two classes for experiments, normal and attack. All mentioned attacks in four classes are merged together to construct the attack class. We select sub-dataset from the whole of 10% KDDCUP99 dataset to evaluate the performance of the proposed algorithm. Sub-dataset is selected in the manner that we ensure all attacks are available in the sub-dataset. Percentage of normal and attack records in our dataset is 70% and 30%, respectively. From normal, DOS, U2R, R2L, and Probe attacks, we use 1700, 400, 38, 80, and 200 records, respectively, in train stage. So, in train stage, we from normal and attack records, we have 1700 and 718 records, respectively. In test stage, we use the same amount of records that are used in train stage, except R2L attack. From U2R attack, we use 15 records in the test stage. Therefore, in test stage, we have 1700 and 695 data records from normal and attack, respectively. All codes needed for performance evaluation are implemented in MATLAB environment. The performance of intrusion detection techniques is evaluated based on two well-known criteria: *detection rate* and *false positive rate*. The detection rate represents the percentage of correctly detected attacks whereas, the false positive rate is the percentage of normal records detected incorrectly as attack. In intrusion detection, ROC curves are used on the one hand to visualize the relation between the TP and FP rate of a certain classifier while tuning it, and on the other hand, to compare the accuracy of two or more classifiers.

Here, in the Figure 1 is shown the ROC curves of Comparison of the three fuzzy clustering algorithms namely, FCM algorithm, Gath-Geva algorithm and Gostafson-Kesel algorithm.

Fig. 1 shows comparison of the intrusion detection rates of the three fuzzy algorithms. It is observed that Gostafson-Kesel fuzzy algorithm with two clusters has better than other algorithms. However, increasing the detection rate, false positive rate increases.

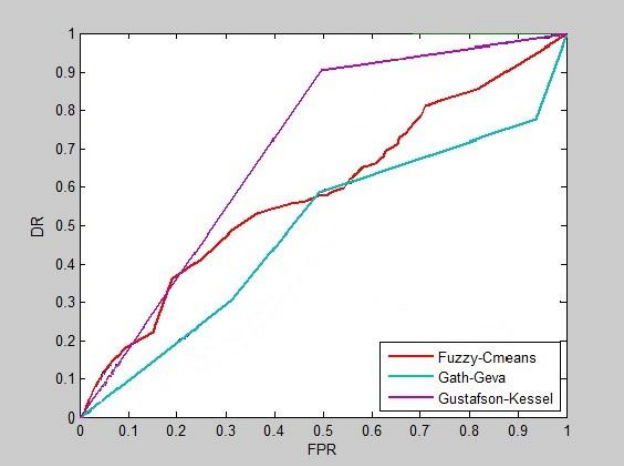

Fig. 1- Comparison of the intrusion detection rates of the three fuzzy algorithms.

CONCLUSION

In this paper, compare of three fuzzy clustering algorithms is developed for intrusion detection systems, and the results are shown in the ROC curve. Fuzzy clustering is a new science that work and study is ongoing in this field because it is considered a lot in different science as a solution. In recent years this method is optimized and the results of optimization are provided as papers. The goal of optimization is obtaining to the minimum number of replicates and clusters with the most similar members. In this paper, the kddcup99 data set is used, although this data set is extremely popular among scholars, but has the disadvantage that the optimization methods of feature reduction as well as feature selection may be helpful.

REFERENCES

[1] C.M. Chen, Y. Chen, H.C. Lin, "An efficient network intrusion detection", *Computer Communications*, vol. 33, pp.477–484, 2010.
[2] Pormohseni, "Review and identify the computer network intrusion detection systems", 2011.
[3] E. Biermann, E. Cloete, L.M. Venter, "A comparison of intrusion detection systems", *Computer and Security*, vol. 2, pp.676-683, 2001.
[4] R. Heady, G. Luger, A. Maccabe, M. Sevilla, "The Architecture of a Network-level Intrusion Detection System", *Technical report*, Department of Computer Science, University of New Mexico, Albuquerque, pp.1-18, 1990.
[5] Bro IDS homepage, Available: *www.bro-ids.org*, last update: 23.07.2012.
[6] A. Ghorbani, W. Lu, M.Tavallaee, "Network Intrusion Detection and Prevention: Concepts and Techniques", *Springer publisher*, 2009, pp 234.
[7] J.P Anderson, "Computer Security Threat monitoring and surveillance", 1980, last update: 05.08.2012. Available: *http://csrc.nist.gov/publications/history/ande80.html*
[8] Ashamed, M. Rezai, "Introduction to Intrusion Detection System", (Part I), *Technical report*, Mashhad University, Iran.

[9] S. Lee, G. Kim, S. Kim," Self-adaptive and dynamic clustering for online anomaly detection", *Expert Systems with Applications*, pp.14891–14898, 2011.
[10] C. Kruegel, F. Valeur, G. Vigna, "Intrusion Detection and Correlation challenges and Solution", *University of California, Santa Barbara, Springer Science USA*, 2005.
[11] M. Ghasemi, M. Khanghandi, "The application of fuzzy logic in recognition scheme", Arak, Iran, 2009.
[12] K. Bharti, S. Shukla, S. Jaim, "Intrusion Detection using Clustering", special Issue of IJCCT2010 for International Confrance , 2010, Vol 1, Issue2,3,4, pp.158-165, 2010 .
[13] S. Garfinkel, G. Spafford, "*Practical unix and internet security"*, OReilly and Associates, Sebastopol, CA, USA, 1996.
[14] Anonymous. "Maximum Security: A Hacker's Guide to Protecting Your Internet Site and Network", Chapter 15, pp.359-362, Indianapolis, 1997.
[15] K. Kendall, "A Database of Computer Attacks for the Evaluation of Intrusion Detection Systems", Bachelor of Science in Computer Science and Engineering and Master of Engineering in Electrical Engineering and Computer Science, pp.124, June 1999.
[16] D. Denning, "An Intrusion Detection Model", IEEE Transactions on software engineering, 1987, pp. 222–232.
[17] J.P. Anderson, "Computer security threat monitoring and surveillance", 1980.
[18] T. Verwoerd, R. Hunt, "Intrusion detection techniques and approaches", Computer Communications, vol.25, (15), 2002, pp.1356–1365.
[19] E. Lundin, E. Jonsson, "Anomaly-based intrusion detection: privacy concerns and other problems", Computer Networks, vol.34 (4), 2000, pp.623–640.
[20] DARPA Intrusion Detection Evaluation Plan, 1999, Available: *http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/docs/ id99-eval-ll.html.*
[21] J.J. Davis, A.J. Clark," Data preprocessing for anomaly based network intrusion detection: A review", Published by Elsevier, 2011, pp.353-375.
[22] F. Soleimanian Gharehchopogh, N. Jabbari, Z. Ghaffari Azar, " Evaluation of Fuzzy K-Means And K-Means Clustering Algorithms In Intrusion Detection Systems", IJSTR,Vol.1, Issue 11, 2012.
[23] V. Faber, "Clustering and the Continuous K-means Algorithm", Los Almas since Number22, 1994, pp: 138-144.
[24] F. Hoppner, F. Klawonn, R. Kruse, T. Runkler, "Fuzzy Cluster Analysis Methods for Classification", Data Analysis and Image Recognition, John Wiley and Sons, 1999.
[25] R. Babuska, H.B. Verbruggen, "Constructing fuzzy models by product space clustering". In H. Hellendoorn and D. Driankov, editors, Fuzzy Model Identification: Selected Approaches, pp.53-90, Springer, Berlin, Germany, 1997.
[26] J.S.R. Jang, C.T. Sun, "Nero fuzzy modeling and control", Proceedings of the IEEE, vol. 83, 1995, pp.378-406.
[27] N.R. Draper, H. Smith, "Applied Regression Analysis", 3rd Edition. John Wiley and Sons, Chichester, 1994.
[28] B James, E. Robert, F.William, "The Fuzzy C-Means Clustering Algorithm", Computers & Geosciences, Vol.10, No. 2-3, 1984, pp.191-203.
[29] I. Gath , B. Geva "Unsupervised Optimal Fuzzy Clustering" published by IEEE , vol. 11, no. 7, pp. 773–781,1989.
[30] J. Abonyi, R. Babuska, F. Szeifer," Modified Gath-Geva Fuzzy Clustering for Identification of Takagi-Sugeno Fuzzy Models".
[31] J.C. Bezdek, J.C. Dunn, "Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions", IEEE Transactions on Computers, 1975, pp.835-838.
[32] J.C. Bezdek, P.F. Castelaz, "Prototype Classification and Feature Selection with Fuzzy Sets", IEEE Trans. on Systems, Man and Cybernetics, Vol.SMC-7, No. 2, February1971, pp. 87-92.
[33] D.E. Gustafson, W.C. Kessel, "Fuzzy clustering with fuzzy covariance matrix", In Proceedings of the IEEE CDC, San Diego, 1979, pp.761-766.
[34] R. Babuska, P. J. van der Veen, U. Kaymak, "Improved covariance estimation for Gustafson-Kessel clustering", IEEE International Conference on Fuzzy Systems, 2002, pp.1081-1085.