Hybrid Fuzzy and Neural Network for Intrusion Detection System

Dr. Manar Younis K. **Dr. Bayda Ibraheem Khaleel**
Department of Computer Science
College of Computer Science and Mathematics
University of Mosul

Received **Accepted**
27 / 05 / 2012 **25 / 06 / 2012**

الخلاصة

مع التطور الكبير لشبكة الانترنيت، زادت الحاجة لاستخدام أنظمة الحماية مثل أنظمة كشف وتصنيف التطفل لحماية الحاسبة والشبكة من الهجمات والوصول الغير مخول به. وهنا تم تطبيق خوارزميات الـ(FCM, CPN) والطريقة الجديدة والتي سميت شبكة CP. المضببة أي (FCPN). لتصنيف ببيانات (kdd cup 99 وNSL_KDD وهي النسخة الجديدة المعدلة لبيانات kdd cup 99.) إلى 5 أصناف أو عناقيد احدها للمرور الطبيعي والأخرى لأنواع الهجمات الرئيسية. وكذلك صنفت هذه البيانات إلى صنفين احدها للمرور الطبيعي والآخر للهجوم واستخدمت هذه الخوارزميات أيضا لكشف التطفل (الشاذ). اذ يحوي كل عنقود يحوي على بيانات متشابه وتختلف عن البيانات التي تحويها العناقيد الأخرى. وفي مرحلة التدريب تم أخذ الفايل (kdd 10%) من بيانات kdd cup 99. والذي يحوي على (494020) سجل من البيانات، ولمرحلة الاختبار تم أخذ الفايل الذي اسمه (corrected kdd) والذي يحوي على (311029) سجل. ومن بيانات NSL_KDD تم أخذ الفايل (NSL-KDD Train) الذي يحوي على (125975) سجل ليتم استخدامه في مرحلة التدريب، والفايل (NSL-KDD Test) الذي يحوي على (22544) سجل لمرحلة الاختبار. وتم حساب نسبة التصنيف والكشف والإنذار الكاذب. وأخيرا في مرحلة التدريب حصلت الخوارزميات (FCM, CPN, FCPN) على نسبة تصنيف 100%. أما في مرحلة الاختبار فقد حصلت الخوارزمية المهجنة الجديدة الـFCPN على نسبة كشف 100% لبيانات kdd cup 99 و (99.703) بالنسبة لبيانات الـNSL_KDD، ومن ثم تمت مقارنة النتائج التي تم الحصول عليها من تطبيق هذه الخوارزميات على هذه البيانات.

Abstract

Along with the development and growth of the internet network, there is an increasing needed to use protection systems such as intrusion classification and detection systems that protect computer and network from attacks and unauthorized access. Fuzzy c-means (FCM), neural network (counterpropagation network CPN, and a new method that called it (Fuzzy counterpropagation network FCPN) algorithms were applied using kdd cup 99 and NSL-KDD which is a new version of kdd cup 99 dataset to classify this dataset into 5 classes or clusters one for normal traffic nd others classes for the main types of attacks. Another type of classification is made on the dataset it was classified into 2 classes one for normal and other for types of attacks and detect a new attack(abnormal). Each cluster will contain dataset more similar to each other within cluster and difference from that in the other clusters. (10% kdd) file from kdd cup 99 was taken in the training stage that contain (494020) records and (corrected kdd) file that contain (311029) records in testing stage. From NSL-KDD was taken (NSL-KDDTrain) file that contain (125973) records used in training stage and (NSL-KDDTest) file that contain (22544) records in testing stage. Classification rate, Detection rate, and false alarm rate were computed. Finally the classification rate obtained is (100%) for FCM, CPN, FCPN algorithms in training stage. With got higher DR(100%) for FCPN to kdd cup 99, and (99.703) is the DR obtained for FCPN to NSL-KDD in testing. and then were made comparisons between results obtained after applying the algorithms on this dataset.

1- INTRODUCTION

The number of intrusion into computer systems is growing because new automated intrusion tools appearing every day, and these tools and different system vulnerability information are easily available on the web[1]. Using an intrusion detection system (IDS) is one way of dealing with suspicious activities within a network[2]. Intrusion detection, an important component of information security technology, helps in discovering, determining, and identifying unauthorized use, and destruction of information and information systems[3]. The goals of intrusion detection are detect as many type of attacks as possible, including those by attackers and those by insiders. Also detect as accurately as possible thereby minimizing the number of false alarms, and also detect the attacks in the shortest possible time[4].

Intrusion detection technique covers basic analysis methods such as anomaly and rule based signature detection techniques. Signature or misuse detection technique uses the signature patterns to be matched with possible malicious content in the packet payload. This is a very

straightforward way to detect any malicious activity but, it can not detect unknown threats and it requires new signature updates every time. Anomaly detection is based on a behavior model that means it differentiates between normal and suspicious traffic. An anomaly detection model is developed based on the learning of normal or usual behavior of the monitored system and so when this model finds an unusual or different pattern, it generates an alert as an intrusion. This technique is able to detect unknown attacks.[5].

There are three types of intrusion detection systems: host-based intrusion detection system (HIDS), network-based intrusion detection system (NIDS), and combination of both types (Hybrid Intrusion Detection System). And the main advantages of IDS is detect many type of attacks with minimizing the number of false alarms, and detect the attacks in the shortest possible time, also IDS monitor and analyze user and system activity.[4][6][7]

The aim of this research is applied fuzzy c-means (FCM) clustering algorithm and counterpropagation network (CPN), and then combined FCM and CPN to obtain fuzzy counterpropagation network, these three algorithms used to classify network intrusion that represented by KDD cup 99 and NSL-KDD dataset. Finally compare between then according performance measures.

2- PREVIOUS WORK

In particular several clustering algorithms and neural networks based approaches were employed for intrusion detection. In 2010 Jawhar and Mehrotra [8] used fuzzy c-means clustering to classified dataset into 2 classes, they used (22133)records, the classification result in training stage is 99.9. Siddiqui[9] used parallel backpropagation neural network and pararllel fuzzy ARTMAP, the detection rate result for parallel BP in training stage is 98.36 and the detection rate in testing stage is 81.73 and false alarm is 1.28. Detection rate for parallel fuzzy ARTMAP in training stage is 80.14 and in testing state detection rate is 80.52 and false alarm is 19.48. Panda and Patra [10] used Sequential information Bottleneck(SIB) clustering algorithm and kdd dataset for attack classification, the detection rate result is 85.5 and false alarm is 34. Al-Sharafat and SH.Naoum [11] used Steady State Genetic Algorithm Based Machine Learning SSGBML, and used kdd 99 dataset, the detection rate for this approach is 97.45 in training state.

3- KDD DATASET

The KDD cup 99 dataset includes a set of 41 features derived for each connection and a label which specifies the status of connection records as either normal or specific attack type. The 41 set of features can be grouped into 4 categories which are[12]:

- Basic Feature: basic feature can be derived from packet header without inspecting the payload.
- Content Feature: Content feature cab be derived by assessing the payload of the original TCP Packet.
- Time based Traffic Feature: These feature can be designed to capture properties that mature over a 2 second temporal windows. Time based feature is suitable to detect the fast attack.
- Host based traffic feature: Utilized historical windows estimated over the number of connection such a 100 connection without considering time. This feature is suitable to detect the slow attack.[12]

And attack types were divided into the following 4 main categories[13][14]

1. Probing: Probing is a class of attacks where scans a network to gather information in order to find known vulnerabilities. An attacker with a map of machines and services that are available on a network can manipulate the information to look for exploits. There are different types of probes: some of them abuse the computer's legitimate features; and some of them use social engineering techniques.
2. Denial of Service: Denial of Service(DOS) is a class of attacks where an attacker makes some computing or memory resource too busy or too full to handle legitimate requests, denying legitimate uses access to a machine. There are different ways to launch a DOS attack: by abusing the computers legitimate features; by targeting the implementations bugs; or by exploiting the system's misconfigurations. DOS attacks are classified based on the services that an attacker renders unavailable to legitimate users.
3. User to root: In this attack, an attacker starts with access to a normal user account on the system by gaining root access. Regular programming mistakes and environment assumption give an attacker an opportunity to exploit the vulnerability of root access. An example of this class of attacks is regular buffer overflows.
4. Remote to user: This attack happens when an attacker sends packets to a machine over a network that exploits the machine's vulnerability to gain local access as a user illegally. There are different types of R2U attacks; the most common attack in this class is done by using social engineering.

4- NSL-KDD dataset

NSL is a new version of KDDcup99 and has some advantages over KDD cup 99, which is also contains 41 features and labeled as either normal or attack as the same in KDD cup 99, the NSL-KDD data set has the following advantages over the original KDD data set[15][16]:

- It does not include redundant records in the train set, so the classifiers will not be biased towards more frequent records.
- There is no duplicate records in the proposed test sets; therefore, the performance of the learners are not biased by the methods which have better detection rates on the frequent records.
- The number of selected records from each difficulty level group is inversely proportional to the percentage of records in the original KDD data set. As a result, the classification rates of distinct machine learning methods vary in a wider range, which makes it more efficient to have an accurate evaluation of different learning techniques.
- The number of records in the train and test sets are reasonable, which makes it affordable to run the experiments on the complete set without the need to randomly select a small portion. Consequently, evaluation results of different research works will be consistent and comparable.

The total number of connection records in training data set of NSL is "kddTrain+.TXT" file that contain (125973) records, and the total number of connection records in testing data set is "kddTest+.TXT" file that contain on (22544) records[16]. Table (1) show the kdd 99 data set used in training and testing stages that contain from normal and attack connection records. Table (2) shows the NSL-KDD data set used in training and testing stages that contain from normal and attack connection records.

Table 1: The number of samples kdd 99data set that were used[13]

Data set	Normal	Dos	Probe	U2R	R2L	Total
Corrected kdd	60593	229853	4166	70	16347	311029
10_precent kdd	97277	391458	4107	52	1126	494020

Table 2: The number of samples NSL-KDD data set that were used[15]

Data set	Normal	Dos	Probe	U2R	R2L	Total
Kdd rain(NSL)	67343	45927	11656	52	995	125973
Kdd test (NSL)	9711	7458	2421	67	2887	22544

5- NEURAL NETWORKS

A neural network contains no domain knowledge in the beginning, but it can be trained to make decisions by mapping exemplar pairs of input data into exemplar output vectors, and adjusting its weights so that it maps each input exemplar vector into the corresponding output exemplar vector approximately. A knowledge base pertaining to vector internal representation (i.e weight values) is automatically constructed from the data presented to train the network. Well-trained neural networks represent a knowledge base in which knowledge is distributed

in the form of weighted interconnections where a learning algorithm is used to modify the knowledge base from a set of given representative cases. Neural networks might be better suited for unstructured problems pertaining to complex relationships among variables rather than problem domains requiring value-based human reasoning through complex issues. Any functional form relating the independent variables (i.e. input variables) to the dependent variables (i.e. output variables) need not be imposed in the neural network model. Neural networks are thought to better capture the complex pattern of relationships among variables than statistical models because of their capability to capture non-linear relationships in data. A neural network is an appropriate method in the misuse detection or anomaly detection[17].

In this research counter propagation network (CPN) was combined with fuzzy c-means (FCM) algorithm in a single system called fuzzy counter propagation network. This system use the input labeled data (normal and attack patterns) to train a neural network model and fuzzy c-means. And in testing phase new patterns were used, these systems compute the detection rate (DR), false alarm rate, and classification rate.

6- PREPROCESSING DATA

From the KDD Cup 99 and NSL-KDD intrusion detection dataset, 41 features were derived to summarize each connection information. In order to train an architecture, several data enumeration and normalization operations were necessary. As a first approach, symbolic variables in the dataset were enumerated and all variables were normalized. Thus, each instance of a symbolic feature was first mapped to sequential integer values[18]. This dataset consist of symbolic and numeric values, all symbolic values were transformed into numeric values such as three types of protocols (tcp, udp, icmp) and 68 type of services in KDD cup 99 and 71 type of services in NSL-KDD and 11 types of flag, each one take value from [1..N], and the standard [0..1] normalization[19] was used for this research according equation(1):

$$X = \frac{x - \min}{\max - \min} \tag{1}$$

7- PERFORMANCE MEASURES

Two indicators were used to measure the accuracy of the methods: detection rate and false alarm rate. The detection rate shows the percentage of true intrusions that have been successfully detected as shown in equation(2). While the false alarm rate is defined as the number of normal instances incorrectly labeled as intrusion by the total number of normal instances as shown in equation(3) [7].

$$Detection\_rate = \frac{number\ of\ correctly\ \det ected\ samples}{total\ number\ of\ samples} \times 100 \tag{2}$$

$$False\ Positive\ rate\ (FPR) = \frac{FP}{TN + FP} = \frac{\#\ normal\ as\ \text{int}\ rusions}{\#\ normal} \times 100 \quad (3)$$

8- FUZZY C-MEANS ALGORITHM

The fuzzy c-means algorithm (FCM) introduced by Bezdek. Fuzzy c-means based on Euclidean distance function. It is a data clustering technique where each data point belongs to a cluster to some degree that specified by membership grade [20]. Let $X=\{x_1,\ldots x_j,\ldots x_n\}$ be the set of n objects and $V=\{ v_1,\ldots v_i,\ldots v_c\}$ be the set of c centroids where $x_j \in R^m$, $v_i \in R^m$, and $v_i \in X$ [21]. It partitions X into c clusters by minimizing the objective function as shown in equation(4):

$$J_m(\mu, v) = \sum_{k=1}^{N}\sum_{i=1}^{c}(\mu_{ki})^m\ d_{ik}(x_k, v_i) \quad (4)$$

where d_{ki} is given by $\| x_i - v_i \|$, c is the number of clusters in X, *m* is a weighting exponent [22]. The cluster centers are then evaluated using the following equation(5) :

$$v_i = \frac{\sum_{k=1}^{N}(\mu_{ki})^m\ X_k}{\sum_{k=1}^{N}(\mu_{ki})^m} \quad (5)$$

And membership matrix μ is update by the following equation(6) :

$$\mu_{ki} = \left[\sum_{j=1}^{c}\left(\frac{d_{ki}}{d_{kj}}\right)^{\frac{2}{m-1}}\right]^{-1} \quad (6)$$

The parameter *m* which is real number greater than 1[23] and it is a weighting exponent on each fuzzy membership and determines the amount of fuzziness of resulting classification[24]. And membership value to the data items for the clusters within a range 0 and 1 [25].

9- COUNTERPROPAGATION NETWORK

The CP network was first developed by Hecht-Nielsen [26], and consisted of combining the Kohonen network with a Grossberg layer[27]. The general form of the CP network can be seen in figure (1). The input nodes of the Kohonen layer are connected to the Kohonen neurons by weights w_{ij} while the Kohonen outputs are connected to the Grossberg layer by the connecting weights v_{ij} [28]. The learning of CPN can be split into two stages, unsupervised and supervised. Unsupervised learning is used during the first stage for clustering the input vectors to separate distinct sets of input data. During the second stage of learning, the weight vector between the kohonen and Grossberg layers are adjusted by supervised learning to reduce the errors between the CPN outputs and the corresponding desired targets. During the first stage, the distances

between the input vector $x=(x_1,....,x_i,x_n)^T$ composed of input nodes and all of the j kohonen nodes with n dimensions are determined to compete for the winner.

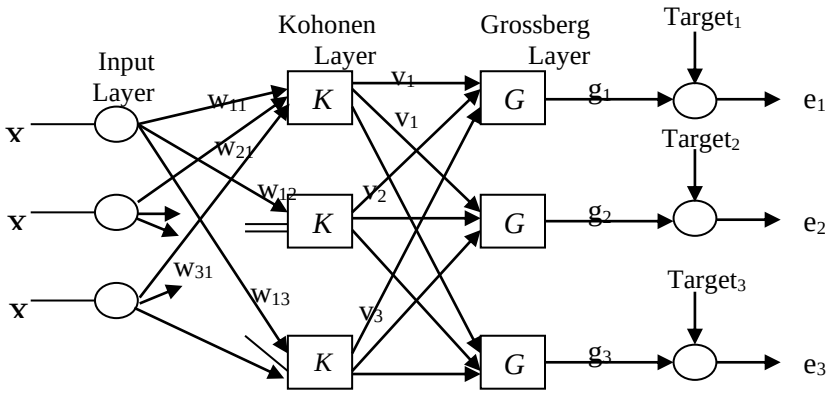

Fig. 1. counter propagation network[27][28]

The training steps of the counter propagation network(CPN) [29][30]as follows:

1. A vector pair (x,y) of the training set, is selected in random.
2. Normalize the input vector x to obtain x' by the equation (7):

$$x'=\frac{x_i}{\sqrt{\sum_j x_j^2}} \tag{7}$$

3. the weights are obtain as equation(8)

$$w=x' \tag{8}$$

namely, the weight vector of the wining kohonen neuron (the j th neuron in the kohonen layer) equals (best approximates) the input vector.

4. In the hidden competitive layer the distance between the weight vector and the current input vector is calculated for each hidden neuron j according to the equation(9)

$$D_j=\sqrt{\sum_{i=1}^{k}(x_{j\_}w_{ij})^2} \tag{9}$$

where k is the number of the hidden neurons and w_{ij} is the weight of the synapse that joins the i th neuron of the input layer with the j th neuron of the kohonen layer.

5. The winner neuron W of the kohonen layer is identified as the neuron with the minimum distance value D_j.
6. the synaptic weights between the winner neuron W and all neuron of the input layer are adjusted according to the equation(10)

$$w(t+1)=w(t)+\alpha[x-w(t)] \tag{10}$$

where α coefficient is known as the kohonen learning rate.

7. The weight between kohonen layer and Grossberg layer v_{ij} obtained at the same way to obtain w_{ij} weight between input layer and kohonen layer as in equation (8) above.
8. Obviously, only weights from non-zero kohonen neurons (non-zero Grossberg layer inputs) are adjusted. Weight adjustment as follows:
$$v_{ij}(t+1) = v_{ij}(t) + \beta\left[T_i - v_{ij}(t)\, k_j\right] \quad (11)$$
T_i being the desired outputs(targets), β is small number that represented the learning rate of Grossberg layer.
9. A major asset of the Grossberg layer is the ease of its training. First the output of the Grossberg layer are calculated as in equation(12)
$$g_i = \sum v_{ij} k_j = v_{ih} k_h = v_{ih} \quad (12)$$
k_j being the kohonen layer outputs and v_{ij} denoting the Grossberg layer weights.

10- HYBRID COUNTERPROPAGATION NETWORK WITH FCM

Counterpropagation developed by Hecht-Nielsen. Can be generalized to design a Fuzzy counterpropagation network, by extending the two layers (kohonen's layer and Grossberg layer) to a fuzzy counterpropagation network. The basic objective of this network is to cluster the input patterns, in each a way that total Euledian distance between each pattern and its nearest cluster centroid is minimum in kohonen layer, and we take the minimum distance output for each winner neuron in kohonen layer and maximum output neuron in Grossberg layer. A novel method is proposed in this research by using fuzzy c-means algorithm in Grossberg layer which is called FCPN, and steps (4,5) in the following algorithm were used to implement the above algorithm which has been applied by using kdd 99 dataset and NSL-kdd. The algorithm for fuzzy counterpropagation is shown below.

1. A vector pair (x, y) of the training set, is selected randomly. It is normalized and used as an input to obtain the weight by the equation(7) and (8) respectively.
2. Compute the distances $d(x_k, w_i)$ from the input pattern x_k to each of the competing neurons w_i.
3. Compute the membership of the winner neuron based on the distance measure $d(x_k, w_i)$.
4. Update the weight associated with each neuron. The weight updation is performed in accordance to the following rule.
$$w_i(t+1) = w_i(t) + \alpha\, z_i(t)\left[x_k - w_i(t)\right] \quad (13)$$
where z_i is the fuzzy scaling function given by:
$$z_i = (\mu_{ik})^m$$
where

$$\mu_{ik} = \left[\sum_{p}^{c} \left(\frac{D_{ik}}{D_{pk}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (14)$$

and $D_{ik} = d(x_k, w_i)$. The scaling function z_i depends on the fuzzy generator m which is a real number greater than 1.

5. Compute the membership between the winner neuron and Grossberg layer based on the distance measure $d(k_j, v_i)$. And update the weight associated with each neuron. The weight updation is performed in accordance to the following rule.

$$v_{ij}(t+1) = v_{ij}(t) + \beta\, z_i(t)\left[T_i - v_{ij}(t) k_j\right] \quad (15)$$

where z_i is the fuzzy scaling function given by:

$z_i = (\mu_{ij})^m$

Where:

$$\mu_{ij} = \left[\sum_{p}^{c} \left(\frac{D_{ij}}{D_{pj}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (16)$$

and $D_{ij} = d(k_j, v_i)$. The scaling function z_i depends on the fuzzy generator m which is a real number greater than 1.

6. Calculate the output of Grossberg as equation(12).

11- EXPERIMENTS AND RESULTS

11.1 *EXPERIMENT 1*

In the First stage we applied fuzzy c-means clustering algorithm FCM, counterpropagation network (CPN), and Fuzzy counterpropagation network(FCPN) on 10%kdd file data set that contains (494020) records. In the first experiment we apply these three algorithms (FCM, CPN, FCPN)to classify this data set into 5 classes or clusters, One for normal and the reset classes for types of attacks {Dos, probe, U2R, R2L}. The input layer in CPN and FCPN contain 41 node according features number in dataset, and Kohonen layer and Grossberg layer consist of 5 node one for normal and others for the main type of attacks. Table(3) shows the result clustering after training these three algorithms (FCM, CPN and FCPN). The results of classification rate from equation (17) obtained is 100% to classify data into 5 classes one class for normal behavior and 4 classes for different types of attacks. Table (4) shows the results after applying these three algorithms. Classification rate that obtained from all these algorithms is 100%.

$$classification\_rate\,(\mathrm{CR}) = \frac{number\ of\ classified\ \ patterns}{total\ number\ of\ \ patterns} \times 100 \quad (17)$$

Table 3: The clustering results after training FCM, CPN and FCPN algorithms to classify data set into 5 clusters

Amount	Type of attack	Samples rate
97277	Normal	19.690903
391458	Dos	79.239302
52	U2R	0.10526
1126	R2L	0.227926
4107	Probe	0.831343

Table 4: Results of the (FCM, CPN and FCPN)algorithms

Type of Clustering algorithms	Iteration number	Time second	CR%
FCM	26	132.6	100%
CPN	10	1428.86	100%
FCPN	5	1164.05	100%

And in the second stage of this experiment "corrected Kdd file" data set that consist of (311029) records were used in the testing stage on fuzzy c-means clustering algorithm. Table (5) show the results of the testing "corrected Kdd" file in FCM with detection rate for each attack type and for normal. In which the normal behavior got the higher detection rate is equal (97.813).

Table 5: Results of testing stage of FCM Algorithm

Type	Input	Output	DR%
Normal	60593	61948	97.813
Dos	229853	164611	71.616
Probe	4166	44212	9.428
U2R	70	0.0	0.0
R2L	16347	35795	45.668

After testing data in counterpropagation network (CPN) algorithm. Normal and Dos, Probe was obtained higher detection rate equal (100%), but R2L got detection rate equal (99.9574), U2R dose not detected. Which are shown in table(6).

Table 6: Results of testing stage of CPN Algorithm

Type	Input	Output	DR%
Normal	60593	60593	100
Dos	229853	229853	100
Probe	4166	4166	100
U2R	70	0.0	0.0
R2L	16347	16417	99.9574

But when applying Fuzzy counterpropagation (FCPN) algorithm on testing dataset (311029) in testing stage. Normal and all attacks got higher detection rate equal (100%) and false alarm rate equal(0.0) shown in table (7). This algorithm FCPN is the best method than other algorithms FCM, CPN.

Table 7: Results of testing stage of FCPN Algorithm

Type	Input	Output	DR%
Normal	60593	60593	100
Dos	229853	229853	100
Probe	4166	4166	100
U2R	70	70	100
R2L	16347	16347	100

Finally table (8) shows the comparisons between three algorithms FCM, CPN, and FCPN for 5 classes. with over all detection rate that obtained for FCM is equal (98.543) and false alarm equal (2.236). As shown in this table, FCPN got the higher detection rate equals to (100%) and low false alarm equals to (0.0), and the CPN was obtained detection rate equals to (99.977) and also low false alarm equals to(0.0), while FCM algorithm got detection rate equals to (98.543) and false alarm equals to(2.236). then the FCPN is the best algorithm.

Table 8. Comparison between FCM, CPN and FCPN Clustering Algorithms

Performance measure	FCM	CPN	FCPN
Normal detection	61948	60593	60593
Attack detection	244548	250366	250436
Detection rate_normal	97.813	100	100
Detection rate_attack	97.649	99.972	100
False_alarm rate	2.236	0.0	0.0
Detection_rate	98.543	99.977	100
Times	2.7 second	330.831 second	329.053 second
Iterations	1	1	1

Figures (2,3, and 4) show the relationship between FCM, CPN, FCPN algorithms with detection rate and false alarm rate and times respectively.

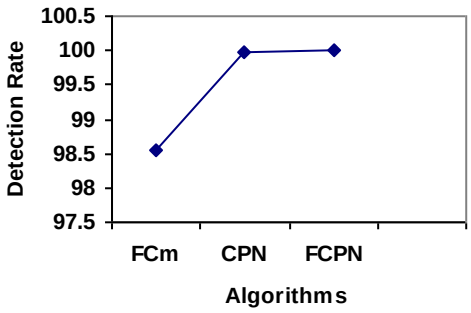

Fig 2. relationship between algorithms with False Alarm Rate

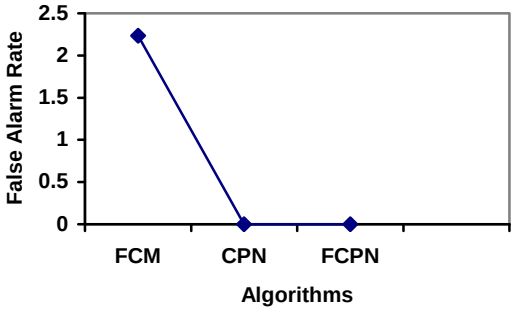

Fig. 3. relationship between algorithms with Detection Rate

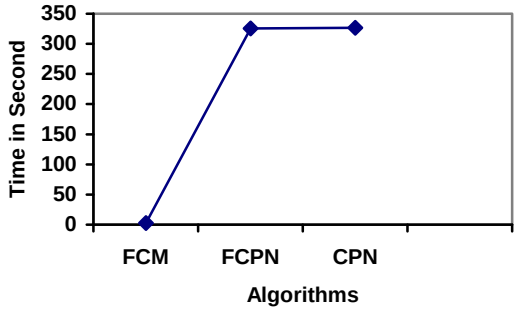

Fig 5: relationship between algorithms with time

The same dataset (494020) records were used after preprocessing it in the training stage to classify it into 2 classes, Table(9) shows the results of experiment after training FCM, CPN, and FCPN on this data set. Table (10) shows the results after applying these three algorithms, Classification rate that obtained from these algorithm is 100%.

Table 9: The clustering results after training FCM, CPN, FCPN algorithms to classify data set into 2 clusters

Amount	Type of attack	Samples rate
396743	Attack	80.309097
97277	Normal	19.690903

Table (10): results of the (FCM, CPN, FCPN) algorithms

Type of clustering Algorithms	Iteration number	Time second	CR%
FCM	16	29.1	100%
CPN	10	1422.64	100%
FCPN	5	1166.78	100%

The ‘corrected kdd‘ file that contain (311029) records were used in the testing state for FCM, CPN, and FCPN, table (11) shows the testing results after applying FCM, CPN, FCPN algorithms with the higher detection rate for each of attack and normal are equal (100%) and low false alarm equal (0.0). but it takes difference times.

Table 11. The results of testing state using FCM, CPN, FCPN Algorithms

Type	Input	Output	DR%
Normal	60593	60593	100
Attack	250436	250436	100

11.2 *EXPERIMENT 2*

In the second experiment we use kdd cup 2009 dataset consist of (kddTrain NSL and kddTestNSL) files. (kdd train NSL) file was used in training stage that contain (125973) records to classify this data into 5 class by using FCM, CPN, FCPN algorithms. Table(12) shows the

clustering results after training this algorithms (FCM, CPN, FCPN). Table (13) shows the results after applying these three algorithms. The result of classification rate obtained is 100% to classify data into 5 classes one class for normal behavior and 4 classes for different types of attacks.

Table 12: The clustering results after training FCM, CPN, FCPN algorithms to classify NSL data set into 5 clusters

Amount	Type of attack	Samples rate
52	U2R	0.041279
67343	Normal	53.458281
11656	Probe	9.789852
995	R2L	0.789852
45927	Dos	36.457812

Table 13: Results of the (FCM, CPN, FCPN) algorithms

Type of Clustering algorithms	Iteration number	Time second	CR%
FCM	26	30.4	100%
CPN	10	195.485	100%
FCPN	5	126.049	100%

The kddTest (NSL) that contain (22544) records was used in the testing stage on FCM algorithm. Table (14) shows the results of the testing "kddTest (NSL)" file in FCM with detection rate for each attack type and for normal behavior. In which the normal behavior got the higher detection rate is equal (100).

Table 14: Results of testing stage of FCM Algorithm on NSL dataset

Type	Input	Output	DR%
Normal	9771	9711	100
Dos	7458	7427	99.584
Probe	2421	0.0	0.0
U2R	67	0.0	0.0
R2L	2887	0.0	0.0

After testing data in counterpropagation network (CPN) algorithm. Normal was obtained higher detection rate equal (100%), Probe, U2R dose not detected. Which are shown in table(15).

Table 15: Results of testing stage of CPN Algorithm on NSL dataset

Type	Input	Output	DR%
Normal	9771	9711	100
Dos	7458	9879	75.493
Probe	2421	0.0	0.0
U2R	67	0.0	0.0
R2L	2887	2954	97.732

And when use FCPN to testing dataset. Detection rate is enhanced and this algorithm detect normal, Dos, Probe, R2L, but it dose not detect U2R attack. Table (16) shows the results of testing stage of FCPN algorithm.

Table 16: Results of testing stage of FCPN Algorithm on NSL dataset

Type	Input	Output	DR%
Normal	9771	9711	100
Dos	7458	7458	100
Probe	2421	2421	100
U2R	67	0.0	0.0
R2L	2887	2954	97.7313

And table (17) shows the comparisons between three algorithms FCM, CPN, and FCPN for 5 classes. with over all detection rate that obtained for FCM is equal (76.0202), while the detection rate that obtained for CPN is equal(88.964), and the detection rate that obtained for FCPN is equal (99.703), and false alarm for all algorithms equal (0.0).

Table 17: Comparison between FCM, CPN and FCPN Clustering Algorithms for 5 Classes of NSL Dataset

Performance measure	FCM	CPN	FCPN
Normal detection	9711	9711	9711
Attack detection	7427	10345	12766
Detection rate_normal	100	100	100
Detection rate_attack	15.748	80.612	99.478
False_alarm rate	0.0	0.0	0.0
Detection_rate	76.020	88.964	99.703
Times	0.2 second	2.7 second	2.6 second
Iterations	1	1	1

Figures (5 and 6) show the relationship between FCM, CPN, FCPN algorithms with detection rate and times respectively.

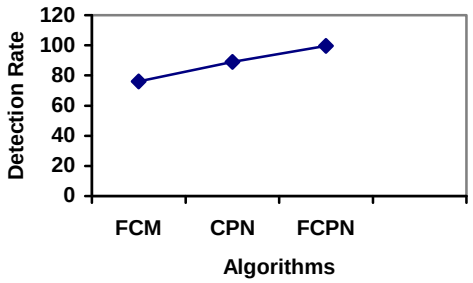

Fig. 5: relationship between algorithms with Detection Rate

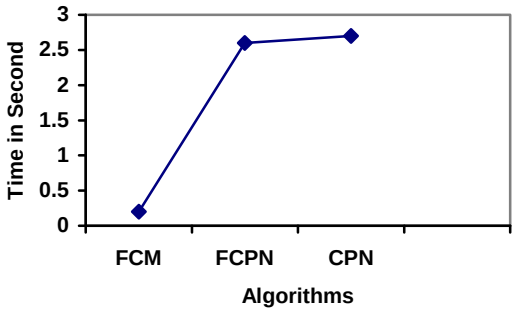

Fig. 6: relationship between algorithms with Time

The same data set (125973) records were used after preprocessing it in the training stage to classify it into 2 classes, Table(18) shows the results of experiment after training FCM, CPN, FCPN on this data set.

Classification rate that obtained from this algorithm is 100%. Table(19) shows the results after applying these three algorithms.

Table 18: The clustering results after training FCM, CPN and FCPN algorithms to classify NSL data set into 2 clusters

Amount	Type of attack	Samples rate
58630	Attack	46.541719
67343	Normal	53.458281

Table 19: Results of the (FCM, CPN and FCPN) algorithms

Type of Clustering algorithms	Iteration number	Time second	CR%
FCM	16	7.5	100%
CPN	10	192.505	100%
FCPN	5	120.932	100%

The ‘kddTest NSL‘ file that contain (22544) records were used in the testing state for FCM, CPN, and FCPN, table (20) shows the testing results after applying FCM, CPN, FCPN algorithms with the higher detection rate for each of attack and normal are equal (100%) and low false alarm equal(0.0). But it takes difference times.

Table 20: The results of testing state using FCM, CPN, FCPN Algorithms

Type	Input	Output	DR%
Normal	**9711**	**9711**	**100**
Attack	**12833**	**12833**	**100**

The detection rate and classification rate of some previous works on kdd data set and algorithms and the hybrid model in this research as shown in table(21).

Table 21: Comparison results of (FCM,CPN and FCPN) algorithms with previous works

	Algorithm type	Dataset	Normal	Attack	Classification rate	Detection rate
	FCM[8]	Training (22133)	*	*	99.9	
	Parallel fuzzy ART MAP [9]	Training set	*	*	*	80.14%
	Parallel fuzzy ART MAP [9]	Testing set	*	*	*	80.52%
	Parallel B.P[9]	Testing data				81.37%
	SIB[10]	Training and testing set 1000				85.5%
	SSGBML[11]	Training data				97.45%
experiment 1	FCM	Training set	100	100	100	
	FCM	Testing set	97.813	97.649		98.543%
(5 classes)	CPN	Training set	100	100	100	
Kdd cup 99	CPN	Testing set	100	99.972		99.977%
	FCPN	Training set	100	100	100	
	FCPN	Testing set	100	100		100%

experiment 1	FCM	Testing &training set	100	100	100	100%
(2 classes)	CPN	Testing &training set	100	100	100	100%
Kdd cup 99	FCPN	Testing &training set	100	100	100	100%
experiment 2	FCM	Training set	100	100	100	
	FCM	Testing set	100	15.748		76.020%
(5 classes)	CPN	Training set	100	100	100	
Kdd cup	CPN	Testing set	100	80.612		88.964%
NSL	FCPN	Training set	100	100	100	
	FCPN	Testing set	100	99.478		99.703%
experiment 2	FCM	Testing &training set	100	100	100	100%
(2 classes)	CPN	Testing &training set	100	100	100	100%
Kdd cup						
NSL	FCPN	Testing &training set	100	100	100	100%

12- CONCLUSION

In this research fuzzy c-means clustering algorithm and neural networks (counterpropagation network and fuzzy counterpropagation network) were applied to classify kdd cup 99 and NSL-KDD data set into 5 classes one for normal behavior and others for types of attacks, and classify the same data set into 2 classes one for normal and other for attacks, and these algorithms satisfied very good results in classification and detection:

- Classification improvement: the applied approaches (FCM, CPN, and FCPN) improved a high classification rate 100% in training stage.
- Architectural framework improvement: the application of these approaches made the intrusion analysis engine more simple and efficient.
- Detection improvement: these approach obtained a high detection rate and low false alarm for kdd cup 99 and NSL-KDD dataset. It has been found that FCPN algorithm is the best approaches.

REERENCES

[1] Gomez J., Dasgupta D., "Evolving Fuzzy Classifiers for Intrusion Detection", proceeding of the 2002 IEEE.

[2] Song D., Heywood M., Zincir-Heywood A., "Training Genetic Programming on Half a Million Patterns: An Example from Anomaly Detection", IEEE Transaction on evolutionary computation, 2005.

[3] Vemuri V., "ENHANCING COMPUTER SECURITY WITH SMART TECHNOLOGY", TK5105.59.E62, 2005.

[4] Sabnani S. V., "Computer Security: A Machine Learning Approach", Royal Holloway, University of London, 2008.

[5] Badii A., Patel D., Bragg H., "Design and Evaluation of Intelligent Data Classifier Based Intrusion Detection System", www.iseing.org/, 2007.

[6] Panda M., Patra M., "SOME CLUSTERING ALGORITHMS TO ENHANCE THE PERFORMANCE OF THE NETWORK INTRUSION DETECTION SYSTEM", journal of theoretical and applied information technology, pp.795-801, 2008.

[7] Chimphlee W., Abdullah A., Sap M., Chimphlee S., Srinoy S., "A rough-fuzzy Hybrid Algorithm for Computer Intrusion Detection", the international Arab journal of information Technology, Vol.4, No.3, 2007.

[8] Jawhar M., Mehrotra M., "Design network intrusion detection system using hybrid fuzzy-neural network". International Journal of Computer Science and Security, Volume(4): Isseu(3), 2010.

[9] Siddiqui M., "high performance data mining techniques for intrusion detection", thesis 2004.

[10] Panda M., Patra M., "A novel classification via clustering method for Anomaly based network intrusion detection system", International Journal of Recent Trend in Engineering, Vol 2, No. 1, November 2009.

[11] Al-Sharafat W., Naoum R., "Adaptive Framework for Network Intrusion Detection by Using Genetic-Based Machine Learning Algorithm", IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.4, 2009.

[12] Abdollah M., Yaacob A., Sahib S., Mohamed I., Iskandar M., "Revealing the Influence of feature Selection for Fast Attack Detection", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.8, 2008.

[13] Anuar N., Sallehudin H., Gani A., Zakari O., "IDENTIFYING FALSE ALARM FOR NETWORK INTRUSION DETECTION SYSTEM USING HYBRID DATA MINING AND DECISION TREE", Malaysian Journal of Computer Science, Vol. 21(2), 2008.

[14] Kdd-cup data set. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[15] Lakhina S., Joseph S., verma B., "Feature Reduction using Principal Component Analysis for Effective Anomaly-Based Intrusion Detection on NSL-KDD", International Journal of Engineering Science and Technology Vol. 2(6), 2010.

[16] http://nsl.cs.unb.ca/NSL-KDD/

[17] Faraoun K. M., Boukelif A., "Neural Network Learning Improvement using the K-Means Clustering Algorithm to Detect Network Intrusion", International Journal of Computation Intelligence 3; 2 2007.

[18] Betanzos A., Marono N., Fortes F., Romero J., Sanchez B., "Classification of computer intrusions using functional networks. A comparative study", ESANN, European Symposium on Artificial Neural Networks, 2007.

[19] Zhang C., Jiang J., Kamel M., 2005, "Intrusion detection using hierarchical neural networks", Pattern Recognition Letters 26, 779-791.

[20] Vlad Z., Ofelia M., Maria T., "fuzzy clustering in an intelligent Agent for diagnosis establishment", inter-eng, 2009.

[21] Maji P., Pal S., "Rough Set Based Generalized Fuzzy C-Means Algorithm and Quantitative Indices ", IEEE, vol.37, no.6, 2007.

[22] Gomathi M., Thangaraj P., "A New Approach to Lung Image Segmentation using Fuzzy Possibilistic C-Means Algorithm", (IJCSIS) international journal of computer science and information security, Vol.7, No.3, 2010.

[23] Acharya T., Ray A., "IMAGE PROCESSING", 2005.

[24] Saad M., Alimi A., "Modified Fuzzy Possibilistic C-Means", proceeding of the international Multi Coference of Engineers and Computer Scientists Vol. I, IMESCS 2009.

[25] Thomas B., Raju G., Wangmo S., "A modified fuzzy c-means algorithm for natural data exploration", World Academy of science, Engineering and technology 49, 2009.

[26] Wasserman P., "Neural Computing Theory and Practice", New York 1989.

[27] Taylor B. J., "Methods and Procedures for the Verification and Validation of Artificial Neural Networks", 2006.

[28] Burks T. F., Shearer S. A., Heath J. R., Donohue K. D., "Evaluation of Neural-network Classification for weed Species Discrimination", Biosystems Engineering, 2005.

[29] Durai S., Saro E., "Image Mapping with Cumulative Distribution Function for Quick Convergence of Counterpropagation Neural Networks in Image Compression", World Academy of science, Engineering and Technology 16, 2006.

[30] Daniel Graupe, "PRINCIPLES OF ARTIFICIAL NEURAL NETWORKS", Advanced Series on Circuits and Systems-Vol. 6, 2007.