Hindawi
Security and Communication Networks

Research Article

A Novel approach for classify MANETs attacks with a neutrosophic intelligent system based on genetic algorithm

Haitham Elwahsh[1*] , Mona Gamal[2], A. A. Salama[3], I.M. El-Henawy[4]

[1] Computer Science Department, Faculty of Computers and Information, Kafrelsheikh University, Kafrelsheikh 33516, Egypt;
[2] Information System Department Faculty of Computers and Information, Kafrelsheikh University, Kafrelsheikh 33516, Egypt;
[3] Department of Mathematics and Computer Science, Faculty of Sciences, Port Said University, Port Said 522, Egypt;
[4] Computer Science Department, Faculty of Computers and Information, Zagazig University, Zagazig, Egypt;
* Correspondence should be addressed to Haitham S. Elwahsh; Haitham.elwahsh@fci.kfs.edu.eg.

Abstract:

Recently designing an effective Intrusion Detection Systems (IDS) within Mobile Ad-hoc Networks Security (MANETs) becomes a requirement because of the amount of indeterminacy and doubt exist in that environment. Neutrosophic system is a discipline that makes a mathematical formulation for the indeterminacy found in such complex situations. Neutrosophic rules compute with symbols instead of numeric values making a good base for symbolic reasoning. These symbols should be carefully designed as they form the propositions base for the neutrosophic rules (NR) in the IDS. Each attack is determined by membership, nonmembership and indeterminacy degrees in neutrosophic system. This research proposes a MANETs attack inference by a hybrid framework of Self Organized Features Maps (SOFM) and the Genetic Algorithms (GA). The hybrid utilizes the unsupervised learning capabilities of the SOFM to define the MANETs neutrosophic conditional variables. The neutrosophic variables along with the training data set are fed into the genetic algorithm to find the most fit neutrosophic rule set from a number of initial sub attacks according to the fitness function. This method is designed to detect unknown attacks in MANETs. The simulation and experimental results are conducted on the KDD-99 network attacks data available in the UCI machine-learning repository for further processing in knowledge discovery. The experiments cleared the feasibility of the proposed hybrid by an average accuracy of 99.3608 % which is more accurate than other IDS found in literature.

Keywords: Security; MANETs; Neutrosophic system; Genetic Algorithm; Intrusion Detection System (IDS).

1. INTRODUCTION

Mobile Ad hoc Network (MANET) is a group of wireless mobile hosts shaping an impermanent system without communication infrastructure. This network may change faster and unpredicted. The exceptional qualities of MANETs give a foe the chance to dispatch numerous assaults against specially ad-hoc networks [1, 2]. Security in MANETs is the most imperative worry for the fundamental usefulness of system. Accessibility of system administrations, privacy and uprightness of the information can be accomplished by guaranteeing that security issues have been met.

MANETs suffer from the unwell effects of security assaults on accounts, in respect of its characteristics like open medium, changing its topology powerfully, absence of focal checking and administration, agreeable calculations and no unmistakable safeguard system. These elements have changed the war zone circumstance for the MANETs versus the security menace. These features make MANETs more impotent to be a victim by an assailant from inside the network. Remote connections likewise make the MANETs more vulnerable to assaults which make it simpler for the assailant to enter the system and access the progressing communication [3, 4 and 5]. Mobile nodes inside the connection area can hear and cooperate in the network.

MANETs must have a protected path for transmission and correspondence, this is very defying, and fundamental issue as there is expanding dangers of assault on the Mobile Network. Security is the cry of the day. To give secure transmission and communication, designers must comprehend distinctive kinds of assaults and their impacts on the MANETs. MANET can suffer from various threats such as routing table over flow, flooding attack, Wormhole attack, Sybil attack, Denial of Service (DoS), Black hole attack, selfish node misbehaving. MANET is more open to these sorts of assaults because the MANETs behavior which base station for network controlling, no mandate facility also the topology changing with limited resources. Therefore, designer's use of Intrusion Detector Learning Software to detect network intrusions and protect a computer network from unauthorized users, including perhaps insiders. Building a predictive model (i.e. a classifier) using intrusion detector is a learning task. The detector should be capable of distinguishing between ``Upnormal'' connections, called intrusions or attacks, and normal connections. The 1998 DARPA Intrusion Detection Evaluation Program was done and controlled by MIT Lincoln Labs. Intrusion Detection Systems (IDS) suffer from uncertainty and imprecise nature. Neutrosophy theory introduced by Smarandache [6] could be utilized to suit the ambiguity nature of the IDS. In [7] Salama presented the principle of Neutrosophic Set (NS) and mathematical Theory, to define any situation by a ternary crisp build. Salama et al. Work [6, 8] formulated a beginning to new fields of neutrosophic theory in computer discipline. The neutrosophic indeterminacy assumption is very significant in many of circumstances such as information fusion (collecting data from various sensors). Also, NS is a conceivable common traditional system that generalize the principle of the traditional set, Fuzzy Set (FS) [9] and Intuitionistic Fuzzy Set (IFS)) [10], etc. NS 'A' determined on universe U. $\boldsymbol{x} = \boldsymbol{x}(\mathbf{T}, \mathbf{I}, \mathbf{F}) \in A$ with $\mathbf{T}, \mathbf{I}$ and $\mathbf{F}$ are defined over the interval $]\mathbf{0}^{-}, \mathbf{1}^{+}[$. $\mathbf{T}$ is the truth-MEMEBERSHIP , **I** is the INDETERMINACY and F is the falsity-MEMEBERSHIP degrees on the set **A**.

Designing a neutrosophic IDS is a proper solution in handling vague circumstances. The neutrosophic IDS is formed of two sub phases: the preprocessing stage and the network attacks classification stage. The preprocessing stage is concerned by formulating the network features in a format appropriate for the classification. The KDD network data [11] is reformatted into neutrosophic form $(\mathbf{x}, \boldsymbol{\mu}_{\mathbf{A}}(\mathbf{x}), \boldsymbol{\sigma}_{\mathbf{A}}(\mathbf{x}), \boldsymbol{\nu}_{\mathbf{A}}(\mathbf{x}))$ where x is the value of feature, $\boldsymbol{\mu}_{\mathbf{A}}(\mathbf{x})$ is the MEMEBERSHIP (MEM), $\boldsymbol{\sigma}_{\mathbf{A}}(\mathbf{x})$ is the INDETERMINACY (I) and $\boldsymbol{\nu}_{\mathbf{A}}(\mathbf{x})$ is the NONMEMEBERSHIP (NON_MEM) degrees of the x in the feature space. The Self Organized Features Maps (SOFMs)[**14**] , machine learning technique, was used to prepare the neutrosophic KDD through learning the MEM, NON_MEM and I functions of the KDD network attacks data [11] downloaded from the UCI repository for further processing in knowledge discovery [12]. After converting the traditional KDD data into a neutrosophic one, the Genetic algorithms (GAs) [13] searching mechanism is utilized in finding a set of neutrosophic (if –then) rules to classify MANETs attacks. The GA initial population is a set of randomly generated individuals. Each individual represents a structure of a neutrosophic (if-then) classification rule. During the GA iterations, the selection, crossover and mutation processes are applied on the populations for generating new fit offsprings. The fitness of the offsprings is quantified by the concept of neutrosophic correlation co-efficient introduced by Salama et al. in [14]. The final population will serve as the neutrosophic rule set for the neutrosophic IDS for the KDD data set. The testing procedure apply new instances of KDD instances (not used during training) to measure the accuracy levels of the obtained neutrosophic IDS. The experiments compared the proposed neutrosophic IDS with a number of well-known classifiers in literature like C4.5 [15], SVM [15], ACO [16], PSO [17] and EDADT [18]. The Comparisons proved the feasibility of the proposed neutrosophic IDS in terms of accuracy level of average 99.3608 %and false alarm rates of average 0.089 % .

The rest of this paper is organized as follows: section 2 presents the theories and overviews. Section 3 proposes the GA in generating the neutrosophic (if –then) rules (inference engine). Experimental results and conclusion are shown respectively in section 4 and 5.

2. THEORIES AND OVERVIEW

2.1 Motivation and Related Work

Mobile Ad hoc Network (MANET) is a system of wireless mobile nodes that dynamically self-organized in arbitrary and temporary network topologies without communication infrastructure. This network may change quickly and unforeseeable. The unique characteristics of MANET give an adversary the opportunity to launch numerous attacks against ad-hoc networks [2]. Security in Mobile Ad Hoc Network is the most important concern for the basic functionality of network [5]. The Security issues have been met when Availability of network services, confidentiality and integrity of the data can be achieved by assuring that. MANET often suffer from security attacks because of its features like open medium, changing its topology dynamically, lack of central monitoring and management, cooperative algorithms and no clear defense mechanism. These characteristics have changed the challenging situation of the MANET against the security threats [3, 4]. Detecting these intrusions is a difficult procedure that is full of uncertainty and indeterminacy problems. Neutrosophic [10] discipline provides a logical base for dealing with uncertainty problems. In neutrosophic theory, truth-membership, falsity and indeterminacy values are independent and calculated separately. Hence, the indeterminacy of the MANET attack could be identified through the intrusion detection process. This research is interested in integrating the neutrosophic concepts with the artificial intelligence search capabilities to produce an accurate Neutrosophic Intrusion Detection System (NIDS).

In literature, many researchers were concerned by the security of the MANET. They proposed many IDS to detect security threats and find a model to prevent their consequences. For Example, Ektefa [15] compared C4.5 & SVM to show the performance of both algorithm and FAR values too. Among these two C4.5 works better compared to other. Since the performances of a classifier are often evaluated by an error rate and it does not suit the complex real problems, in particular multiclass. Holden [16] has proposed hybrid PSO algorithm that can deal with nominal attributes without going for the both conversion and nominal attribute values. To overcome the drawback (features) that the PSO/ACO algorithm lacks. The proposed method shows simple rule set efficiently to increase in accuracy. Likewise Ardjani [17] applied SVM with PSO as (PSO-SVM) to optimize the performance of SVM. 10-Fold cross validation is done to estimate the accuracy. It utilizes the advantage of minimum structural risk with global optimizing features. The result shows better accuracy with high execution time. Since there [19] is an existence of multidimensional data set, it is necessary to extract the features and also to remove the redundant and inconsistent features that affects classification. Based on this, information gain and genetic algorithm has been combined to select the significant features. This method shows better accuracy when features are selected than individually applied.

2.2 Genetic Algorithms

Genetic algorithms (GAs) [13] [20] [21] [22], utilize the Darwin's evolution theory of life. Massive machines [23] and transporters [1] are obvious and frequent applications of the GAs. Through the progress of the GA, the fit individuals are passed to the next generation. This guarantee the survival and reproduction of the best individuals over the worst. Upon assigning a problem, GAs evolve to find the optimal Solution. The process begins with a population of random collection of solutions (chromosomes). During selection, the new population is formed from the best individuals. The individual fitness should be the criteria of the selection. The chance of reproducing fit individuals is guaranteed. The fitness function is defined by the issued problem. The new population (having the best individuals) goes through the crossover or mating process. The motivation of producing better populations during the GA process is the obvious hope of finding the optimal solution. The selection- reproduction processes are repeated until the optimal solution is found or reach the end of iterations with an acceptable error rate. The algorithm is declared below.

```
Begin
    Initialize a population
    Evaluate fitness for each individual
    While (! stop condition) do
            Proceed the cross over process.
            Proceed the mutation process.
            Evaluate new individual.
            Select individual to replace and their replacement.
            Update stop condition.
    End
    Return best solution
    End
```

3. DESIGNING THE PROPOSED METHOD

Intrusion Detection System (IDS) is an essential security part for any online network nowadays. An intrusion is "a collection of actions that try to comprehend the privacy, integrity or availability for various resources". Intrusion can likewise be characterized as "a collection of actions imagine to get unapproved assets, abuse rights, cause finish frameworks and systems smashed, diminish running intensity, or refuse any assistance". In this manner, IDS might be a framework to monitor events in PCs or systems and examinations and checking the frameworks uprightness and privacy. IDS could be figured as an arrangement of if-then rules that depict the potential intrusions of the network or systems. Finding the ideal arrangement of these rules is a vital search problem. These algorithms switch the problem in a particular space into a model by utilizing a chromosome-like data structure and develop the chromosomes using selection, recombination, and mutation operators. GA use as critical solving strategy and give ideal solution of the problem GA works on the Darwinian principle of reproduction. It is change set of individual objects, which each correlating fitness value into new generation of population and after that apply crossover and mutation function. The proposed hybrid combines SOFM and GA algorithms to produce the neutrosophic rules in two phases. The first phase sets the neutrosophic variables by creating the membership, non-membership and indeterminacy functions for the neutrosophic subsets of the variables. The implementation for the first stage is done by using SOFM from a prior research [12]. The outcome of this stage is passed to GAs [4] along with the training data to first randomly generate initial population, thereafter the neutrosophic correlation coefficient is used as a fitness function to pick out the most fit rules with regard to the training data. Afterwards, the test data is utilized to check for the precision of the rules created.

3.1 Formatting Neutrosophic KDD features using SOFM

In order to build a neutrosophic IDS, the system should be based on neutrosophic variables. The regular features in the KDD data set cannot be used in neutrosophic processing. Hence, reformatting the KDD features into neutrosophic ones is a preprocessing step in the intrusion discovery system. Self-Organized Feature Maps (SOFM) are unsupervised artificial neural networks that were used to define the neutrosophic variables [12]. SOFMs capabilities cluster inputs using self-adoption techniques. These capabilities were utilized in generating neutrosophic functions for the subsets of the variables. The SOFM are used to define the membership, non-membership and indeterminacy functions for the KDD data set features. The algorithm for generating the neutrosophic features definitions is cleared below.

Input: input_data vectors(Training_data set), Input_dim, output_dim,
Output: neutrosophic variable membership, nonmembership and indeterminacy functions
//membership function generation

1. Training_Data←Read_data(membership_data)
2. Membership_data← SOFM(Trainig_data, Input_dim, output_dim)
3. Draw (Membership_data)
4. Training_Data←Read_data(nonmembership_data)
5. Non_Membership_data← SOFM(Trainig_data, Input_dim, output_dim)
6. Draw (Non_Membership_data)
7. Indeterminancy← Calculate_ind(Membership_data, Non_Membership_data)
8. Draw (Indeterminancy)

End

Function SOFM
Input: Trainig_data, Input_dim, output_dim
Output: Output _Function
Initialize_SOFM (input_neurons, output_neurons)
Randomly_Initialize_SOFM_Weights ()
While Error>threshold **Do**
Foreach Record in **Training_Data**
Input_Record ();

Function Update_weights
Input: Winning_neuron$\mathbf{q_j}$
Output: Update_weights

1. Find (Winning_neuron$\mathbf{q_j}$)
2. $\eta_{qj}[t] = \begin{cases} \mu\,[t] j \epsilon N_q \\ 0 \qquad j \notin N_q \end{cases}$
3. $w_j[t+1] = w_j[t] + \eta_{qj}[t]\big(x_n[t] - w_j[t]\big)$

Winning_neuron$\mathbf{q_j}$=$\mathbf{q}(\boldsymbol{x_n}) = \min\limits_{\forall j} \lVert \boldsymbol{x_n} - \boldsymbol{w_j} \rVert$; Update_weights (Winning_neuron$\mathbf{q_j}$); **Endforeach** **Error =Calculate_ErrorRate ();** End while Retrieving_phase (); Output Function←Network_Weights) End fun	4. Output (Update_weights) 5. End fun
SOFM algorithm for generating the membership , non-membership and indeterminacy functions of the neutrosophic variable	

These definitions are used during the GA search in fitness calculation (the neutrosophic correlation co-efficient). The procedure of generating the neutrosophic IDS classification rules is introduced in the next section.

3.2 creating neutrosophic rules utilizing Genetic Algorithms

The neutrosophic knowledge based system is composed of a set of neutrosophic rules. (Figure 1). This procedure is in charge of designing the neutrosophic conditional rules via applying the processing power of GAs at random initial population. After that, utilizing the neutrosophic correlation co-efficient as a fitness function, it picks the most proper performers from the population. At the last stage, the population becomes the set of neutrosophic rules required for the neutrosophic inference engine of the IDS.

R1 $\underbrace{\mathbf{S_1, (\mu_A(x), \sigma_A(x), \nu_A(x))}}_{\mathbf{A_1}}$ $\underbrace{\mathbf{S_2, (\mu_A(x), \sigma_A(x), \nu_A(x))}}_{\mathbf{A_2}}$ $\underbrace{\mathbf{S_3, (\mu_A(x), \sigma_A(x), \nu_A(x))}}_{\mathbf{A_i}}$ …….. …….. $\mathbf{d_1,}$ $\underbrace{\mathbf{MIN_i\, (\mu_A(x), \sigma_A(x), \nu_A(x))}}$

R2 $\underbrace{\mathbf{S_1\, (\mu_A(x), \sigma_A(x), \nu_A(x))}}_{\mathbf{A_1}}$ $\underbrace{\mathbf{S_1\, (\mu_A(x), \sigma_A(x), \nu_A(x))}}_{\mathbf{A_2}}$ $\underbrace{\mathbf{S_2\, (\mu_A(x), \sigma_A(x), \nu_A(x))}}_{\mathbf{A_i}}$ …….. …….. $\mathbf{d_2,}$ $\underbrace{\mathbf{MIN_i\, (\mu_A(x), \sigma_A(x), \nu_A(x))}}$

Figure 1: Neutrosophic rules of the knowledge based system

An individual represents the possible solution or the possible neutrosophic rule which is composed of a set of conditional propositions (neutrosophic attributes) and the consequence decision attribute. Each neutrosophic attribute and the decision attribute will occupy one gene within the individual. To demonstrate the neutrosophic format, each gene will be represented by $\boldsymbol{A} \in$ S, $(\boldsymbol{\mu_A(x)}, \boldsymbol{\sigma_A(x)}, \boldsymbol{\nu_A(x)})$ where A is the neutrosophic feature that belongs to the subset S with degrees of membership $\boldsymbol{\mu_A(x)}$, non- membership $\boldsymbol{\nu_A(x)}$ and indeterminacy $\boldsymbol{\sigma_A(x)}$. The GA individual is presented in Figure 2.

$\boldsymbol{A_1} \in \mathbf{S_2}, (\boldsymbol{\mu}_{\mathbf{A}}(\mathbf{x}). \boldsymbol{\sigma}_{\mathbf{A}}(\mathbf{x}). (\mathbf{x}))$	$\boldsymbol{A_2} \in \mathbf{S_3}, (\boldsymbol{\mu}_{\mathbf{A}}(\mathbf{x}). \boldsymbol{\sigma}_{\mathbf{A}}(\mathbf{x}). \boldsymbol{\nu}_{\mathbf{A}}(\mathbf{x}))$	$\boldsymbol{A_3} \in \mathbf{S_2}, (\boldsymbol{\mu}_{\mathbf{A}}(\mathbf{x}), \boldsymbol{\sigma}_{\mathbf{A}}(\mathbf{x}), \boldsymbol{\nu}_{\mathbf{A}}(\mathbf{x}))$		$\boldsymbol{A_5} \in \mathbf{S_1}, (\boldsymbol{\mu}_{\mathbf{A}}(\mathbf{x}), \boldsymbol{\sigma}_{\mathbf{A}}(\mathbf{x}), \boldsymbol{\nu}_{\mathbf{A}}(\mathbf{x}))$	$\boldsymbol{d_1} \in \mathbf{S_3}$

$$\text{If } (A_1 \in s_2 \wedge A_2 \in s_3 \wedge A_3 \in s_2 \wedge A_5 \in s_1) \text{ then } d_1 \in s_3$$

Figure 2: the GA individual neutrosophic rules of a system has 5 feature and a dependent class

Note that A1, A2, A3, A5 are the neutrosophic attributes and s2, s3, s2, s1 are the neutrosophic subsets of the attributes. ^ is the logical and operator. Moreover, the previous neutrosophic conditional rule does not rely on A4 as a proposition and produces d1 as a decision within the neutrosophic subset s3. The flow chart for the generating neutrosophic rules procedure is provided in figure 3. This procedure is composed of 7 phases with an iterated loop until the predefined number of iterations is performed. The first phase generates a collection of random neutrosophic rules according to the previous format (figure 2) which forms the initial population. The second and third phases are the fitness evaluation and the selection process that determines the most fit individuals in the pool according to the neutrosophic correlation coefficient as a fitness function. The fourth phase is the crossover which produces new offspring from the fit individuals selected during the selection phase. The fifth phase is the mutation which switches one of the genes randomly to help increase the performance but this is accomplished under very rare circumstances. The sixth and seventh phases recalculate the fitness of the new offspring's and replace the children in place of their parents. These phases are repeated for a number of iterations from the third to the seventh. After completion, the final fit generation will form the set of neutrosophic conditional rules for the neutrosophic IDS. The GA procedure is straight forward and mostly used in the same way for most of classification applications. The main module that differentiate one application from the other is the fitness function calculation. In order to integrate the neutrosophic concepts within the GA, the neutrosophic correlation co-efficient is used to find the rules with the highest dependency between the neutrosophic conditional features and the decision attribute. The co-efficient equations are illustrated in the next section.

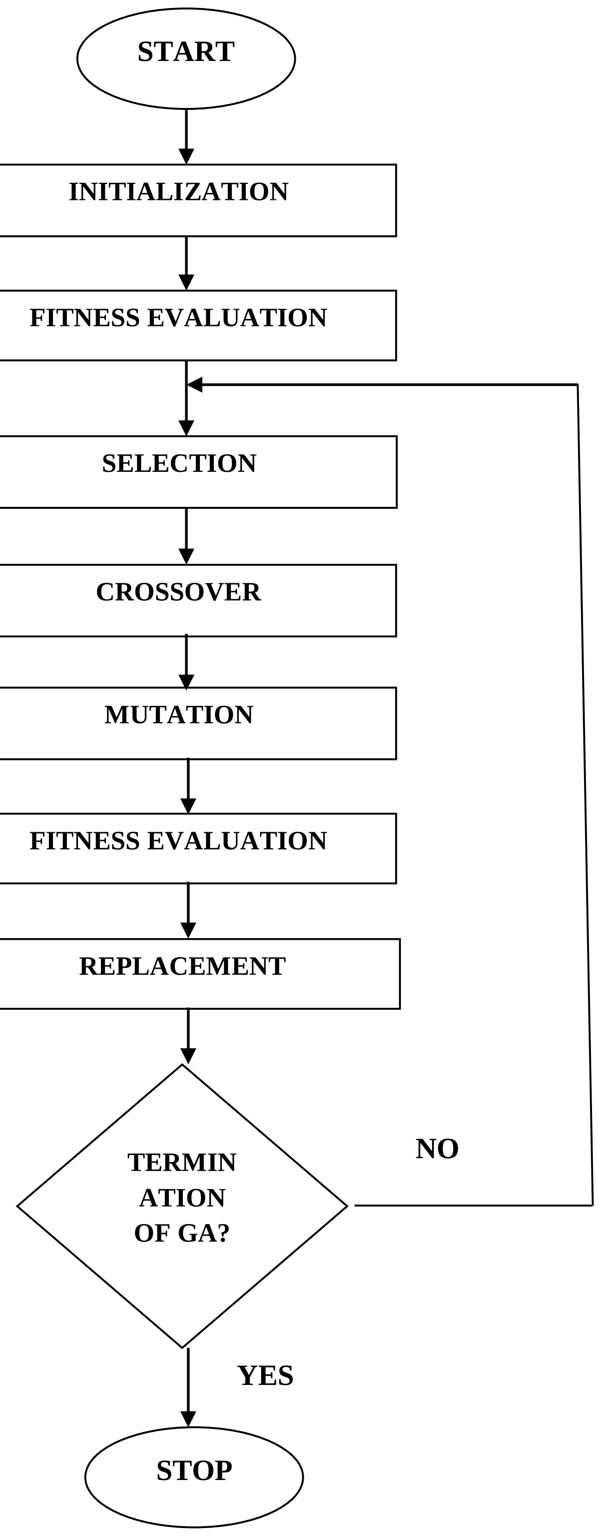


Figure 3: Genetic Algorithm Process

3.2 Fitness function

The neutrosophic correlation coefficient is used to measure the fit rules. The neutrosophic correlation coefficient measures the degree of relation between propositions and the decision attribute. The generated neutrosophic rules that maximizes this correlation will be the most fit rules in the population and will be selected and passed to the next generation during the GA process. Some operations on neutrosophic sets introduced and studied by Salama et al in 2012 [14]. For S and Y are two neutrosophic sets in a finite space x={x1, x2,…………, xn } , the correlation of neutrosophic sets S and Y is defined as follows:

$$\boldsymbol{C}\ (\boldsymbol{S},\boldsymbol{Y}) = \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \left\lfloor \left(\boldsymbol{\mu}_{\mathbf{S}}(\mathbf{x}_{\mathbf{i}}).\ \boldsymbol{\mu}_{\mathbf{Y}}(\mathbf{x}_{\mathbf{i}}) + \boldsymbol{\sigma}_{\mathbf{S}}(\mathbf{x}_{\mathbf{i}}).\ \boldsymbol{\sigma}_{\mathbf{Y}}(\mathbf{x}_{\mathbf{i}}) + \boldsymbol{\nu}_{\mathbf{S}}(\mathbf{x}_{\mathbf{i}}).\ \boldsymbol{\nu}_{\mathbf{Y}}(\mathbf{x}_{\mathbf{i}})\right)\right\rfloor \quad \text{(Eq.1)}$$

and the correlation coefficient of S and Y given by

$$\boldsymbol{R}(\boldsymbol{S},\boldsymbol{Y}) = \frac{\boldsymbol{C}(\boldsymbol{S},\boldsymbol{Y})}{(\boldsymbol{T}(\boldsymbol{S}).\boldsymbol{T}(\boldsymbol{Y}))^{½}} \qquad \text{(Eq.2)}$$

where $\boldsymbol{T}\ (\boldsymbol{S}) = \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \left\lfloor (\boldsymbol{\mu}^{2}{}_{\mathbf{S}}(\mathbf{x}_{\mathbf{i}}) + \boldsymbol{\sigma}^{2}{}_{\mathbf{S}}(\mathbf{x}_{\mathbf{i}}) + \boldsymbol{\nu}^{2}{}_{\mathbf{S}}(\mathbf{x}_{\mathbf{i}}))\right\rfloor$, $\boldsymbol{T}\ (\boldsymbol{Y}) = \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \left\lfloor (\boldsymbol{\mu}^{2}{}_{\mathbf{Y}}(\mathbf{x}_{\mathbf{i}}) + \boldsymbol{\sigma}^{2}{}_{\mathbf{Y}}(\mathbf{x}_{\mathbf{i}}) + \boldsymbol{\nu}^{2}{}_{\mathbf{Y}}(\mathbf{x}_{\mathbf{i}}))\right\rfloor$; $|\boldsymbol{R}(\boldsymbol{S},\boldsymbol{Y})| \le \mathbf{1}$

'S' refers to the conditional propositions in a neutrosophic if-then rule like srv_serror_rate type, the Srv_serror_rate type, and the flag [12] where 'Y' refers to the class attributes like there is an attack or not.

4. EXPERMINTAL RESULTS

The proposed hybrid aims to identify the attacks that take place in the network to classify them correctly and increase the detection rate of attacks, figure 4 indicate the whole system. The hybrid consists of a preprocessing step and two major phases. The preprocessing step utilizes the WEKA [24] data mining tool to get the most important attributes from the KDD-99 data set. The first phase is the neutrosophic variables definition which converts the normal data into neutrosophic variables utilizing the Self Organized features maps (SOFM) [12]. The neutrosophic definition of the variables along with the training KDD-99 data are fed into the classification process. The second phase is the neutrosophic IDS building. The proposed system utilizes the evolutionary capabilities of The Genetic Algorithms (GA) to find the appropriate classification (if-then) rules. Simulation of the proposed system is implemented by C# environment on Dell Inspiron 15.6" Laptop - Intel Core i5, Memory (RAM): 8.00 GB, System type: 64-bit operating system and Windows edition: windows 10.

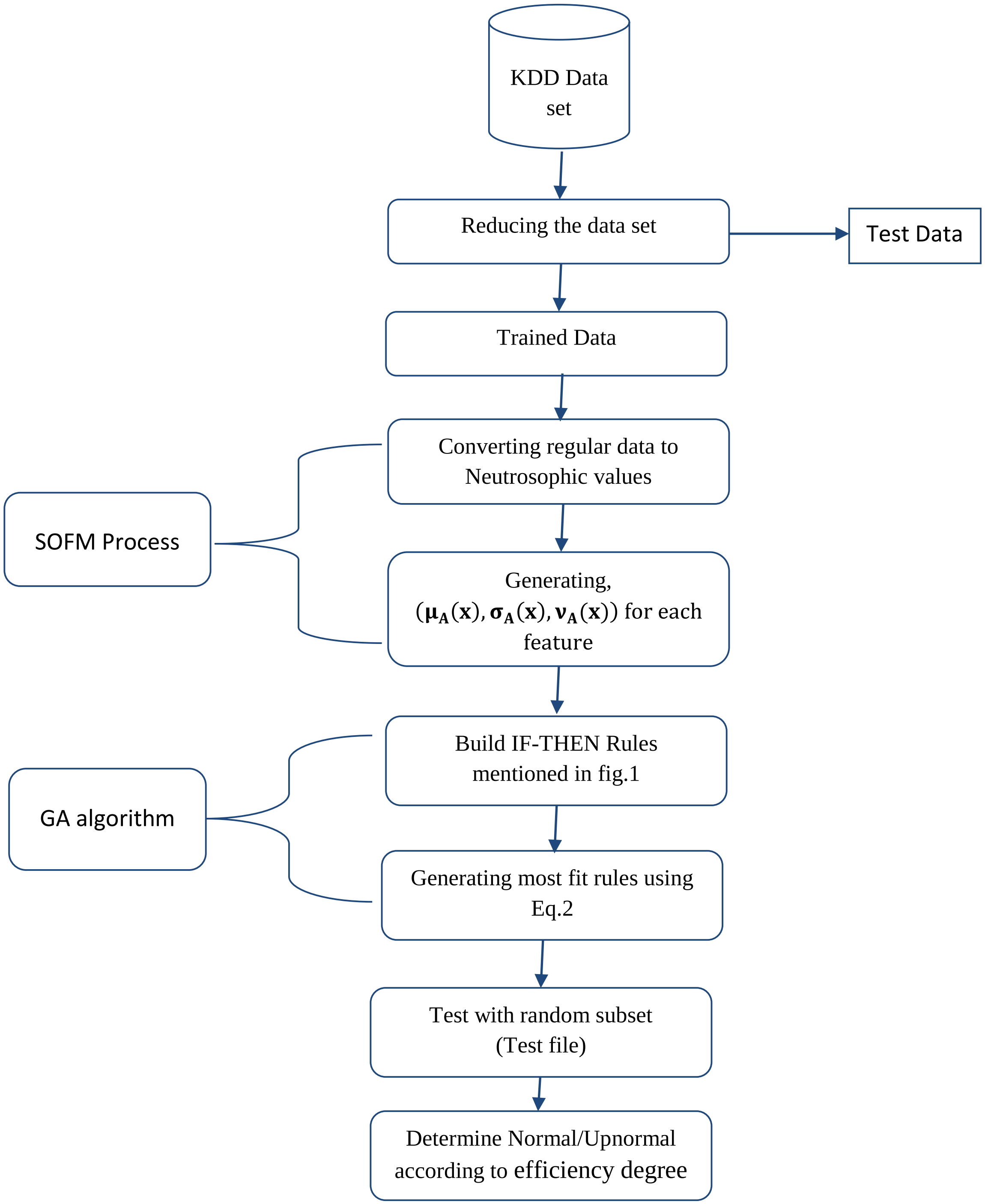


Figure 4: A frame work design for the whole system

The original data set is composed of 42 attributes; hence a reduction preprocessing step is required. The preprocess stage is implemented using Waikato Environment for Knowledge Analysis (Weka) [24]. The reduction algorithm used is the Attribute Evaluator 'FuzzyRoughSubsetEval' and search method 'HillClimberWithClassifier'. After the reduction process, KDD-99 data file contain 25 features and 1721 instance as in figure 5 which red color Upnormal (attacks) and the blue is normal.

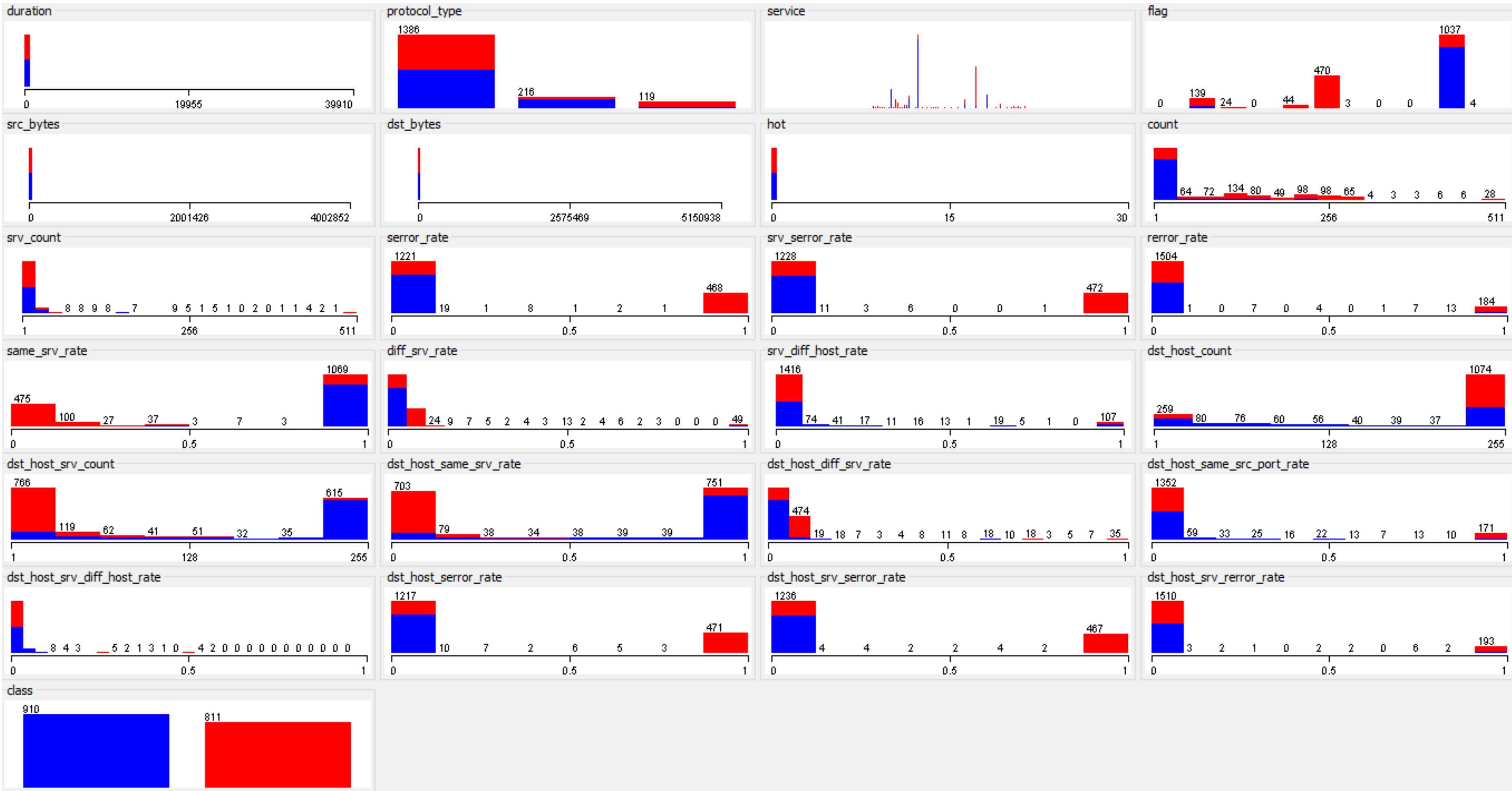


Figure 5: 25 features red color Upnormal (attacks) and the blue is normal.

Through the neutrosophic variable definition phase, The SOFM is applied to the KDD-99 data set to define the MEMEBERSHIP, INDETERMINACY and NONMEMEBERSHIP functions. The technique used for neutrosophic variables definition is illustrated in our previous work in [12]. Figure 6 shows the result of the neutrosophic features definition phase.

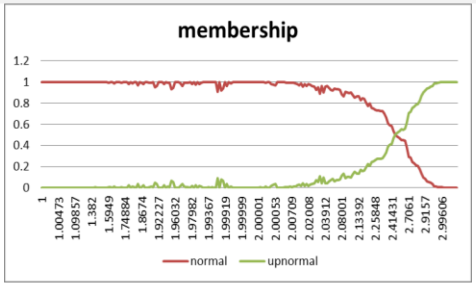
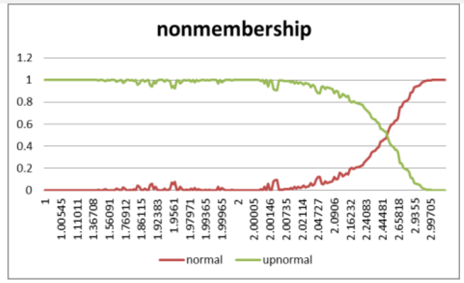
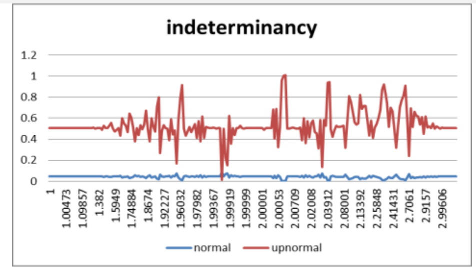
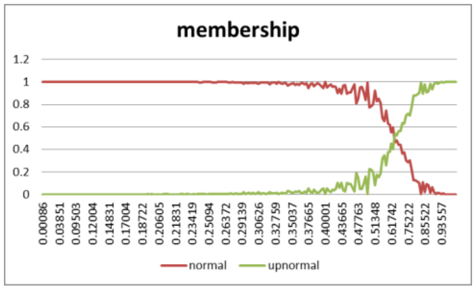
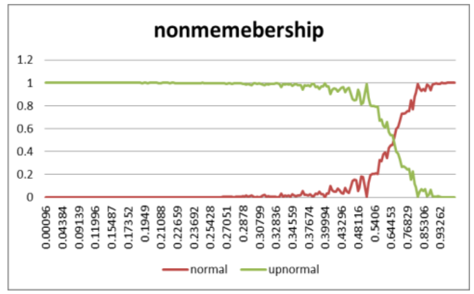
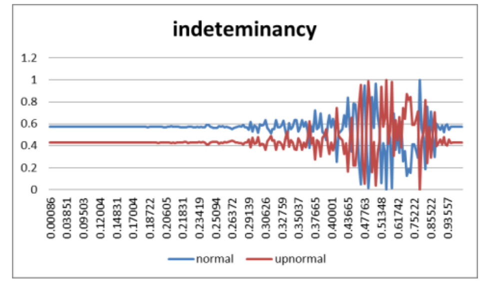
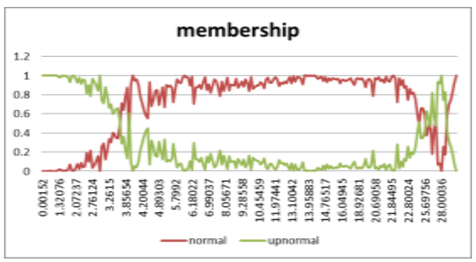
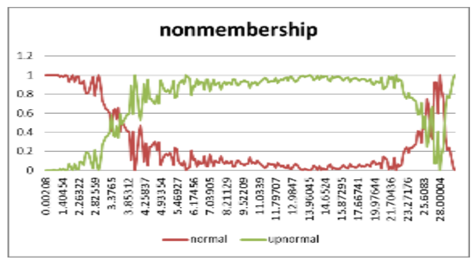
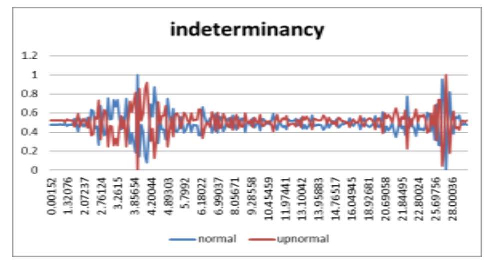
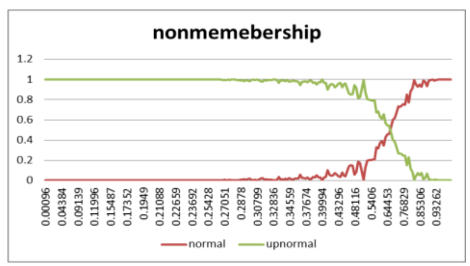
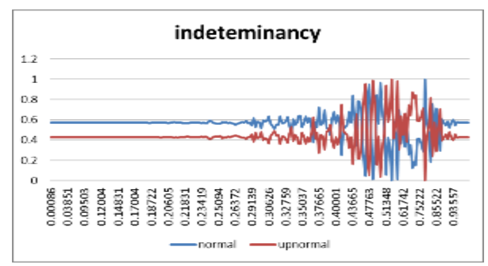
NO	ATTRIBUTE	MEMEBERSHIP	NONMEMBERSHIP	INDETEMINANCY
1	Protocol_Type			
2	srv_serror_rate			
3	Hot			
5	Srv_serror_rate			

Figure 6: the neutrosophic functions of membership, non-membership and *indeterminancy for a number of KDD-99 data set*

During the second phase, The IDS classification pattern which detects threats in the MANET network is implemented by an artificial intelligent algorithm (GA). Each feature from the KDD-99 data set will have three different definitions for MEMEBERSHIP, INDETERMINACY and NONMEMEBERSHIP assumption for the variable values. These features along with a random sub set of KDD-99 (training data) are passed to GA program to build the set of if-then rules which represent the neutrosophic IDS. The generated output file contains the most fit rules according the steps in GA pseudo code. The fitness function of the GA is the neutrosophic correlation coefficient which is calculated according to Eq. 2. The GA program simulation is implemented in 3 dimensions "MEMEBERSHIP, INDETERMINACY and NONMEMEBERSHIP" compared with other techniques or algorithms used only two dimension "MEMEBERSHIP and NONMEMEBERSHIP". The genetic algorithm parameters like cross over rate, mutation rate, number of population and number of iterations are assigned to the values illustrated in table1.

Table1: GA parameters in experiments

cross over rate	0.6
mutation rate	0.90
number of population	500
number of iterations	50

At the end of the GA iterations, the final file will contain the most fit (if-then) rules. Each one will have MEMBERSHIP, INDETERMENANCY AND NONMEMEBERSHIP values for each feature. These rules will be the inference engine for the neutrosophic IDS. The inference methodology used mimics the (min-max) mamdani inference methodology [25]. The output rules file generated have the most appropriate neutrosophic rules which indicate whether a given instance is a Normal connection or an attack. On applying a new instances (network data) to the system, the program select the MIN $(\boldsymbol{\mu_A}(\mathbf{x}), \boldsymbol{\sigma_A}(\mathbf{x}), \boldsymbol{\nu_A}(\mathbf{x}))$ value from all features in each rule in assumption that all feature are interconnected by 'AND' gate. Also, assuming that all rules are connected by 'OR' gate, the program select the rule with MAX $(\boldsymbol{\mu_A}(\mathbf{x}), \boldsymbol{\sigma_A}(\mathbf{x}), \boldsymbol{\nu_A}(\mathbf{x}))$ value to be the matched one. Then, the result of that matched rule will be compared to the actual KDD-99 data set to calculate the number of accurate instances percentage and the false rate percentage. During experiments, the KDD99 data set is divided randomly into two equal data sets (training and testing). The neutrosophic IDS is built using the training data set, then its accuracy is measured by the new instances in the test data set. The neutrosophic IDS reached an average accuracy 99.3608% which indicates that the proposed technique is more accurate than the previous algorithms used in this area [20, 26 and 27], this appears in the table 2, the figures 7 and 8 below. The results shown below in Table 1 represents the accuracy, sensitivity and specificity values for the proposed neutrosophic genetic algorithm against C4.5, SVM, C4.5 +ACO, SVM + ACO, EDADT , SVM + PSO and C4.5 + PSO algorithms [18]. In light of values got, the precision of C4.5 is 93.23%, the precision of SVM is 87.18%, the precision of C4.5 + ACO is 95.06%, the precision of SVM + ACO is 90.82%, the precision of C4.5 + PSO is 95.37%, the precision of SVM + PSO is 91.57% and the accuracy of Improved EDADT is 98.12%. It is obvious that the neutrosophic IDS generated by GA takes highest precision percentage when compared to all seven classification based algorithms. Fig. 7 indicate the corresponding chart for the result obtained in Table 2. Fig. 8 shows the performance of existing and proposed neutrosophic Intrusion Detection system (IDs) algorithm based on false alarm rate (FAR). Thus the proposed neutrosophic Intrusion Detection System (IDS) Algorithm effectively detects attack with less false alarm rate.

Table2: Performance of proposed neutrosophic genetic algorithm vs. existing algorithms

Algorithms	**Accuracy (%)**	**FAR (%)**
C4.5	93.23	1.65
SVM	87.18	3.2
C4.5+ACO	95.06	0.87
SVM+ACO	90.82	2.42
C4.5+PSO	95.37	0.72
SVM+PSO	91.57	1.94
EDADT	98.12	0.18
proposed neutrosophic genetic algorithm	99.3608	0.089

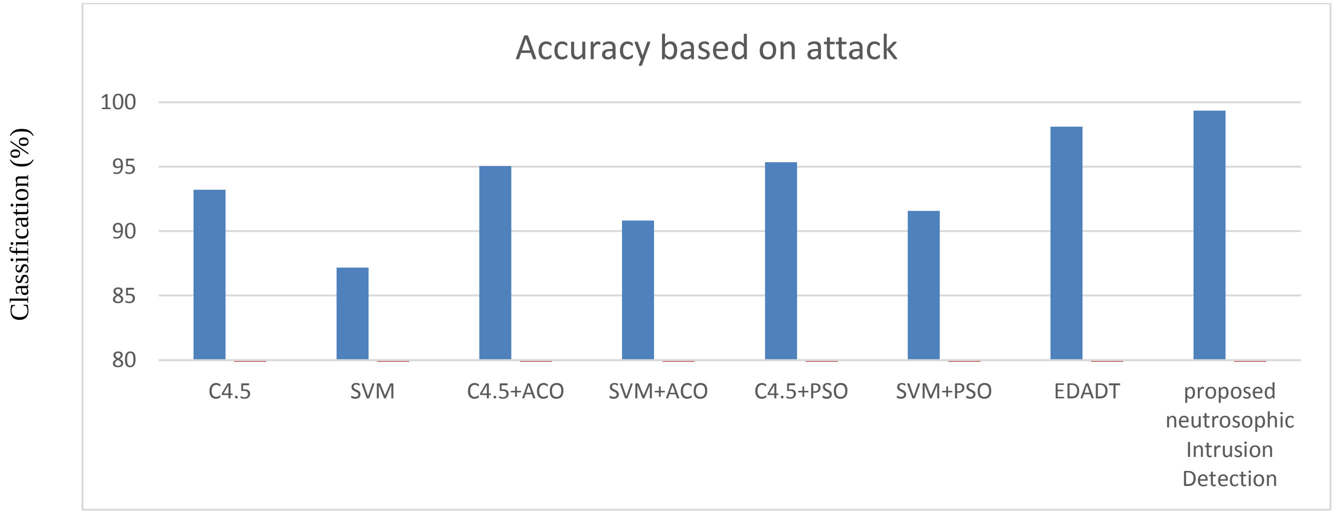


Figure 7: Results of neutrosophic genetic algorithm vs. existing algorithms

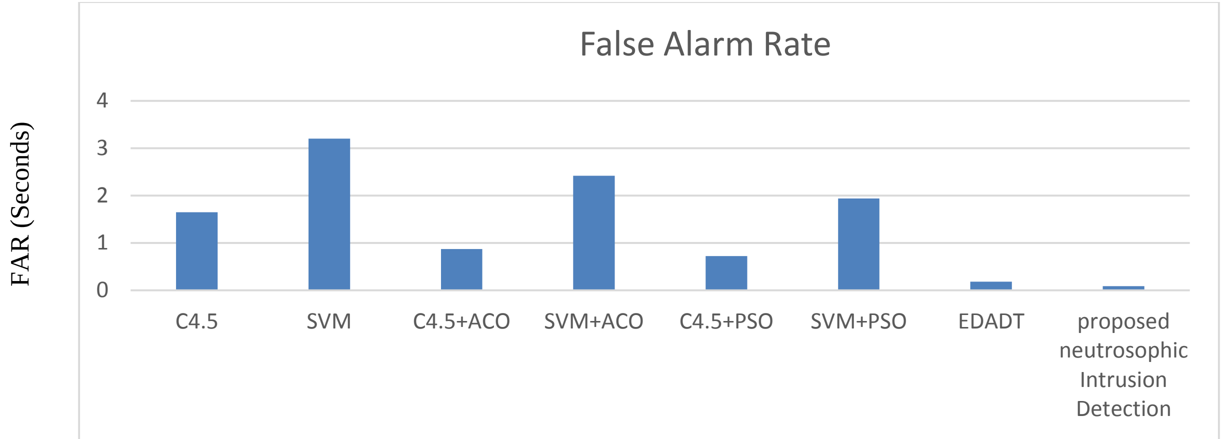


Figure 8. False alarm rate of proposed neutrosophic genetic algorithm vs. existing algorithms.

5. Conclusion and future work

The nature of the Mobile Adhoc Networks (MANETs) puts it under attacks from inside and outside the network. The security has two mechanisms, first prevent mechanism like cryptography and hash function, second reactive mechanism like intrusion detection system. Such systems are full of indeterminacy and uncertainty. Hence, building a classification pattern for the neutrosophic variables to detect threats in MANETs is a vital propose. This research indicates the integration of neutrosophic correlation coefficient into the genetic algorithm for upgrading an effective intrusion detection system. The proposed hybrid can increase the detection rate and reduce the false alarm rate in MANETs networks. Our experimental results prove that the proposed algorithm solves and detects the attacks in an effective manner compared with other existing works. Therefore, it will pave the way for an effective means for intrusion detection with better accuracy and reduced false alarm rate. In future, the performance of the system can be utilized in increasing the security of the system and predicting the intruders in MANETs networks by enhancing issues such as lack of resource consumption information to achieve an automatic Intrusion Detection System.

Acknowledgments

We are very thankful to prof. Florentin Smarandache from USA for the valuable comments and suggestions which improve this paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] Christopher M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, Inc. New York, NY, USA, **1995**, ISBN No. 0198538642.

[2] D. P. Agrawal, Q.-A. Zeng, Introduction to Wireless and Mobile Systems, Brooks/Cole Publishing, 2002, 438 pages, ISBN No. 0534-40851-6.

[3] Elwahsh.H, Hashem.M, Amin.M, Secure service discovery protocols for ad hoc networks, Communications in Computer and Information Science, 2011, Volume 131 CCIS, pp. 147–157, DOI.org/10.1007/978-3-642-17857-3_15.

[4] K. Biswas, Md. Liaqat Ali, Security threats in Mobile Ad Hoc Network, Blekinge Institute of Technology, Sweden, 2007.

[5] P.V.Jani, Security within Ad Hoc Networks, Position Paper, PAMPAS Workshop, London, 2002.

[6] A. A. Salama, Florentin Smarandache, S. A. ALblowi, New Neutrosophic Crisp Topological Concepts, Neutrosophic Sets and Systems, 2014.

[7] A. A. Salama, Smarandache, Neutrosophic Crisp Set Theory, Educational, Education Publishing 1313 Chesapeake, Ohio 43212 USA, Columbus, 2015.

[8] A. A. Salama, Basic Structure of Some Classes of Neutrosophic Crisp Nearly Open Sets and Possible Application to GIS Topology, Neutrosophic Sets and Systems, 2015, Volume 7, pp. 18-22.

[9] Mitra. S, Pal. S.K., Self-organizing neural network as a fuzzy classifier, IEEE, 1994, Volume 24, pp 385–399.

[10] K. Atanassov, Intuitionistic Fuzzy Set, Fuzzy Sets and Systems, 1986, Volume 20, pp. 87-96.

[11] KDD Cup 1999 Data, https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[12] Haitham Elwahsh, Mona Gamal, A. A. Salama, Modeling Neutrosophic Data by Self-Organizing Feature Map: MANETs Data Case Study, Procedia Computer Science, 2017, Volume 121, pp 152-159, DOI.org/10.1016/j.procs.2017.11.021.

[13] L.A. Zadeh, Fuzzy Sets, Inform and Control, 1980, Volume 8, pp. 338–353.

[14] I.M. Hanafy, A.A. Salama, K. Mahfouz, Correlation of neutrosophic Data, International Refereed Journal of Engineering and Science (IRJES),2002 Volume 1, PP.39-43.

[15] Mohammadreza Ektefa, Sara Memar, Fatimah Sidi, Lilly Suriani Affendey, Intrusion detection using data mining techniques, IEEE, International conference on information retrieval and knowledge management, 2010, DOI.10.1109/INFRKM.2010.5466919.

[16] Holden Alex Nicholas, Freitas A, A hybrid PSO/ACO algorithm for discovering classification rules in data mining. Journal of Artificial Evolution and Applications, 2008; pp 1–11, DOI.org/10.1155/2008/316145.

[17] Fatima Ardjani, Kaddour Sadouni, Mohamed Benyettou, Optimization of SVM multiclass by particle swarm (PSO-SVM), IEEE, 2nd International Workshop on Database Technology and Applications, 2010, DOI. 10.1109/DBTA.2010.5658994.

[18] G.V. Nadiammai, M. Hemalatha, Effective approach toward Intrusion Detection System using data mining techniques, Egyptian Informatics Journal, 2013.

[19] Nour Moustafa, Jill Slay, A Hybrid feature selection for network intrusion systems: Central points, Security Research Institute, Edith Cowan University, Australian Information Warfare and Security Conference, 2015, pp. 5-13.

[20] R. Becker, S. Eick, A. Wilks, Visualizing network data, IEEE Transactions on Visualization and Computer Graphics, 1995, Volume 1, pp 16-28, DOI.10.1109/2945.468391.

[21] Kohonen. T., 1990. The self-organizing map, Proceedings of the IEEE, 1990, Volume 78, Issue 9, pp 1464 – 1480, DOI: 10.1109/5.58325.

[22] Kohonen. T, Self-Organised formation of topologically correct feature map, Biological Cybernetics, 1982, Volume 43, pp 56–69.

[23] Wenke Lee, Salvatore J. Stolfo, A framework for constructing features and models for intrusion detection systems, ACM Transactions on Information and System Security (TISSEC), 2000, Volume 3, pp 227 – 261, DOI.10.1145/382912.382914.

[24] Machine Learning Group at the University of Waikato https://www.cs.waikato.ac.nz/ml/weka/.

[25] Madjid Tavana, Farshad Azizi, Farzad Azizi, Majid Behzadian, "A fuzzy inference system with application to player selection and team formation in multi-player sports, Sport Management Review, 2013, Volume 16, pp 97-110,

[26] Ehsan Amiria, Hassan Keshavarzb, Hossein Heidaric, Esmaeil Mohamadid, Hossein Moradzadehe, Intrusion Detection Systems in MANET: A Review, Procedia - Social and Behavioral Sciences, 2014, pp. 453 – 459.

[27] R. Thanuja, A. Umamakeswari, EFFECTIVE INTRUSION DETECTION SYSTEM DESIGN USING GENETIC ALGORITHM FOR MANETs, ARPN Journal of Engineering and Applied Sciences, 2016, Volume 11.