UNIVERSITY OF MISKOLC

FACULTY OF MECHANICAL ENGINEERING AND INFORMATICS

Fuzzy Automaton-based Early Detection Model

Summary of PhD dissertation

Mohammad Almseidin

Msc in Computer Science

'JÓZSEF HATVANY' DOCTORAL SCHOOL

OF INFORMATION SCIENCE, ENGINEERING AND TECHNOLOGY

ACADEMIC SUPERVISOR

Dr. habil. Szilveszter Kovács

Miskolc, 2019

Contents

1. Introduction 3

1.1 Aims of Research 5

2. Scientific Results 6

2.1 Investigating the Capabilities of the FRI in the IDS Application Area 6

2.2 FRI and SNMP-MIB for Emerging Network Abnormality 13

2.3 Fuzzy Automaton Based Detection Model Architecture 19

2.4 Implementation of the Fuzzy Automaton Based Intrusion Detection 22

3. Contribution and Future Research Direction 29

4. Summary 31

1. Introduction

Nowadays, network administrators face stressful environments with an overload of network traffics. Network traffic needs to be analyzed and investigated to detect abnormalities. The IDS has benefited from the rapid growth of technology; however, intruder techniques have also adapted to the intrusion detection mechanisms' new technological developments. Intruders have continued to advance their techniques and alter their behaviors to avoid detection by recent detection mechanisms. As a result, the danger of attacks has become increasingly more difficult to combat.

Computer and network security systems face different types of sophisticated attacks. One type of sophisticated attack is the multi-step attack. The multi-step attack [1, 2] is an attack composed of several prerequisite steps leading up to the final step which launches an attack targeting the victim's security hole. The attackers follow this technique to avoid detection. The prerequisite steps resemble normal behavior and serve as a subterfuge to facilitate the execution of the final step of the attack. The multi-step attack is a constant challenge for the intrusion detection system because intruders may implement complex attack scenarios, composed of several prerequisite steps, all aimed at executing their final attack [3]. Often, there is a causal relationship between the attack steps and forecasting the next step of attack [4].

There is an increasing need to design and implement an efficient IDS detection mechanism capable of handling different attack scenarios. The IDSs face several challenges including being able to detect multi-step attacks and the boundary problem (applying the binary decisions in the detection mechanism) [5]. In terms of the multi-step attacks, there is a causal relationship between the prerequisite steps which allows for administrators to be able to predict the next step of the attack [4]. Therefore, the multi-step attacks consist of different preliminary phases that can be distinguished from one another. On the other hand, implementing an efficient detection mechanism is also challenged by the boundary problem because there are no clear boundaries and no convincing threshold for defining normal and intrusion traffics [6]. The fuzzy system extends the binary decision to the continuous space, smoothing the boundaries and offering a solution to the boundary problem. Additionally, the results generated by the fuzzy systems are more comprehensible [5].

This work proposes a novel detection method for the multi-step attack built upon Fuzzy Rule Interpolation (FRI) based fuzzy state machine. In that respect, the FRI method instruments the fuzzy state machine to be able to act on a not fully defined state transition rule-base, by offering interpolated conclusion even for situations that are not explicitly defined. The proposed detection method was able to detect the multi-step attack even within the early stages of the attack. Furthermore, it had the ability to extend the binary decision to continuous space. The proposed detection method was performed using fuzzy automaton. The reasoning part of the proposed detection method adopts the FRI method instead of classical reasoning methods. This is done in order to decrease the total number of fuzzy rules required to define the state transition rule-base (simplification) and to offer interpolated results, even when the knowledge representation is not complete.

Also, this work proposes a novel method to detect the abnormality within the network by combining the FRI reasoning with the Management Information Base (MIB) parameters. In that respect, there is no need to deal with raw traffic processing, which is time-consuming, and difficult to compute. This method also eliminates the need for creating a complete fuzzy rule base. The MIB parameters reflect the normal and abnormal nature of the network traffics.

1.1 Aims of Research

One aim of the research was to investigate the capabilities to implement the FRI (FIVE) method in the IDS application area. This investigation is practiced by producing two novel detection models built-upon FRI (FIVE) method. This method not only allows the intrusion detection system to be used in continuous spaces but makes it possible to use a sparse fuzzy rule base. This way, the overall rule base size significantly smaller. Because of its fuzzy rule base knowledge representation nature, it can be easily adapt expert knowledge, and also be suitable for predicting the level of degree for threat possibility.

Another, somewhat distinct goal, to design a novel detection model for the multi-step attack built upon the Fuzzy Rule Interpolation (FRI) based fuzzy state machine which allows the usage of sparse intrusion state transition rule-based, and permitting the system's state to have a degree of the membership function. The fuzzy rule interpolation-based fuzzy automaton (fuzzy state machine) extended with a capability to be suitable for detecting and preventing the multi-step attack in stages, where the planned attack is not fully elaborated. Furthermore, to implement and evaluate the suggested model in practice and comparing it with other detection methods, in order to highlight and discuss the difference between the proposed detection model and others.

2. Scientific Results

2.1 Investigating the Capabilities of the FRI in the IDS Application Area

Fuzzy rule interpolation methods can serve deducible (interpolated) conclusions even in case if some situations are not explicitly defined in fuzzy rule-based knowledge representation. This property can be beneficial in partial heuristically solved applications; there the efficiency of expert knowledge representation is mixed with the precision of the machine learning method. The implementation of IDS Model-based fuzzy rule interpolation was divided into three main steps:

- To identify observable features suitable for IDS and the way they can handle the intrusion boundaries problem.
- To implement the FRI-IDS model as a detection mechanism for DDOS attacks.
- To compare the FRI-IDS model with other literature's results, which had used the same test-bed environment with different classification algorithms for detecting DDOS attacks.

2.1.1 FRI-IDS Model Generation

Typically, in the classical fuzzy system, the inferring of consequences could not be deduced in case if some situations were not explicitly defined in a fuzzy rule-based. Therefore, the inferring of the consequences of the fuzzy system required a completed fuzzy rule base. In the case of the sparse rule base which was not covered all of the possible situations, FRI methods offer the capability to generate the possible inference, even in case of lack definitions and information of existing knowledge representation. This benefit could be beneficial in partial heuristically solved applications. The FRI-IDS model was a description of the problem domain (IDS application area). It constraints of the key features (relevant features) to detect the intrusion. There are four major components needed to implement the FRI-IDS model as a detection mechanism:

- Setup the input and output of the FRI-IDS model. (The input parameters of the FRI-IDS model were the relevant features of the dataset which are mentioned in Table 13, the output of the FRI-IDS model supposed to be the level of attack instead of binary decision).
- Setup the fuzzy sets for each input/output of the FRI-IDS model.
- Setup the fuzzy rules for all the possible events of normal and intrusion.
- Testing and validating the FRI-IDS model.

The inference engine of the FRI-IDS model was performed by the Fuzzy Interpolation based on the Vague Environment method (FIVE). It was introduced by Kovacs in [7, 8, 9] in 1996. The FIVE method serves the deducible conclusions even in case if some situations are not explicitly defined in fuzzy rule-based knowledge representation. It is produced to serve many application areas such as IDS solution, which is served a crisp observation and at the same time required a crisp conclusion. It is worth mentioning that since using the FRI (FIVE) method as an inference engine there is no need for an additional defuzzification step.

The architecture of the FRI-IDS model was shown in Fig. 1. starts by data filtration phase where the network traffics (training data) analyzed in order to extract and determine the relevant

features. During the data filtration phase, the irrelevant features were removed. It should be known that the existence of irrelevant features could decrease the performance of the FRI-IDS model because it could give an incorrect indication about the existence of attacks. In the modeling phase, the sparse fuzzy model identification [15] was performed, it was introduced by Johanyák in 2008. The modeling phase had several actions, this includes the estimation of fuzzification and membership functions, fuzzy rules generation besides deducing the consequences and tuning methods.

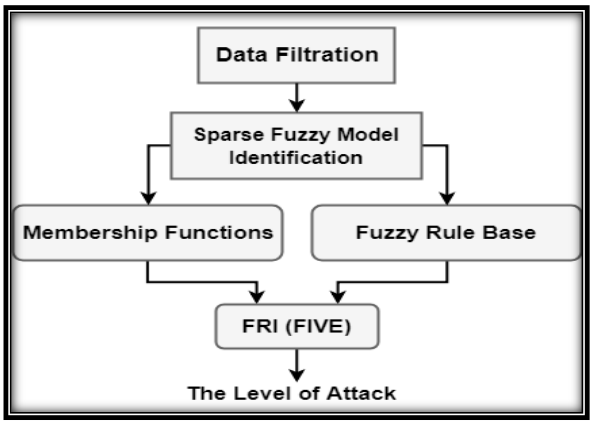

Fig. 1. The Architecture of The FRI-IDS Model

2.1.2 Data Filtration Phase

The dataset was divided into a training part and testing part. The training data consisted of 10000 records with 5000 normal cases and 5000 intrusion cases. The test data consisted of 10000 records with 5000 of normal cases and 5000 of intrusion cases. In order to increase the efficiency of IDS, it is important to identify the observable features that are relevant to detect intrusions from the network traffic data [11].

These are some of the relevant features: utilization, packet rate, byte rate, pkt size, and pkt delay. For the sake of reducing the possible number of fuzzy rules and low complexity system, the highest three relevant features according to the IG algorithm were used as input parameters of the FRI IDS model. These relevant features as found in Table 13 are the packet rate, byte rate, and utilization. The anomaly-based and misuse based detection techniques detect the attacks based on the predefined rule base (i.e. rules for normal and intrusions). For this, sorting the normal and intrusions cases of the training data is required [12]. Algorithm 1 presents the sorting and feature extraction of the training data.

Algorithm 1: Sorting and Feature Extraction

Input: The training data
Input: Two pools of the dataset (normal and intrusion)
The training data two pools of the dataset (normal and intrusion)

1: **while** Termination Condition Not Met **do**
2: Classify whole test-bed dataset into" normal" and "attack" class
3: Check for missing entry for all records
4: Extract the suitable features for IDS based o IG algorithm
5: Remove all irrelevant features
6: Store the values of normal pool
7: Store the values of intrusions pool
8: **end while**

The outputs of sorting and feature extraction algorithm were two pools of normal and intrusion records. These two pools consisted of only the relevant observable features that are suitable for the FRI-IDS model (packet rate, byte rate, and utilization) features, where all other values are removed.

2.1.3 Modeling Phase

The part of fuzzy modeling considered as one of the important parts of the fuzzy system. The FRI-IDS model was constructed by using the sparse fuzzy model identification [15]. The training data of the FRI-IDS model had three input parameters (packet rate, byte rate, and utilization). These parameters were chosen according to the IG algorithm to infer the DDOS attack. In order to generate the optimized fuzzy rules and fuzzy sets, the Rule Base Extension using Default Set Shapes (RBE-DSS) method which is introduced by Johanyák in [13] was applied. According to [13, 14] the main steps of RBE-DSS method can be summarized as follows:

- In the early stage of modeling the fuzzy system, the RBE-DSS method generates two rules that covered (fit) the minimum and maximum of the output.
- In the next step, the hill-climbing tuning algorithm started. It is adjusting the previous parameter values one by one. For each iteration, the fuzzy system is evaluated with different parameter values based on training data. The retrieved parameter values ensure that the fuzzy system belongs to the better performance index for the later iterations.
- The performance index is computed in each iteration to compare the obtained results with different parameter values. The relative root main square error was chosen as a performance index for tuning the FRI-IDS model.
- On the assumption, the increasing of fuzzy system performance appeared too slow or interrupted (i.e. fuzzy system obtained the local minimum) then, the new fuzzy rule generated to increase the possibilities of fuzzy system enhancement.
- The new fuzzy rule created where the difference between the value of actual output and computed output is maximum.
- The tuning process stopped either in case if the predefined performance index value

obtained or when the number of the iterations is reached.

As a result of applying the RBE-DSS method, Fig. 2. shows the support of antecedent fuzzy sets of the tuned FRI-IDS model based on the training data.

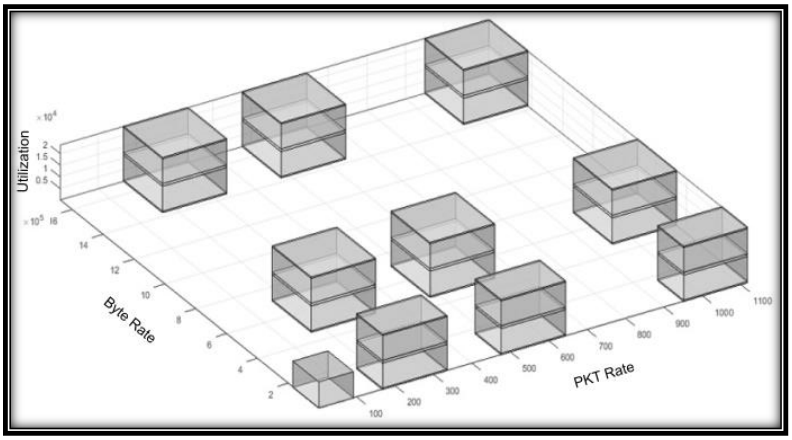

Fig. 2. Support of The Antecedent Fuzzy Sets of FRI-IDS Model

It is worth mentioning that, the generated fuzzy rules by the RBE-DSS method were sparse, these fuzzy rules for the fuzzy interpolation and if it is implemented for classical fuzzy reasoning there is no result could be obtained. Out of 28 fuzzy rules were generated in order to detect the DDOS attack based on the training data. Table 1 presents the generated fuzzy rules, Subsequently, of the modeling phase, the obtained fuzzy sets were represented by trapezoidal membership functions. The byte rate and utilization input parameters have three trapezoidal membership functions and the packet rate input parameter has four trapezoidal membership functions. Table 2 presents the optimized values of fuzzy sets for the FRI IDS model based on the training data.

Table 1. The Obtained Fuzzy Rules

No.	Packet Rate	Byte Rate	Utilization	Consequences
1	L	L	L	FA
2	L	L	M	FA
3	L	L	H	FA
4	L	M	L	FA
5	L	M	M	FA
6	L	M	H	FA
7	L	H	L	A
8	L	H	M	A

No.	Packet Rate	Byte Rate	Utilization	Consequences
9	L	H	H	A
10	M	L	L	FA
11	M	L	M	FA
12	M	L	H	FA
13	M	M	L	FA
14	M	M	M	FA
15	M	M	H	FA
16	M	H	L	A
17	M	H	M	A
18	M	H	H	A
19	H	L	L	A
20	H	L	M	A
21	H	L	H	FA
22	H	M	L	A
23	H	M	M	A
24	H	M	H	A
25	H	H	L	A
26	H	H	M	A
27	H	H	H	A
28	VL	L	L	A

Table 2. The Obtained Fuzzy Set Parameters Of FRI-IDS Model

Packet Rate	**Very Low**	**Low**	**Medium**	**High**
	[1 1 35.92 91.78]	[166.81 222.66 278.51 334.36]	[475.73 531.58 587.43 643.28]	[950.67 1006.52 1062.37 1118]
Byte Rate	**Low**	**Medium**	**High**	
	[55 55 83268.03 167136.28]	[461330.38 545198.63 629066.88 712935.13]	[1425835.73 1509703.98 1593572.23 1677420]	
Utilization	**Low**	**Medium**	**High**	
	[3 3 594.18 11235.33]	[594.18 11235.33 12417.68 23058.83]	[12417.68 23058.83 23650 23650]	

FRI-IDS model serves crisp values and at the same time generates a crisp conclusion. Therefore, each observation within the training data in the example of this work presented as a fuzzy singleton. Fuzzy systems had the capability to extend the binary decision to the continuous

truth value which is more readable and easier to be understood and analyzed. Suppose that, there are two observations within the training data, the first observation had the following crisp values (packet rate= 200, byte rate = 55943 and utilization = 11560). The second observation had the following crisp values (packet rate= 900, byte rate = 1190251 and utilization = 22029). The inferred consequence of the FRI-IDS model for the previous two observations was shown in Fig. 3. and Fig. 4. respectively where the first observation presents the normal event and the second observation shows the DDOS attack event. FRI-IDS model can serve the interpolated conclusions even in case if some observations are not covered directly by fuzzy rules as Fig. 4. presented. The interpolated conclusions mean that the FRI-IDS offers the ability to generate the required comprehensive alert, and present it based on the FRI-IDS output membership functions (Normal or DDOS attack).

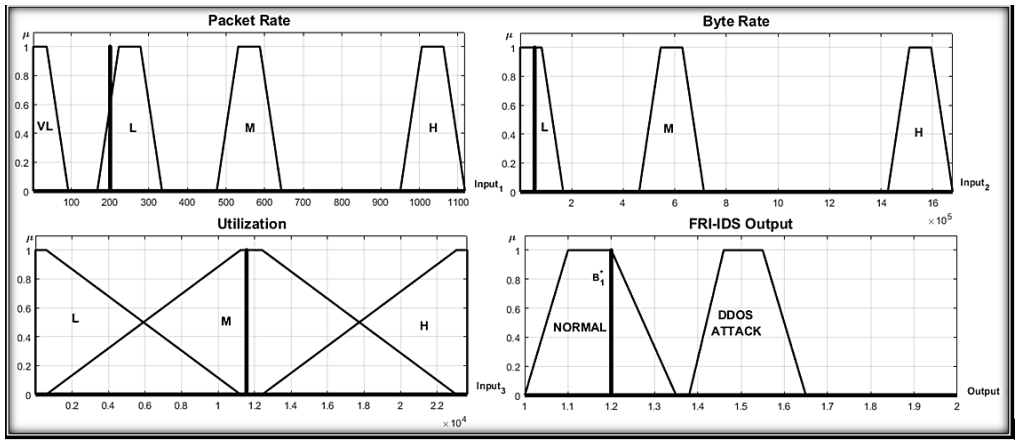

Fig. 3. FRI-IDS Output Response in Case of Normal

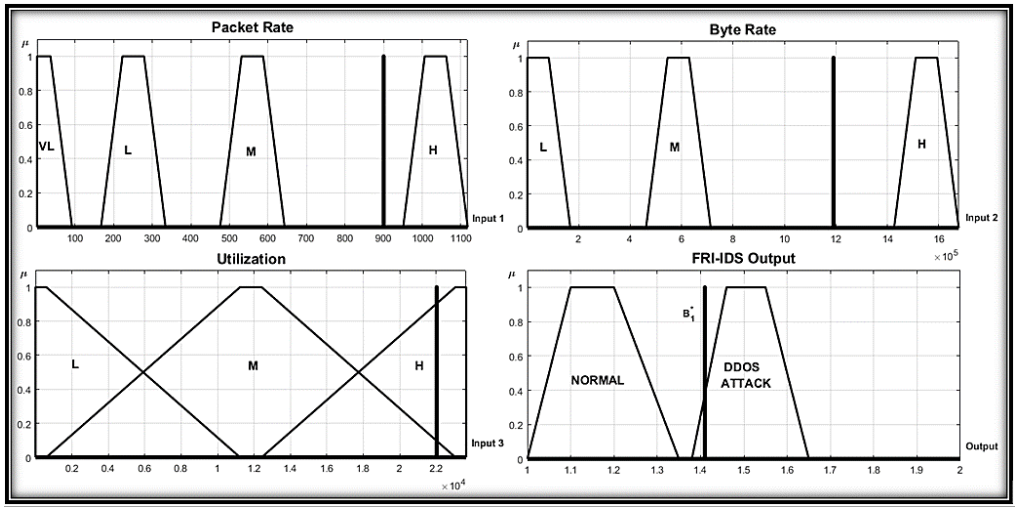

Fig. 4. FRI-IDS Output Response in Case of Attack

2.1.4 Experiments and Discussion

This subsection illustrates the testing and validating of the FRI-IDS model using the test-bed

dataset. Thereupon, all experiments were conducted using MATLAB [15] and FRI toolbox [16]. The inference engine of the FRI-IDS model was performed using the FIVE method. Out of 28 fuzzy rules were generated in order to detect the DDOS attack. It is worth mentioning that the FRI-IDS was tuned using RBE-DSS. The tuning process stopped when the predefined performance index value obtained or when the number of iterations is reached. The tunning process of FRI-IDS stopped criteria was the maximum number of iterations. The code of FIVE method and other FRI methods can be used through the FRI toolbox which can be downloaded freely from [16]. The overall process of testing and validating the FRI IDS model was shown in Fig. 5.

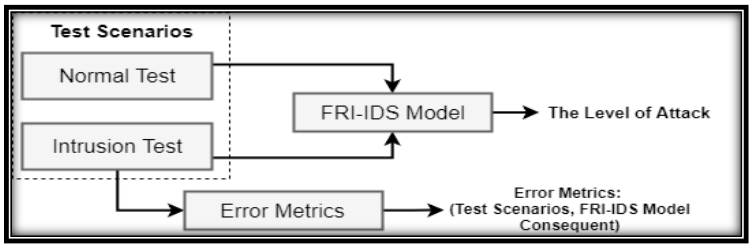

Fig. 5. The Testing and Validating Process of FRI-IDS Model

The FRI-IDS model was tested and evaluated based on two test scenarios:

- The first test titled as the normal test scenario where 5000 instances of normal cases is used as input parameters of the FRI-IDS model. The extracted 5000 instances of normal cases presented as a matrix of the normal test.
- The second test scenario was titled as an intrusions test scenario where 5000 instances of intrusion cases are used as input parameters of the FRI-IDS model. The extracted 5000 instances of intrusion cases presented as a matrix of intrusion test.

The previous two matrices of normal and intrusion test scenarios were chosen as a two-input testing file of the FRI-IDS model. The evaluation of the FRI-IDS model carried through the computed error metrics of the test scenarios. The inferred consequence of the FRI-IDS model was compared along with the actual values of normal and intrusion. According to [17], the error metrics of the test scenarios were extended to the following performance metrics: TP, FP, TN, and FN parameters.

Furthermore, the previous mentioned performance metrics offer the capability to compare the FRI-IDS model results with other algorithms that have been implemented for detecting DDOS attacks. As a result of the test scenarios of the FRI-IDS model, 10000 cases were tested successfully. These cases were split as 5000 of normal cases and 5000 of intrusion cases. During the normal test scenario, only 3 records of normal cases were inferred incorrectly by the FRI-IDS model. Besides, during the intrusion test scenario, 332 of intrusion cases were inferred incorrectly by the FRI-IDS model. The obtained results besides the error metrics values presented in Table 3.

Table 3. The Result of The Test Scenarios Cases

	Normal	Intrusion	Total
Normal	4997	3	5000
Intrusion	332	4668	5000
Total	5329	4671	10000

According to the obtained results and the error metrics values of Table 3, the confusion matrix parameters of the FRI-IDS model presented in Table 4.

Table 4. Confusion Matrix Of FRI-IDS Model

Alert Response	Intrusion Packet Prediction	Normal Packet Prediction
Intrusion	TPR = 0.93	FNR = 0.06
Normal	FPR = 0.0006	TNR = 0.999

The implemented experiments have demonstrated that the FRI-IDS model obtained 96.65% as an overall detection rate. The computed performance metrics concluded that the FRI-IDS model obtained an acceptable value for the detection rate, and it decreases effectively the false positive rate. Decreasing the false positive rate helps to reduce a large amount of IDS alerts. To summarize the aforementioned results, the FRI-IDS model could be a suitable approach to be implemented as a detection mechanism for the following reasons:

- The FRI-IDS model offers an extension of the binary decision problem to continuous truth value, in which the inferred consequence like “the level of intrusion”, which makes the response result more readable and clearly analyzed rather than a binary decision.
- It is difficult to identify a clear boundary between normal and intrusion packets. Therefore, the fuzzy system effectively smooths the abrupt break of normal and intrusion.
- FRI methods can serve deducible (interpolated) conclusions even in case if some situations are not explicitly defined in fuzzy rule-based knowledge representation.
- The implemented experiments show that the FRI-IDS model obtained an accepted value for the detection rate and false-positive rate.

2.2 FRI and SNMP-MIB for Emerging Network Abnormality

The MIB parameters are characterized by the wealth of beneficial information they can offer for defining abnormality and reflect the normal and abnormal nature of the network traffics. The MIB dataset applied as an example in this work is introduced by Al-Kasassbeh et al. [18]. This dataset was originally generated to target DoS attacks. A DoS attack is blocking legitimate user requests for services the server can provide. Such attacks can be carried out by flooding the chosen server with a high volume of traffic, thereby consuming all of the server's resources and, consequentially, preventing the server from responding to genuine requests. These attacks can be generated either from a local or remote node in a different network. DoS attacks are usually difficult to assess and prevent [19], making them one of the most challenging type threats. An even more severe threat is the DDoS attack, which is a type of flooding attack that is generated from

various nodes simultaneously [20].

SNMP-MIB data are rich sources providing clear statistical information about the current network device status. The SNMP-MIB is a widely deployed protocol in most network devices, and available without any additional new hardware or software investment. By reading the MIB data, some of the major challenges of intrusion detection can be avoided. The dataset we used contains 4998 connection records. This MIB dataset has been collected from a router. The dataset has 34 MIB variables from 5 MIB groups in MIB-II. The groups are IF, IP, TCP, UDP, and ICMP.

2.2.1 IDS Model-based FRI and SNMP-MIB

The proposed IDS approach is adapting the FIVE method as an inference engine. The concept of the vague environment, introduced by Klawon in [21], the vague environment can be defined on the basis of similarity or indistinguishability of the elements. The concept of the vague environment can be expressed by a scaling function (s). The proper scaling function (s) which describes all the fuzzy sets of a fuzzy partition, should be implemented to produce a vague environment. According to [7, 8], the scaling function (s) is suitable for describing the shapes of all fuzzy set of a fuzzy partition. In the vague environment, the level of similarity between two fuzzy sets illustrates the fuzzy membership function M(x). In the vague environment, two values are ε-distinguishable if their distance is greater than ε:

$$\varepsilon > \delta_s(X_1, X_2) = \left| \int_{X_2}^{X_1} s(x)dx \right| \tag{14}$$

Likewise, $\delta_s(X_1, X_2)$ represents the vague distance for the values X1 and X2.

The general structure of the proposed detection approach, as it is shown in Fig. 6., is initiated by the data-cleaning stage. This stage is responsible for assembling the required information using the SNMP agents. This information is then forwarded to the SNMP manager which consists of a repository of MIB parameters. During the data cleaning stage, the MIB parameters were evaluated to determine their relevant parameters. The cleaning stage aims to reduce a large number of MIB parameters by eliminating those that are irrelevant.

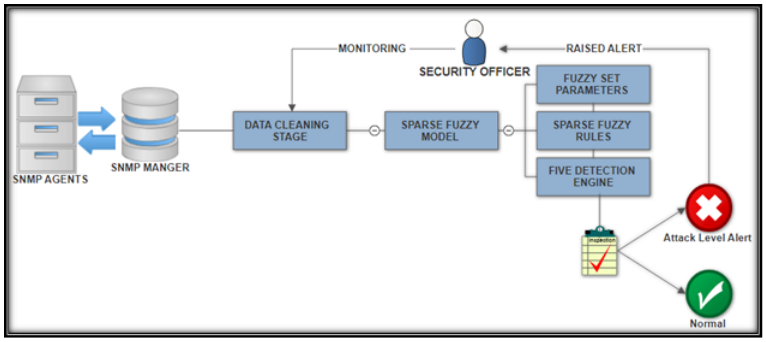

Fig. 6. The General Structure of The Proposed Detection Approach

The imported MIB dataset consists of a large number of MIB parameters. To simplify the process, the top five relevant MIB parameters [22, 18] were chosen as input parameters for detecting the abnormality in the proposed detection approach. These MIB parameters are the IP-Out-Discards and IP-In-Discards from the IP MIB group. IF-In-Discards and IF-Out-Discards from the interface MIB group and ICMP-Out-Dest-Unreachs from the ICMP MIB group. The training part is designed to generate two repositories. The first repository consists of only the intrusion traffics, and the second repository includes only the normal traffics. Consequently, 2970 instances of normal and abnormal traffic were stored in two repositories. These instances had only the top five relevant MIB parameters.

The detection stage consists of several operations including fuzzification, sparse rule base generation and adapting the inference engine. The proposed detection approach was designed and constructed using the SFMI [15]. Before constructing the proposed detection approach, the top-five relevant MIB parameters were forwarded to the SFMI. The fuzzy rule generation and fuzzy sets optimization are the necessary modeling steps for constructing the proposed detection approach. The fuzzy rule generation and the fuzzy sets optimization were adapted using the RBE-DSS method which is introduced by Johanyák in [13]. As a result of adapting the RBE-DSS method, the FRI methods effectively reduce the total number of fuzzy rules, having 245 fuzzy rules for the FIVE method to detect the abnormality based on five MIB parameters. Table 5 presents a sample of the sparse rule base that was generated by the RBE-DSS method.

Table 5. The Sparse Rule-base on The MIB Parameters

No.	IFoutDiscards	IFInDiscards	IPOutDiscards	IPInDiscards	ICMPOutDestUnreachs	Consequence
1	Low	Low	Very Low	Very Low	Low	Normal
2	Low	Low	Very Low	Very Low	Medium	Normal
3	Low	Low	Very Low	High	High	Normal
4	Low	Low	Medium	Medium	Medium	Normal
5	Low	Medium	Very Low	Very Low	Low	abnormal
6	Medium	High	Very Low	Very Low	Low	abnormal
7	High	Low	Very Low	Very Low	Low	abnormal
8	High	High	Very High	Very Low	Low	abnormal
9	Medium	Low	Very High	Very Low	Low	Normal
10	Medium	High	Very Low	High	Medium	abnormal
11	Medium	High	Medium	Medium	High	abnormal
12	Medium	Medium	Very High	Very Low	Low	Normal
13	High	Low	Medium	Medium	Low	abnormal

In the RBE-DSS method, the trapezoidal membership functions were chosen to apply during the fuzzy set parameters optimization. The ICMP-Out-Dest-Unreachs, IF-Out-Discards and IF-In-Discards MIB parameters have three membership functions. These membership functions are classified into the following linguistic terms: Low, Medium and Large. The ip-In-Discards MIB parameter has four membership functions classified into the following linguistic terms: Very Low, Low, Medium and Large. Finally, the IP-Out-Discards MIB parameter represents five membership functions which are classified into the following linguistic terms: Very Low, Low, Medium, Large and Very Large.

The RBE-DSS method optimizes the values of fuzzy set parameters to the maximum performance of the fuzzy IDS. Fig. 7. presents the proposed detection approach's antecedent partitions. The proposed detection approach could offer the conclusion (detection result) even in situations where some MIB parameters are not explicitly defined in the generated fuzzy rule base.

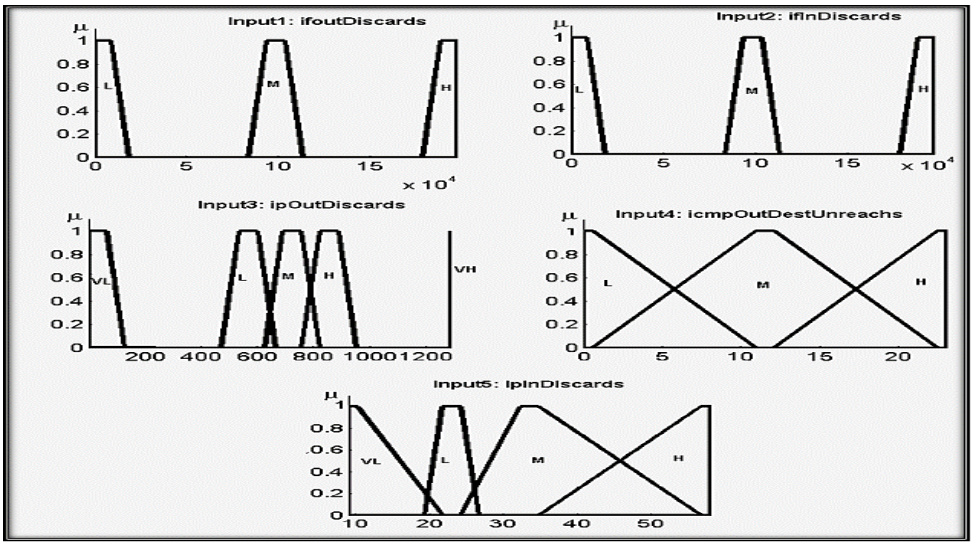

Fig. 7. The Antecedents Partitions of The Proposed Detection Approach

2.2.2 Simulation and Results

The total number of training data consisted of 2998 instances of normal and abnormal traffics. The training data were used to construct and optimize the proposed detection approach. The rest of the MIB dataset consisted of 1998 instances that were used for the validation process. It is worth mentioning that, every observation within the SNMP-MIB dataset was presented as a fuzzy singleton. The proposed detection approach was able to generate intelligible results due to its fuzzy nature, subsequently allowing the degree of abnormality to be determined.

Fig. 8. presents the output response of the proposed detection approach in the case of the abnormal instance with the parameters which are listed in Table 6. The data conclude that the degree of abnormality has been determined. Subsequently, the results, which are now more concise, serve to help administrators for better understanding the current security status.

Table 6. Abnormal MIB Parameters Example

MIB Parameters	Value
IF-Out-Discards	7270
IF-In-Discards	7270
IP-Out-Discards	1287
IP-In-Discards	9
ICMP-Out-Dest-Unreachs	0

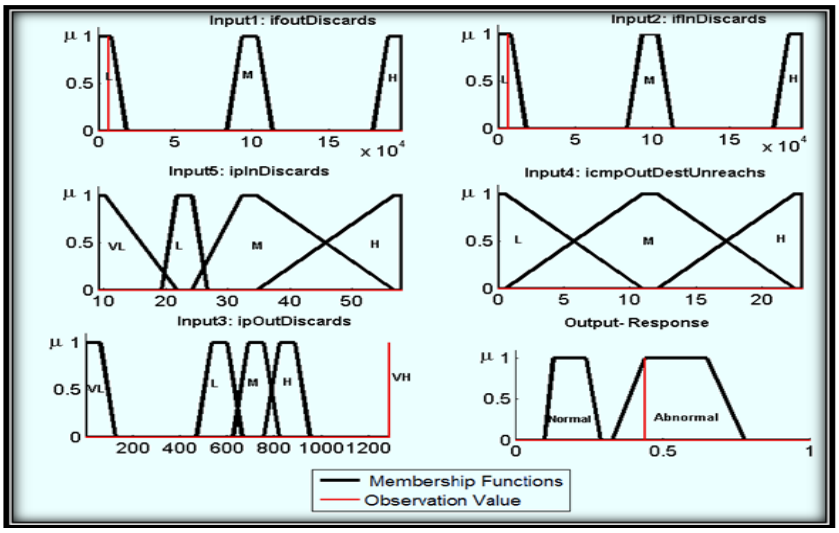

Fig. 8. The Output Response of The Proposed Detection Approach

The proposed detection approach was evaluated in a two-phase process. The first phase evaluated the normal repository and the second phase was evaluated the abnormal repository. A total of 1998 MIB parameter instances were tested and evaluated. Fig. 9. displays the results from both phases (normal and abnormal) of the detection approach's evaluation process.

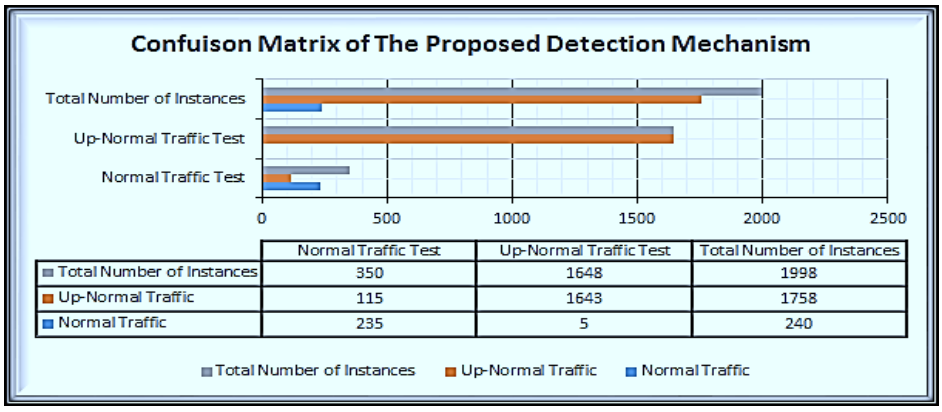

Fig. 9. The Confusion Matrix of The Evaluation Process

It was concluded that five instances of normal traffic were inferred incorrectly and 115 instances of abnormal traffic were inferred incorrectly. The obtained results have been carefully analyzed and investigated to highlight the strengths of the proposed detection approach. Table 7 presents the performance metrics for the proposed detection approach.

Table 7. The Performance Metrics For The Proposed Approach

Performance Parameter	Value	Formula
Sensitivity	0.9346	TPR = TP / (TP + FN)
Specificity	0.9792	SPC = TN / (FP + TN)
Precision	0.9970	PPV = TP / (TP + FP)
False Positive Rate	0.0208	FPR = FP / (FP + TN)
False Negative Rate	0.0654	FNR = FN / (FN + TP)
Accuracy	0.9399	ACC = (TP + TN) / (P + N)

To summarize the aforementioned results, the performance of the proposed detection approach achieved satisfactory values and, at the same time, supports the idea that implementing the fuzzy rule interpolation methods for the reasoning part together with the SNMP-MIB parameters could be a promising approach in the IDS application area. Moreover, the results obtained from the proposed detection approach were compared with other literature results [22] in which the same MIB dataset and the same number of relevant MIB parameters (top-5) were applied in combination with neural network, support vector machine and Bayesian network algorithms. Fig. 10. compares the results between the proposed detection approach and other algorithms (neural network, support vector machine and Bayesian network).

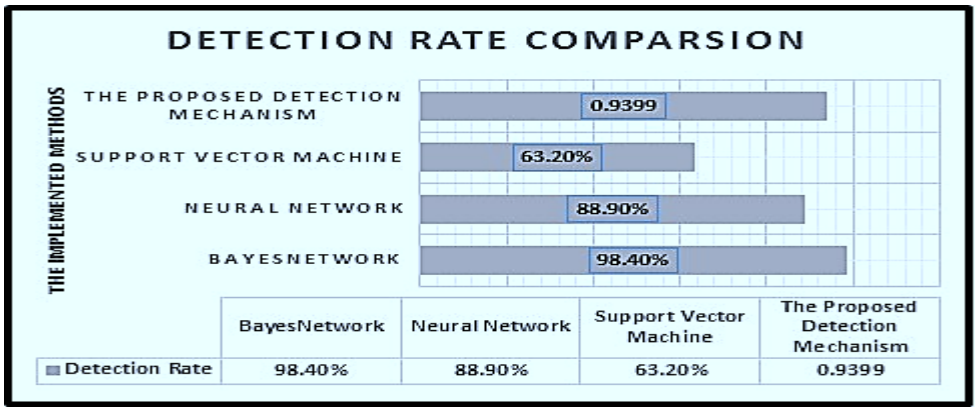

Fig. 10. The Detection Rate Comparison Results

Consequently, the implemented experiments demonstrated that the proposed detection approach achieved an acceptable accuracy rate. Moreover, it effectively reduced the false positive rate parameter. The conventional detection approaches focus on adapting the typical data mining algorithms, or the classical fuzzy reasoning methods, to be used with raw network traffics. Although FRI methods have been implemented in the IDS application area, these methods are still

under investigation. Nevertheless, current research has yielded satisfactory results. The strength of FRI methods is derived from the combination of the fuzzy concept and interpolation techniques. Therefore, the FRI methods could pose an effective solution for the boundary problem and could also handle the deficiencies of the knowledge-base representation. The strength of the proposed detection approach is based on combining the MIB parameters with the fuzzy rule interpolation reasoning method. Thus, there is no need to deal with raw traffics which are time-consuming and difficult to compute. Furthermore, this method eliminates the need for the complete fuzzy rule base.

2.3 Fuzzy Automaton Based Detection Model Architecture

Automata theory [23] is defined as the analytical study of abstract systems to solve computational problems. The integration between the fuzzy system and automaton theory results in a fuzzy automaton. This incorporation offers the ability to handle the computational challenges for both discrete and continuous spaces. The fuzzy automaton implemented based on the strengths of two paradigms, the automata, and the fuzzy system. Fuzzy systems are being implemented more frequently in different application areas. Fuzzy systems present comprehensive approximate reasoning results for the system's computational problems. Furthermore, they provide the required extension of the binary decision problem to the continuous truth value [5]. The general definition of the fuzzy automaton [24] is presented as a 6-tuple, illustrated in Equation 16.

$$\tilde{F} = (Q, \Sigma, \delta, R, Z, \omega) \tag{16}$$

Where Q is the finite set of the system states, Q= $\{q_0, q_1, \dots, q_k\}$. Σ is the finite set of the input symbols, Σ = $\{x_0, x_1, \dots, x_n\}$. δ is the fuzzy transition function, it is used to map the current system state to the next system state based on the finite set of inputs, $\delta: Q \times \Sigma \times Q \to (0,1]$. R shows the initial system state $\tilde{F}, R \in Q$. Z presents the finite set of output, $Z = \{Z_0, Z_1, \dots, Z_k\}$. Finally, ω presents the output mapping function which is responsible for mapping the fuzzy states into the output set, $\omega: Q \times \Sigma \to Z$.

In the fuzzy automaton, the system states, inputs and outputs are all presented as fuzzy sets. The predefined fuzzy states had a degree of membership values. Contrary to other state machines (deterministic, non-deterministic and probabilistic), the transition function was interpreted as a fuzzy transition function. In addition, the transitions between different states occurred based on the predefined fuzzy rules. In [25], the general definition of the fuzzy automaton was extended as shown in Equation 17.

$$\tilde{F} = (S, X, \delta, P, Y, \omega) \tag{17}$$

Where S is the finite set of fuzzy system states, $S = \{m_{s_1}, m_{s_2}, \dots, m_{s_k}\}$. X is the finite set of dimensional input values, $X = \{x_0, x_1, \dots, x_n\}$. δ is the fuzzy transition function, it is used to map the current state to the next state based on the finite set of inputs, $\delta: S \times X \to S$. P shows the

initial fuzzy state of $\tilde{F}, P \in S$. Y is the finite set of output dimensional vectors, $Y = \{Y_0, Y_1, \dots, Y_k\}$. Finally, ω presents the output mapping function which is responsible for mapping the fuzzy states based on input values to the output set, $\omega: S \times X \to Y$.

The fuzzy automaton detection approach consists of six major components. These components are listed as follows:

- Setting up the finite fuzzy system states (S).
- Setting up the initial system state (P), assumed to be in the normal state.
- Defining the possible system input values (X). These values depend on which type of multi-step attack could be detected. The input values of the fuzzy automaton detection approach presented as a set of system observations $(i_1, i_2, \dots, i_n)$.
- Defining the fuzzy state-transition function δ which is used to map the current system state to the next system state based on the observations, $\delta: S \times X \to S$.
- Defining the finite set of system outputs, $Y = \{Y_0, Y_1, \dots, Y_k\}$.
- Defining the output mapping function $(\omega: S \times X \to Y)$ which is responsible for mapping the fuzzy states based on input values to the output.

The suggested fuzzy automaton detection mechanism adapts FRI (FIVE) method [7, 8, 9]. The FRI (FIVE) method is used to simplify the rule definition and to interpolate the missing state-transition rules. Contrary to the classical reasoning methods, the FRI methods offer the interpolated conclusion even when some situations are not explicitly defined [26]. Fig. 11.shows the general architecture of the fuzzy automaton detection mechanism using the previous six major components.

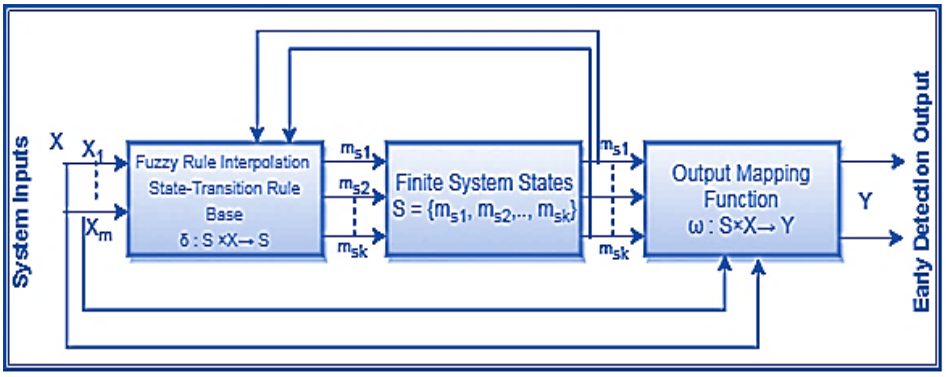

Fig. 11. The Fuzzy Automaton Detection Mechanism Architecture

The suggested fuzzy automaton based detection method consisted of four system states: $S = \{N, A, P, C\}$. These states are similar to those used in [27] which was defined as follows:

- Normal(N): the system behavior is in normal mode and there are no attempts to attack.
- Attempt(A): there are different attempts to gather information about the system in legal ways (different probe tools are launched).

- Prerequisites(P): malicious activity has commenced and the multi-step attack is in the process of launching its final step of the attack.
- Compromise(C): the multi-step attack has been completed successfully. The system is completely infected.

Fig. 12. presents the graph of the system states within the fuzzy automaton detection mechanism. The graph is fully connected to indicate that the transition (between states) may occur from any system to any system state.

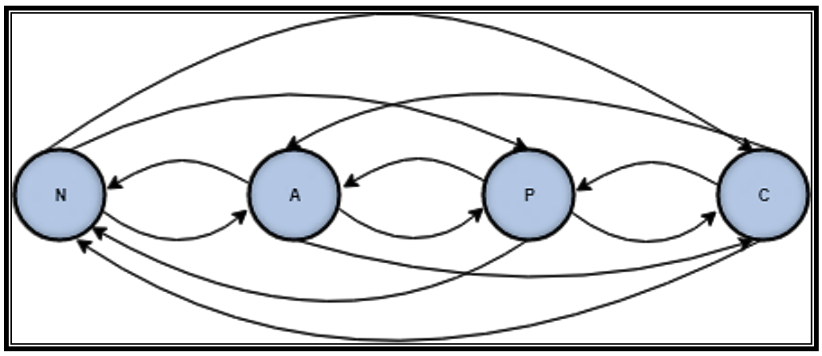

Fig. 12. System States of The Fuzzy Automaton Detection Mechanism

The fuzzy automaton detection mechanism focuses on the initial steps of the multi-step attack to prevent the launch of any further attack steps. Suppose that, there is a multi-step attack with n+m steps to be launched successfully. The fuzzy automaton detection mechanism focuses on predicting the multi-step attack penetrations within the period step (1) and step (n). Fig. 13. presents the concentration intents of the fuzzy automaton detection mechanism. The multi-step attack may be detected early because it built upon different preliminary phases that can be distinguished from one another.

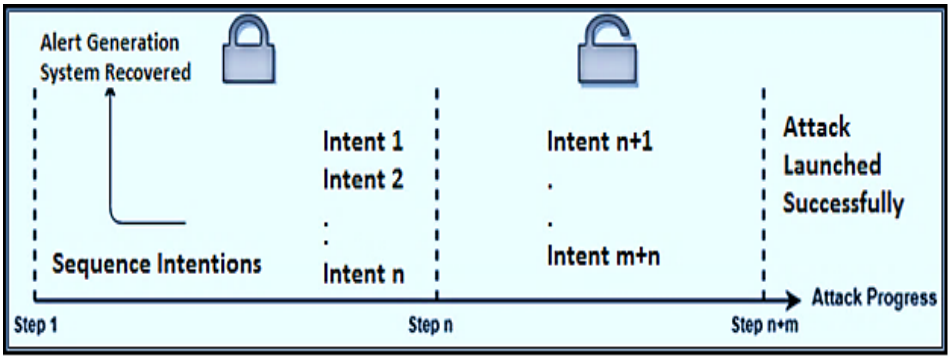

Fig. 13. The Concentration Intents of Fuzzy Automaton Detection Mechanism

2.4 Implementation of the Fuzzy Automaton Based Intrusion Detection

To evaluate the proposed detection model in practice, the DARPA 2000 attack scenarios dataset LLDOS1.0 was used [28]. It seems to be a proper benchmark for the multi-step attack. It consisted of different multi-step attack scenarios. One of the benefits of using the DARPA 2000 dataset is that it contains a detailed truth table that allows for the obtained results to be checked. Moreover, most of the IDS detection approaches have applied this dataset for testing and evaluating processes [29]. This work extracts the first attack scenario which was a DDOS multi-step attack.

According to the extracted DDOS multi-step attack scenario, the attacker aimed to install the DDOS multi-step attack on any computer within the target network. The attack was based on five steps [30]. It lasted three hours and was performed for these subnets 172.16.112.0/24, 172.16.113.0/24, 172.16.114.0/24 and 172.16.115.0/24. Consequently, there were three hosts infected by the DDOS multi-step attack. These hosts were 172.16.115.20, 172.16.112.50 and 172.16.112.10. Table 8 illustrates the five sequence steps of the first DARPA attack scenario.

Table 8. The Sequence Steps of The DARPA Attack Scenario

Step	Name	Time
1	IP Sweep	09:45 - 09:52
2	Sadmind	10:08 - 10:18
3	Break-In	10:33 - 10:34
4	Installation	10:50
5	Launching	11:27

- Step (1): The attacker sends a large number of Internet Control Message Protocol (ICMP) echo requests in this sweep and waits for the echo replay to obtain the live IP addresses (hosts).
- Step (2): The result of the step (1) is the list of live hosts. Every live host in the previous step was probed to define the hosts running the sadmind service. The sadmind investigation was applied using sadmind exploit software and ping command.
- Step (3): The result of step (2) is the list of live hosts running the sadmind service. The break-in script was executed for every live host. Break-in script tries the sadmind remote to root access. During the period (10:33 to 10:34) there were 6 break-in attempts.
- Step (4): The result of step (3) is the list of infected hosts (three hosts were infected). Herein, the break-in script executed the remote to root successfully. Therefore, the attacker had the required access to install the DDOS multi-step attack for these infected hosts.
- Step (5): The attacker launched the DDOS multi-step attack using the TELNET login.

The simulated DDOS multi-step attack scenario lasted for a total of (11836 seconds). Table 9 presents the DDOS multi-step attack phases according to the simulation time.

Table 9. The Phases During The DDOS Multi-step Attack

Attack States	Description	Time in Seconds
IP Sweep	Step 1 of Attack	1500 - 1920
Sadmind	Step 2 of Attack	2880 - 3480
Break-in	Step 3 of Attack	3650 - 5200
Installation	Step 4 of Attack	5400 - 6500
launching	Step 5 of Attack	7620 - 11836

The DARPA attack scenario dataset LLDOS1.0 was reformulated by extracting the values of the main features and labeling the data according to the existing literature results [27, 30]. The fuzzy automaton detection mechanism's fuzzy system states are defined as follows: $S = N, A, P, C$. The initial state of the fuzzy automaton detection mechanism is assumed to be in the normal state (N). The N state indicates there are no attack attempts or privacy violations; the system is in normal mode. The A state indicates that there are some attempts to gather and probe for information using IP Sweep and sadmind. The P state indicates that malicious activity has been initiated by running the break-in and installation scripts. The C state indicates that the system has been completely infected; the multi-step attack has been launched successfully.

The fuzzy automaton based intrusion detection mechanism's input parameters (the set of observations) are the reformulated DARPA attack scenario. Due to a large number of extracted features and for the sake of simplification, one-eighth of the total number of features was selected as an input parameter. The state transition rule base definition is based on expert heuristic. An efficient tool for easily creating an expert fuzzy rule-base is the Fuzzy Behavior Description Language (FBDL) [31]. It is a declarative language providing a simple structure for defining the state-transition rule-base size in a humanly readable form, closely resembling the original verbal form. Regarding the fuzzy declarative language, there are two conditions used to define the state transition rule-base:

- Each rule-base should have a unique name.
- The name of the rule-base must be the same as the name it's consequent.

The aim behind using the FBDL is to present a simple form for defining the state transition rule base which could be more readable and understandable by a human.

It is worth mentioning that, in the classical reasoning methods, the size of the state-transition rule-base grows exponentially with the number of the inputs (observations). For this reason, the proposed detection mechanism adapts the FRI, as it can effectively reduce the size of the state-transition rule-base. The fuzzy automaton detection mechanism has continuous states which are presented as a vector of membership values. These states were defined in the fuzzy declarative language as follow:

```
Universe "Normal State"
Description "The Degree of Normal State"
    "Low"  0 0
    "High" 1 1
End
Universe "Attempt State"
Description "The Degree of Attempt State "
    "Low"  0 0
    "High" 1 1
End
```

```
Universe "Prerequisite State"
Description "The Degree of Prerequisite State "
    "Low"  0 0
    "High" 1 1
End
Universe "Compromise State"
Description "The Degree of Compromise State "
    "Low"  0 0
    "High" 1 1
End
```

The applications of the FRI methods are beneficial in the IDS application area [5]. Using FRI methods, expert knowledge can be used as the basis of fuzzy rules. In the suggested FRI fuzzy automaton detection mechanism, the rules are not strict; the expert can sort some of the known cases only. Most important cases and scenarios can be sufficiently defined by using the proposed fuzzy declarative language. The description contains the definition of ranges (as universes) and the rules (in the form of rule-bases). The definition of the universes describes non-linear scaling on the considered input and output dimensions. Experts must define language symbols which may be similar to the domain-specific terms. Therefore, the FRI method formalizes the expert knowledge to the form, which can be interpreted and evaluated automatically by the inference engine. Using the language symbols allows the results to more closely resemble the natural language equivalent.

The universe definitions of the observations of the proposed detection mechanism are defined based on the expert knowledge and presented in the fuzzy declarative language as follows:

```
 Universe "Pure_A2B "
"VSmall"   0 31
"Small"     31 69
"Medium"  161 69
"Large"     2430 616
"VLarge"   8845 2430
End
Universe "Pure_B2A "
"Low"   0 380
"High"  380 780
End
```

```
Universe "Total_A2B "
"Small"   1 8400
"Large"   8400 17693
End
Universe "Mss_Request"
"VSmall"   1 1987
"Small"     2200 2700
"Medium"  3200 5350
"Large"     6500 8000
End
```

The state-transition rule-bases were defined based on expert knowledge. Fourteen state

transition rules were constructed. For example, the attempt state rule definitions and the prerequisite state rule definitions presented as follows:

Rulebase "Attempt_State"

Rule

"High" when

"Mss_Requested" is "Medium" and
"Pure_A2B" is "Meduim"

end

Rule

"High" when

"Pure_A2B" is "Small" and
"Mss_Requested" is "Medium"

end

Rule

"High" when

"Pure_B2A" is "High" and
"Mss_Requested" is "Medium"

end

Rule

"Low" when

"Mss_Requested" is "Small"

end

Rulebase "Prerequisite_State"

Rule

"High" when

"Mss_Requested" is "Large" and
"Pure_A2B" is "VSmall"

end

Rule

"High" when

"Mss_Requested" is "Medium" and
"Pure_A2B" is "Small" and

"Pure_B2A" is "Low"

End

Rule

"Low" when

"Mss_Requested" is "VSmall"

end

Rule

"Low" when

"Mss_Requested" is "Large"

end

End

According to the way, as the FRI (FIVE) method calculates the conclusion, the evaluation process of rule bases can be described in a bottom-up manner. In the first step, the inference engine calculates the observations' distances from the defined symbols on the given universes. Subsequently, the rules' distances are evaluated. In the considered configuration, the rule's distance is the normalized Euclidean norm of the included symbol distances. The measure of the rule-base was obtained by the Shepard interpolation (inverse distance weighting) of the rule distances and their consequent values.

The proposed detection mechanism generated intelligible results due to its fuzzy nature, subsequently allowing the degree of the system state to be determined and for the system to be in more than one state at the same time. Table 10 presents the proposed detection mechanism's output response in case of intrusion instances. Unlike DFSM and HMMs, the system states within the proposed detection mechanism are presented as a vector of membership values. This could benefit administrators because it helps them to understand the current security status and to mitigate future risks by forecasting the upcoming system state.

Table 10. The Output Response of The Suggested FRI Fuzzy Automaton Based IDS

Input Parameters			
	Instance 1	**Instance 2**	**Instance 3**
Mss request	4200	6300	3869
Pure A2B	110	96	141
Pure B2A	614	750	688
Total A2B	10536	9365	12369
The Proposed Detection Method Output			
	Output 1	**Output 2**	**Output 3**
Normal	0.270791	0.085362	0.126221
Attempt	0.919518	0.212365	0.932641
Prerequisite	0.446831	0.926831	0.482133
Compromise	0.157381	0.357381	0.198752

The fuzzy automaton detection mechanism was tested and evaluated in the following durations of the DDOS multistep attack: (15-1062 seconds), (1800-2786 seconds), (37505191 seconds) and (8210-10342 seconds). These durations were chosen to verify the performance of the fuzzy automaton detection mechanism in order to detect the DDOS multi-step attack in its early stages before it posed a severe risk.

The fuzzy automaton detection mechanism was evaluated using 5639 observations. The first detection was obtained by the fuzzy automaton detection mechanism at 1800 seconds, 2 minutes before the attacker completes the works in step 1. The second detection was obtained at 3750 seconds, 24 minutes before the attacker completes the works in step 3. The third detection was at 8210 seconds, 60 minutes before the attacker completes the works in step 5. Thus, early detection of the multi-step attack gives administrators time to take the necessary actions to mitigate any future risk from this type of attack. The IDS detection mechanism's standard performance measure is typically performed using both the Receiver Operating Characteristic (ROC) and the confusion matrix [29], where the ROC shows the trade-off between sensitivity and specificity [32]. In keeping with the standard measure of most other IDS detection mechanisms, Fig. 14. shows the evaluation performance, with the ROC curve, for the fuzzy automaton detection mechanism states.

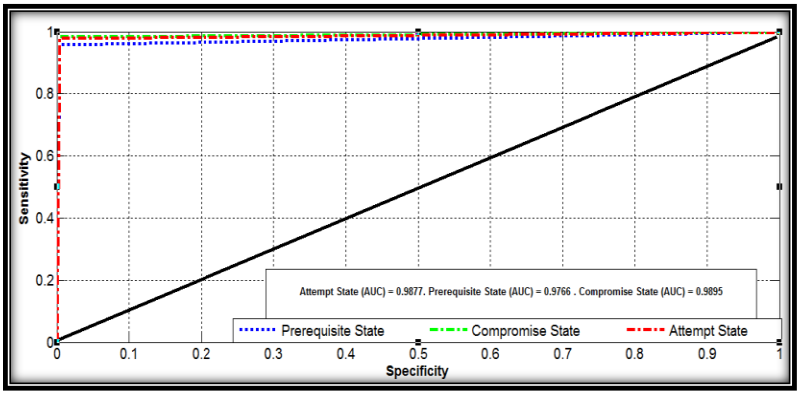

Fig. 14. The ROC Curve for The Fuzzy Automaton Detection States

Table 11 illustrates the confusion matrix obtained during the evaluation process. The results reflected that the fuzzy automaton detection mechanism obtained a 97.836% overall accuracy rate. Furthermore, the implemented experiments demonstrated that the fuzzy automaton detection mechanism was able to detect the DDOS multi-step attack within its early stages, using the DARPA dataset. Therefore, the early detection of the multi-step attack could be beneficial for the administrator to perform the required mitigation actions.

Table 11. The Confusion Matrix of The Evaluation process

	Normal	**Attempt**	**Prerequisite**	**Compromise**	**Overall Observations**	**Precision**
Normal	1045	2	2	0	1049	**99.619%**
Attempt	13	985	8	0	1006	**97.913%**
Prerequisite	0	58	1312	1	1371	**95.697%**
Compromise	0	0	38	2175	2213	**98.283%**
Truth Overall	1058	1045	1360	2176	5639	
Overall Accuracy	**97.836%**					

For summarizing the results of the benchmark based tests, it can be stated, that the suggested FRI fuzzy automaton based IDS could be a promising mechanism for detecting multi-step attacks. The FRI fuzzy automaton based detection mechanism can be characterized by the following key points:

- The fuzzy automaton detection mechanism offers the system states as a vector of membership values.
- Unlike the DFSM, the system can be in more than one state at the same time, thereby allowing the fuzzy automaton detection mechanism to follow more than one path of system states changes.
- Adapting the FRI (FIVE) method offers interpolated results even when lacking knowledge-based representation. In other words, The FRI (FIVE) method interpolates the results even when some of the state transition rules are missing.
- The fuzzy automaton detection mechanism produces verbal detection results which can be more easily understood by administrators.
- The fuzzy system extends the binary decision to the continuous space, smoothing the boundaries and offering a solution to the boundary problem in addition to generating more comprehensible results.
- The fuzzy automaton detection mechanism can detect the DDOS multi-step attack within its early stages, using the DARPA dataset. Thus, early detection could help the administrator mitigate this type of attack.
- The proposed detection mechanism's strength is based on combining fuzzy automaton and FRI based reasoning. Thus, the fuzzy system effectively smooths the decision boundary

between normal and intrusion traffics, avoiding the binary decision. And the FRI based implementation is eliminating the need for the complete state-transition rule-base definition.

The main characteristics of the proposed FRI fuzzy automaton IDS and the other state machine detection mechanisms are compared in Table 12.

Table 12. The Main Characteristics of Some Widely Used Detection Methods

	HMM Detection Mechanism	**DFSM Detection Mechanism**	**Fuzzy Automaton Detection Mechanism**
Binary Decision	Yes	Yes	Approximated
System State	Discrete	Discrete	Continuous
Uncertainty	Not Applicable	Not Applicable	Applicable
Rules	Statistical	Knowledge Base	Knowledge Base

The proposed FRI fuzzy automaton IDS eliminates the boundary decision problem, which is considered as a constant challenge because in real situation there are no clear boundaries between the normal and intrusion traffics. In addition, implementing the FRI (FIVE) method instead of the classical reasoning methods for the reasoning part helps to reduce the total number of state-transition rules (simplification) and offers interpolated results even if the knowledge representation is incomplete.

Contribution and Future Research Direction

This dissertation contributes to the field of fuzzy systems (especially fuzzy rule interpolation), intrusion detection system and also fuzzy state machine.

A novel method, IDS model-based fuzzy rule interpolation along with a novel method to detect abnormalities by combining the Fuzzy Interpolation based on the Vague Environment (FIVE) FRI reasoning with the Management Information Base (MIB) parameters have been constructed. This method not only allows the intrusion detection system to be used in continuous spaces but makes it possible to use a sparse fuzzy rule base. This way, the overall rule base size significantly smaller. Because of its fuzzy rule base knowledge representation nature, it can be easily adapt expert knowledge, and also be suitable for predicting the level of degree for threat possibility. Furthermore, the combining of the (FIVE) FRI reasoning with the Management Information Base (MIB) parameters is a promise detection method to mitigate the network intrusions. See Thesis I. and Thesis II. below, and also the fifth and sixth chapters for a detailed description of these methods. The incorporated fuzzy rule interpolation method, FIVE, has been successfully constructed specifically for the intrusion detection systems, taking the performance of these methods to a higher level.

Furthermore, this dissertation proposes a novel model for detecting the multi-step attacks. The proposed model was built upon the fuzzy rule interpolation-based fuzzy state machine. Based on the proposed model which has a simple rule-based knowledge representation format and where the completeness of the rule-base is not required. This model does not only allows the intrusion detection system to be used in continuous spaces, but also makes it possible to be in more than one state at the same time. Additionally, the proposed model interpolates the results even when some of the state transition rules are missing. Details on the structure and implementation of the model can be found in chapters seven and eight, and also see thesis III and thesis IV. below.

Future research and investigation into the possibilities of adapting the IDS based fuzzy rule interpolation in the Internet of Things (IoT) based smart environments seems promising. The IoT paradigm has recently evolved to incorporate different application areas. In the IoT environment, several heterogeneous devices are connected via different types of sensors. These wireless sensors beside the IPv6 added advantage to extend the IoT environment to serve many application areas. The heterogeneous devices within the IoT environment may have a different level of security. Some of the IoT devices have little or no security embedded into them. This deficiency could affect the availability of the IoT connected network and made several security flaws. Moreover, attackers continuously targeted the modern aspects of technology, and trying abusing these technologies using complex attack scenarios such as Botnet attacks. Due to the limited computing and storage capabilities of IoT devices and the specific protocols used, typical IDS may not be suitable for IoT environments. Therefore, the IDS based fuzzy rule interpolation could be a suitable alternative to mitigate IoT-related security attacks. This is due to its fuzzy nature, and its ability to render results even when faced with only partially completed fuzzy rules. Moreover, the results are rendered in

a human-readable form.

Another direction for future research lies in investigating the potential for adapting the fuzzy rule interpolation for the Cyber Forensics system. The field of digital forensic is growing dramatically in parallel with the rise of computer crimes. Network forensics is a digital forensic in networked environments. It consists of a large amount of network traffic. All this traffic needs to be analyzed and investigated to define the digital evidence; however, not all of this information is useful as evidence. Therefore, the irrelevant information needs to be removed by an expert. The expert also is responsible to define some rules to decide which information could be used as digital evidence. Thus, there is an emerged need to design an automated system for analyzing the network forensics, and at the same time had the ability to deal with the issues associated with the deficiencies of the knowledge-based representation. Furthermore, generating the digital evidence with a level of severity could be beneficial for clarifying computer crimes. Also, it helps the expert to understand the current digital evidences in a more readable form. Therefore, the fuzzy rule interpolation reasoning methods could be suitable for use as an expert system capable of providing the forensics experts with the necessary information to successfully reduce the time, and cost of analyzing the network forensics, and generating the digital evidence in case of lacking knowledge-based representation.

3. Summary

Thesis I.: [5]

The FIVE based fuzzy rule interpolation model can be used in the IDS as a suitable inference method. Furthermore, the FRI inference system has yielded promising results when implemented as an IDS detection mechanism. Additionally, during the studies test application, the FRI inference system effectively decreased the rate of false positive values. Moreover, because of it's tendency for fuzzy rule based knowledge representation, it can easily adapt to expert knowledge, and be suitable for predicting the potential threat level.

Thesis II.: [33]

The FIVE based fuzzy rule interpolation for SNMP-MIB data based intrusion detection achieving acceptable results in an IDS detection mechanism. I concluded that, using this method there is no need to deal with raw traffic processing, which is time-consuming, and difficult to compute. The MIB parameters reflect the normal and abnormal nature of the network traffics. Furthermore, expert knowledge can be easily adapted by eliminating the need for creating a complete fuzzy rule base.

Thesis III.: [34]

The FRI (FIVE) based fuzzy automaton could be a suitable model for detecting and preventing the multi-step attacks. By offering interpolated conclusion even for situations that are not explicitly defined, the FRI method instruments the fuzzy automaton to be able to act on partly defined state transition rule-base. The integration of fuzzy state machine and fuzzy rule interpolation allows a discretely defined state-machine to act on continuous universes and handle uncertainty in applications like intrusion detection systems.

Thesis IV.: [34]

The FRI (FIVE) based fuzzy automaton, with an expert-defined state-transition rule-base given in Fuzzy Behavior Description Language (FBDL), was suitable for detecting and preventing the multi-step attacks in stages. The state-transition fuzzy rule-base is required for the FRI (FIVE) based fuzzy automaton intrusion detection model, and can be easily defined using the FBDL.

Author's Publication

5. Mohammad Almseidin and Szilveszter Kovacs. Intrusion detection mechanism using fuzzy rule interpolation. *Journal of Theoretical and Applied Information Technology*, 96(16):5473–5488, 2018. **Scopus Indexed [Q3].**

33. Mohammad Almseidin, Mouhammd Al-kasassbeh and Szilveszter Kovacs. Fuzzy Rule Interpolation and SNMP-MIB for Emerging Network Abnormality. *International Journal on Advanced Science, Engineering and Information Technology*, vol. 9, no. 3, pp. 1-10, 2019. **Scopus Indexed [Q2].**

34. Mohammad Almseidin, Imre Piller , Mouhammd Al-kasassbeh and Szilveszter Kovacs. Fuzzy Automaton as a Detection Mechanism for the Multi-Step Attack. *International Journal on Advanced Science, Engineering and Information Technology*, vol. 9, no. 2, pp. 17-31, 2019. **Scopus Indexed [Q2].**

35. Ibrahim M Obeidat, Nabhan Hamadneh, Mouhammd Alkasassbeh, Mohammad Almseidin, and Mazen Ibrahim AlZubi. Intensive pre-processing of kdd cup 99 for network intrusion classification using machine learning techniques. *International Journal of Interactive Mobile Technologies*, 13(1), 2019. **Scopus Indexed [Q3].**

36. Mohammad Almseidin, Maen Alzubi, Szilveszter Kovacs, and Mouhammd Alkasassbeh. Evaluation of machine learning algorithms for intrusion detection system. *In Intelligent Systems and Informatics (SISY), 2017 IEEE 15th International Symposium on*, pages 000277– 000282. IEEE, 2017. [Online]. Available: https://doi.org/ 10.1109/SISY.2017.8080566.

37. Mouhammad Alkasassbeh and Mohammad Almseidin. Machine learning methods for network intrusion detection. *International Journal of Computer and Information Engineering* 12 : 8 p. 1 (2018).

38. Mohammad Almseidin, Mouhammd Al-kasassbeh and Szilveszter Kovacs, Detecting Slow Port Scan Using Fuzzy Rule Interpolation. *In IEEE International Conference on New Trends in Computing Sciences (2019).*

References

[1] C Yuan. Research on multi-step attack detection method based on gct. Master's Thesis, Jilin University, Jilin, 2010.

[2] Yanxue Zhang, Dongmei Zhao, and Jinxing Liu. The application of baum-welch algorithm in multistep attack. The Scientific World Journal, 2014.

[3] Salem Benferhat, Tayeb Kenaza, and Aicha Mokhtari. A naive bayes approach for detecting coordinated attacks. In Annual IEEE International Computer Software and Applications Conference, pages 704–709. IEEE, 2008.

[4] Can Chen and BQ Yan. Network attack forecast algorithm for multi-step attack. Computer Engineering, 5(37):172–174, 2011.

[5] M. Almseidin and S. Kovacs. Intrusion detection mechanism using fuzzy rule interpolation. Journal of Theoretical and Applied Information Technology, 96(16):5473–5488, 2018.

[6] R Shanmugavadivu and N Nagarajan. Network intrusion detection system using fuzzy logic. Indian Journal of Computer Science and Engineering (IJCSE), 2(1):101–111, 2011.

[7] Szilveszter Kovács. New aspects of interpolative reasoning. In Proceedings of the 6th. International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Granada, Spain, pages 477–482, 1996.

[8] Szilveszter Kovács and László T Kóczy. The use of the concept of vague environment in approximate fuzzy reasoning. Fuzzy Set Theory and Applications, Tatra Mountains Mathematical Publications, Mathematical Institute Slovak Academy of Sciences, Bratislava, Slovak Republic, 12:169–181, 1997.

[9] Szilveszter Kovacs and Laszlo T Koczy. Approximate fuzzy reasoning based on interpolation in the vague environment of the fuzzy rulebase. In Intelligent Engineering Systems, 1997. INES'97. Proceedings., 1997 IEEE International Conference on, pages 63–68. IEEE, 1997.

[10] Zs Cs Johanyák. Sparse fuzzy model identification matlab toolox-rulemaker toolbox. In Computational Cybernetics, 2008. ICCC 2008. IEEE International Conference on, pages 69–74. IEEE, 2008.

[11] Rufai Kazeem Idowu, Zulaiha Ali Othman, et al. Denial of service attack detection using trapezoidal fuzzy reasoning spiking neural p system. Journal of Theoretical & Applied Information Technology, 75(3), 2015.

[12] Swati Dhopte and NZ Tarapore. Design of intrusion detection system using fuzzy class-association rule mining based on genetic algorithm. International Journal of Computer Applications, 53(14), 2012.

[13] Zsolt Csaba Johanyák and Szilveszter Kovács. Sparse fuzzy system generation by rule base extension. In Intelligent Engineering Systems, 2007. INES 2007. 11th International Conference on, pages 99–104. IEEE, 2007.

[26] Zsolt Csaba Johanyák and Szilveszter Kovács. Polar-cut based fuzzy model for petrophysical properties prediction. Scientific Bulletin of Politehnica University of Timisoara, Romania, Transactions on Automatic Control and Computer Science, 57(67):195–200, 2008.

[15] SN Sivanandam, Sai Sumathi, SN Deepa, et al. Introduction to fuzzy logic using MATLAB, volume 1. Springer, 2007.

[16] Zsolt Csaba Johanyák, Domonkos Tikk, Szilveszter Kovács, and Kok Wai Wong. Fuzzy rule interpolation matlab toolbox-fri toolbox. 2006.

[17] Mohammad Almseidin, Maen Alzubi, Szilveszter Kovacs, and Mouhammd Alkasassbeh. Evaluation of machine learning algorithms for intrusion detection system. In Intelligent Systems and Informatics (SISY), 2017 IEEE 15th International Symposium on, pages 000277–000282. IEEE, 2017.

[18] Mouhammd Al-Kasassbeh, Ghazi Al-Naymat, and Eshraq Al-Hawari. Towards generating realistic snmp-mib dataset for network anomaly detection. International Journal of Computer Science and Information Security, 14(9):1162, 2016.

[19] Mohiuddin Ahmed, Abdun Naser Mahmood, and Jiankun Hu. A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60:19–31, 2016.

[20] Jelena Mirkovic and Peter Reiher. A taxonomy of ddos attack and ddos defense mechanisms. ACM SIGCOMM

Computer Communication Review, 34(2):39–53, 2004.

[21] Frank Klawonn. Fuzzy sets and vague environments. Fuzzy Sets and Systems, 66(2):207–221, 1994.

[22] Mouhammd Alkasassbeh. An empirical evaluation for the intrusion detection features based on machine learning and feature selection methods. Journal of Theoretical and Applied Information Technology, 95(22), 2017.

[23] Rahul Kumar Singh and Ajay Guide Kumar. Conversion of Fuzzy Regular Expressions to Fuzzy Automata using the Follow Automata. PhD thesis, Thapar University, 2014.

[24] Mansoor Doostfatemeh and Stefan C Kremer. New directions in fuzzy automata. International Journal of Approximate Reasoning, 38(2):175–214, 2005.

[25] Szilveszter Kovács, Dávid Vincze, Márta Gácsi, Ádám Miklósi, and Péter Korondi. Ethologically inspired robot behavior implementation. In Human System Interactions (HSI), 2011 4th International Conference on, pages 64–69. IEEE, 2011.

[26] Szilveszter Kovács. Fuzzy rule interpolation. In Encyclopedia of Artificial Intelligence, pages 728–733. IGI Global, 2009.

[27] Alireza Shameli Sendi, Michel Dagenais, Masoume Jabbarifar, and Mario Couture. Real time intrusion prediction based on optimized alerts with hidden markov model. JNW, 7(2):311–321, 2012.

[28] DARPA Datasets. Mit lincoln laboratory, darpa intrusion detection evaluation data sets, 2000.

[29] Joel Branch, Alan Bivens, Chi Yu Chan, Taek Kyeun Lee, and Boleslaw K Szymanski. Denial of service intrusion detection using time dependent deterministic finite automata. In Proc. Graduate Research Conference, pages 45–51, 2002.

[30] André Årnes, Fredrik Valeur, Giovanni Vigna, and Richard A Kemmerer. Using hidden markov models to evaluate the risks of intrusions. In International Workshop on Recent Advances in Intrusion Detection, pages 145–164. Springer, 2006.

[31] Imre Piller, Dávid Vincze, and Szilveszter Kovács. Declarative language for behaviour description. In Emergent Trends in Robotics and Intelligent Systems, pages 103–112. Springer, 2015.

[32] Aleksandar Lazarevic, Vipin Kumar, and Jaideep Srivastava. Intrusion detection: A survey. In Managing Cyber Threats, pages 19–78. Springer, 2005.