Pragmatic Miner to Risk Analysis for Intrusion Detection (PMRA-ID)

Samaher Al-Janabi(✉)

Department of Computer Science, Faculty of Science for Women (SCIW),
University of Babylon, Babylon, Iraq
samaher@uobabylon.edu.iq

Abstract. Security of information systems and their connecting networks has become a primary focus given that pervasive cyber-attacks against information systems are geometrically increasing. Intrusion Detection and Prevention Systems (IDPS) effectively secure the data, storage devices and the systems holding them. We will build system consist of five steps: (a) description the orders that required to archives the event by five fuzzy concepts as input and three fuzzy concepts as output, then save it in temporal bank of orders, (b) Pre-processing that order by convert from the description to numerical values and compute the Membership function for that values. (c) applied the association data mining techniques on these database after compute the correlation among their features, this lead to generation thirty two rules but not all this rules is salsify the confidence measures (i.e., we take only the rules that satisfy the purity 100%) (d) Building the Confusion matrix for all the samples using in training processing (e) Testing the Pragmatic Miner to Risk Analysis (PMRA) model and verification from the accuracy of their results by press new samples to model not used in training stage then compute the values of error and accuracy measures, in addition of correct. The existing systems employing firewalls and encryptions for data protection are getting outdated. IDPS provides a much improved detection system that can prevent the intrusions to attack the system. However, as effective as it is in preventing intrusions, which can disrupt the retrieval of desired information as the system sometimes perceives it as an attack. The base aim of this work is to determine a way to risk analysis of IDPS to an acceptable level while detecting the intrusions and maintaining effective security of a system. Experimental results clearly show the superficiality of the proposed model against the conventional IDPS system.

Keywords: Information security · Intrusion Detection and Prevention
Fuzzy descriptive · Data mining · Error and accuracy measures · Risk analysis

1 Introduction

One of the obligations of government, corporations, private business, financial and healthcare institutions, the military and several other organizations is to gather a detailed information of their customers, products, researchers, employees and their financial status. Informatics takes the key responsibility for providing the security, privacy and confidentiality of the primary digital data. Collection, processing and

A. Mohamed et al. (Eds.): SCDS 2017, CCIS 788, pp. 263–277, 2017.
https://doi.org/10.1007/978-981-10-7242-0_23

storage of most of the information on computers, necessarily takes place before transmission of the information across networks. Likewise, protection of a patient's medical history is very taxing.

The growth in the use of information management systems has become more powerful and widely distributed [1, 15] allowing the number of threats facing computers networks and the files stored in them grow and diversify, as a result. Researchers are encouraged to intensify and focus on intrusion detection mechanisms so that novel attacks can be detected through anomaly detection since ways of attacks are increasing and misuse detection functions often fail to detect no signature of novel attacks [3].

Intrusion Detection and Prevention system (IDPS) can perform an early detection of malicious activities and therefore, an organization can easily monitor events occurring within its computer network or system. Then, it can shield itself from misuse of its information, analyze the network for signs of intrusion and can prevent serious damage to already protected systems [10].

IDPS can disclose abnormal patterns of behaviors through the establishment of baselines of normal usage patterns and anything that differs from the norm is considered as a possible intrusion [10]. Unfortunately, the false alarms challenge this system, which causes inadvertent system behaviour and unnecessary data processing [7], This paper is similar with [7] in stage of using Fuzzy Logic as pre-processing stage for raw data but it different with it in terms of its use of data mining techniques in a constructive rules and then it'll take the rules that satisfy t the percentage of confidence 100% and ignore the rest. Also, the current research evaluated the results reached by a precision and error scales.

IDPS triggers a positive alarm when it mistakenly interprets a benign activity as a malicious one. On the other hand, sometimes IDPS does not detect an attack or intrusion; still it flags a false positive alarm. Once the basic criteria is used to distinguish IPSs from IDSs, the intrusion prevention system of the former tries to prevent the success of detected threats, unlike the latter [2, 12]. The IPS can change the content of the attack or change the security environment into a countermeasure. It can also alter the configuration of other security controls in order to stop an attack by denying access to the attacker or disabling the target so that the attack cannot proceed. The host-based firewall settings can be reconfigured so that the incoming attack is completely hindered. Some IPSs can take away the malicious parts of the attack and render those as impotent to attack [5].

IDPSs fail to provide a complete and an accurate detection of attacks because they do not employ proper risk analysis and assessment techniques [5]. This paper presents a new way of detection and prevention of intrusions based on risk analysis and risk assessment in a way that the false alarm rate will be minimized in an IDPS. This approach employs a Pragmatic Miner-risk analysis technique to analyze the generated alarms. The Pragmatic Miner technique, also known as Pragmatic Miner to Risk Analysis (PMRA) computes the significance and impact severity of the detected activities. By doing so, the system will adapt itself while evaluating an activity should be regarded as an attack attempt or a possible normal behavior.

The organization of the paper is as follows: Sect. 2 contains the latest trends in the PMRA researches. Section 3 discloses the most significant limitations in the preexisting intrusion detection and prevention methods. Section 4 reviews the proposed architectural model of the PMRA system. Section 5 presents the mechanics and

experimental results obtained from the implementation of PMRA. The result proves that the proposed architecture is helpful in order to predict the vulnerabilities and array countermeasures that can be used against reasonable and manageable alarm rates. Finally, Sect. 6 contains a brief discussion and the conclusion of this paper, with suggested possible future plans that can yield better results.

2 Related Works

IDPS has been a major topic of interest in terms of research work to develop sophisticated systems. In this section, some of the latest research that relate to soft computing techniques regarding false alarm rates in IDPS are presented. [9] used the Self-Organizing Map (SOM) neutral networks to develop a two-stage classification system. By this way, it is possible to reduce the false alerts to a level, which can be over 50% of the entire false alarm occurrences. Conversely, [6] used a technique for mining the data based on a Growing Hierarchal Self-Organized Map (GHSOM). The GHSOM can modify its design if necessary, based upon the aspects of the input alarm data. This map reorganizes the alarms according to the incoming data and classifies them as true or false, providing aid to the network administrator.

[8] proposed a post-processing filter that reduces the false positives via a "network-based intrusion detection system." This system employs specific feature of detections that align itself with true attacks to filter alarms and limit false positives.

In another approach, [11] proposed a New Intrusion Detection Method Based on Antibody Concentration (NIDMBAC) to reduce the false alarm rate. Definitions of self, non-self, antigen and detector in the intrusion domain are utilized. An antigen recognition scheme is employed as a part of the clone proliferation. The intrusion detectors are processed to alter the number of antibodies in relation to the overall intensity of the antigen intrusion. In this case, the probabilistic calculation method was used for the intrusion alarm production based on the correlation between the antigen intrusion intensity and the antibody concentration factors. The analysis and results in this proposed method provided better results compared to the traditional methods.

The strategy proposed by [1] is based on statistical analysis, detection of both attacks and normal traffic patterns with respect to a hybrid statistical approach, which utilizes a data mining and decision classification techniques. However, the outcomes of the statistical analysis can be adjusted in order to minimize the miscalculation of false positive alarms and to distinguish between real attacks and false positive alarms of the traffic data.

Each of the above-mentioned schemes provides some solution against intrusion and reduces the rate of false alarm. Yet none of them alone was successful in taking a complete security risk of intrusions into account as a serious issue and is able to reduce the false positive alarms. They all failed to propose a uniform and robust architecture which can secure the whole system and be an aid for the administrator.

In this proposed work, it is assumed that unwanted false positive alarms can be eliminated by using a combination of soft computing techniques ranging from fuzzy logic (FL) for objective risk analysis, knowledge-based systems for determining intrusion patterns, artificial intelligence for machine and reinforced learning. The application of FL is considered appropriate for performing for good risk analysis and

risk assessment. FL ensures that a set of complex related variables are grouped when making a final decision risk assessment and reducing unwanted false positive alarms and protecting the system not just effectively, but also efficiently, which is what PMRA is. In short, by employing PMRA, information and systems can be protected effectively without undue false positives.

3 Limitations of the Current Systems

The occasions when IPS accidentally or sincerely confirms a good activity as a malicious one or vice-versa hence flagging, a false alarm is not palatable to the users. Anomaly detection may not be able to detect an attack sometimes, but can still trigger a high false alarm. For instance, a legitimate system behavior may sometimes can be seen as an abnormal operation; hence, an alarm is triggered with respect to that. Furthermore, anomaly-based IDS systems are prone to false positive alarms, where the prevailing attack is based on changes to the evaluation of "normal" and false attack. Therefore, the application of risk analysis on the detected attacks and the measurement of their exposure factor on impact can help to confirm the validity of the alerts, as well as help to reduce the false alarms to a minimal acceptable level.

Information and computer technology is complicated and dynamic in nature. This nature consequently does not only pose a big problem to limit false alarms, but also for the usage of IDPS, that is also complex. Formulation of a Collaborative-IDPS (CIDPS) with soft computing elements is the suggestion in 2013 to overcome these complexities. This however does not, overcome the other challenges, such as newly injected attacks. Figure 1 shows the CIDPS architecture. The management function flows from the CIDPS in the intermediate section/layer of various components like the fuzzy risk manager, which controls the triggering of false alarms. The knowledge and multi-agent manager manages the intrusions from host computers and network elements that in turns provides the right feeds for the countermeasure operations. It also allows the monitoring and enabling the IDPS to detect hardware and software seamlessly and automatically.

The automatic manager includes four types of agents to cover the four segments of computing: Self-configuration, self-healing, self-optimization and self-protection. Self-configuration is important as the automatic manager makes the rules at runtime. Self-healing allows the system to operate in a cycle to detect the anomalous behavior while a solution is developed. The self-optimization allows the search without compromising other resources. Self-protection allows for the detection of bad functions and updates the knowledge base (KB) to limit its future recurrences. The checker monitors related resources through the consultation of sub ontology. It also detects the abnormal behavior. The ontology is normally updated so that it will be able to identify any non-expected change. At the same time, the checker reports the status to the analyzer agent that determines the system's current state by modeling, often complex behavior of the data to predict future anomalies. An estimated risk tool is used to consult the KB to find the best recourse and take the most appropriate actions. In addition, the KB is also updated to aid in future analysis. The planner structures actions should make sure the goals are achieved and should produce a series of changes that will help the element under protection. The executor performs the healing action, such as updating the policies by following the instructions given to it.

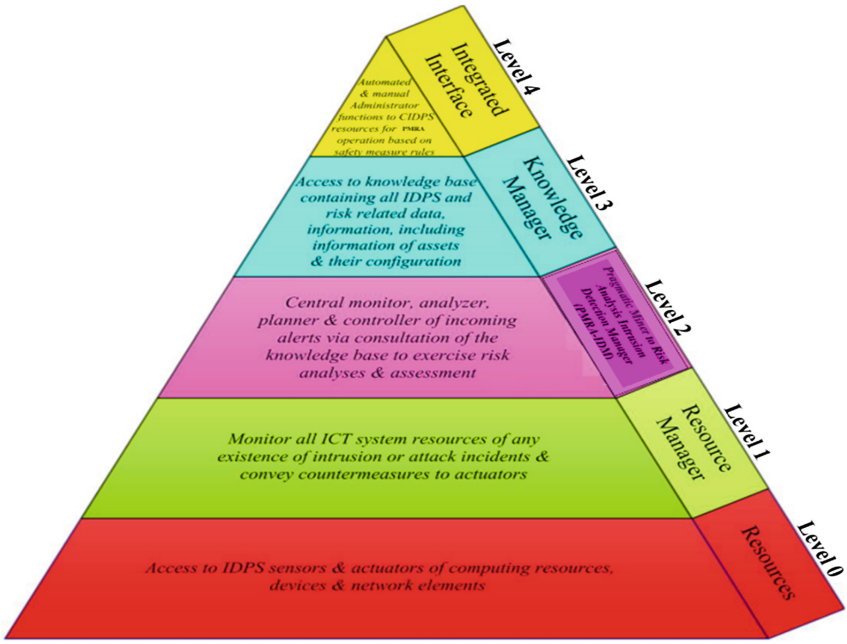

Fig. 1. Information and processes management layers integrating from the lowest to the highest layers in the proposed PMRA model

Furthermore, as soon as a threat is spotted, the affected systems are inspected deeper by the vulnerability scanner. The vulnerability is later assessed, while the scanner makes a real-time picture of all the ongoing attacks in order to determine the possible impact of the attack on the target system. The domain ontology, including high-level concepts (attacks, vulnerabilities and incidents, etc.) is acquired and the risk calculator then assigns a critical rating to the assets. After this step, any intrusion prevention solutions can be taken to evade intrusions, ensure appropriate system operation and limit operational overheads. For example, in the instance that intrusion prevention rules cannot be used on certain systems, a range can be disabled on a specific IP address, reducing false positive alarms. Furthermore, this dynamic protection/prevention allows the system to maintain a constant state of monitoring, assessment and optimization. A proper analysis of the false alarm reduction strategy requires the actual risk exposure to the attacked assets to be quantified.

4 Proposed Pragmatic Miner to Risk Analysis (PMRA) Model

To minimize the exposure of vulnerable resources on information systems and network services is the main purpose of using IDPS [10, 13, 14]. This paper proposed the framework as a way of reducing false alarm rates in intrusion detection systems by the

implementation of Pragmatic Miner to Risk Analysis (PMRA) model. PMRA can compute the significant and the impact severity of each suspected activity in terms of malicious objects. In the process, the system will be capable of making more accurate decisions whether an activity or an object is an attack attempt or a normal system behavior that is misinterpreted by the detection mechanism. Figure 1 describes a five layered organization of the model. Each of the layers has a distinct function and services.

The hierarchy in Fig. 1 illustrates the five layers of the proposed PMRA model. These layers are explained in the following paragraph, while Table 1 shows a description to those management layers.

Table 1. Description of the five management layers of the proposed PMRA model

LAYER	MODULE	FUNCTIONALITY	
0	Resources	All the computing resources, transmission network, network elements, monitoring devices and IDPS sensors and actuators accessed by the resource manager that come under its purview and responsibilities.	
1	Resource Manager	The resource manager manages the total number of sensors under its purview that monitors the ICT system resources of any existence of intrusion or attack incidents.	
2	Pragmatic Miner to Risk Analysis Intrusion Detection Manager (PMRA-IDM)	Monitor	This sub-layer is responsible for monitoring, gathering incoming security of related information and analyzing same for indications of the potential incidents, malicious activities or policy violations.
		Analyzer	This sub-layer is responsible for evaluating and computing the risk of the detected attacks by using Pragmatic Miner technique as well as the predefined policies, rules and standards. On the other side, the guidelines and information from the knowledge-base are administered and controlled by the knowledge manager layer. In conventional systems, managing the knowledge-base would be provided by the system administrator in an ad hoc manner from time of initiation and subsequent updates.
		Planner	This sub-layer provides a way to observe and analyze the problems to better determine if any modifications need to be made, considering the risk assessment obtained from the analyzer module.
		Controller	This sub-layer provides the mechanism to schedule and carry out the necessary changes to protect the elements, which are under attack.
3	Knowledge Manager	The knowledge manager is not only the source of all pertinent knowledge via the knowledge-base that harvest and host details of data, information and existing rules and defines new rules as the operational case of the overall system warrants but also gives the general information of all assets under its control. This knowledge and information are formed by facts, beliefs, rules, norms and contracts. All the essential experiences, learning and knowledge are stored in the knowledge-base as a central repository, which is populated with and include the: 1) rules, policies, guidelines, results from the previous IDPS actions. 2) The security management functions, associated formulas and algorithms that leverage through the integrated interface layer by the system administrator functions (where the knowledge, information and data are the key inputs used in the risk analysis and risk assessment processes).	
4	Integrated Interface	The integrated interface is a unique bridging point between the system administrator functions (both in an automated and manual mode of operation) and the CIDPS. The learned experiences and system operation knowledge are important in defining and updating the system policies, rules and guidelines in the knowledge-base for subsequent consumption by various CIDPS/PMRA components in order to be continuously computed and operated with the latest information.	

5 Deployment and Analysis of the Pragmatic Miner to Risk Analysis (PMRA) Model

The fundamental parts of PMRA are (Preprocessing stage using FL, Mining Stage using Classify association rules, Evaluation stage using accuracy and error measures). Fuzzy linguistic variables or fuzzy expressions also termed as input and output parameters. Low, medium, high, very high, and critical respectively are the membership functions that were used for each input variable. The output functions have three variables (countermeasures) and they are 'Avoidance', 'Transference' and 'Acceptance'. Table 2 shows the characteristics of the input and output variables.

Table 2. Shows the input and output variables.

INPUT	
Residual risk	Low, Medium, High, Very High, Critical
Explore risk	
OUTPUT	
Countermeasures	Avoidance, Transference, Acceptance

The first input is the RR that can be defined as an "indicator of risk conveyed by an asset" and the second input is the ER that can be defined as an "indicator to risk generated by the attack." The ranges of inputs are divided into five classes (fuzzy sets) for each of the residual and exposed risks. The ranges of risk extend from 'Critical' to 'Low' with 'Very High', 'High' and 'Medium' falls in between of them.

5.1 Membership Functions for Input and Output Pragmatic Miner Model

As far as "fuzzifications" are concerned, relevant events determine the type of membership functions that should be used in the experiment [4]. The trapezoidal-shaped membership function is used to describe the fuzzy sets for input and output variables.

Furthermore, three basic membership functions for the countermeasure output are defined in the fuzzy sets First, the membership function denotes 'Avoidance' that is a high-risk exposure requiring some action to eliminate the threat. Second, denotes the 'Acceptance' that is a low- risk exposure that not necessarily requires action against the threat. Finally, the 'Transference' as the final step requires an expert judgment that needs to be taken by the system administrator.

X is a function of the trapezoidal curve x. The curve also depends on a, b, c and d as shown below.

$$f(x;\, a,\, b,\, c,\, d) = \begin{Bmatrix} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \\ \frac{d-x}{d-c}, & c \le x \le d \\ 0, & d \le x \end{Bmatrix} \tag{1}$$

If the expression is compressed, it becomes;

$$f(x;\, a,\, b,\, c,\, d) = \max\left(\min\left(\frac{x-a}{b-a},\ 1,\ \frac{d-x}{d-c}\right),\ 0\right) \tag{2}$$

The Parameters a and d trace the "feet" the trapezoid while the parameters b and c trace the "shoulders".

5.2 Structure of the Pragmatic Miner Rules

Practically, residual and exposed risks determine the type of response the IDPS will trigger. For example, if the residual and exposed risks are very high, then the correct response is to apply safeguards in order to reduce the effects of the attack. However, if they are very low, the correct response to the trigger is first to understand the effects and acknowledge the risk without any attempts to control or mitigate them as illustrated in Table 3.

Table 3. Description risk analysis matrix

Types of risks		Exposed risk (ER)				
		Critical	Very high	High	Medium	Low
Residual risk (RR)	Critical	Avoidance	Avoidance	Avoidance	Transference	Transference
	Very high	Avoidance	Avoidance	Transference	Transference	Transference
	High	Avoidance	Transference	Transference	Transference	Transference
	Medium	Transference	Transference	Transference	Transference	Acceptance
	Low	Transference	Transference	Transference	Acceptance	Acceptance

Obviously, the matrix in Table 2 makes the fact that the upper and lower parts of the triangle are equivalent (meaning that they lead to the same result). Therefore, the cross between RR-VERY HIGH and ER-HIGH leads to the same transference as between ER-VERY.HIGH and RR-HIGH. There is a possibility of twenty-five combinations of inference rules in the fuzzy sets as shown in Table 2. Table 4 shows that a set of fifteen rules constructed based on the actual experimental qualitative analysis and the characteristics of the input and output variables (i.e., this result refute the hypothesis that appear in (Qassim and Mohd-Zin [7]).

5.3 Defuzzification

The term "defuzzification" implies the conversion of a fuzzy quantity into a precise value. It is the exact opposite process of fuzzification, which is the conversion of a precise value to a fuzzy quantity [4, 16]. The union of the output of each rule is used to develop the resultant membership functions.

Table 4. The description rules

1.	IF (RR is Critical) and (ER is Critical) Then (countermeasure is Avoidance)
2.	IF (RR- Critical) and (ER_ is Very High) OR (ER- is Critical) and (RR _ is Very High) Then (countermeasure is voidance)
3.	IF (RR is Critical) and (ER is High) OR(ER is Critical) and (RR is High) Then (countermeasure is Avoidance)
4.	IF (RR is Critical) and (ER is Medium) OR(ER is Critical) and (RR is Medium) Then (countermeasure is Transference)
5.	IF (RR is Critical) and (ER is Low) OR (ER is Critical) and (RR is Low) Then (countermeasure is Transference)
6.	IF (RR is Very High) and (ER is Very High)OR (ER is Very High) and (RR is Very High) Then (countermeasure is Avoidance)
7.	IF (RR is Very High) and (ER is High) OR (ER is Very High) and (RR is High) Then (countermeasure is Transference)
8.	IF (RR is Very High) and (ER is Medium) OR (ER is Very High) and (RR is Medium) Then (countermeasure is Transference)
9.	IF (RR is Very High) and (ER is Low) OR (RR is Very High) and (ER is Low) Then (countermeasure is Transference)
10.	IF (RR is High) and (ER is High) Then (countermeasure is Transference)
11.	IF (RR is High) and (ER is Medium) OR (ER is High) and (RR is Medium) Then (countermeasure is Transference)
12.	IF (RR is High) and (ER is Low) OR (ER is High) and (RR is Low) Then (countermeasure is Transference)
13.	IF (RR is Medium) and (ER is Medium) Then (countermeasure is Transference)
14.	IF (RR is Medium) and (ER is Low) OR (ER is Medium) and (RR is Low) Then (countermeasure is Acceptance)
15.	IF (RR is Low) and (E R is Low) OR (ER is Low) and (RR is Low) Then (countermeasure is Acceptance)

However, a close ascent of the countermeasure values clearly reveals and confirms that the pragmatic miner model can be used to predict countermeasure values under consideration.

Table 5 shows some of samples while Table 6 shows the correlation matrix that is calculated after the defuzzification carried out on the original database. Statistical analysis can be used to determine the active attributes of each class, while some attributes in linguistic terms can be expressed through the use of FL.

Each class of the acquired pattern classified can be inspected after classification has been made through PMRA. Statistical analysis is not only the best method of inspection on each class, but also can help one to generate the rules that control each class, attribute. This can be done through the use of the measures of central tendency and dispersion. Although there are many ways to measure the variability of data, in this work, we will use measures of dispersion, Standard deviation gives the average distance with which each element deviated from the mean. However, besides standard deviation, there are other important techniques prevails which are discussed below.

Range refers to the difference between the highest outcome in the data and the lowest outcome; it can be calculated by using the formula Xmax−Xmin. However, this range only uses two values from the entire set of data, making it unreliable. It cannot take into consideration the fact that extreme values can be very large and at the same time, many elements may be very close to each other. Looking at the range of this set of data 1, 1, 2, 4, 7, the range is 7−1 = 6, since 1 is the lowest while 7 is the highest outcome.

There are a few books indicate that the statistical range is the same as mathematical range. Therefore, the interval over which a data occurred is more important than a single number. From the above example, the range is from 1 to 7 or [1, 7]. Most of these measures will be used in this work.

Table 5. Sample of generation database

Residual risk						Exposed risk				Countermeasures
RR_Low	RR_Medium	RR_High	RR_Very High	RR_Critical	ER_Low	ER_Medium	ER_High	ER_Very High	ER_Critical	
3.518423248	28.97838246	44.70769534	82.51340374	99.51084659	6.491937775	11.49547943	44.86631873	87.0750865	96.75724227	1
5.99512786	25.09239558	58.91119317	76.77269645	98.93912277	2.940579222	24.30951448	51.02016885	73.14588889	97.7655921	2
7.847275971	36.99791336	58.27590313	82.44395231	95.28481309	3.423143681	14.22235216	56.42649048	88.63707334	97.60051147	3
6.542029219	39.83819901	47.78704478	80.22383222	93.82694231	7.774794347	28.97202866	51.60354076	75.4583978	92.04814068	1
5.861745268	19.68527951	69.54887876	84.23679779	96.53790496	0.638851319	35.24259225	67.1518145	73.75239924	93.4298592	2
9.867828263	20.02176915	58.01326439	74.17035904	94.39298796	7.544338868	33.46747283	54.07889779	88.95311385	92.96186557	3
8.547584561	21.43947004	54.77429516	81.71850608	95.30333806	6.37563549	27.17654586	53.17090207	74.36135939	97.42440043	1
2.752772132	31.05953781	46.05527981	78.49724646	98.56625556	9.969384472	12.98368693	67.63718875	76.14097827	97.70969963	2
8.592575547	23.38343425	55.25247732	78.0142526	97.34718736	3.271051575	21.30597164	62.80206931	87.14585915	98.92727356	3
0.605905878	12.93375694	57.94187465	85.94892174	93.8641164	4.9916625	32.75716133	55.61123628	72.14424367	96.15211615	1
7.7861302	32.35252307	62.0362408	83.88373015	96.88168811	0.218471278	19.52531705	69.96555856	85.30930202	98.62832459	2
5.036512958	13.05588823	64.94809088	87.64271193	99.18226391	6.539912209	33.92214584	63.46221583	80.82448607	99.47560229	3

Table 6. Correlation matrix after the Defuzzification stage

Variables	RR_Low	RR_Medium	RR_High	RR_Very High	RR_Critical	ER_Low	ER_Medium	ER_High	ER_Very High	ER_Critical
RR_Low	1.000	0.040	0.067	0.205	0.006	0.208	−0.025	−0.038	−0.049	−0.003
RR_Medium	0.040	**1.000**	−0.068	0.094	0.025	−0.088	−0.102	−0.094	−0.023	0.084
RR_High	0.067	−0.068	**1.000**	−0.082	0.114	0.172	0.027	0.048	0.016	0.073
RR_Very High	0.205	0.094	−0.082	**1.000**	0.054	0.092	0.006	−0.064	−0.153	0.094
RR_Critical	0.006	0.025	0.114	0.054	**1.000**	−0.055	0.054	0.194	0.067	0.098
ER_Low	0.208	−0.088	0.172	0.092	−0.055	**1.000**	−0.028	−0.019	0.075	−0.201
ER_Medium	−0.025	−0.102	0.027	0.006	0.054	−0.028	**1.000**	0.152	−0.124	−0.117
ER_High	−0.038	−0.094	0.048	−0.064	0.194	−0.019	0.152	**1.000**	0.007	−0.029
ER_Very High	−0.049	−0.023	0.016	−0.153	0.067	0.075	−0.124	0.007	**1.000**	−0.070
ER_Critical	−0.003	0.084	0.073	0.094	0.098	−0.201	−0.117	−0.029	−0.070	**1.000**

The real rules of a database, which contains 100 samples and ten inputs (i.e., five features related to RR and five features related to ER), are selected. Because the accuracy is one of the main goals of this work as explained in Table 5, therefore the only rules that satisfy 100% of accuracy are selected as shown in Table 7. The confusion matrix of the estimating samples is shown in Table 7 and the final decision surface of the PMRA System is shown in Fig. 2.

5.4 Pragmatic Miner Accuracy and Error

After the formation of classify association rules, five experimental tests were done from separate experiments and the proposed mining model is used to recognize the abnormality of the system at the same conditions as shown in Table 2, so that the investigation of mining accuracy and error can be accomplished. The error is computed in order to measure the gap between the predicted and the measured values. The individual error percentage can be determined by dividing the absolute difference between the predicted and the measured values is given by:

$$e_i = \left(\frac{|A_m - A_p|}{A_m} \right) * 100\% \tag{3}$$

where, e_i is the individual error, Am is the measured value and Ap is the predicted value.

The accuracy, however, measures the closeness of the predicted value to the measured value. The average of the individual accuracies is the model accuracy as shown in Eq. 3.

$$a = \frac{1}{N} \sum_{i=1}^{N} \left(1 - \frac{|A_m - A_p|}{A_m} \right) * 100\% \tag{4}$$

where a is the model accuracy and N are the total number of the tested data sets.

Table 7. The actual rules for PMRA

Purity	Rules
34.00%	
37.08%	If ER_High in [41.3, 67.624[then Countermeasures = 1 in 37.1% of cases
81.82%	If ER_High in [67.624, 69.901[then Countermeasures = 2 in 81.8% of cases
37.18%	If ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 37.2% of cases
72.73%	If ER_Very High in [87.212, 89.855[and ER_High in [41.3, 67.624[then Countermeasures = 1 in 72.7% of cases
35.94%	If RR_Critical in [91.123, 98.704[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 2 in 35.9% of cases
64.29%	If RR_Critical in [98.704, 99.915[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 64.3% of cases
35.00%	If ER_Very High in [71.414, 86.173[and RR_Critical in [91.123, 98.704[and ER_High in [41.3, 67.624[then Countermeasures = 1 in 35% of cases
100.00%	If ER_Very High in [86.173, 87.212[and RR_Critical in [91.123, 98.704[and ER_High in [41.3, 67.624[then Countermeasures = 2 in 100% of cases
40.00%	If ER_Low in [0.227, 7.714[and ER_Very High in [71.414, 86.173[and RR_Critical in [91.123, 98.704[and ER_High in [41.3, 67.624[then Countermeasures = 2 in 40% of cases
60.00%	If ER_Low in [7.714, 9.975[and ER_Very High in [71.414, 86.173[and RR_Critical in [91.123, 98.704[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 60% of cases
66.67%	If RR_Low in [0.452, 8.611[and RR_Critical in [98.704, 99.915[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 66.7% of cases
50.00%	If RR_Low in [8.611, 9.186[and RR_Critical in [98.704, 99.915[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 2 in 50% of cases
72.73%	If RR_Low in [0.452, 7.955[and RR_Critical in [98.704, 99.915[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 72.7% of cases
100.00%	If RR_Low in [7.955, 8.611[and RR_Critical in [98.704, 99.915[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 1 in 100% of cases
100.00%	If RR_High in [47.424, 58.355[and RR_Low in [8.611, 9.186[and RR_Critical in [98.704, 99.915[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 100% of cases
100.00%	If RR_High in [58.355, 69.285[and RR_Low in [8.611, 9.186[and RR_Critical in [98.704, 99.915[and ER_Very High in [71.414, 87.212[and ER_High in [41.3, 67.624[then Countermeasures = 2 in 100% of cases
88.89%	If RR_Low in [0.033, 6.778[and ER_Very High in [87.212, 89.855[and ER_High in [41.3, 67.624[then Countermeasures = 1 in 88.9% of cases
100.00%	If RR_Low in [6.778, 9.762[and ER_Very High in [87.212, 89.855[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 100% of cases
100.00%	If ER_Low in [1.448, 7.651[and RR_Low in [0.033, 6.778[and ER_Very High in [87.212, 89.855[and ER_High in [41.3, 67.624[then Countermeasures = 1 in 100% of cases
66.67%	If ER_Low in [7.651, 9.179[and RR_Low in [0.033, 6.778[and ER_Very High in [87.212, 89.855[and ER_High in [41.3, 67.624[then Countermeasures = 1 in 66.7% of cases

100.00%	If RR_High in [52.354, 58.22[and ER_Low in [7.651, 9.179[and RR_Low in [0.033, 6.778[and ER_Very High in [87.212, 89.855[and ER_High in [41.3, 67.624[then Countermeasures = 1 in 100% of cases
100.00%	If RR_High in [58.22, 60.819[and ER_Low in [7.651, 9.179[and RR_Low in [0.033, 6.778[and ER_Very High in [87.212, 89.855[and ER_High in [41.3, 67.624[then Countermeasures = 3 in 100% of cases
88.89%	If ER_Critical in [91.718, 98.217[and ER_High in [67.624, 69.901[then Countermeasures = 2 in 88.9% of cases
50.00%	If ER_Critical in [98.217, 98.653[and ER_High in [67.624, 69.901[then Countermeasures = 1 in 50% of cases
100.00%	If RR_Low in [0.861, 9.198[and ER_Critical in [91.718, 98.217[and ER_High in [67.624, 69.901[then Countermeasures = 2 in 100% of cases
50.00%	If RR_Low in [9.198, 9.581[and ER_Critical in [91.718, 98.217[and ER_High in [67.624, 69.901[then Countermeasures = 2 in 50% of cases
100.00%	If RR_Very High in [76.721, 79.65[and RR_Low in [9.198, 9.581[and ER_Critical in [91.718, 98.217[and ER_High in [67.624, 69.901[then Countermeasures = 2 in 100% of cases
100.00%	If RR_Very High in [79.65, 82.58[and RR_Low in [9.198, 9.581[and ER_Critical in [91.718, 98.217[and ER_High in [67.624, 69.901[then Countermeasures = 3 in 100% of cases
100.00%	If RR_Low in [3.089, 6.221[and ER_Critical in [98.217, 98.653[and ER_High in [67.624, 69.901[then Countermeasures = 1 in 100% of cases
100.00%	If RR_Low in [6.221, 9.353[and ER_Critical in [98.217, 98.653[and ER_High in [67.624, 69.901[then Countermeasures = 2 in 100% of cases

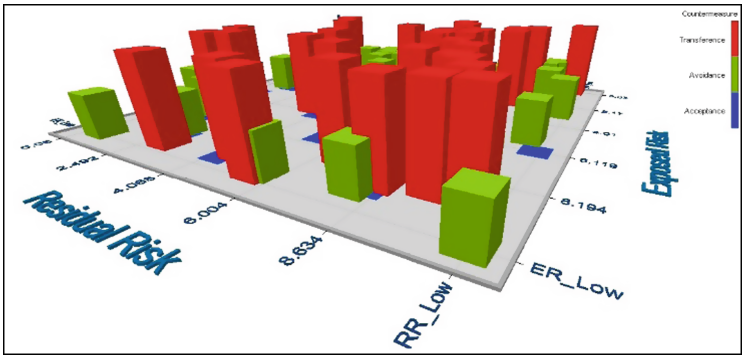

Fig. 2. The final decision surface of PMRA system

The model accuracy for PMRA_ID was determined after calculating the error of the data set. Table 8 shows the experimental condition, countermeasure results, and the miner model predicted values.

As we can observe from Table 9 that the highest percentage of error in the PMRA model prediction is 0.32%. This indicates and confirms that the PMRA prediction countermeasure results are low and very close to the real experimental where is this number countermeasures values. It also shows that the average accuracy of proposed PMRA model is 90.11%. The value of the accuracy shows that the proposed model can predict the vulnerability of a system as it can be observed from the graph trend lines.

Table 8. Confusion matrix for all the samples

From\To	*Avoidance*	*Transference*	*Acceptance*	Total	% Correct
Avoidance	10	16	8	34	29.41%
Transference	0	32	1	33	96.97%
Acceptance	0	11	22	33	66.67%
Total	10	59	31	100	

Table 9. The accuracy and error of the pragmatic miner model

Risk parameters (INPUTS)		Countermeasure parameter (OUTPUT)			STATISTICS				
Residual risk	Expose risk	1st epoch	2nd epoch	3rd epoch	Average	Standard deviation (σ)	Measured	Error %	Proposed pragmatic miner model
90.00	70.00	90.00	100.0	70.0	86.67	15.28	76.00	0.18	76.80
10.00	40.00	80.00	90.00	70.0	80.00	10.00	68.00	0.18	78.30
70.00	50.00	67.00	70.00	100.0	79.00	18.25	98.00	0.32	99.07
98.00	98.00	90.00	100.0	69.00	86.33	15.82	95.40	0.06	96.80
67.00	43.00	80.00	79.00	98.00	85.67	10.69	99.00	0.19	99.60
							Average accuracy of miner model = 90.11%		

6 Discussion and Conclusion

The combination of risk analysis mechanism with a developed PMRA for online IDS through the modification of an FL Controller with mining algorithm detects Distributed Denial of Service (DDoS) attack with 90.11% accuracy and that is superior to FL Controller IDS and D-SCIDS by themselves. The main parameters used to compute MF is a = 0.02, b = 0.05, c = 0.08, d = 1.2.

The calculation of Discretization, feature selection and accuracy are simultaneously handled in this work. This reduced the cost of computation and built the detection in a detailed manner. Observation has shown that the detection of continuous attack attribute by FLC when the same parameters are applied to all the attributes causes the classified association rules accuracy to vary widely. Conversely, the best result of classification accuracy is obtained when FLC is combined with the object Risk Analysis for different attributes in a different class.

Because of the increased level of computer information attacks, the necessity to provide an effective intrusion detection and prevention methods had increased. IDPS suffered from a several weaknesses including post event detection, overwhelming false alarms and a centralized analysis of intrusion. This paper introduced a centralized and automatic system called as PMRA, as a reliable substitute for conventional IDPS. The experimental result showed PMRA was more effective and consistent than the other IDPSs.

References

1. Anuar, N.B., Papadaki, M., Furnell, S., Clarke, N.: Incident prioritisation using analytic hierarchy process (AHP): Risk Index Model (RIM). Secur. Commun. Netw. **6**, 1087–1116 (2013). https://doi.org/10.1002/sec.673
2. Bajpai, S., Sachdeva, A., Gupta, J.P.: Security risk assessment: applying the concepts of fuzzy logic. J. Hazard. Mater. **173**, 258–264 (2010). https://doi.org/10.1016/j.jhazmat.2009.08.078
3. Catania, C.A., Garino, C.G.: Automatic network intrusion detection: current techniques and open issues. Comput. Electr. Eng. **38**(5), 1062–1072 (2012). https://doi.org/10.1016/j.compeleceng.2012.05.013
4. Chen, P.Y., Kataria, G., Krishnan, R.: Correlated failures, diversification and information security risk management. MIS Q. **35**, 397–422 (2011)
5. Liao, H.J., Lin, C.H.R., Lin, Y.C., Tung, K.Y.: Intrusion detection system: a comprehensive review. J. Netw. Comput. Appl. **36**(1), 16–24 (2013). https://doi.org/10.1016/j.jnca.2012.09.004
6. Mansour, N., Chehab, M., Faour, A.: Filtering intrusion detection alarms. Clust. Comput. **13**, 19–29 (2010)
7. Qassim, Q., Mohd-Zin, A.: Strategy to reduce false alarms in intrusion detection and prevention systems. Int. Arab J. Inf. Technol. (IAJIT) **11**(5) (2014)
8. Spathoulas, G.P., Katsikas, S.K.: Reducing false positives in intrusion detection systems. Comput. Secur. **29**, 35–44 (2010). https://doi.org/10.1016/j.cose.2009.07.008
9. Tjhai, G.C., Furnell, S.M., Papadaki, M., Clarke, N.L.: A preliminary two-stage alarm correlation and filtering system using SOME neural network and K-means algorithm. Comput. Secur. **29**, 712–723 (2010). https://doi.org/10.1016/j.cose.2010.02.001
10. Whitman, M.E., Mattord, H.J.: Principles of Information Security. Cengage Learning, Boston (2011)
11. Zeng, J., Li, T., Li, G., Li, H.: A new intrusion detection method based on antibody concentration. In: Huang, D.-S., Jo, K.-H., Lee, H.-H., Kang, H.-J., Bevilacqua, V. (eds.) ICIC 2009. LNCS, vol. 5755, pp. 500–509. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04020-7_53
12. Zhou, Y.P., Fang, J.A.: Intrusion detection model based on hierarchical fuzzy inference system. In: The 2th International Conference on Information and Computing Science, ICIC 2009, vol. 2, pp. 144–147. IEEE (2009). http://dx.doi.org/10.1109/ICIC.2009.145
13. Al-Janabi, S., Al-Shourbaji, I., Shojafar, M., Shamshirband, S.: Survey of main challenges (security and privacy) in wireless body area networks for healthcare applications. Egypt. Inform. J. **18**, 113–122 (2017)
14. Al-Janabi, S., Al-Shourbaji, I.: A study of cyber security awareness in educational environment in the Middle East. J. Inf. Knowl. Manag. **15**, 1650007 (2016)
15. Ahamad, S.S., Al-Shourbaji, I., Al-Janabi, S.: A secure NFC mobile payment protocol based on biometrics with formal verification. Int. J. Internet Technol. Secur. Trans. **6**, 103–132 (2016)
16. Folorunso, O., Ayo, F.E., Babalola, Y.E.: Ca-NIDS: a network intrusion detection system using combinatorial algorithm approach. J. Inf. Priv. Secur. **12**, 181–196 (2016)