A Combined Approach to DoS Attack Detection System

Archana Salaskar
Computer Engineering Department, Savitribai Phule
Pune University,
Dhole Patil College of Engineering,
Wagholi, Pune

R.N. Phursule
Computer Engineering Department,
Savitribai Phule
Pune University
JSPM's Imperial College Engineering & Research,
Wagholi, Pune

ABSTRACT

In the network system attack like Denial of service (DoS) is forthcoming damaging attack. The performance of online servers degrades within seconds. Intensive computation on the target server is imposed due to this attack. Target Server gets flooded with large useless packets. The fatality server can be forced out of service From a few minutes to even several days. Eventually crucial business services running on the target fatality causes work down on the target fatality. So for the researchers it is very challenging task. The development of network-based detection mechanisms is the focus of the solution of this kind of attack. Based on these mechanisms in the existing detection systems, traffic transmitted over the protected networks get monitored. Mainly two methods are introduced for detection mechanism namely Misuse based and Anomaly based detection systems. But to enhance the detection rate they are not sufficient. In the proposed system the features which are directly associated with DoS attacks are extract by monitoring the network traffic. To generate geometrical triangular area measurements for normal profiles on the basis of these features the multivariate correlation analysis (MCA) model is used. To detect any unknown DoS attack in the network, these models are used as references. And furthermore to detect attack anomaly detection method is used. Only MCA and anomaly based system is not sufficient for accurate attack detection. So the inventive work behavioral based rule model integrated with MCA and anomaly, as a hybrid model used to enhance the accuracy of DoS attack detection. In proposed inventive model behavioral rules are generated for suspected packets and ultimately detection accuracy as well as detection rate get increased.

Keywords

Denial of Service Attack (DoS), Multivariate Correlation, Triangle area, network traffic characterization. Behavioral Rule.

1. INTRODUCTION

The tremendous growth of computer networks, particularly of the Internet has created security problems. Nowadays is Denial of Service attacks is one of the greatest threats that network security faces. To defence against Denial of Service attacks is becoming an important challenge and difficult task. A lot many Intrusion detection systems have been developed in order to defend computer networks against the continuous evolution of various types of threats, including DoS attacks. DoS are one type of insistent and threatening intrusive behaviour to online servers. DoS attacks severely degrade the availability of a sufferer. Sufferer can be a host, a router, or an entire network. They enforce rigorous computation tasks to the sufferer by exploiting its system susceptibility or flooding it with huge amount of useless packets. The fatality can be forced out of service from a few minutes to even several days. Serious damages causes to the services running on the fatality. Therefore, effectual detection of DoS attacks is essential to protect online services. Work on DoS attack detection mainly focuses on the development of network-based detection mechanisms. Monitoring of traffic transmitting over the protected networks is done detection systems based on these mechanisms. These mechanisms release the protected online servers from monitoring attacks and ensure that the servers can dedicate themselves to provide quality services with minimum delay in response. In addition, network-based detection systems are loosely coupled with operating systems running on the host machines which they are protecting. As a result, the configurations of network-based detection systems [9]are less complicated than that of host-based detection systems. Intrusion detection techniques generally divided into two types anomaly detection and misuse detection. A previous knowledge on intrusions is necessary to detect Misuse detection. And it tries to detect attacks based on specific patterns or signatures of known attacks. Misuse detection systems are very accurate in detecting known attacks. But their basic drawback is that network attacks are under a continuously changing and this needs for an up-to date knowledge base of all attacks. In reverse unknown attacks can be detected by using anomaly detection. What is normal is defined by anomaly detection techniques and any deviations from the normal behaviors that are considered to be intrusions. The DoS attack detection system presented in existing system employs the principles of MCA and anomaly based detection. They provide known and unknown attacks with capabilities of accurate characterization for traffic behaviors and detection. To enhance and to speed up the process of MCA a triangle area technique is used. To eliminate the bias from the raw data a statistical normalization technique will used. Existing Multivariate Correlation Analysis approach has been approved. And it is used for Daniel of Service attack detection. But only this Multivariate Correlation analysis is not sufficient to achieve maximum accuracy. Existing system also has high complexity. To deal with above issues proposed system architecture enforces behavioral based attack detection combined with existing multivariate correlation analysis based and anomaly based attack detection system. In the Behavioral based system approach[8], system extracts all the behaviours of malicious samples. These behaviours are saved and used after multivariate correlation analysis to increase the accuracy of the system.

2. RELATED WORK

Various techniques have evolved to detect different kinds of DoS attacks successfully. These techniques have shown some constraint such as they are applicable for certain network traffic. Zhiyuan Tan, Aruna Jamdagni, Xiangjian He, Priyadarsi Nanda1, and Ren Ping Liu proposed technique which runs analysis on original feature space (first-order statistics) and extracts the multivariate correlations between the first-order statistics [2]. The extracted multivariate correlations, namely second-order statistics, preserve significant discriminative information for accurate characterizations of network traffic records, and these multivariate correlations can be the high-quality potential features for DoS attack detection. The effectiveness of the proposed technique is evaluated using KDD CUP 99 dataset and experimental analysis shows encouraging results.

Shuyuan Jin, Daniel S. Yeung proposed technique which discusses the effects of multivariate correlation analysis on the DDoS detection and proposes a covariance analysis model for detecting SYN flooding attacks [3]. The simulation results show that this method is highly accurate in detecting malicious network traffic in DDoS attacks of different intensities. This method can effectively differentiate between normal and attack traffic. Indeed, this method can detect even very subtle attacks only slightly different from normal behaviors. The linear complexity of the method makes its real time detection practical. Mihui Kim, Hyunjung Na, Kijoon Chae, Hyochan Bang, and Jungchan Na proposed a technique which presents a combined data mining approach for modeling the traffic pattern of normal and diverse attacks [4]. This approach uses the automatic feature selection mechanism for selecting the important attributes. And the classifier is built with the theoretically selected attribute through the neural network. And then, our experimental results show that our approach can provide the best performance on the real network, in comparison with that by heuristic feature selection and any other single data mining approaches. Aikaterini Mitrokotsa, Christos Douligeris proposed a technique which presents an approach that detects Denial of Service attacks using Emergent Self-Organizing Maps [5]. The approach is based on classifying “normal” traffic against “abnormal” traffic in the sense of Denial of Service attacks. The approach permits the automatic classification of events that are contained in logs and visualization of network traffic. Extensive simulations show the effectiveness of this approach compared to previously proposed approaches regarding false alarms and detection probabilities.

Masayoshi Mizutani [9] and others, the behavior rule based intrusion detection model is introduced. This behavioural rule based model uses correlations of packet/payload data patterns and communication patterns. Scenario-based intrusion detection method has similar features based on state transition machine, However scenarios of compromise consist of not only sequential events but also random order events and certain scenarios have to be described complicated correlations between communications.

3. PROPOSED WORK

The proposed work consist of implementation of innovative intrusion detection system for denial of service attack detection with maximum accuracy using multivariate correlation analysis integrated with behavioural based systems. This approach monitors the network traffic and extracts the features. These features are mapped with the reference normal features taken from KDD Cup 99[10]. The extracted normal features are further applied to triangular area map generation where geometrical areas of all the normal features are computed. Along with the triangular map areas, another model known as behavioural based model extracts the behavioural patterns of normal profiles and stores them into database. This phase is training phase, which mainly consists of computation and storage of geometrical TAM (Triangular Area Map) as well as normal behavioural patterns. These entities are further used to classify any unknown profile with high accuracy in testing phase.

3.1 Key Notations

Following Table is used as key notations used in algorism.

Table 1: Key Notation

NOTATIONS	DEFINITIONS
X	A arbitrary network traffic dataset
Xob	Observed traffic record
XnTl	The set of the lower triangles of the TAMs of legitimate traffic records in the training dataset Xn.
TAM ob,i,l	The lower triangle of the TAM of the Observed traffic record
TAMn,i,l	The lower triangle of the TAM of the ith legitimate traffic record in the training dataset Xn
aTAMn	The expectation of the lower triangles of the TAMs of the legitimate traffic records in the training dataset Xn
EDn,i	The Euclidian distance between the ith training legitimate traffic record TAMn,i,l and the mean aTAMn,l of the overall legitimate training records in the training dataset Xn
μ	EDn,i between training legitimate traffic records TAMn,i,l and the expectation aTAMn,l of the legitimate records.
σ	The standard deviation of the Euclidian distances EDn,i.
TAMob,l	The TAM of the observed test traffic record
EDob	The Euclidian distance between observed test traffic record TAMob,l and the mean aTAMn,l of the legitimate training traffic records.
Threshold	A Pre-defined threshold for distinguishing attack from legitimate traffic.
α	A parameter ranged from 1 to 3.
NPro	Normal Profile

3.2 Proposed Algorithms

In the Proposed system Normal Profiles are generated for legitimate traffic. In Existing system Mahalabonis Distance and covariance matrix is used but due to high time complexity in proposed system Euclidian distance[5] is used in Multivariate correlation analysis. There are basically main three algorithms are used in proposed system, namely Normal Profile generation, Attack Detection and Behavioral Rules generation algorithms.

a) **Algorithm For Normal Profile Generation:**
This algorithm is used to generate Normal Profiles for the legitimate traffic.

```
1. Input: XnTl
2. Output: NPro
3. For i from 1 to g
       Construct EDn,i with TAMn,i,l & aTAMn,l
4. End for
```

```
5.  Construct μ with EDn,i
6.  Construct σ with EDn,i.
7.  Generate NPro with μ, σ and aTAMn,l
8.  Output Npro
```

b) Algorithm For Attack detection

```
1. Input: Xob; Npro; aTAMn,l ; α
2. Output: Normal or Attack
3. Generate TAM ob,i,l
       Construct EDob,i with TAMob,i,l & aTAMn,l
4. If ( μ- σ * α ) <= EDob,i <= ( μ+ σ * α ) then
5. Output : Return Normal
6. Else
7. Output : Return Attack
```

c) Algorithm for Behavioral Rules

```
1. Compute Normal Profile
2. If above detection algorithm gives output as a
   Attack then
3. Attack
4. Else do following
5. Compute behavioral rules for Normal Profile
     i.   Define first priority attribute,
     ii.  Define second priority attribute,
     iii. Define third priority attribute,
     iv.  Get values for first priority attribute from
          Normal Profile,
     v.   Get values for second priority attribute
          from Normal Profile,
     vi.  Get values for third priority attribute from
          Normal Profile,
6. Compute behavioral rules for Observed traffic
7. Map the step 5 & 6.
```

4. SYSTEM IMPLEMENTATION

The entire system is based upon comparative analysis between existing and proposed system. The proposed system is evaluated using standard dataset called as KDD Cup 99[10].

Fig 2 illustrate the architecture of the proposed system. In the Multivariate correlation analysis Feature Normalization is done on provided data between -3 to 3 range by using statistical method as,

$$xi = \frac{vi - \mu}{\sigma}$$

Where,

xi is Normalized attribute

vi is actual value of attribute,

μ is mean of n values of given attributes,

σ is its standard deviation.

Feature normalization is done to remove bias from original data. For this statistical method is used here.

This MCA approach uses triangular area for extracting correlation between the features within an observed traffic record.

If $X = \{x1, x2, x3 \ldots\ldots, xn\}$ where $xi = [f1\, f2 \ldots\ldots fn]^T$, (1<=i<=n) represents the ith n dimensional traffic record. Here concept of triangular area is applied to extract the geometrical correlation between jth and kth features in the vector xi. Vector xi is first projected on (j,k)th 2D Euclidian. And on the Cartesian coordinate system triangle formed by the original and projected points, area Trjk for ith is defined as

$$Trjk = (\| (fj, 0) - (0,0) \| \times \| (0, fk) - (0,0) \|)/2$$

In the decision making training and testing is there. First in the training phase normal profile generation of legitimate traffic is done. Furthermore Triangular Area Map TAM is generated for individual record. A TAM is constructed and all Triangular area is mapped with respect to their indexes. Eventually getting the TAMi which is symmetric matrix having elements of zero on the main diagonal. So for further calculation choice is taken as lower triangle. The lower triangle of TAMi is converted into a new correlation vector TAMi,l as

$$TAMi, l = [Tr2{,}1\; Tr3{,}1 \ldots\ldots Tr4{,}2 \ldots\ldots . Trm, m - 1]^T$$

If g is set of legitimate training traffic records Xn is

$$Xn\ of\ TAMl = \{TAMn, l, 1; TAMn, l, 2; \ldots\ldots TAMn, l, g\}$$

Here Euclidian distance EDn is used to measure the dissimilarity between the traffic records.

The mean of Edn is considered as a μ. And standard deviation σ . We get normal distribution

$$Normal\ Distribution = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2} \backslash 2\sigma^2$$

With the help of N(μ, σ^2) of normal training traffic records and TAMn,l are stored in the normal profile for attack detection.

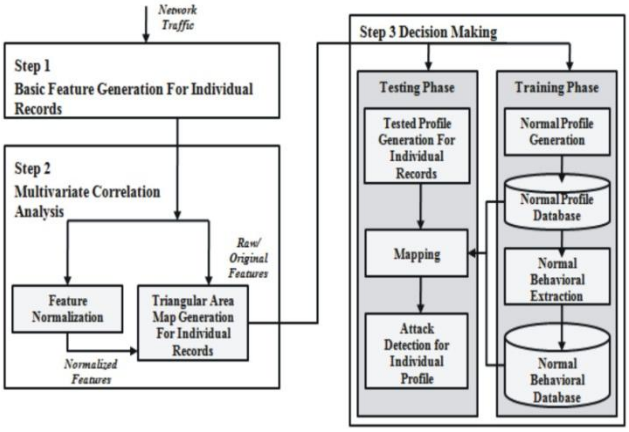

Network Traffic
Step 1
Basic Feature Generation For Individual Records
Step 2
Multivariate Correlation Analysis
Raw/ Original Features
Feature Normalization
Triangular Area Map Generation For Individual Records
Normalized Features
Step 3 Decision Making
Testing Phase
Training Phase
Tested Profile Generation For Individual Records
Mapping
Attack Detection for Individual Profile
Normal Profile Generation
Normal Profile Database
Normal Behavioral Extraction
Normal Behavioral Database

Fig 2 System Architecture

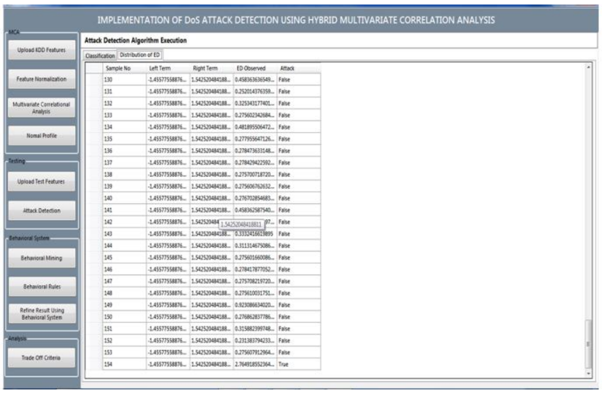
In the test phase, tested profiles are generated for observed traffic records. To detect attack for observed traffic TAMob,l is generated as discussed above. And Edob is also calculated using TAMob,l and aTAMn,l. Tested profiles are handed over the attack detection phase which compares the individual tested profiles with the respective stored normal profiles. A threshold based classified is employed here in attack detection module.

$$Threshold = \mu + \sigma * \alpha$$

Where α is normal distribution ranged from 1 to 3. For more accuracy and to enhance the attack detection rate innovative behavioral rule based module is used. From normal profiles by providing priorities behavioral rules are generated. If some attack records are left by attack detection mechanism. With the help of these rules and mapping results high accuracy and as well as high detection rate in proposed system.

5. RESULT AND DATASET

Proposed system is evaluated using KDD cup 99. KDD'99[10] has been the most wildly used data set for the evaluation of anomaly detection methods. This data set is built based on the data captured in DARPA'98 IDS evaluation program. KDD training dataset consists of approximately 4,900,000 single connection vectors each of which contains 41 features and is labeled as either normal or an attack, with exactly one specific attack type. Applying feature normalization[14], normal profile are generated for the legitimate traffic. Test records are applied to the testing module to get the attack result. Fig 3 shows results for attack detection.

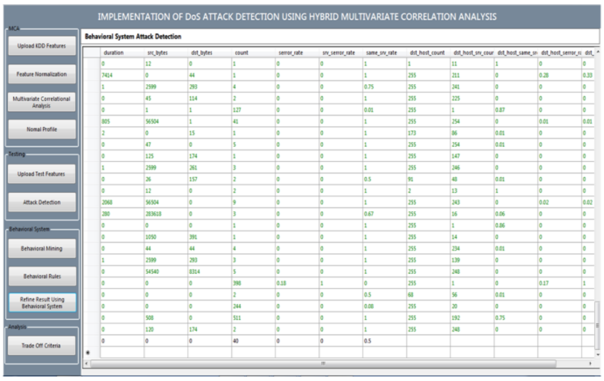


Attack Detection Algorithm Execution

Sample No	Left Term	Right Term	ED Observed	Attack
130	-1.45577558876...	1.542520484188...	0.458363636549...	False
131	-1.45577558876...	1.542520484188...	0.252014376359...	False
132	-1.45577558876...	1.542520484188...	0.325343177401...	False
133	-1.45577558876...	1.542520484188...	0.275602342684...	False
134	-1.45577558876...	1.542520484188...	0.481895506472...	False
135	-1.45577558876...	1.542520484188...	0.277955647126...	False
136	-1.45577558876...	1.542520484188...	0.278473633148...	False
137	-1.45577558876...	1.542520484188...	0.278429422592...	False
138	-1.45577558876...	1.542520484188...	0.275700718720...	False
139	-1.45577558876...	1.542520484188...	0.275606762632...	False
140	-1.45577558876...	1.542520484188...	0.276702854683...	False
141	-1.45577558876...	1.542520484188...	0.458362587540...	False
142	-1.45577558876...	1.542520484 1.54252048418811	97...	False
143	-1.45577558876...	1.542520484188...	0.3332416619895	False
144	-1.45577558876...	1.542520484188...	0.311314675086...	False
145	-1.45577558876...	1.542520484188...	0.275601660086...	False
146	-1.45577558876...	1.542520484188...	0.278417877052...	False
147	-1.45577558876...	1.542520484188...	0.275708219720...	False
148	-1.45577558876...	1.542520484188...	0.275610031751...	False
149	-1.45577558876...	1.542520484188...	0.923086634020...	False
150	-1.45577558876...	1.542520484188...	0.276862837786...	False
151	-1.45577558876...	1.542520484188...	0.315882399748...	False
152	-1.45577558876...	1.542520484188...	0.231383794233...	False
153	-1.45577558876...	1.542520484188...	0.275607912964...	False
154	-1.45577558876...	1.542520484188...	2.764918552364...	True

Fig.3 Applying testing data to Attack detection algorithm results with classification

Figure 4 shows results for proposed system .

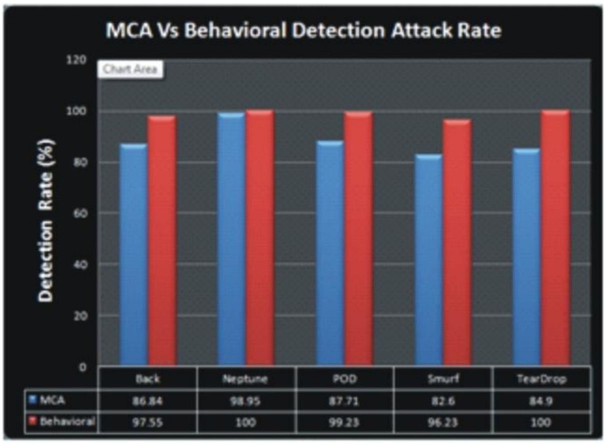


Behavioral System Attack Detection

duration	src_bytes	dst_bytes	count	serror_rate	srv_serror_rate	same_srv_rate	dst_host_count	dst_host_srv_cour	dst_host_same_sr	dst_host_serror_r	dst_
0	12	0	1	0	0	1	1	11	1	0	0
7414	0	44	1	0	0	1	255	211	0	0.28	0.33
1	2599	293	4	0	0	0.75	255	241	0	0	0
0	45	114	2	0	0	1	255	225	0	0	0
0	1	1	127	0	0	0.01	255	1	0.87	0	0
805	56504	1	41	0	0	1	255	254	0	0.01	0.01
2	0	15	1	0	0	1	173	86	0.01	0	0
0	47	0	5	0	0	1	255	254	0.01	0	0
0	125	174	1	0	0	1	255	147	0	0	0
1	2599	261	3	0	0	1	255	246	0	0	0
0	26	157	2	0	0	0.5	91	48	0.01	0	0
0	12	0	2	0	0	1	2	13	1	0	0
2068	56504	0	9	0	0	1	255	243	0	0.02	0.02
280	283618	0	3	0	0	0.67	255	16	0.06	0	0
0	0	0	1	0	0	1	255	1	0.86	0	0
0	1050	391	1	0	0	1	255	14	0	0	0
0	44	44	4	0	0	1	255	234	0.01	0	0
1	2599	293	3	0	0	1	255	139	0	0	0
0	54540	8314	5	0	0	1	255	248	0	0	0
0	0	0	398	0.18	1	0	255	1	0	0.17	1
0	0	0	2	0	0	0.5	68	56	0.01	0	0
0	0	0	244	0	0	0.08	255	20	0	0	0
0	508	0	511	0	0	1	255	192	0.75	0	0
0	120	174	2	0	0	1	255	248	0	0	0
0	0	0	40	0	0	0.5					

Fig 4 Applying testing data to innovated algorithm results with behavioural rules

Applying behavioral rules and after mapping results are much accurate. And hence MCA and anomaly based with behavioral rule based system gives better results than only MCA based and anomaly based. So with this proposed hybrid model more accuracy in detection as well as increase in the attack detection rate of the system is observed.



	Back	Neptune	POD	Smurf	TearDrop
MCA	86.84	98.95	87.71	82.6	84.9
Behavioral	97.55	100	99.23	96.23	100

Fig 4 Result analysis for Existing system and Proposed system

From the results analysis graph Fig 4 shows that the proposed system has better accuracy and high detection rate compare with existing system.

6. ACKNOWLEDGMENTS

This is a great pleasure & immense satisfaction to express my deepest sense of gratitude & thanks to everyone who has directly or indirectly helped me in completing my work successfully. I express my gratitude towards project guide Prof. Rajesh Phursule and Head, Department of Computer Engineering, Dhole Patil College Of Engineering, Wagholi, Pune who guided work in scheduled time.

I would like to thanks our Principal Dr. S. B. Allampallewar, for allowing us to pursue my project in this institute.

I also thanks to ME Project Head,and our PG Coordinator for their guidance and for being a constant source of support.

7. CONCLUSIONS

Proposed system is combined approach which is combination of MCA based and anomaly based and Behavioral rule based. To reduce high time complexity of existing system in the MCA based attack detection mechanism instead of Mahalabonis distance method used in existing system alternative Euclidian Distance method is used in Proposed System.

To improve accuracy and to get high detection rate compare to existing system, in the Proposed system behavioral Rule Based model is innovated. This hybrid DoS attack detection system is motorized by the triangle area based MCA technique and the anomaly based detection technique.

The previous technique extracts the geometrical correlations hidden in individual pairs of two distinct features within each network traffic record, and offers more accurate characterization for network traffic behaviors. Evaluation has been conducted using KDD Cup 99 data set to verify the effectiveness and performance of the proposed DoS attack detection system.

The results have revealed that when working with normalized data, and Modified MCA our detection system achieves approximately 80 percent detection accuracy. But with innovative model our proposed system achieves approximately100 percent accuracy. Ultimately getting high detection rate compared with exiting system. Normalization of the data can be done using statistical normalization technique to eliminate the bias from the data.

The proposed approach is applicable only for offline dataset. It is tested for online dataset. In the future scope, as users/attacker may attack online, same approach can be applicable for that with some more modification. We can use more classy classification techniques to improve the detection rate.

8. REFERENCES

[1] Zhiyuan Tan, Aruna Jamdagni, Xiangjian He,Priyadarsi Nanda, Ren Ping Liu , ‘A System for Denial-of-Service Attack Detection Based on Multivariate Correlation Analysis’, IEEE Transactions On Parallel And Distributed Systems, Vol. 25, No. 2, February 2014

[2] Shuyuan Jin,Daniel S. Yeung, ‘A Covariance Analysis Model for DDoS Attack Detection’. IEEE Communications Society 0-7803-8533-0/04/$20.00 (c) 2004 IEEE Hong Kong RGC project research grant number B-Q571

[3] Mihui Kim, Hyunjung Na, Kijoon Chae, Hyochan Bang, and Jungchan Na, ‘A Combined Data Mining Approach for DDoS Attack Detection’. ICOIN 2004, LNCS 3090, pp. 943–950, 2004 Springer-Verlag Berlin Heidelberg 2004

[4] Aikaterini Mitrokotsa, Christos Douligeris, ‘Detecting Denial of Service Attacks Using Emergent Self-Organizing Maps’. 2005 IEEE International Symposium on Signal Processing and Information Technology

[5] Zhiyuan Tan1; Aruna Jamdagni1; Xiangjian He1, Priyadarsi Nanda1, and Ren Ping Liu, ‘Multivariate Correlation Analysis Technique Based on Euclidean Distance Map for Network Trac Characterization’.

[6] Lata1, Indu Kashyap , ‘Study and Analysis of Network based Intrusion Detection System’,International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 5, May 2013

[7] Ajoy Kumar, Eduardo B. Fernandez,’ Security Patterns for Intrusion Detection Systems’, 1 st LACCEI International Symposium on Software Architecture and Patterns (LACCEI-ISAP-MiniPLoP’2012), July 23-27, 2012, Panama City, Panama

[8] Punit Gupta , ‘Behavior Based IDS for Cloud IaaS’ , International Journal of Software and Web Sciences (IJSWS)

[9] S. Jin, D.S. Yeung, and X. Wang, “Network Intrusion Detection in Covariance Feature Space,” Pattern Recognition, vol. 40, pp. 2185-2197, 2007

[10] HhhhM. Tavallaee, E. Bagheri, L. Wei, and A.A. Ghorbani, “A Detailed Analysis of the KDD Cup 99 Data Set,” Proc. IEEE Second Int’l Conf. Computational Intelligence for Security and Defense Applications,

[11] V. Paxson, “Bro: A System for Detecgting Network Intruders in Real-Time,” Computer Networks, vol. 31, pp. 2435-2463, 1999.

[12] Ghhh D.E. Denning, “An Intrusion-Detection Model,” IEEE Trans.Software Eng., vol. TSE-13, no. 2, pp. 222-232, Feb. 1987.

[13] S.J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P.K. Chan, “Cost-Based Modeling for Fraud and Intrusion Detection: Results from the JAM Project,” Proc. DARPA Information Survivability Conf. and Exposition (DISCEX ’00), vol. 2, pp. 130-144, 2000.

[14] W. Wang, X. Zhang, S. Gombault, and S.J. Knapskog, “Attribute Normalization in Network Intrusion Detection,” Proc. 10th Int’l Symp. Pervasive Systems, Algorithms, and Networks (ISPAN), pp. 448-453, 2009.