Two-stage orthogonal network incident detection for the adaptive coordination with SMTP proxy

Ruo Ando, Yoshiyasu Takefuji
Graduate school of media and governance, Keio University, 5322 Endo Fujisawa,
Kanagawa, 252 Japan
{ ruo, takefuji }@sfc.keio.ac.jp

Abstract. The number of Security incidents that impose serious burden on the network system and its administration is still increasing rapidly. It is frequently pointed out that new intrusion types of which detection systems are unaware are hard to detect even for the skilled administrator with current in-service IDS. Also the more intelligent, automated and prompt method is required to protect and recover the system that could be attacked and infected. In this paper we present the adaptive detection and coordination system, two-stage detection, which consists of anomaly and misuse detection combined by lightweight neural networks to synchronize with the more specific data control of proxy server. The proposal method is able to correct false positive of AID module for the unusual changes in the system by the subsequent misuse detection learning labeled data. Therefore, it is possible to detect new type of attacks while maintaining a low false positive rate. Another feature of our model is to set delay line in the protection system for IDS to synchronize with proxy server for more effective data control. By This technique, the other servers can coordinate with each other to invalidate the unseen malicious code and search out the machines infected faster than prior protection method. Empirical experiments show that our model and deployment can be effective in reducing the false positive rate and in adaptive coordination with SMTP proxy server.

Keywords
AID; MID; clustering; classification; double-layer signature matrix; forwarding delay time; synchronization with detection intervals; adaptive coordination.

1 Introduction

1.1 Anomaly and misuse detection

Intrusion detection system is kind of alarm deployed on computer system and network to detect activity called misuse, that is something unauthorized action such as leaking confidential data. Recent increasing of the number of attacks against computer systems is rapid enough to pare off the effectiveness of human response. The more effective, automated and intelligent detection method is researched in many fields to take some measures for the unseen incidents. Generally, researches are objective to construct a system treating attacks with automatic response.

Among attacks, network intrusion includes far-flung activity to threaten the stability or assurance of information stored in the system linked to network. Network Intrusion detection system (NIDS) works on network to analyze the packets that passing through the firewall. NIDS monitors traffic of internal and external malicious activity so that administrator of system can take corrective countermeasure.

Almost traditional IDS is adapting signature-based detection methods that collate patterns in packet data, and comparing the patterns to dataset of signatures afforded manually by experts. Because signature database is now maintaining by system administrator, this methods is not able to find unknown types of intrusion unless the new signature is updated to the database.

Intrusion detection techniques are generally classified into two categories: anomaly detection and misuse detection. Misuse detection is performed by looking for the behavior of a known exploit scenario, which is usually described by a specific sequence or data. Current signature based methods is classified in misuse detection but lacks the scalability to extract features from attacks observed to detect even derivatives of conventional incidents by itself. Misuse learning algorithms on labeled data generally cannot detect new intrusion as it is. In addition, processing labeled data in order to find variation is usually so expensive. On the other hand, anomaly detection is performed by the inspection for the state that deviates from the baseline or normal state defined before. Profiling algorithms for AID on unlabeled data is frequently causing false positive because the audit data can be very large. And the output of this method is inclined to depend much on the numbers and features of data to train.

1.2 Tradeoffs between clustering and classification

There are two major data mining techniques applied for intrusion detection, clustering or classification. Clustering is the automated, unsupervised process that allows one to group together data into similar characteristics. Classification is the method to learn to assign data to predefined classes. These two methods are applied for two detection styles as we mentioned before, anomaly detection and misuse detection. Anomaly detection uses clustering algorithms because the behavior to find is unlabeled, with no external information sources. Misuse detection adapts classification algorithm because the activity to analyze requires that detector know how classes are defined.

The tradeoff about the accuracy and range of detection exists between clustering and classification. Classification deal with predefined data, so it affords detection of weaker signal and figure out accurate recognition. But in some cases, it may be biased by incorrect data to train and it is not able to detect new type of attacks in the sense that the attack does not belong to any category defined before. Clustering is not distorted by previous knowledge, but therefore needs stronger signal to discovery. At the same time it can deal with unlabeled attacks because the training doesn't specify what the detection system is trying to find while clustering go too far to perceive the activity that is not included incident affair.

1.3 Two-stage orthogonal incident detection

The thrust of this paper is to propose a new style of intrusion detection process called two-stage orthogonal intrusion detection, to address these tradeoffs. This algorithm has adjustment function based on misuse detection processing labeled data for the outputs of anomaly detection module. Our method is able to correct misjudges, as which anomaly detection module recognizes the unusual changes in the system and network, by adapting the sequent misuse detection trained on labeled data. As a result, it is possible to detect new type of attacks while maintaining a low false positive rate.

A previous approach to the method of this paper includes probabilistic or induction frameworks that learns the observed data. Some conventional misuse detection was applied for data of attacks to train using classification and anomaly detection was processing the normal or usual data. The other research is to compare the anomaly detection and misuse detection in monitoring system or network behavior. In those approach, the tradeoffs between misuse classification and anomaly clustering remains. In our model, instead of using misuse detection model to detect or specify the attacks, we use the adapting signal processing for misuse detection to train data the case of normal and attack so that it has adjustment function for the output of anomaly detection. We give up the processing incidents caused by one or few packets such as BOF exploit or information leak so that our model is able to validate the same traffic data in two-stage anomaly and misuse frameworks.

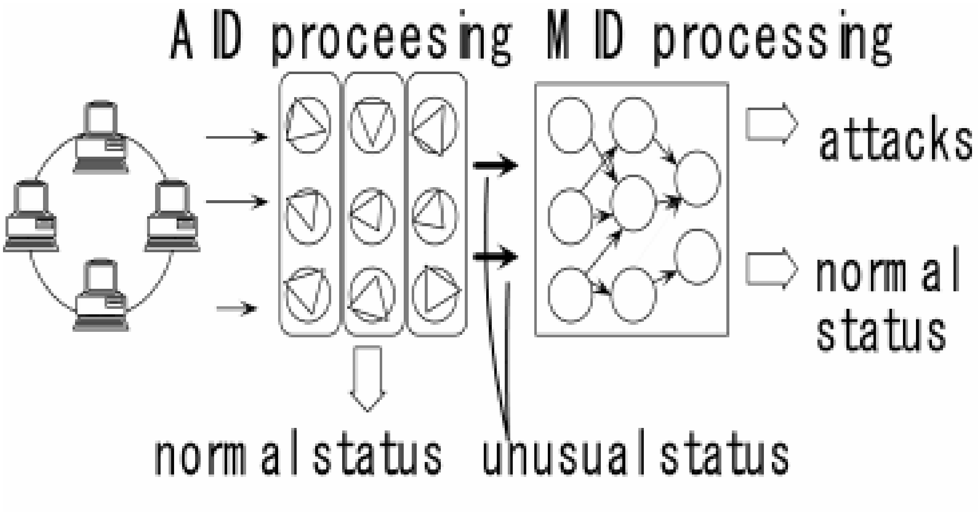

Fig. 1. Two stage incident detection consists of AID discrimination and MID recognition.

2 Related work

One of the most classic researches of intrusion detection was [1] in 1980. Since then, many data mining techniques have been proposed for more accurate detection as one active field [8][13]. In [3], the elements central to IDS are: resources to be guarded, models with features extracted as normal or legitimate, and techniques that identify monitoring behavior that are abnormal or intrusive. It is well-known method as anomaly detection [20][23], which has been implemented in the systems that extract the rules for usual behavior to discover possible intrusions. Misuse detection, on the other side of detection, is the method to attempt to identify instances by comparing actual activity against the expected behavior of attacker.

Anomaly detection to model normal behavior is implemented in the statistical techniques [17][18][21] and machine learning to recognize anomalous user and program [14][15][25][27]. Among anomaly detection techniques, clustering is popular and well studied method in this field [28]. There could be categorized anomaly detection into two types: profiling-emphasized [12][16][26][35] and real-time detection aiming type [4][19]. Misuse detection to discover exploitation that is recognized by a specific pattern or sequence of the events data observed is also performed in the expert systems[5]. And state-transition analysis [9][36], logic induction [7] and keystroke dynamics monitoring is proposed. In [6][10], discussion is mainly based on how to classify the predefined data.

In this paper, we adapt two-stage processing based on neural networks. Early work on the applying neural network for intrusion detection was due to Henning et al [24]. Ghosh and Schwartzbard have proposed the application of neural for anomaly and misuse detection[2]. They use artificial neural networks for anomaly detection in order to detect unseen behavior and for misuse detection in order to detect variations of known attacks. In recent years, there are various techniques in applying neural networks. Richard and Robert have improved the detection performance by adapting keyword selection and neural networks [34]. The advantages of neural networks are pointed in the potential to process the limited, incomplete and nonlinear data sources by Cannaby[33].

3 Methodology

In this paper, we aim to validate the applicability of two-stage orthogonal network incident detection. We first collect and generate data whose features and tendencies incidents causes in the network. Steps to generate quantitative profiles and signatures consists of AID clustering and MID classification. And then we discuss how the output of AID modules is adjusted by MID module. Finally, we investigate the some cases in which conventional data mining model causes false positive and compare system performances with the other popular IDS systems.

3.1 Collecting and generating data

The traffic data for AID processing is supplied by DARPA 2000 intrusion detection evaluation set and some experiment to generate traffic data in our research laboratory. We generate packets loaded the specific character strings in order to cause IDS (i.e. snort) alert. We also use nmap portscanning tool in our laboratory segment as supplement for DARPA dataset. In this paper, we do not deal with the incidents caused by one or few packets such as BOF exploit or information leak. About these attacks, we can take some measure by ex-post specification discussed in 1.4.

3.2 Profile matrix

Intrusion detection system is basically regarded as the system to detect activities to compromise the integrity, confidentiality and availability of the system monitored. Variety of data mining method has been researched to construct more automated, intelligent and accurate detection system. Among these, Clustering is the major technique to figure out the pattern hard to recognize manually and group related data together.

To discover groups from the input traffic data instances, we adapted the fast-convergence clustering algorithm called maximum neural network. As we discuss in 4.1, among unsupervised clustering algorithms such as kohonen model, maximum neural network has advantages that it decreases E of objective function rapidly. To create clusters, traffic is monitored and aggregated according to some conditions such as Input / output packets, the number of sessions connected, variance of STMP traffic in regular intervals.

3.3 Double-layer signature matrix

Generally, classification is applied for misuse detection which deal with labeled data to train. In this paper we adapt classification method for both normal and attack case. As we mentioned before, classification rely on predefined data, so it affords processing of weaker signal and figure out accurate recognition not to alert. The main purpose of classification of our model is not rather detect misuse than drop the unusual change into normal classes in order not for IDS to call an alert.

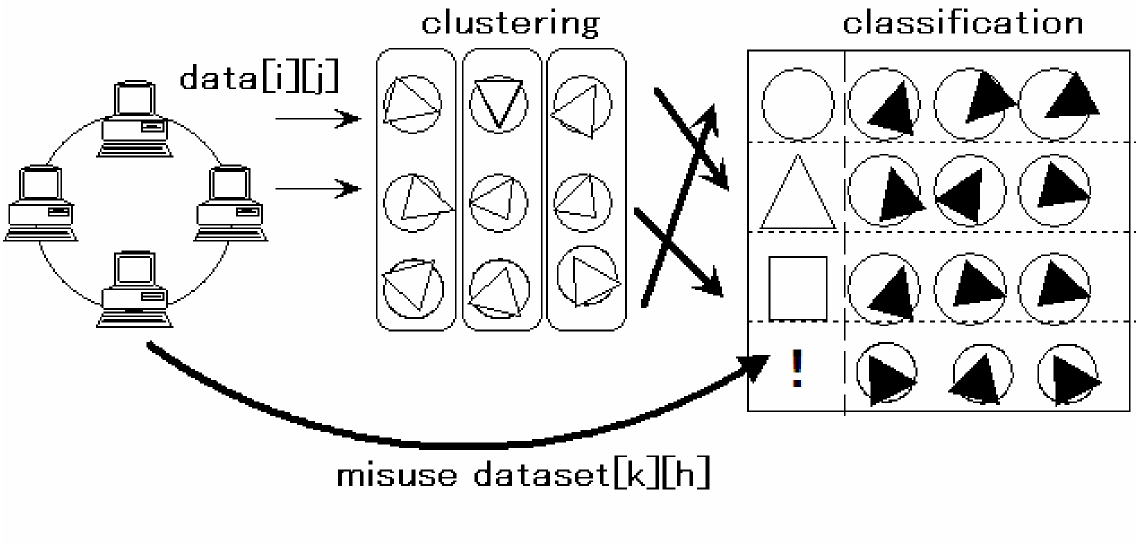

Fig. 2. Generating double-layer signature matrix

To assign instances to predefined classes, we use the lightweight classification algorithm called functional link neural network. Functional link model is simplified from multi-layer network to work out the extra neuron instead of dropping the hidden layer. If the convergence of processing traffic data is expected to some extent, we can take advantage of using this algorithm. In classification, misuse pattern preliminary generated is added to each vectors of profile matrix.

3.4 Two-stage incident detection

After figuring out hybrid signature matrix, we monitor network and aggregate packets data in regular intervals. Anomalies discriminating and misuse recognition is processing the same data aggregated from monitoring network. Fig.3 is the schematic diagram of ordinary signature-based detection system. As we discussed section 1, this model drops the packets of unknown attack. Packets whose feature does not math the predefined patterns have never recorded because logging, alert and data dump is executed after the process of detector. In this approach, the maintenance of signature becomes complicated and the number of alert logs where the flaw has already been fixed can be enlarged while missing unknown attacks.

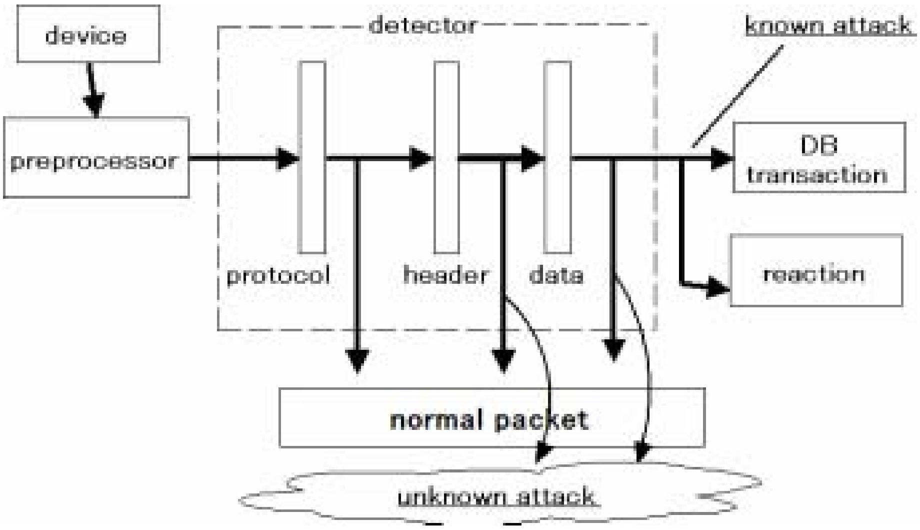

Fig. 3. Schematic diagram of typical intrusion detection model

The schematic diagram of our model is shown in Fig. 4. The interruptive data dump can be executed after the process of aggregation, AID discrimination and MID recognition. In this method, automated profiling and generating hybrid signature matrix are able to afford the function similar to signature and rule set of prior detection system. And it this enables administrators to complete configuration just by setting the several parameters while catching data of unknown attacks.

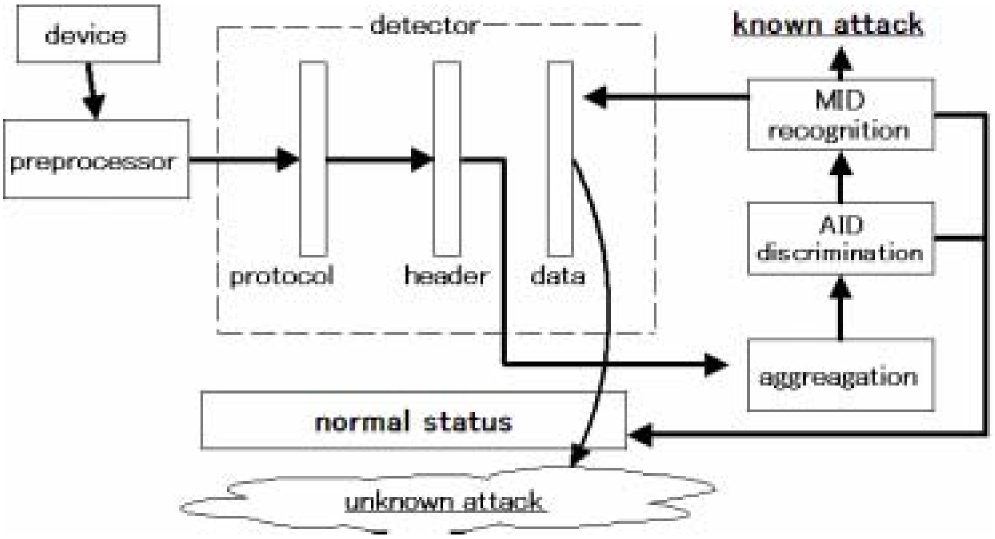

Fig. 4. Schematic diagram of proposal intrusion detection model

3.5 Response and adaptive proxy operation

Intrusion detection system is basically regarded as the system to detect activities to compromise the integrity, confidentiality and availability of the system monitored. This system itself does not work as a part of protecting or controlling network. Another technique to protect network is the application gateway using proxy. Application gateway takes advantages in operation cost effectiveness, information hiding and application-level security. On firewall, proxy is performing all data transaction between inside systems and outside internet in order to perform more specific rules and therefore tightly control (and sometimes modify) data passing through.

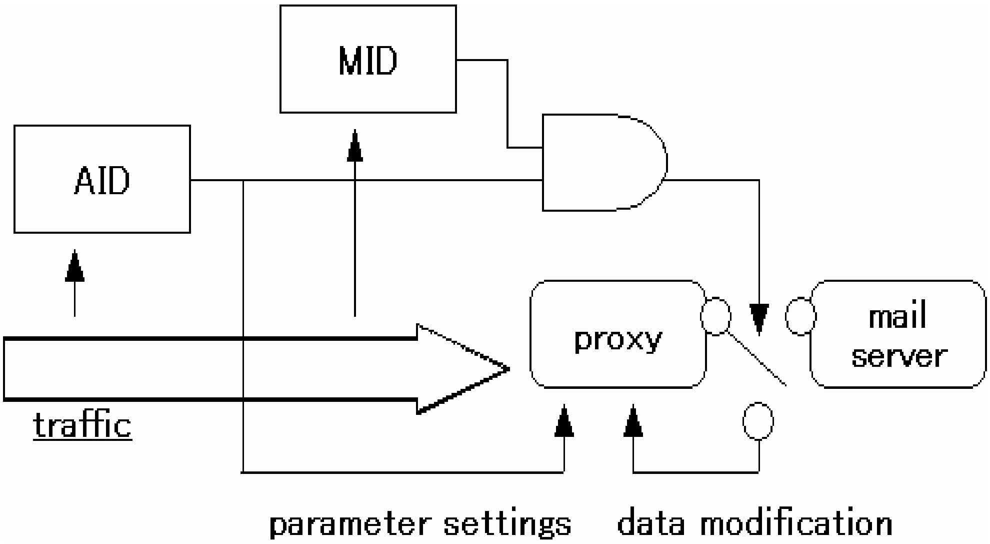

Fig. 5. Response. Application gateway synchronizes with IDS.

In proposal method of this paper, application gateway generates delay time to synchronize with IDS. The amount of time between receiving request and sending data in proxy server should be set same as the unit interval of intrusion detection. Consequently we can take more specific inspection or control of forwarding data according to the output of IDS.

Fig.5 show the coordination between IDS and SMTP proxy. Among proxies, SMTP proxy should have almost all function similar to main server because of its character of that protocol In anomaly detection, the out-of-range state dropped by misuse module is recognized as unusual event. But these false positive of AID can work as trigger to coordinate with proxy server because subsequent misuse detection are combined with AID and telling us that the current state of network is unusual. Besides, in the proposal method we can recognize the pattern that is discriminated according to the current state of network by the output of AID module. Therefore it is possible to execute adaptive operation by adaptive parameter settings of SMTP transactions according to how the network is going.

Misuse detection module is connected to the proxy in order to activate the data modification of proxy. When the security incident emerges in the network, our model can invalidate the malicious code passing through the proxy and execute more detailed inspection. To change the extension of all attachment forwarded when IDS is alerting is effective to prevent the spread of damages of malicious code that has not been unexampled before. Besides, this method affords the flexibility to select proper method according to the recognition of network status while variety of data mining techniques has been researched to detect computer viruses.

4 Algorithms and attacks

In this section, we discuss the algorithms applied and survey attacks detected. Both clustering and classification is applied in neural network because this algorithm has advantages of affording lightweight processing. And then, the four attacks detected are described.

4.1 Maximum neural network

Maximum neural network is one of the algorithms for unsupervised clustering. This paper use a two-dimensional maximum neural network that has M clusters N neurons figuring out M*N processing elements. To categorize N neurons into M clusters with P features, the feature vector is defined as follows

$$X_n=(x_{1k},x_{2k},....x_{pk}),\bar{X}_l=(\sum_{k=1}^{N} X_k V_{kl})/nl \tag{1}$$

We define the distance between the element K and cluster L according to Euclid geometry.

$$R_{kl}=(X_k-\bar{X}_l)^2 \tag{2}$$

The energy function extended for the two dimensional case is given by:

$$E=\sum_{k=1}^{N}\sum_{l=1}^{M} R_{kl}V_{kl} \tag{3}$$

To minimize E, we set the rate of change of the internal state as:

$$dU_{kl}/dt==-R_{kl}V_{kl} \tag{4}$$

In this algorithm, only one neuron will fire in each cluster and the output of other neurons in the same cluster become zero so that a stable state will always represent a valid solution. The condition of equilibrium state for the maximum neural model is written as:

$$\text{if } U_{km} = \max[U_{kl}(t);\forall l] \ V_{km}(t+1)=1$$

$$\textit{else } V_{km}(t+1)=0 \qquad (5)$$

This function takes advantage in that every state of convergence corresponds to a feasible solution.

In discrimination analysis, Mahalanobis distance is the useful way of calculating similarity in units of standard deviation from the group mean. In this measurement, the figured out circumscribing ellipse formed around the cluster defined the one standard deviation boundary of that group.

Let the vector from traffic monitoring:

$$\bar{x}_i' = (\bar{x}_{1i}, \bar{x}_{2i}, ...\bar{x}_{pi}) \qquad (6)$$

And the corresponding means from an appropriate normative group is written by:

$$x' = (x_1, x_2 ... x_p) \qquad (7)$$

Consequently, Mahalanobis D square is as follows:

$$D_i^2 = (x - \bar{x}_i)C^{-1}(x - \bar{x}_i)'$$

$$= \sum_{r=1}^{p}\sum_{s=1}^{p}(x_{ri} - \bar{x}_r)C^{rs}(x_{si} - \bar{x}) \qquad (8)$$

where C is the covariance matrix.

4.2 Functional neural network

A standard neural network typically contains on lots of simple computational elements or nodes arranged in on or more stages between input and output as shown in Fig. 6. The inputs to a node are linearly weighted before the sum before calculating sum by some nonlinear function, which gives to the network its nonlinear approximation ability.

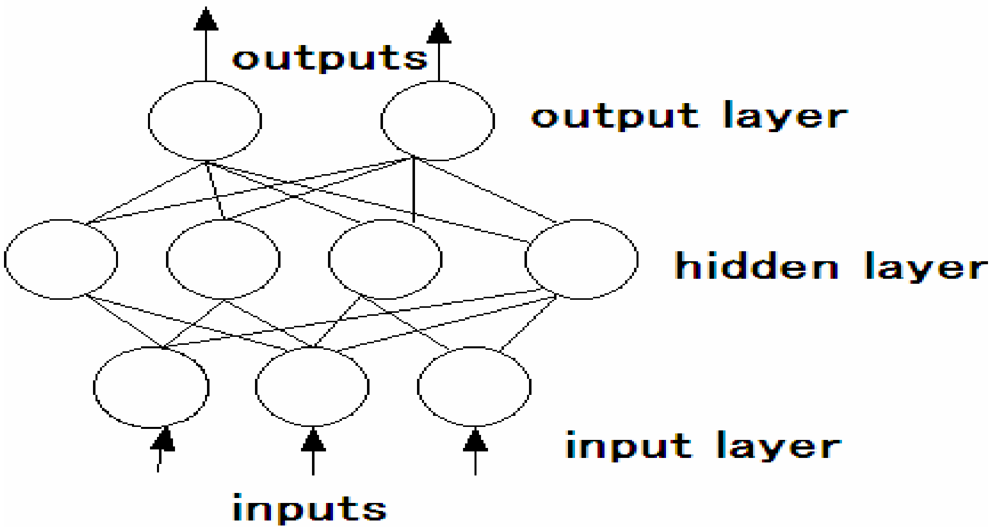

Fig. 6. Multi layer neural network

If nonlinear processing rules are adapted, the learning rate is often slow and local minimum may case problems. The functional link model shown in Fig.7 that eliminates all layers between input and output by using single step of processing is one way to avoid the nonlinear learning. The simplicity of the network, the time it takes to complete a model trained by prediction is so fast measured in milliseconds. Another benefit of functional link neural network is flexibility when the learning time is based on the numerous processing elements necessary for computing.

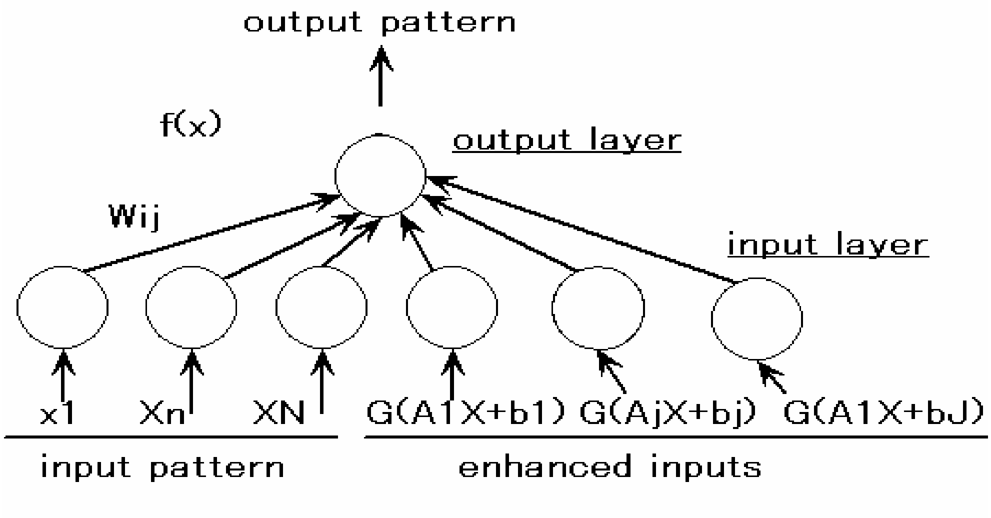

Fig. 7. Functional link neural network. This is simplified from multi layer network in order to reduce computing cost and system resources.

Let the rate of change in weight between input and output layer:

$$\Delta_p w_{ij} = \eta \delta_{pj} O_{pi} \tag{9}$$

where и is learning coefficient and O is the output. And δ is written by

$$\delta_{pj} = (t_{pj} - O_{pj}) f'(net_{pj}) \tag{10}$$

where t is teacher data and net is the sum of neuron outputs from input layers to output layers. We define f' as differential function of sigmoid equitation:

$$f'(x) = \frac{e^{-x}}{(1+e^{-x})^2} \tag{11}$$

The update rule is used the rate of weight change. This leads to the following rule:

$$w_{ij}(t+1)s = w_{ij}(t) + \Delta w_{ij}(t) \tag{12}$$

It is known that functional link model could be faster than the other multi-layer learning algorithm in the problem where the convergence is predicted to some extent.

4.3 Attacks detected

The incidents processed in the experiment of this paper are as follows:

[1] Denial of service

Denial of service is an attempt to preventing legitimate traffic or connection in order to deprive the services of a resource a user and organizations would normally expect to On the Internet, a denial of service attack is an incident in which a user or organizations is deprived of the services of a resource they would normally expect to have. Denial of Service includes flooding network, disruption of connection and dropping services of specific service or person. This attack is observed as packets flow that is of invalid sequence, size and frequency distribution.

[2] Port and vulnerability scanning

Port scanning is a common method for searching communication channels that is available. This technique is used by attackers and administrators to identity hosts or networks whose ports status is different individually. Also vulnerability scan is the process of searching for known security employing tools that find out security flaws usually based on a exposed exploits databases. These probing are characterized by the rapid flow with packets of which sequence is irregular and size is very small.

[3] Virus attack

Computer viruses can be observed like DoS attacks while replicating across a network. Some of major viruses spread mainly via SMTP(25), netbios-ssn(139) and HTTP(80).

As one point of view in monitoring traffic, there is a traffic of which variance is constant from these ports in a affected system because virus eventually replicate itself in various ways.

5 Experimental results and adaptive proxy operation

Our system is built to run on UNIX system to process DARPA training dataset and complementary generated data in our laboratory network. On creating profiles, we also use packet data captured by TCPDUMP. Each stage of anomaly discrimination and misuse recognition was tested on the same traffic data. The performance of any intrusion detection techniques should be evaluated for the detection ability and false positive rate at the same time. In this section we discuss the results of our approach in three cases where false positive tend to be caused frequently.

5.1 Unusual increase of session and traffic

Fig.8 and Fig.9 show the output of IDS in the case of unexampled increasing of session connected. As we expected before, MID recognizes the events as *not intruded* while AID call an alert that it is unusual events. These cases are relevant to the trend changes of network operation mode, which conventional anomaly detection causes false positive because the current status diverges from data profiled past.

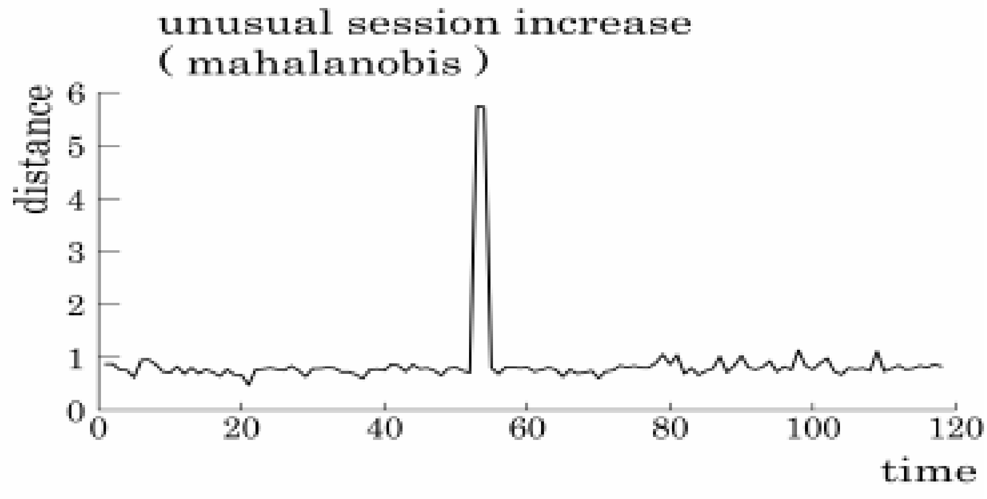

Fig. 8. AID output of unusual session increase (mahalanobis distance)

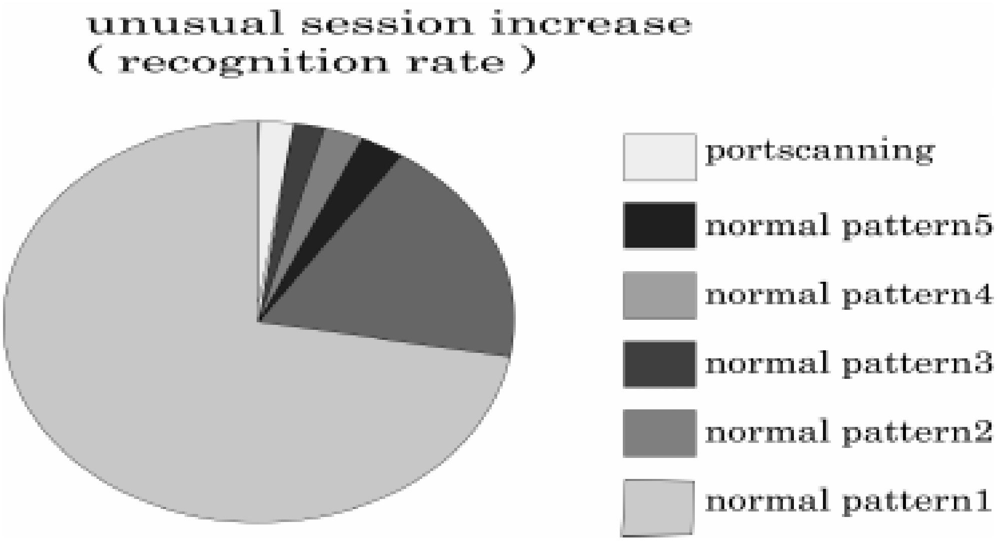

Fig. 9. MID output of unusual session increase (recognition rate of each pattern)

For network administrators, rapid traffic increase is the observable event even if it is not caused by intrusion. The output below is in the case where traffic is generated intentionally by network tools for investigation such as remote controller and streaming software.

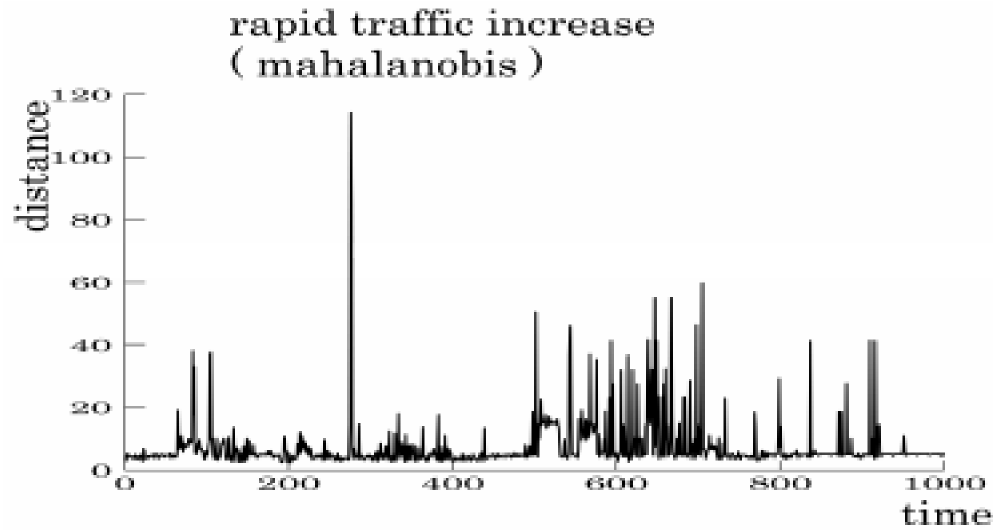

Fig. 10. AID output of rapid traffic increase

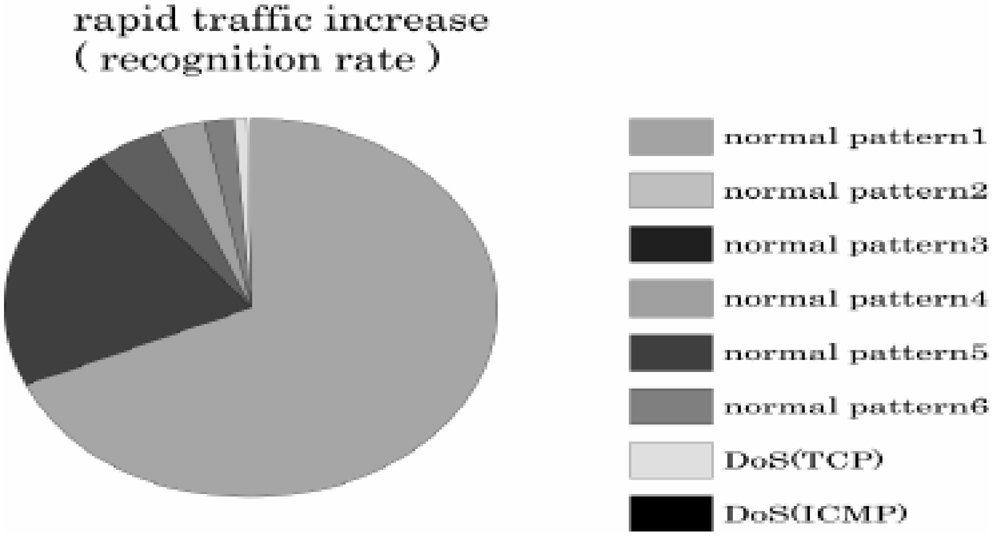

Fig. 11. MID output of rapid traffic increase

In this paper, we process clustering and classification on the same data set. Conventionally, classification method is applied for misuse detection and never contains normal cases. By using double-layer signature matrix that is generated by learning usual events we can some kind of the orthogonal nature between the results of clustering and classification. As we discussed before, there are some tradeoffs between clustering and classification. Clustering is able to detect the large number of attacks because the unsupervised learning and discrimination does not require predefined data. All the clustering algorithms can do is to find just unusual or abnormal event compared with the profiles from system monitored. To proceed to specify the events alerted, we should process weaker signal. Classification algorithm can handle weaker signal to figure out accurate recognition. System administration sometimes requires the detecting the unexampled attacks and taking some action faster while preventing misadjusted action by false recognition of current network status. The proposal model generates the double-layer signature in order to make the outputs of two algorithms compatible. We believe our system helps to address some bottlenecks of current system administration.

5.2 Adaptive coordination with SMTP proxy

As we mentioned in section 3.3, proposal model generates the double-layer signature matrix and execute two-stage incident detection. More specific recognition of network status and the further reduction of probability of false positive afford the adaptive coordination between IDS and the other services. Fig.12 show that two-stage detection system can be applied for the switch condition between parameter settings for the change of network status and active defense such as nullification of malicious code in application gateway.

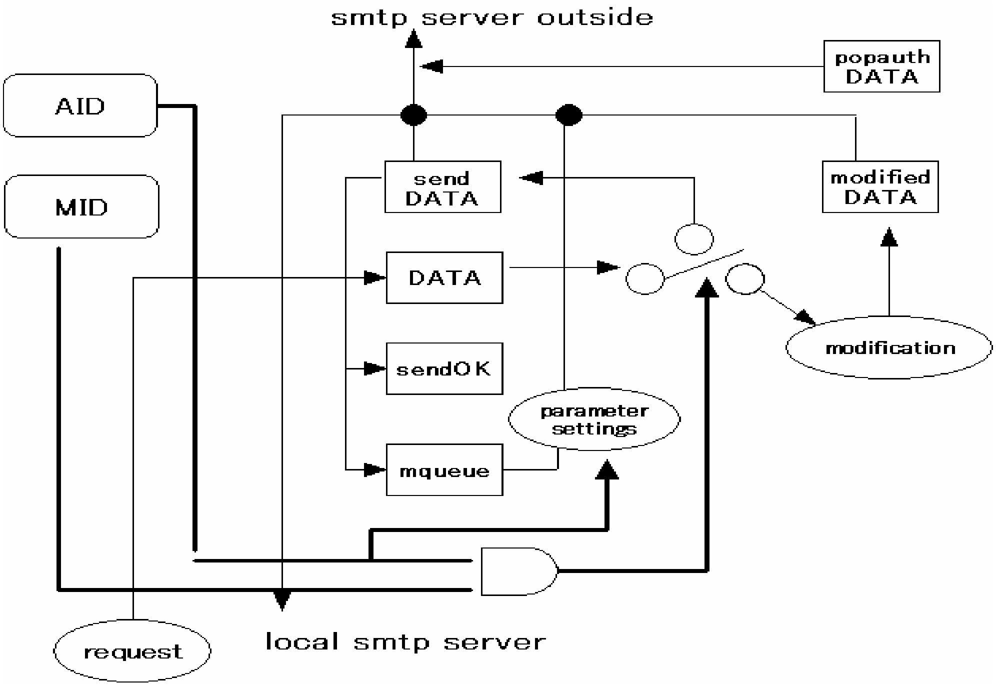

Fig. 12. Adaptive coordination with SMTP proxy

Table.1 shows that adaptive configuration of SMTP retry intervals. In the conventional system, these parameters are fixed regardless of any changes of the status of network. In proposal method, the reconfiguring of intervals is automated dynamically in order to utilize network resources more effectively. If there is no signal from MID module, our system controls parameters according to the recognition of network status of AID module.

	All traffic	Mail traffic	Unreach packets	Control packets	DATA(1)	DATA(2)	retry
cluster 1	15.31209	0.388798	0.75893	1.04656	0.607327	0.38473	2 times
cluster 2	3.70198	0.823086	1.48317	1.94138	4.47806	5.44776	1 time
cluster 3	73.05008	0.214085	0.153997	0.30955	0.038767	0.01159	2 times
cluster 4	2.17613	4.49294	3.98463	4.45534	3.30673	1.66963	1 time
cluster 5	5.33767	6.23943	5.83315	7.24659	17.5283	25.8547	1 time
cluster 6	2.19785	27.5897	15.51418	8.67507	3.60182	2.35259	no retry
cluster 7	209.463	1.95136	98.486	2.59399	1.9094	0.830959	no retry
cluster 8	0.559268	0.584629	1.09401	1.49928	2.13005	2.14854	2 times
cluster 9	3.27737	2.70642	3.62287	4.4212	10.0689	11.7759	1 time
cluster 10	23.75423	4.08378	8.93687	11.5061	43.3434	53.7749	1 time

Table.1. An example of automated configuring according to the network status

6 Conclusions

In this paper, we have proposed an intrusion detection system with two-stage orthogonal method to address the conventional tradeoffs between clustering for AID and classification for MID. Another advantage of the proposal method is that the other servers can synchronize with detection intervals to enables more specific data control to detect the unexampled attacks and protect the systems monitored faster. The advantages pointed out in our discussion are as follows:

[1] Adjustment function of classification using double-layer signature matrix offers the ability to keep the rate of AID false positive reasonably low while detecting numerous unlabeled attacks.

[2] The drastic reduction of false positive of AID enables intrusion detection system to be applied for switching condition between parameter reconfiguring and protective data modification of proxy.

[3] The delay time set in the proxy server according to the detection intervals affords the synchronized coordination with the servers to protect the system from the unseen malicious attacks faster.

Empirical experiments in the case where false positive is frequently caused show that the proposal method is functional with a recognition rate of attacks less than 10%, while finding the unusual status. And also the results obtained from the experiments in that adaptive synchronous coordination of proxy servers enables protection system to take some effective action after IDS is alerting.

7 Further work

Further work is planned for more precise updating the profiles. The usage trend of network seldom keeps consistency sometimes even for a single day. The AID clusters should be updated as the new machines are acquired, the old machine is release out, or new system is service-in. The issue of the renewal timing of profile must be addressed in the future. Additional attention should be given to adapting another algorithms and data. Concerning clustering, the other neural models such as kohonen model and LQV may also posses advantages. And besides the packet monitor, processing MIB information is expected to figure out more vivid profiling of the system.

References

[1] James P.Anderson. Computer security threat monitoring and surveillance,Technical report, James P. Anderson Co., Fort Washington, PA, April 1980.

[2] Anup K. Ghosh and Aaron Schwartzbard, A Study in Using Neural Networks for Anomaly and Misuse Detection, in: proceedings of USENIX Technical Program - Abstract - Security Symposium 99.

[3] Wenke Lee, Salvatore J. Stolfo, Data Mining Approaches for Intrusion Detection , in: Proceedings of the 7th USENIX Security Symposium

[4] W. Lee, S. Stolfo, P. Chan, E. Eskin, W. Fan, M. Miller, S. Hershkop, and J. Zhang , Real time data mining-based intrusion detection, in: Proceedings of Second DARPA Information Survivability Conference and Exposition, pp. I85-100, 2001.

[5] James Cannady, Artificial Neural Networks for Misuse Detection, in: Proceedings of the 1998 National Information Systems Security Conference (NISSC'98) October 5-8 1998. Arlington, VA.

[6] Ulf Lindqvist and Erland Jonsson. How to Systematically Classify Computer Security Intrusions. In: Proceedings of the 1997 IEEE Symposium on Security & Privacy, pages 154-163, Oakland, California, May 4-7, 1997. IEEE Computer Society Press.

[7] Calvin. Ko, Logic Induction of Valid Behavior Specifications for Intrusion Detection, in: Proceedings of the 2000 IEEE Symposium on Security and Privacy. 2000

[8] T.Lunt et al. Knowledge-based Intrusion Detection, in: Proceedings of 1989 Governmental Conference Artificial Intelligence Systems. March, 1989.

[9] K. Ilgun, R. A. Kemmerer, and P. A. Porras, State Transition Analysis: A Rule-Based Intrusion Detection ApproachIEEE Transactions on Software Engineering, 21(3). March, 1995.

[10] S.W.Shieh, Virgil D. Gligor, A Pattern-Oriented Intrusion-Detection Model and Its Applications IEEE Symposium on Security and Privacy 1991 p327-342,1991.

[11] U. Lindqvist and P. A. Porras, Detecting Computer and Network Misuse Through the Production-Based Expert System Toolset (P-BEST), in: Proceedings of the 1999 IEEE Symposium on Research in Security and Privacy. 1999.

[12] Fawcett, T. and F. Provost, Combining Data Mining and Machine Learning for Effective User Profiling, In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96),1996.

[13] Wenke Lee, Salvatore J. Stolfo, Kui W. Mok ,A Data Mining Framework adaptive intrusion detection, , in: Proceedings of the 7th USENIX Security Symposium.

[14] Marina Thottan, Chuanyi Ji, "Proactive Anomaly Detection Using Distributed Intelligent Agents", IEEE Network, Special Issue on Network Management, vol. 12,1998,Sept./Oct., pp. 21 – 27,1998

[15] R. Sekar, T. Bowen, and M. Segal. On Preventing Intrusions by Process Behavior Monitoring, in: Proceedings of the Workshop on Intrusion Detection and Network Monitoring, 1999..

[16] Makoto Iguchi and Shigeki Goto," Detecting Malicious Activities through Port Profiling", IEICE TRANS. INF. &YSYT. Vol.E82-D,No4,April 1999

[17] David Marchette, A Statistical Method for Profiling Network Traffic ,in: Proceedings of first USENIX Workshop on Intrusion Detection and Network Monitoring,1999.

[18] David Wagner and Drew, Intrusion detection via Statistic Analysis. IEEE Symposium on Security and Privacy 2001: 156-,2001.

[19] Linda lankewicz and mark benard, Real Time Anomaly Detection Using a Non parametric Pattern Recognition Approach, in: Proceedings of 7th Annual Computer Security Applications Conference, San Antonio, 1991.

[20] Winkler, J. R., A Unix Prototype for Intrusion and Anomaly Detection in Secure Networks, in: proceedings of the 13th National Computer Security Conference, Washington, D. C., Oct. 1990, p115 – 124,1990.

[21] DuMouchel, W., and Schonlau, M., A fast computer intrusion detection algorithm based on hypothesis testing of command transition probabilities. in:Proceeding of The Fourth International Conference of Knowledge Discovery and Data Mining, August 27-31, New York, 189-193.,1998.

[22] G. E.Liepins and H. S. Vaccaro, Anomaly Detection : Purpose and Framework, in: Proceedings of the 12th National Computer Security Conference, pages 495-504, October 1989, 1989.

[23] H. S. Vaccaro and G. E. Liepins. Detection of anomalous computer session activity, in: Proceedinges of the IEEE Symposium on Research in Security and Privacy [OAK], 280-289, 1989.

[24] K. Fox, R. Henning, J. Reed, and R. Simonian, A neural network approach towards intrusion detection, Tech. Rep., Harris Corporation, July 1990.

[25] K. Chen, S.C. Lu and H.S. Teng, Adaptive Real-Time Anomaly Detection Using Inductively Generated Sequential Patterns, Proceedings of the IEEE Symposium on Research in Security and Privacy, Oakland, CA, pp. 278-295, May 1990.

[26] Learning Program Behavior Profiles for Intrusion Detection ,Anup K. Ghosh, Aaron Schwartzbart, Michael Schatz,Proceedings 1st USENIX Workshop on Intrusion Detection and Network Monitoring,1999.

[27] Ghosh, Anup, K. Wanken, James and Charron, Frank. "Detecting Anomalous and

Unknown Intrusions Against Programs”, Proceedings of the 14th IEEE Annual Computer Security Applications Conference, pp. 259-267, 1998.

[28] Leonid Portnoy, Eleazar Eskin and Salvatore J. Stolfo, Intrusion Detection with unlabeled data using clustering, in: Proceedings of ACM CSS Workshop on Data Mining Applied to Security (DMSA-2001). Philadelphia, PA: November 5-8, 2001.

[29] D. Roverso, Neural Ensembles for Event Identification, in: Proceedings of Safeprocess'2000, The 4th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, June 14, 2000, Budapest, HU,2000

[30] C. Jirapummin, N. Wattanapongsakorn and P. Kanthamanon, Hybrid Neural Networks for Intrusion Detection System, The 2002 International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2002), pages 928-931, Phuket, Thailand, 16-19 July 2002.

[31] John A. Marin, Daniel Ragsdale, A Hybrid Approach to Profile Creation and Intrusion Detection (2001) ,DARPA Information Survivability Conference and Exposition, 2001

[32] R. Lippman and S. Cunningham, Improving intrusion detection performance using keyword selection and neural networks, in:Proceedings of the Second International Workshop on Recent Advances in Intrusion Detection September 7-9, 1999, West Lafayette, Indiana,1999.

[33] James Cannady, Artificial Neural Networks for Misuse Detection, in: Proceedings of the 1998 National Information Systems Security Conference (NISSC'98) October 5-8 1998. Arlington, VA., 1998

[34] S. Mukkamala, G.Kakarla, A.H. Sung,S.Veeramachaneni, Intrusion Detection: Comparison of Support Vector Machines and Neural Networks, IASTED Artificial Intelligence and Soft Computing conference, 2002

[35] D.Roverso, Neural Ensembles for Event identification, in: Proceedings of Safeprocess'2000, The 4th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, June 14, 2000, Budapest, HU,2000.

[36] S.T. Eckmann, G. Vigna, and R.A. Kemmerer, STATL: An Attack Language for state-based intrusion detection, Journal of Computer Security, vol. 10, no. 1/2, pp. 71-104, 2002

[37] Yingjiu Li, Ningning Wu, X. Sean Wang, Sushil Jajodia, Enhancing Profiles for Anomaly Detection Using Time Granularities, in Journal of Computer Security, IOS press, Vol. 10, Nos. 1,2, 2002, pages 137-157 ,2002

[38] .Y. H. Pao, and Y. Takefuji, "Functional-link net computing: theory, system architecture and functionalities," IEEE Computer, 25, 5, 76-79, 1992

[39] Takefuji Y.- Lee K.C.- Aiso H., An artificial maximum neural network : a winner - take - all neuron model forcing the state of the system in a solution domain, Biological Cybernetics, 67 : pp 243 -251, Springer- Verlag Ltd, London, 1992.