An Artificial Immune Model for Network Intrusion Detection

Jungwon Kim and Peter Bentley
Department of Computer Science, University Collge London
Gower Street, London, WC1E 6BT, U. K.
Phone: +44-171-380-7329, Fax: +44-171-387-1397
email:{J.Kim, P.Bentley}@cs.ucl.ac.uk

ABSTRACT: This paper investigates the subject of intrusion detection over networks. Existing network-based IDS's are categorised into three groups and the overall architecture of each group is summarised and assessed. A new methodology to this problem is then presented, which is inspired by the human immune system and based on a novel artificial immune model. The architecture of the model is presented and its characteristics are compared with the requirements of network-based IDS's. The paper concludes that this new approach shows considerable promise for future network-based IDS's.

KEYWORDS: artificial immune system, network intrusion detection, evolutionary algorithm, computer security

1. INTRODUCTION

An intrusion detection system (IDS) is an automated system for the detection of computer system intrusions. Early IDS's operated at the *host level*, whereas contemporary systems tend to be *network-based* [6]. *Host-based* IDS's monitor a single host machine using the audit trails of a host operating system and *network-based* IDS's monitor any number of hosts on a network by scrutinising the audit trails of multiple hosts. Even though various approaches have been developed and proposed, no network-based IDS has satisfied all its requirements [5].

This paper proposes a novel approach to building a network-based IDS, which is inspired by a human immune system. Kim and Bentley [5] carefully studied the several salient features of human immune systems and showed the possibility and advantages of adopting these features for network intrusion detection. This paper presents a more specific artificial immune model, which actually monitors a real-network, and describes the main components of this model.

The paper is structured as follows: section 2 categorises existing network-based IDS's into three types. It summarises each approach and identifies limitations. Section 3 presents the architecture for a new network-based IDS, using an artificial immune model. The characteristics of this system are analysed and compared with the requirements for network-based IDS's in section 4, and the paper ends with conclusions drawn from this work.

2. TAXONOMY OF NETWORK-BASED IDS'S

According to the overall architecture, we categorise network-based IDS's into three groups: *monolithic*, *hierarchical* or *co-operative*.

2.1 MONOLITHIC APPROACH

The monolithic approach employs a central intrusion detection server and simple host audit programs running on multiple local hosts. Monitored local hosts transfer their collected audit trails to an intrusion detection server and then this server performs audit trail analysis. Most network-based IDS's which have been developed until now use this approach and run in real small-scale networks [6]. However, such methods show some critical deficiencies in their scalability, robustness and configurability. Firstly, as a network size grows, a huge number of audit trails needs to be transferred from local hosts to a central server. This causes severe degradation of the network performance and it is difficult to guarantee scalability. Secondly, if a central intrusion detection server is subverted or fails, the overall IDS

becomes crippled. Thirdly, a single intrusion detection server should *uniformly* configure itself to the *various* local requirements of each host.

2.2 HIERARCHICAL APPROACH

The *hierarchical approach* was proposed to overcome the problems of the monolithic approach. It was designed to monitor large-scale networks, which have more than several thousand hosts. It defines a number of hierarchical monitoring areas and each IDS monitors a single area. Instead of transferring all the collected audit data from local hosts to a central IDS, each single IDS at any level of monitoring area performs local analysis and sends its local analysis results up to the IDS at the next level in the hierarchy. Thus, IDS's at higher levels only need to analyse transferred local reports collectively. The Graph-based Intrusion Detection System (GrIDS) [10] and Event Monitoring Enabling Responses to Anomalous Live Disturbances (EMERALD) [8] project propose this approach to monitor large-scale networks and they are still in progress. The hierarchical approach seems to show better scalability by allowing local analyses at distributed local monitoring areas. However, other problems raised from the monolithic approach still remain. When the topology of the current network is changed, it causes a change of network hierarchy and the whole mechanisms to aggregate local analysis reports must be changed [6]. In addition, when a monitor residing at the highest level is attacked or crashed, then all network-wide co-ordinated intrusions, which are identified only by the global analysis of local results collected from distributed monitors at lower levels, easily escape detection.

2.3 CO-OPERATIVE APPROACH

The *co-operative* approach attempts to distribute the responsibilities of a single central server to a number of co-operative host-based IDS's. Each IDS is responsible for monitoring only a small aspect of a local host and a number of IDS's operate concurrently and co-operate with each other. Moreover, they can make a coherent inference and make a global decision. The difference of this approach from the hierarchical approach is that there is no hierarchy among distributed local IDS's. Therefore, the failure and subversion of any IDS does not always prevent the detection of co-ordinated attacks. The Co-operative Security Managers (CSM) project [12] and the Autonomous Agent For Intrusion Detection (AAFID) project [1] proposed this approach. In these proposals, it is claimed that most of problems encountered by the two approaches previously mentioned would be resolved. These projects are still in progress and the validity of this claim remains unproven. In particular, this approach raises a different problem, namely the maintenance of efficiency. It places too many overheads on monitored local hosts such as many communication mechanisms, auditing mechanisms and analyses of audit trails and these can be a significant encumbrance to them.

To summarise, various architectures of network-based IDS's have been proposed and here they have been grouped into three different approaches. However, each approach shows different problems and no network-based model completely resolves the encountered problems.

3. ARTIFICIAL IMMUNE MODEL OVERVIEW

The human immune system has been successful at protecting a human body against a vast variety of foreign pathogens or organisms [11]. This remarkable property is attractive to computer security researchers and artificial intelligence researchers. Based on the studies by immunologists, a growing number of computer scientists have proposed several different computer immune models [3]. The main idea of these models is distinguishing self, which is normal, from non-self, which is abnormal. In this paper, with respect to network intrusion detection, we view the normal activities of monitored networks as self and their abnormal activities as non-self. Many sophisticated network intrusions such as sweeps, co-ordinated attacks and Internet worms are detected by monitoring the anomalies of network traffic patterns [9]. Most network-based IDS's monitor network packets and their identified anomalies show critical signatures of these network intrusions [6], [11]. Thus, the artificial immune model is designed for distinguishing normal network activities from abnormal network activities and expected to detect various network intrusions[1].

The overall architecture of the novel artificial immune model developed as part of this work is presented in Figure 1. The artificial immune model for network intrusion detection consists of a primary IDS and secondary IDS's. For a

[1] Most network-based IDS's operating in real network environments monitor the audit trails generated by a local host together with the network activities. This kind of approach is more reliable at detecting various intrusions. Even though the artificial immune model proposed in this paper restricts its monitoring scope to network activities, it should be extended by monitoring local audit trails and this extension might be possible by employing one suggestion, a host-based computer immune system, introduced in [7].

human body, at the bone marrow and the thymus, various detector cells, called antibodies, are continuously generated and distributed to secondary lymph nodes, where antibodies reside to monitor living cells. The distributed antibodies monitor all living cells and detect non-self cells, called antigens, invading into secondary lymph nodes. For the artificial immune model, the primary IDS, which we view as the bone marrow and thymus, generates numerous detector sets. The architecture shown in Fig.1 is assumed to monitor a single network domain. Therefore, all the input network packets transferred to a monitored single network domain firstly arrive at the first router[2]. Each individual detector set describes abnormal patterns of these network traffic packets. It is unique and transferred to each local host. We view local hosts as secondary lymph nodes, detectors as antibodies and network intrusions as antigens. At the secondary IDS's, which are local hosts, detectors are background processes which monitor whether non-self network traffic patterns are observed from network traffic patterns profiled at the monitored local host. The primary IDS and each secondary IDS have communicators to allow the transfer of information between each other.

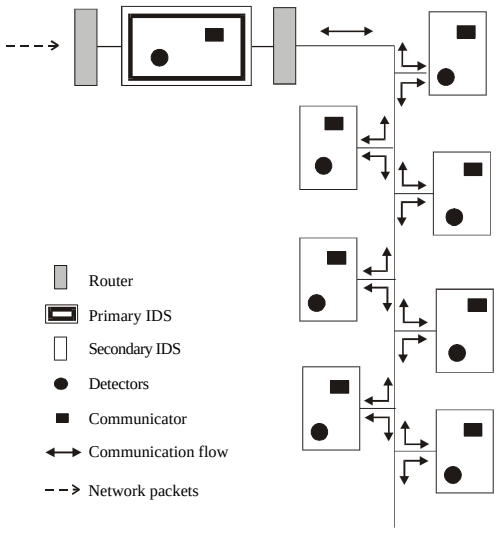

Figure 1 Physical Architecture of the Artificial Immune Model.

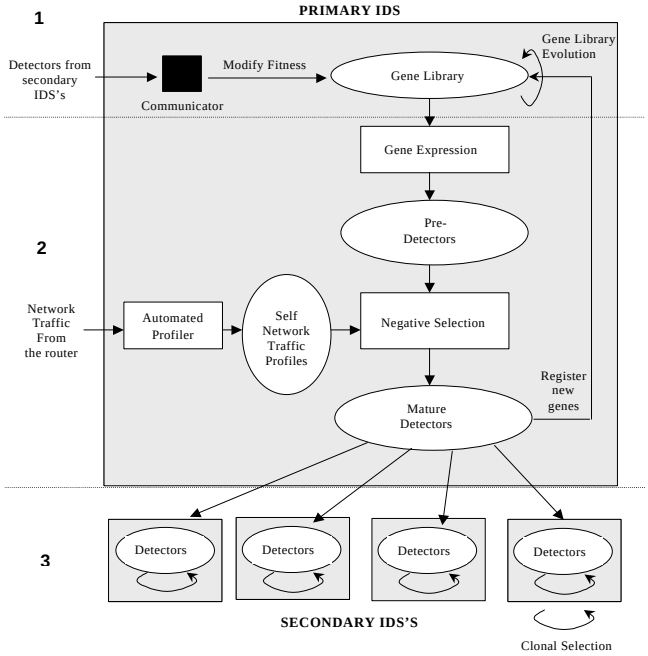

Figure 2. Conceptual Architecture of the Artificial Immune Model

Kim and Bentley [5] identified three main goals for designing an effective network-based IDS's: being distributed, self-organising and lightweight. Furthermore, they showed that the several sophisticated mechanisms of the human immune system allow it to satisfy these three goals. For the proposed artificial immune system, these mechanisms are embedded in three evolutionary stages: gene library evolution, negative selection and clonal selection. While the currently existing computer immune models focus on the use of a single significant stage according to their perceived purpose [3], [4], [6], the new artificial immune model proposed in this paper combines these three significant evolutionary stages into a single methodology. The overall conceptual architecture of the proposed artificial immune model is shown in Figure 2. In Figure 2, stage one indicates gene library evolution, stage two presents negative selection and stage three shows clonal selection. The functions in each stage and how these three stages operate together for performing network intrusion detection are described in the following two sub-sections: Primary IDS and Secondary IDS's.

3.1 PRIMARY IDS

The primary IDS performs the first two evolutionary processes: gene library evolution and negative selection. At the gene library evolution stage, it aims to gain general knowledge on effective detectors. At the negative selection stage, it aims to generate a number of diverse detectors, which do not match self, and transfer a number of unique detector sets to distributed local hosts. In order to achieve these tasks, it contains the following components (shown in Figure 2).

[2] This assumption can be extended for monitoring large-scale networks which include a number of different domains. It is achieved simply by installing a single primary IDS on each domain and monitoring each domain independently.

At the first stage, a gene library is generated and maintained by an evolution process[3]. The *gene library* of the artificial immune model stores the potential genes of detectors and diverse genetic mechanisms generate new detectors. The potential genes are the selected fields of profiles to describe anomalous network traffic patterns. They are selected after understanding the detailed mechanisms of network protocol and their security holes [9]. The initial genes might be set by the values of these fields that are observed when a previously known intrusion is simulated. They can be described by the number of packets, bytes, specific errors, etc of typical network services for a specific short period or one connection time [6], [9]. If a new detector, which is generated from initial genes and transferred to a local host, detects anomalous network traffic activity, the genes comprising this detector will be added to the gene library. But, if the genes are already stored in the gene library, the fitness values of these genes are increased. If this process continues, the size of the gene library will grow. However, if the size of the gene library is limited, whenever the size is above a fixed length, the genes that have lowest fitness values will be removed from the gene library. This mechanism drives the artificial immune model to perform *gene library evolution*. This process allows the artificial immune model to learn knowledge of currently existing intrusions regardless of whether they were detected previously or not, making it self-organising. Furthermore, its self-organising feature allows it to be lightweight. This is because it does not have to contain all the information of intrusions that have been detected so far. Instead, it holds only the smaller and limited number of genes which currently survive.

At the second stage, the *gene expression* process generates various *pre-detectors* via rearrangement of selected genes, the selection of various gene-joining points, mutation of genes, which are randomly selected from the gene library. These mechanisms can lead to the generation of a vast number of possible pre-detectors from combinations of genes [11]. This process permits the artificial immune model to detect numerous intrusions using a smaller number of detectors, making it lightweight. The *automated profiler* produces a self network traffic profile of raw network traffic packets transferred from the first router. However, the raw network traffic volume is huge and the normal activity patterns are hidden. The automated profiling component reduces the huge volume of raw network packets into a self profile. The fields of the *self network traffic profile* are identical to those of the generated pre-detectors. In other words, specific values of these fields can determine whether the observed network activities are normal (the self-profiles), or anomalous (the pre-detectors). However, some pre-detectors can be false detectors because they have novelty generated via mutation in the gene expression process. These false pre-detectors are removed by the *negative selection* process, which matches them to a self network profile produced by an automated profiler. If the field values of pre-detectors match the field values of the self network traffic profiles, we can consider these new pre-detectors as false detectors which wrongly identify self as anomalies, and thus they are eliminated [4]. This process removes false pre-detectors by presenting self without any global information about self and hence it shows the property of self-organisation.

Finally, the surviving detectors from negative selection become *mature detectors*. Before each detector set is transferred to an individual local host, the genes made up of mature detectors are newly registered in the gene library. Unique sets of detectors and self network traffic profiles are selected from these mature detectors based on each network connections in order to transfer them to local hosts. This selection guarantees the uniqueness of individual detector sets. These unique detector sets detect network intrusions independently in a local host level [7] and permit the artificial immune model to be distributed. The selected detector sets and self network traffic profiles are transferred to the second router and it distributes them to their corresponding secondary IDS's.

In order to perform above processes, the primary IDS needs to communicate with the secondary IDS's. For example, the former needs to transfer mature detectors to the latter and the latter needs to send newly found useful genes to the former. The *communicator* controls any type of communication between the primary IDS and the secondary IDS's.

3.2 SECONDARY IDS

The secondary IDS's perform the last evolutionary process: clonal selection. Its main tasks are detecting various intrusions with a limited number of detector sets and cloning the identical detectors that are performing well, producing memory detectors and driving the gene library evolution in the primary IDS. These tasks are achieved by the operations of several components: self network profiles, unique detector sets, network traffic anomaly detection, clonal selection of detectors, memory detectors and a communicator.

[3] It should be noted that this evolutionary process is a simulation of the natural evolutionary process for gene libraries. In nature, the DNA (gene libraries) of an organism cannot change within the lifetime of that organism. Evolution operates on populations of organisms, evolving gene libraries based on which organisms survive (i.e., how effective their immune systems are, throughout their lives). This is clearly computationally expensive, so in this model we treat the gene library as a population in itself and evolve it with a single artificial immune system. However, unlike gene library evolution, the other two evolutionary processes within the model operate in a conceptually similar manner to natural immune systems.

In order to perform *network traffic anomaly detection*, the detectors of *unique detector sets* and *self network profiles* transferred from the primary IDS are compared. First of all, the match strength between the field values of a detector and the self profile is measured. When this strength is over a pre-defined threshold, this process informs it to the communicator. This approximate binding helps make the artificial immune model lightweight. This is because one detector can bind to a number of different intrusions if only their match strength is over the threshold [7].

After detecting anomalies, the secondary IDS's perform *clonal selection*. When a new detector detects an abnormal network traffic activity, this detector remains as a *memory detector* in a secondary IDS and clones itself. The cloned detectors can be transferred to other hosts. They act as misuse detectors. They detect quickly the same intrusions in the future, which have previously detected. Furthermore, the genes of this detector will be added to the gene library in the primary IDS if they do not exist in the gene library or the fitness values of these genes will be increased otherwise. This drives the gene library evolution in the primary IDS. As the anomaly detection of detectors in local hosts continues, each local host will have more memory detectors and the number of detectors that need to be transferred to each local host will decrease. This process allows the model to be self-organised and lightweight. Instead of having the predefined information about specific intrusions, it self-organises the fittest detectors by detecting the currently existing intrusions. In addition, the evolved gene library and memory cells decrease the efforts to create various new detectors, helping to make the model lightweight.

The final decision of whether a network intrusion has occurred is made according to the collective decisions from several local hosts. The artificial immune model employs the agent communication mechanism suggested by Balasubramaniyan et al. [1]. When suspicious activity is detected by anomaly detection process at any secondary IDS, it sends a signal to a *communicator*. The communicator increases the risk level and sends a signal to the communicators in other hosts and the primary IDS. Other communicators, which receive the signal, increase the risk level. If suspicious activities are found from several hosts within a short time, the risk level in each host and the primary IDS will be rapidly increased. When this risk level becomes above a certain threshold, a communicator can inform the breach of network intrusion to a security officer through a user-interface.

3.3 SUMMARY OF ARTIFICIAL IMMUNE MODEL

The artificial immune model described above consists of the primary IDS and the secondary IDS's. It combines three evolutionary stages. *Gene library evolution* simulates the first stage of evolution, which learns knowledge of currently existing antigens. This process allows the model to be lightweight and self-organising. *Gene expression* and *negative selection* form the second stage of evolution, generating diverse pre-detectors and selecting mature detector sets by eliminating false pre-detectors in a self-organising way. The transfer of unique detector sets to the secondary IDS's also occurs at this stage, making the model distributed. *Clonal selection* is the third stage of evolution, detecting various intrusions with a limited number of detector sets using approximate binding, and generating memory detectors. This generality and efficiency results in the model being lightweight. In addition, this process drives the gene library evolution in the primary IDS. These three processes are co-ordinated across a network to satisfy the three goals for designing effective IDS's: being distributed, self-organising and lightweight [5].

4. DISCUSSION OF ARTIFICIAL IMMUNE MODEL

To provide an indication of the advantages of this approach, the new artificial immune model suggested in this paper is now analysed with respect to the requirements of a network-based anomaly detector. Kim and Bentley [4] described the seven requirements of a competent network-based IDS. The proposed artificial immune model is assessed with respect to these seven requirements.

The proposed artificial immune model is distributed by using a unique detector set in a local secondary IDS for detecting local intrusions and employing communications among secondary IDS's for detection network intrusions. This distributed feature allows the model to be robust, configurable, extendible and scalable. Firstly, the artificial immune model is ***robust***. The failure of any detector set residing at any local host does not cripple an overall artificial immune system even though it may cause some minor degradation of detection accuracy. Each detector set can still detect network intrusions even after the failure of the primary IDS. This is because each local host already has detector sets, which were transferred before the failure. Besides, if an intruder breaks through a local host and gains the information about how detectors describe anomalous behaviour, this intruder might attempt to use this information to disguise his or her activities. However, the uniqueness of each detector set makes this kind of attempt difficult. Secondly, it is ***configurable***. Even though detectors are generated in the primary IDS, their usefulness is proved at a

local level by employing clonal selection in each secondary IDS. Furthermore, this local level clonal selection drives the gene library evolution in the primary IDS. In other words, the generated detectors co-evolve to detect various intrusions and this co-evolution is led by the self profiles and existing intrusions in each local level. Therefore, the artificial immune model configures local requirements in a self-organised way disregarding various requirements of other hosts. Thirdly, it is ***extendible***. When a new local host is added to a network, it simply needs to generate another detector set for the new host and install a secondary IDS consisting of an automated profiler, anomaly detection process, clonal selection process and a communicator without considering other hosts. These components are totally independent from the components at other secondary IDS's and thus they ensure that the artificial immune model is easy to extend. Fourthly, it is ***scalable.*** At initial stages, an artificial immune system might need to generate a large detector set. However, as it detects anomalies more and more, each local host will be equipped with more and more memory detectors and eventually will require very few new detectors to be transferred. Nevertheless, this requires the occurrence of a number of various intrusions within a practically short time. Therefore, the overall artificial immune mechanisms may be simulated by presenting a number of intrusions for a short time and this is used for the initial learning process before the launch of real intrusion monitoring by the artificial immune model.

In addition, the artificial immune model is self-organising by performing gene library evolution, negative selection and clonal selection. This property of self-organisation makes the model both ***adaptable*** and capable of ***global analysis***. Firstly, the negative selection process allows detectors to consider dynamically the self information at any moment. The clonal selection and the gene library evolution generate various detector sets that are the fittest for the recently encountered intrusions. Therefore, the newly generated detectors always dynamically learn knowledge about currently existing intrusions and self. Furthermore, when a new intrusion is detected, these new abnormal patterns will be registered to the gene library of the primary IDS and remain as the memory detectors at the secondary IDS's. Therefore, the artificial immune model still can be highly adaptive. Secondly, global analysis is achieved via the communication between the primary IDS and the secondary IDS's and this communication mechanism is simple and autonomous, which does not require a global communication controller.

Finally, the artificial immune model is lightweight by detecting various intrusions using approximate binding and memory cells, performing gene library evolution and gene expression[4]. This lightweight feature provides good ***efficiency***. Firstly, the approximate binding permits one detector to detect a number of different intrusions. Consequently, the model needs to generate a much smaller number of detectors than the number of intrusions that are expected to be detected. Secondly, as mentioned above, clonal selection generates memory detectors within local hosts. As the number of memory detectors increases, the number of new detectors required will decrease, resulting in a reduction of computation time. More importantly, as the detection of intrusions continues, a gene library collects useful genes. Through gene library evolution, these genes define detectors that have already proved their usefulness by identifying anomalies. Since such detectors use only the most useful features of the profile at any one time, this removes the need for each local host to perform feature selection during profiling. This feature certainly reduces the overheads of local monitored hosts compared to the co-operative approach. The final example of efficiency in the system is provided by the gene expression process. This process allows the artificial immune model to generate a huge number of detectors from a small number of genes in the gene library.

5. CONCLUSION

This paper investigated the existing network-based IDS's. They were categorised into three different approaches: monolithic, hierarchical and co-operative and problems were identified for each approach. In order to resolve these problems, a novel artificial immune model was presented. This model combines the three evolutionary stages: gene library evolution, negative selection and clonal selection into a single methodology. These three processes are co-ordinated across a network to satisfy the three goals for designing effective IDS's: being distributed, self-organising and lightweight. Analysis of the characteristics of this unified evolutionary approach show that, unlike existing approaches, the proposed artificial immune model does satisfy the requirements of network-based IDS's. Consequently, algorithms based on this model show considerable promise for future IDS's.

A network-based IDS utilising the artificial immune model presented in this paper is being implemented in order to prove the validity of this approach. Current work is focusing on building initial self profiles and detectors from normal

[4] Even though the novel evolutionary approach of the artificial immune model allows the secondary IDS's to be lightweight, it may impose some more work on the primary IDS. To resolve this problem, it may be designed as a parallel array of the primary IDS's [2]. For example, the first router which receives all network input packets outside a network domain can split network packets into groups of flow based on each connections. Then a number of different flow groups can be sent to each primary IDS. Each primary IDS will have the identical components that have been introduced in this paper and it generates specific detector sets and self profiles based on each connection. The specific detector sets and self profiles generated by an individual primary IDS are sent to the second router and this router can transfer them to a specific secondary IDS (a local host) within a domain.

and abnormal TCP/IP packets, which were collected from a real network environment. For the short-term future work, a more efficient encoding scheme to represent detectors and self network profiles and their matching function will be investigated.

6. ACKNOWLEDGEMENT

This work has been partially supported by the Korea International Collaboration Research Funds (I-03-002), the Ministry of Science and Technology, Korea. J. Kim would like to acknowledge K. Carlberg at SAIC for his useful comments.

7. REFERENCES

[1] Balasubramaniyan, J. S. et al., 1998, "An Architecture for Intrusion Detection using Autonomous Agents", Department of Computer Sciences, Purdue University, Available at http://www.cs.purdue.edu/coast/coast-library.html
[2] Carlberg, K.(SAIC), Dec.1998, personal communication.
[3] Dasgupta, D.; Attoch-Okine, N., 1997, "Immunity-Based Systems: A Survey", *Proceeding of the IEEE International Conference on Systems, Man and Cybernetics*, Orlando.
Available at http://www.msci.memphis.edu:80/~dasgupta/publications.html
[4] Forrest, S.; Hofmeyr, S.; Somayaji, A., 1997, "Computer Immunology", *Communications of the ACM*, vol.40, No.10, pp.88-96. Available at http://www.cs.unm.edu/~forrest/papers.html
[5] Kim, J.; Bentley, P., 1999, "The Human Immune System and Network Intrusion Detection", submitted to EUFIT'99.
[6] Mykerjee, B., et al, 1994, "Network Intrusion Detection", *IEEE Network*, Vol.8, No.3, pp.26-41.
[7] Somayaji, A.,; Hofmeyr, S.; Forrest, S., 1997, " Principles of a computer immune system", *Proceeding of New Security Paradigms Workshop, Langdale, Cumbria*, pp.75-82.
[8] Porras, P. A.; Neumann, P. G., 1997, "EMERALD: Event Monitoring Enabling Responses to Anomalous Live Disturbances", *Proceeding of 20th National Information System Security Conference*.
Abailable at http://www.csl.sri.com/emerald/downloads.html
[9] Porras, P. A.; Valdes, A., 1998, "Live Traffic Analysis of TCP/IP Gateways", *Proceeding of ISOC Symposium of Network and Distributed System security*. Available at http://www2.csl.sri.com/emerald/downloads.html
[10] Staniford-Chen, S., et al., 1996, "GrIDS -- A Graph-Based Intrusion Detection System for Large Networks", *Proceeding of the 19th National Information Systems Security Conference.* http://seclab.cs.ucdavis.edu/papers.html
[11] Tizard, I. R., 1995, *Immunology: Introduction*, 4th Ed, Saunders College Publishing.
[12] White, G. B.; Pooch, U; Fisch, E. A, 1996, "Cooperating Security Managers: A Peer-Based Intrusion Detection System", *IEEE Network*, Vol.10, No.1, pp.20-23.