MAPPING HUMAN IMMUNE SYSTEM to NETWORK INTRUSION DETECTION SYSTEM

Miss.Farha Rafiq Qureshi,
Computer Technology,RTM Nagpur University,Rajiv Gandhi college of engg. Research and technology ,Chandrapur,India(+91)
farha.r.qureshi@gmail.com

***Abstract*: This paper describes the difference between the human immune system and network intrusion detection systems. The paper begins by briefly introducing existing intrusion detection systems (IDS's).An overview of the human immune system is presented and its salient features that can contribute to the design of competent network-based IDS's are analyzed. The analysis shows that the actions of the human immune system satisfy all the identified design goals. Consequently, the paper concludes that the design of a novel network-based IDS based on the human immune system is promising for future network-based IDS's.**

***Keywords*: computer immune systems, human immune systems, network intrusion detection.**

I. Introduction

An intrusion detection system (IDS) is used for detecting computer system intrusions. The main goal of an IDS is to detect unauthorised use, misuse and abuse of computer systems by both system insiders and external intruders.As one novel approach, a few computer scientists have proposed simple computer immune models for intrusion and computer virus detection [1], [7], [6], [8].This paper aims to unravel the significant features of the human immune system, which would be successfully employed for a novel network intrusion detection model. Several salient features of the human immune system, which detects intruding pathogens, are carefully studied and the possibility and the advantages of adopting these features for network intrusion detection are reviewed and assessed.

This paper is structured as follows: the following section briefly describes existing IDS's. Section 3 outlines the requirements of network-based IDS's. Section 4 introduces three network-based IDS design goals to satisfy these requirements. Then, in section 5, an overview of the human immune system is presented. Section 6 analyses the significant features of human immune systems and compares these with the design goals of network-based IDS's.

Finally, this paper presents the conclusion drawn from this work and future work.

II. Intrusion Detection Systems

Early IDS's operated at the host level, whereas contemporary systems tend to be network-based [2]. 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 and network traffic. Both host-based IDS's and network-based IDS's mainly employ two techniques: anomaly detection and misuse detection [2].

This paper focuses on presenting the analogy between human immune systems and network-based IDS's. Somayaji et al. [8] present more general principles and suggest various possibilities for a computer immune system. In contrast, this paper concentrates on the design of competent network-based IDS's, and analyses the several outstanding features of the human immune system with this specific problem in mind.

III. Requirements Of Network-Based IDS's

Functions that are required to design a competent network-based IDS's is distilled into seven points:

A. Adaptability

Computer system environments are not static. Therefore, the normal activities of networks and intrusions are also continuously changing according to this environment. it should be dynamically adjusted in order to detect dynamically changing network intrusions [4], [8].

B. Robustness

If intruders already know the existence of an IDS and can subvert it, then the effort to develop the IDS was futile. Therefore it should have multiple detection points, which are robust enough against the attack and any system faults on IDS's [4], [1].

C. *Configurability*

Hosts in a network environment are heterogeneous. They may have different security requirements. Along with hosts, there are routers, filters, DNS, firewalls, or various network services may have various security requirements. Therefore it should be able to configure itself easily to the local requirements of each host or each network component [4], [8].

D. *Scalability*

In the case of the monolithic IDS's, the audit trail collection procedure is distributed and its analysis is centralized [2]. However, it is very difficult to forward all audit data to a single IDS for analysis without losing the data. Even if it scales for all audit data correctly, it may cause severe network performance degradation. To avoid this it is necessary to achieve reliable scalability to gather and analyze the high-volume of audit data correctly from distributed hosts [4]

E. *Efficiency*

A single IDS is expected to perform data gathering, monitoring, data manipulation and decision making. It may impose a large overhead on a system and could place a particularly heavy burden on CPU and I/O, resulting in system and network performance degradation. it should be simple and lightweight enough to impose a low overhead on the monitored host systems and network [4], [7], [8].

F. *Extendibility*

When a new host that has a different format of audit and different operating system is added to an existing network environment then the data handled by the host is not simple to monitor in a consistent manner with existing IDS's. For this reason it should be easy to extend the scope of IDS monitoring by and for new hosts easily and simply regardless of operating systems [4], [8].

G. *Global Analysis*

Many network intrusions often exploit the multiple points of a network. Thus, from a single host, they might appear to be just a normal mistake. But if they are collectively monitored from multiple points, they can be identified as a single attack attempt. So as to detect network intrusions, it should collectively monitor multiple events generated on various hosts to integrate sufficient evidence and to identify the correlation between multiple events [4], [2].

IV. The Design Goals Of Network-Based IDS's

The requirements identified above can be used to derive three main design goals of an effective network based IDS. They are being distributed, self-organizing and lightweight.

A. *Distributed*

A distributed network-based IDS distributes its responsibilities to a number of components. Many independent intrusion detection processes monitors only a small aspect of the overall system. They work concurrently and co-operate with each other. If a network-based IDS is distributed, it will satisfy the following requirements.

1)Scalability: Distributed IDS's are more scalable than IDS's based on a single central server because audit data collection and its analysis takes place in the same place, at a monitored local host, the high volume of audit data is distributed amongst many local hosts.

2)Configurability: A single intrusion detection process can be simply tailored to local requirements of a specific host without considering the various requirements of other hosts.

3)Robustness: For a distributed network-based IDS, the failure of one local intrusion detection process does not affect overall IDS, though it causes minimal degradation in overall detection accuracy.

4)Extendibility: Intrusion detection processes are independent and thus existing processes do not need to be modified even when a new intrusion detection process is added. Therefore when a new host with different operating system is added to a network, it is easy to add a new intrusion detection processes on this new host.

B. *Self-Organization.*

The second goal is being self-organizing. Without a central controller having predefined information, a self-organizing network-based IDS automatically learns intrusion signatures which are previously unknown and/or distributed. This is achieved through the interaction with changing network environments, various security requirements and other intrusion detection processes. If a network-based IDS is self-organizing, it will satisfy the following requirements.

1)Adaptability: It is highly adaptive because there is no need for manual update of its intrusion signatures.

2)Global analysis: The overall intrusion detection system simply provides the global analysis. This is because it is self organizing from the interactions among a large number of various intrusion detection processes.

C. *Lightweight*

The third design goal is being lightweight. A lightweight network-based IDS does not impose a large overhead on a system or place a heavy burden on CPU and I/O. If a network-based IDS is lightweight, it will satisfy the last requirement.

1)Efficiency: By placing minimal work on each component of the IDS, the main jobs that should be performed by local hosts and networks are not adversely affected by the monitoring.

V. OVERVIEW OF HUMAN IMMUNE SYSTEMS

[12] The overall human immune system is implemented through the interactions between a large number of different types of innate and acquired cells rather than the function of one particular human organism. From a large number of different cells, lymphocytes (white blood cells), play a central role. Their main mechanism is distinguishing self cells, which are the cells of human body, from non-self cells, which are dangerous foreign cells. Each lymphocyte is specialized in reacting to a limited number of structurally related harmful foreign cells, known as antigens. Lymphocytes have the specific binding areas, called receptors, which have complementary shapes to the determinants of antigens, called epitopes. A specific antigen is recognized by its epitopes binding to lymphocyte antibody receptors.

Lymphocytes are classified into two main types: B-cells and T-cells. B-cells are antibody secreting cells and T-cells kill antigens or help or suppress the development of B-cells. Both B-cells and T-cells have their own unique genetic structures. Both B-cells and T-cells are expressed by several chains of DNA (gene libraries) and each chain has a variable domain and a constant domain. The genes in a variable domain are highly variable from one to another and this determines the specific binding area to antigens. The genes in the constant domain are invariable and show various biological effects when B-cell antibody receptors bind to antigen epitopes. B-cells and T-cells are developed in the bone marrow and the thymus respectively. At the bone marrow and the thymus, several gene libraries uniquely corresponding to domains of B-cells and T-cells contain the candidate genes to express B-cell and T-cell receptors. A specific receptor is generated by selecting gene segments randomly from gene libraries and joining them. Furthermore, in order to generate diverse receptors, they adopt a progressive series of genetic operators during their development processes [11].These include gene rearrangements, choosing different joining sites, somatic mutation, class switching and others (the details of these genetic operators are presented in [10]).Gene Library(DNA) Pre-Detectors (maturing B-Cells) Pre-Detectors (maturing T-Cells) Negative Selection Negative Selection B-Cells T-Cells Bone Marrow Thymus.[12]

Development of B-cells and T-cells and Clonal selection .Before leaving the bone marrow and the thymus, maturing B-cells and T-cells have to pass the last test, negative selection. In B-cell and T-cell development process, totally new cell receptors can be generated via various genetic operators. Therefore, it leaves the possibility for randomly generated receptors to bind to self cell epitopes. To prevent this, when maturing B-cells and T-cells bind to self cells circulating through the bone marrow and the thymus, they are killed instead of being released into a body. Figure. 1 (left) shows the development of B-cells and T-cells in the bone marrow and the thymus. Mature B-cells and T-cells that pass the negative selection are released from the bone marrow and thymus. Both B-cells and T-cells continuously circulate around the body in the blood and encounter antigens for activation and evolution. The antibodies of B-cells, which recognize harmful antigens by binding to them, are activated directly or indirectly. When B-cell antibody receptors bind to antigen epitopes with strong affinity above a threshold, they are directly activated. On the other hand, B-cell antibody receptors can bind to antigen epitopes with weak affinity below a threshold. In this case, B-cells need the help of T-cells and Major-Histocompatibility Complex (MHC) molecules to be activated. MHC molecules have two important functions to help B-cell activation. Firstly, they bind to the fragments of antigens specially hidden inside cells, (not visible on the cell surface) and secondly, they transport these fragments to the B-cell surface. When B-cell antibody receptors bind to antigen epitopes with weak affinity, MHC molecules try to find some hidden antigen inside cells. When MHC molecules find them, they transport them on the surface of B-cells. The receptors of T-cells are genetically structured to recognize the MHC molecule on the B-cell surface. Thus, T-cells can bind to MHC molecules on B-cell surfaces. When the T-cell binds to MHC molecule with strong affinity, it sends a chemical signal to the B-cell which allows it to activate, grow and differentiate [11].What makes the T-cells determine the B-cell activation? One major difference between B-cells and T-cells is that only B-cells perform somatic mutation, which is a very high rate of mutation, to increase its diversity when they are developed in the bone marrow. Hence, Bcells have more various and new receptors that T-cells. In addition, the thymus is centrally located while the bone marrow is distributed. Thus, most of self cells pass through the thymus and hence the negative selection in the thymus is more reliable than that in the bone marrow. Therefore, the final decision of B-cell activation with weak affinity is made by the T-cells. With or without the assistance of T-cells, B-cells are activated and this activation is immediately followed by clonal selection. The activated B-cells are divided into a number of clones that have the same antigen-binding properties as parent B-cells or mutated antigen-binding properties. On the other hand, if any antigen cannot activate B-cells within a limited time, they rapidly die off. Therefore, based on the existing antigens, only the fittest B-cell antibodies survive. Because antigens constantly change, the efficiency of detection is maintained by the evolution of B-cell antibodies via clonal selection. Furthermore, when antigens activate B-cells, they produce memory cells for the reoccurrence of same antigens in the future. Because of these memory cells, antigens that have been identified previously are detected much quicker (known as the secondary response). Immune systems let antigens and anti-antibodies compete to bind to antibodies and the winning anti-antibodies can suppress binding between antigen and antibody. The inhibition of idiotope antibody against antigen contributes to regulate an appropriate level of immune responses. Immunologist, Jern, proposed immune network theory [5], [9] based on understanding the role of the idiotope antibody. He views immune systems as a functional network of lymphocytes and the network at any moment has the dynamic state of internal interactions of antibodies and antigens. The continuous chain of differentiation by antigens and suppression by idiotope antibodies can form a large-scaled network. When this network finally reaches the equilibrium status between suppression and stimulation, it determines the overall immune system.[12]

VI. Human Immune System Features For Network-Based Ids's

By performing a careful analysis of the complex capabilities of human immune systems summarized above, it is possible to identify several significant features for network-based intrusion detection. Upon investigation, it becomes clear that specific features can act together in order to satisfy each of the three design goals of competent network-based IDS's: being distributed, self-organizing and lightweight.

A. *Distributed*

The human immune system is distributed. The following mechanisms allow the human immune system to detect antigens in a truly distributed way.

1)Immune Network: The human immune system is implemented through the interactions between a large number of different types of cells. Instead of employing a central coordinator, human immune systems sustain the appropriate level of immune responses by maintaining the equilibrium status between antibody suppression and activation using idiotope antibodies [5], [9].

2)Unique Antibody Sets: The human immune system generates various groups of antibodies to detect different antigens. Its evolution mechanism through natural selection of gene libraries and clonal selection maintains a number of different sets of antibodies. Therefore, each antibody set is unique and independent. These properties do not require any central coordinator and they allow the human immune system to detect antigens in a local antibody level [8].

B. *Self-Organization*

The overall immune response is composed of three evolutionary stages: gene library evolution generating effective antibody, negative selection eliminating inappropriate antibodies and clonal selection cloning well-performing antibodies. These three stages are self-organizing rather than being directed by a central organ or predefined information.

1)Gene Library Evolution: Antibodies recognize antigens by the complementary properties that only antigens, not self cells, show. Thus, some knowledge of antigen properties is required to generate competent antibodies. The human immune system learns this knowledge by its evolution over time and hence provides us with efficient and 'knowledge rich'

DNA. Because of this evolutionary self-organization process, our gene libraries act as archives of information on how to detect commonly observed antigens [10].

2)Negative Selection: As the second stage, this eliminates inappropriate and immature antibodies, which bind to self. The important constraint that the immune system has to satisfy is not to attack self cells. Instead of having any global information about self cells, this constraint satisfaction is performed in the thymus and bone marrow by presenting self cells, and removing any antibodies which attack these cells [4], [3].

3)Clonal Selection: As the third stage, this process clones antibodies performing well. In contrast, antibodies performing badly die off after a given life time. Thus, according to currently existing antigens, only the fittest antibodies survive. Similarly, instead of having the predefined information about specific antigens, it self-organizes the fittest antibodies by interacting with the currently existing antigens [3], [10].

C. *Lightweight*

The human immune system is lightweight. The following mechanisms allow it to be lightweight and are focused on three ideas:

- How a vast number of antigens can be detected with a smaller number or antibodies
- How the known antigen information can be reused efficiently and
- How numerous antibodies can be generated with a limited number of genes. Approximate binding, memory cells and gene expression provide the answers to these questions respectively.

1)Memory Cells: Memory cells store the genetic information of previously detected antigen epitopes and respond efficiently and quickly when they meet the same antigens in the future [8], [10]. Because memory cells have a longer life span than ordinary antibodies, they retain immunity without the need to create the same antibodies again.

2)Approximate Binding: The immune response activates when the affinity of antibody and antigen binding is above a certain threshold. In other words, a single antibody can detect any number of antigens as long as their affinity is above the threshold. This approximate binding contributes to increase the generality of immune systems [1].

3)Gene Expression: The immune system maintains antibody diversity in order to ensure the effective detection of a wide range of antigens. In an antibody development process, known as gene expression, several genetic mechanisms are employed to generate diverse antibodies from the gene libraries. The main idea of these mechanisms is that a vast number of new antibodies can be generated from new combinations of gene segments in the gene libraries [3], [10].

In summary, this analysis shows that the human immune system is distributed through its immune network and unique antibody sets. It is self-organizing because of the three evolutionary processes of gene library evolution, negative selection and clonal selection. It is lightweight because of the generality of approximate binding and gene expression, and the efficiency of memory cells. Since the human immune system is distributed, self-organizing and lightweight, it clearly fulfils the design goals for network-based intrusion detection systems. Perhaps most importantly, the mechanisms used by human immune systems satisfy the three goals in an elegant and highly optimized way and this motivates future research harnessing such processes. Because of this study, it is thought that the application of computer immune systems to network-based intrusion detection is likely to provide significant benefits over other approaches.

VII. Conclusion And Future Work

This paper has investigated network-based IDS's and provided a set of general requirements for them by a careful examination of the literature. Based on these requirements, three principal design goals were identified. After sketching the simplified human immune system, their salient features that can contribute to build a competent network-based intrusion detection system were analyzed. This analysis show that the human immune system is equipped with a number of sophisticated mechanisms, which satisfy the three identified design goals. Consequently, the design of a novel network-based IDS based on the human immune system is promising for future network-based IDS's.

One more application that can be made using an artificial immune system. Let there be two-robots. Both the robots programmed such that they move along two near concentric tracks. The outer robot (referred to as the Helper) is capable of guiding the inner robot (referred to as the Learner) back to its track in case of misalignment. If the Learner gets misaligned at a particular spot, it initially attempts to recover its path. It sends a message to the Helper if it is unable to do so. The Helper in turn, realigns itself to a position from where it can guide the panicked Learner to revert to its original track and then resume normal functioning.

References

[1] Forrest, S.; Hofmeyr, S; Somayaji, A, 1997, "Computer Immunology", *Communications of the ACM,* Vol.40, No.10,pp.88-96.

[2] Mykerjee, B.; Heberlein, L. T.; Levitt, K. N., 1994, "Network Intrusion Detection", *IEEE Network,* Vol.8, No.3,pp.26-41.

[3] Paul, W. E., 1993, "The Immune System: An Introduction", in *Fundamental Immunology* 3rd Ed., W. E. Paul (Ed),Raven Press Ltd.

[4] Balasubramaniyan, J. S. et al., 1997, "Software Agents for Intrusion Detection", Department of Computer Sciences,Purdue University, available at http://www.cs.purdue.edu/coast/coast-library.html

[5] Dasgupta, D.; Attoch-Okine, N., 1997, "Immunity-Based Systems: A Survey", *Proceeding of the IEEE International Conference on Systems, Man and Cybernetics,* Orlando, October.

[6] Kephart, J. O., 1994, "A Biologically Inspired Immune System for Computers", *Artificial Life IV, Proceeding of the Fourth International Workshop on the Synthesis and Simulation of Living Systems,* pp.130-139.

[7] Forrest, S. et al., 1996, "A Sense of Self for Unix processes", *Proceedings of 1996 IEEE Symposium on Computer Security and Privacy,* Los Alamos, CA, pp.120-128.

[8] Somayaji, A.; Hofmeyr, S.; Forrest, S., 1997, " Principles of a Computer Immune System", *Proceeding of New Security Paradigms Workshop, Langdale, Cumbria,* pp.75-82.

[9] Farmer, J. D.; Packard, N. H.; Perelson, A. S., 1986, "The Immune System, Adaptation and Machine Learning", *Physica* 22D, pp.182-204.

[10] Tizard, I. R., 1995, *Immunology: Introduction,* 4th Ed, Saunders College Publishing.

[11] E-rich & K knight : Artificial intelligence , tata mc graw hill.

[12] The Human Immune System and Network Intrusion Detection paper by Jungwon Kim and Peter Bentley.