Intrusion Detection System using Artificial Immune System

Inadyuti Dutt
Dept. of Computer Applications,
B. P. Poddar Institute of Management & Technology, MAKAUT,
137, Poddar Vihar, Kolkata, West Bengal-700052

Samarjeet Borah
Dept. of Computer Applications,
Sikkim Manipal Institute of Technology, SMU
SMIT, Majitar, Sikkim-737132

Indrakanta Maitra
Sr. System Analyst,
B. P. Poddar Institute of Management & Technology, MAKAUT,
137, Poddar Vihar, Kolkata, West Bengal-700052

ABSTRACT
Nature and natural organisms have always inspired researchers and scientists for solving real world issues. And Computer security is no exception. Artificial Immune System inspired from natural Immune System works efficiently for detecting intrusion in a network. Two layers of defenses: innate system and adaptive system are implemented in this proposed methodology where the innate system mimics the natural Innate Immune System to form the first line of defense. The adaptive system imitates the Adaptive Immune System by incorporating the T-cell and B-cell defensive mechanisms. The results exhibit that the proposed methodology works efficiently for detecting intrusion after inducing malicious attacks on the network system.

Keywords
Intrusion Detection System (IDS), Immune System (IS), Artificial Immune System (AIS)

1. INTRODUCTION
The extensive growth of Internet and increasing availability of tools and tricks for attacking networks has enabled security providers to rethink on the security of our computer systems and data. The government, military and commercial bodies have increased their reliance on Internet technologies for their day to day business in the past few decades. This has created innumerable challenges for combating external attacks. Such attacks that are external to these bodies are deliberate in action against data, software or hardware and that destroy, degrade, disrupt or deny access to a network computer system are called Cyber attacks. In an attempt to guard against the unknown cyber attacks, much effort has been given in researching and developing Intrusion Detection System (IDS), which tries to filter out the cyber attacks from the network traffic. IDS are software tools meant specifically for strengthening the security of information and communication systems. An IDS dynamically monitors logs and network traffic, applies detection algorithms to identify intrusions in a network. Intrusion detection is based on the assumption that intrusive activities are noticeably different from normal system activities and thus are identifiable. Intrusion detection is not used to replace prevention-based techniques such as authentication and access control; instead, it is intended to complement existing security measures. Intrusion detection system is therefore considered as a second line of defense for computer network systems to detect actions that bypass the security monitoring and control component of the system.

The majority of IDS can be classified as host based i.e. HIDS or network based i.e. NIDS whilst the former collects data from a host (for example, system calls, system and application log files), whereas the latter collects the data directly from the network. Both HIDS and NIDS can use two different approaches in detecting attacks. Anomaly detection is one of the two approaches which consider that an intrusion will always reflect some deviations from normal patterns. It can be based on the current resources being utilized by the system and capturing the deviation from the normal consumption. Another approach is referred to as the misuse detection, which is based on the knowledge of system vulnerabilities and known attack patterns. It is concerned with finding intruders who are attempting to break into a system by exploiting some known vulnerability. Ideally, a system security administrator should be aware of all the known vulnerabilities and eliminate them.

Nature and its organisms have always inspired researchers and scientists for solving problems of real world. And computer security is no exception. Nature inspired approaches are being applied to various domains of computer security including cryptology, secure protocol design and intrusion detection. Artificial Immune System (AIS) uniquely mimics the Natural Immune System (IS). Forrest et. al. [1] first introduced this concept of computer immunology based on the principles of Human Immune System (HIS).

HIS provides an excellent metaphor for applying Artificial Immune Systems (AIS) to the intrusion detection problem. Similar to human beings, who are exposed to trillions of previously unknown antigens of varying types, the computer networks are exposed to multitude of attacks. HIS silently tackles all these pathogens without us realizing that our immune system is constantly defending against the intruders. HIS is highly distributed, dynamic, self organized, adaptive and robust information processing systems. These are the appropriate type of qualities that are required for the security of modern computing environments. The proposed methodology would focus on these attributes of HIS to develop IDS which would be distributed, dynamic, self-organized and adaptive by nature. The proposed system would try to learn the unknown vulnerabilities and make use of those to prepare an adaptive system that would organize itself automatically when such vulnerabilities actually take place.

2. LITERATURE REVIEW
The review of literatures related to immune-inspired based IDS depict that human immune system has innate coherence with the intrusion detection system in computer system. As mentioned by D. D. DasGupta in his paper [2], the usage of immunological principles in designing resource consumption

based IDS, where there are different levels of abstraction being created in order to get user-level, system-level, packet-level and network-level information respectively. He designed a general framework and used multi-agents to communicate among these levels of abstraction. His works are strictly based on LINUX environment and used system calls to collect the resource usages. In his other paper with U. Aickelein [9], he utilized the approach of multi-agents called Manager Agents and Monitor agents for accessing the different levels of abstraction. Monitor agents are going to communicate within the user, packet and network –levels to determine any abnormal behaviors in the system whilst the Manager agent is going to manage and convey the alarm. U. Aickelein in his paper [3], [4], he used special repository servers for maintaining the input data arriving from the system calls based on resource usage and used immunological principles for matching of antigens or unknown behaviors. In the rest of the papers [5], [6], [7], [8], [10] [11], and [12] some immunological principles like self, non-self tissues or danger theory etc. have been utilized in order to detect the anomaly in the behavior of the system.

The above literature reviews reveal that the researchers have not used the collective information based on resource consumption of the system as well as the users' specific activities. Studies explore that there exists some limitations in handling intrusion detection using the current techniques. The techniques utilized are mostly multi-agent based that are time consuming while communicating among themselves.
There are some open issues in solving intrusion detection problem with respect to the huge data of resource consumption. The techniques are primarily based on recognition of self, non-self tissues or passing danger signals. So it is inevitable to explore this limitation in order to provide more information in solving the issues using other immune-based approaches.

3. BACKGROUND

3.1 Human Immune System

Human immune system (HIS) has its origin in the study of how the body protects itself against infectious diseases caused by micro-organisms such as bacteria, viruses, protozoa, fungi and other parasitic organisms. The first initial physical barrier is the skin which sets itself as a strong protector from these micro-organisms. Skin protects by secreting sweat, tear and saliva that try to trap these organisms.

In its most complex forms, the immune system consists of two branches: the innate system that uses certain strategies to provide rapid, general response when alerted by certain invasion of infectious organisms. It becomes the first-line of defense, and the adaptive immune system takes time to develop specific defensive responses to such invasion of bacteria or viruses. The adaptive immune system has a persistent immune memory which recollects in order conserving the responses specific to such invasion.

3.1.1 First Line of Defense-Innate Immune System

Mast cells and basophils are innate immune cells that when activated secrete histamine, which can be an important inflammatory mediator produced in response to initial tissue damage as a result of infection. Mast cells are tissue resident while basophils reside in the blood stream. Innate Immune System consists of both cellular and humoral elements. The cellular elements are phagocytes (neutrophils) and macrophages. They generate the initial responses to foreign bodies before being engulfed by the phagocytes. The natural killer (NK) cells are another type of innate cells that have the ability to detect and target intracellular infection of body cells by viruses.

Immune System	
Innate Immune System	Adaptive Immune System
i. Macrophages ii. Basophils iii. Eosinophils iv. Neutrophils v. Natural Killer Cells vi. Dendritic Cells	i. T-Cells a. CD4+ T-cell b. CD8+ T-cell ii. B-cells iii. Natural Killer T-cells

Fig. 1 Immune System

3.1.2 Second Line of Defense-Adaptive Immune System

All immune cells originate in the bone marrow but an important set of immune cells (T lymphocytes) undergo maturation in an organ known as the Thymus. Thymus and bone marrow become the Primary Lymphoid tissues. The Secondary Lymphoid tissues are the lymph nodes, spleen, Mucose-Associated Lymphoid Tissues (MALT). They are the important sites for generating Adaptive Immune responses.

The key component to adaptive immune system is the lymphocyte. There are several categories to lymphocytes namely T-lymphocytes (T-cells) and B-lymphocytes (B-cells). Both the types originate from the bone marrow. T-cells mature in the thymus whilst the B-cells mature in the bone marrow. Both these lymphocytes have certain unique characteristics. Both have the ability to recognize a specific molecular target and also have the ability to differentiate between body (self-tissue) and foreign (non-self tissue) tissues respectively. The self-tissues are identified and weeded-out. T-cells on maturation can take two different forms namely Helper T-cells (CD4 +T-cells) and Cytotoxic T-cells (CD8 +T-cells).

The adaptive immune system can also be categorized into Cellular and Humoral Adaptive Responses. The cellular adaptive T-cells are directed towards pathogens that have colonized body cells or that have become malignant. The humoral adaptive response B-cells target pathogens (antigens) that are free in the bloodstream or present at mucosal surfaces. Adaptive immunity uses many kinds of receptor to coordinate its activities. T-cells carry T-cell receptors (TCR) and B-cells carry B-cell receptors (BCR).

4. PROPOSED METHODOLOGY

The proposed methodology in Fig. 2 is based on the concepts of Artificial Immune System which mimics the natural Human Immune System. The proposed system employs HIDS based approach where the IDS is placed on each host-based computer. Each such host-based computer has two layers of defense namely, the first line of defense and second line of defense imitating the original HIS. The first line of defense is referred to as Innate Immune Detection System and the second line of defense as Adaptive Immune Detection System.

The Innate Immune Detection system monitors each file that arrives on the host computer. The file can be of any extension and of any type. As soon as a file arrives into the host computer, the Innate Immune Detection System tries to detect whether the file is of the extension type .exe, .bat or not. If the

file extension type does not belong to the above type then the file is accepted if and only if its extension is of type .doc, .pdf. The files with anonymous extensions or types .exe or .bat are immediately send as alarming ones to the second line of defense i.e. Adaptive Immune Detection System by changing their file extensions to .doc or .txt type. This again would help to minimize the vulnerability of the file to multiply or reside in the hard disk.

The Adaptive Immune System is invoked when it receives an alarm signal for a file with its extension being changed from its original one by the Innate Immune Detection System. A B-Cell_Activation_Module is invoked in order to detect the file's vulnerability. The file with its properties like original name, extension, time of arrival, is stored in a repository table called B-cell table. The file is parsed to detect the vulnerability. If the file is found to be malicious after parsing then a T-Cell_Activation_Module is invoked. The T-cell Activation Module mimics the original T-cells of HIS. It deletes the file by targeting the malicious file. Even if the file is eradicated from the host computer, the repository table, B-cell table keeps track of all the files that have arrived in the host-based system.

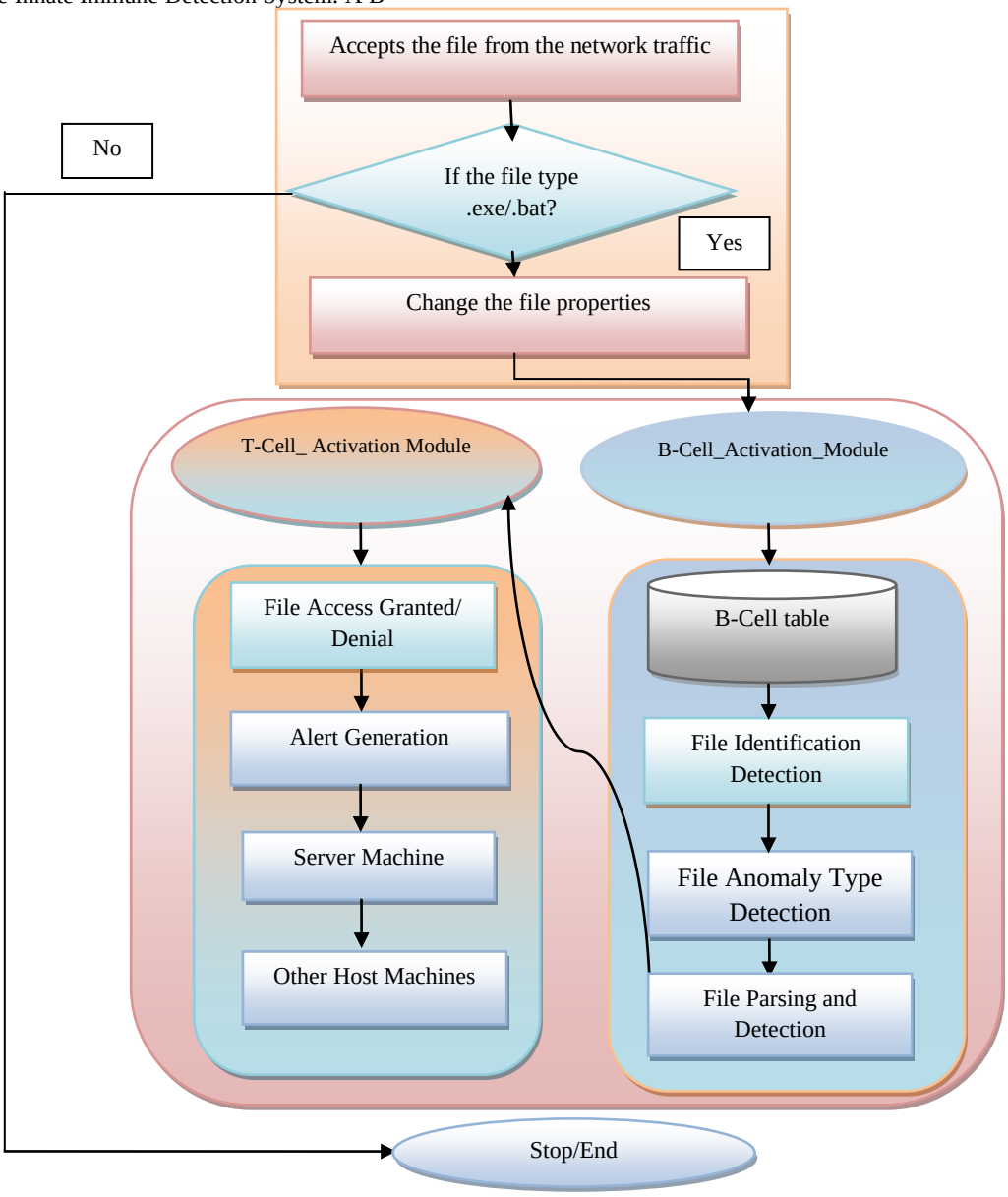

Fig. 2 Schematic diagram of the Proposed Methodology

4.1 Proposed Algorithm

The Proposed Methodology is based on the following algorithm:

Step 1: Accept the file from the network traffic (the file can be any of the type .pdf, .txt, .exe, .bat, .text .vbs etc.)

Step 2: Detect the file type using the file extension for determining whether the file is valid /invalid file using enlisted invalid file extension

Step 3: Accept the file name with extension and store it to the master table with the following details:

Step 3.1: Original file name, original path, date of arrival, modify, date of modify, delete, date of delete, date of modified, status, duplicate filename, duplicate path, extension.

Step 3.2: Accept the filename with extension of rename and store it to B-Cell table if the file is considered as invalid in Step2. The fields Duplicate Filename, Duplicate path and extension would be created.

Step 4: Accept the file in B-cell table if the file in Step 3.1 is found to be invalid and malicious. The B-Cell table has the following entries: original file name, original path, date of arrival, date of modified, status, duplicate filename, duplicate path, extension. The entries are saved for further analysis.

Step 5: Accept the file from the B-Cell table for parsing.

Step 6: Detect the file type: virus or worm or malware.

Step 7: If file type is found to be malicious then

Step 7.1: File access denied

Step 7.2: Else, File access granted

Step 8: If file type is found to be malicious in Step 7.1 then alert is generated to the Server.

Step 9: Generate alert from the Server to the other host machines and Stop.

5. RESULT AND ANALYSIS

The proposed methodology was implemented in Java with 40 host Pentium V machines connected to a Server. When the malicious attacks by virus or worms or malware were induced the proposed methodology exhibits the following results as in Fig. 3. The malicious files are induced ranging from 1-1000 as inputs on the X-axis where as the detection rate in % is on the Y-axis. The malicious files are induced in the simulated environment for detection of vulnerability of the proposed methodology.

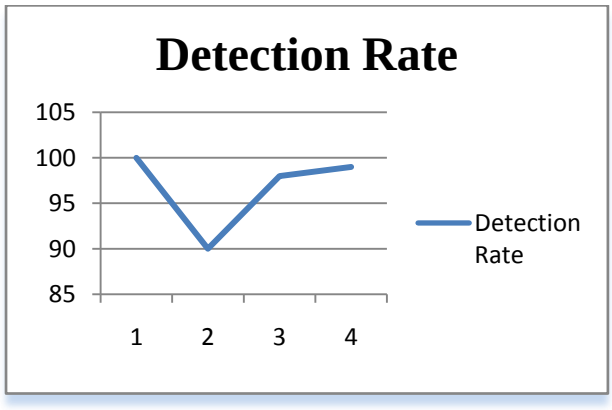

Fig. 3 Depicts Detection Rate (%) w. r. t Number of Malicious Inputs

Table 1 Tabular details of the Number of Inputs vs. Detection Rate

Number of Input files at a time (X-Axis)	Detection Rate (Y-Axis) %
1	100
10	90
100	98
1000	99

6. CONCLUSION

The proposed methodology attempts to incorporate the immune system using Innate System and Adaptive System. The files are initially accepted from the host-based system and differentiated if being .exe or .bat files. As soon as such files are detected, their properties are changed and sent to the Adaptive Immune System. The T-cell and B-cell defensive mechanisms are used to detect the vulnerability of the files. The results exhibit that the proposed methodology works efficiently for detecting intrusion after inducing malicious attacks on the host-based system.

7. REFERENCES

[1] S Forrest, SA Hofmeyr, A Somayaji, 1997 Computer Immunology Communications of the ACM, Vol. 40, Issue 10, 88-96.

[2] D. Dasgupta, 2007 “Immuno-inspired autonomic system for cyber defense”, Information Security, Tech. Rep., Volume 12, Issue 4, pp. 235-241.

[3] Jamie Twycross, Uwe Aickelin, 2008 “An Immune-Inspired Approach to Anomaly Detection”, Handbook of Research on Information Security and Assurance, Information Science Reference, Hershey, New York.

[4] Jamie Twycross, Uwe Aickelin, Amanda Whitbrook, 2010 “Detecting Anomalous Process Behaviour using Second Generation Artificial Immune”, International Journal for Unconventional Comp., Volume 6, Issue 3–4, pp. 301–326.

[5] Tarek S. Sobha, Wael M. Mostafab, 2011 “A Cooperative Immunological Approach for Detecting Network Anomaly”, Applied Soft Computing, , Volume 11, pp. 1275–1283.

[6] Carlos A. Catania, Carlos García Garino, 2012 “Automatic Network Intrusion Detection: Current techniques and open issues”, Elsevier, Computers & Electrical Engineering, Volume 38, Issue 5, Pages 1062–1072.

[7] Chung-Ming Ou, 2012 “Host-based Intrusion Detection Systems adapted from Agent-based Artificial Immune Systems”, Neurocomputing, Volume 88, pp. 78–86.

[8] Manoj Rameshchandra Thakur, Sugata Sanyal, 2012 “A Multi-Dimensional approach towards Intrusion Detection System”, International Journal of Computer Applications (0975 – 888), Volume 48, Issue 5.

[9] U. Aickelin, D. Dasgupta, 2014 Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques, Artificial Immune Systems, Chapter 13, pp. 1-29.

[10] Man-Ki Yoon, Sibin Mohany, Jaesik Choiz, and Lui Sha, 2015 “Memory Heat Map: Anomaly Detection in Real-Time Embedded Systems Using Memory Behavior”, DAC ’15, ACM Digital Library, ISBN: 978-1-4503-3520-1.

[11] Farhoud Hosseinpour, Sureswaran Ramadass, Andrew Meulenberg, Payam Vahdani Amoli and Zahra Moghaddasi, 2013 “Distributed Agent Based Model for Intrusion Detection System Based on Artificial Immune System”, International Journal of Digital Content Technology and its Applications (JDCTA), Volume 7, Number 9.

[12] Meng-Hui Chen, pei-ChannChang, Jheng-LongWu, 2016 “A Population-based Incremental Learning Approach with Artificial Immune System for Network Intrusion Detection”, Engineering Applications of Artificial Intelligence”, (51)171–181.