IMMUNE BASED EVENT-INCIDENT MODEL FOR INTRUSION DETECTION SYSTEMS: 
A NATURE INSPIRED APPROACH TO SECURE COMPUTING

A thesis submitted

to Kent State University in partial
fulfillment of the requirements for the
degree of Masters of Science.

by

Swetha Vasudevan

August, 2007
Thesis written by
Swetha Vasudevan
B.E, Madras University, India, 2001
M.Sc, Salford University, England, 2004
M.S, Kent State University, U.S.A, 2007

Approved by

Dr. Michael Rothstein, Advisor, Department of Computer Science

Dr. Robert Walker, Chair, Department of Computer Science

Dr. Jerry Feezel, Dean, College of Arts and Sciences
# TABLE OF CONTENTS

1. LIST OF FIGURES .................................................................................. iv

2. ACKNOWLEDGEMENTS ........................................................................ vi

3. INTRODUCTION ................................................................................... 1

4. WORKING OF AN IDEAL INTRUSION DETECTION SYSTEM ................. 4

5. UNDERSTANDING HUMAN IMMUNE SYSTEM ....................................... 11

6. REVIEW OF EXISTING LITERATURE .................................................... 20

7. THE DANGER THEORY ....................................................................... 26

8. APPLICATION OF DANGER THEORY TO INTRUSION DETECTION .......... 36

9. IMMUNE BASED EVENT-INCIDENT MODEL FOR INTRUSION DETECTION 39

10. CONCLUSION ..................................................................................... 58

11. BIBLIOGRAPHY ................................................................................ 62
LIST OF FIGURES

FIG 1. DEFENSE IN DEPTH ................................................................. 5
FIG 2. TYPES OF INTRUSION DETECTION ARCHITECTURE .................. 6
FIG 3. IDEAL INTRUSION DETECTION SYSTEM ................................. 9
FIG 4. IMMUNE SYSTEM ..................................................................12
FIG 5. STRUCTURE OF LYMPHOCYTE .............................................13
FIG 6. NEGATIVE SELECTION .........................................................16
FIG 7. LYMPHOCYTE-ANTIGEN BONDING .......................................17
FIG 8. CO-STIMULATION ...............................................................19
FIG 9. APOPTISIS ...........................................................................27
FIG 10. NECROSIS ..........................................................................28
FIG 11. BURNET’S ONE SIGNAL MODEL ...........................................29
FIG 12. BRETSCHER AND COHN’S TWO SIGNAL MODEL ...............30
FIG 13. LAFFERTY AND CUNNINGHAM’S MODEL…………………………………….31

FIG 14. JANEWAY’S MODEL………………………………………………………32

FIG 15. MATZINGER’S DANGER SIGNAL MODEL……………………………….34

FIG 16. ARCHITECTURE OF EVENT-INCIDENT MODEL………………………….44

FIG 17. COMPARISON……………………………………………………………….46

FIG 18. ATTACK RESISTANCE……………………………………………………….51

FIG 19. TOPOLOGY……………………………………………………………………..55

FIG 20. SCENARIO………………………………………………………………………..57
ACKNOWLEDGEMENTS

This thesis would not have been possible without the constant support and encouragement I received from my academic advisor, Dr. Michael Rothstein. From the first vague proposal of this topic to later queries on focus and connection, he was always eager to entertain my ideas and help me solve conceptual difficulties. I sincerely thank him for his consistent efforts and true desire to keep me on track.

I would also like to thank my committee members, Dr. Johnnie Baker and Dr. Austin Melton for serving in my defense committee despite their overwhelmingly busy schedule.

Special thanks to the staff at the Kent State Department of Computer Science, in particular Marcy Curtiss who helped and encouraged me every step of the way during my time in the department.

Finally, I would like to express my deepest gratitude to my parents. Their support and unwavering confidence in my ability helped me achieve my academic dreams.
CHAPTER 1

1. Introduction

Why are Computer Security scientists interested in the Human Immune System? Even though Computer Security and our Immune System are totally unrelated, conceptionally they work the very same. Intrusion Detection Systems (IDS) try to detect active sequence of related events that deliberately try to cause harm. And so does our Immune System which helps our body ward off harmful microorganisms. The primary function of our Immune System is to be able to distinguish between our body’s own cells from that of foreign cells and neutralize potentially pathogenic organisms. They achieve this by employing Defense in Depth mechanism that establishes defense at three unique levels namely, the External line of defense, Innate or non-specific immunity and Adaptive or specific immunity [11]. The multiple layers work in parallel, hence, even if one layer fails, the other layer takes over to halt the spread of infection. An infection will have to break through each layer without being detected in the process.

Today we have numerous types of Intrusion Detection Systems in the market. None of which has lived up to its expectations. A prevailing challenge with Computer Security Systems is being able to accurately determine the difference between normal system activity (Event) and potentially harmful activity (Incident). With the threat domain growing exponentially large, it is an urgent necessity to have a system that is authentic in its ability to detect and respond to attacks.
Researchers hope that studying the functioning of Human Immune System and understanding the mechanisms these systems employ could give new ways of looking at issues that are yet to be solved in the field of Computer Security. The *uniqueness*, *distributed detection and subsequent elimination of foreign activities*, *anomaly detection*, *self replication*, *learning*, *memory*, *adaptability*, *and a methodology of protecting itself from attacks* are some of the few characteristics of Immune System that fascinate computer scientists [9]. Implementing such properties in a single system would be any researcher’s dream as it would result in a system that is accurate, scalable, robust, flexible and resilient to subversion.

The idea of using immunological principles in Computer Security has been in existence for over a decade. Today it has branched out into four distinct areas, each of which endeavors to emulate the magnificent properties of defense system that our body naturally exhibits to protect itself from foreign invaders:

- Techniques implementing the concept of *Negative Selection* (Forrest et al.)
- Automated Computer Virus Detection by employing *Decoy Programs* (Jeffrey Kephart)
- Multi-level Intrusion Detection mechanism using mobile *Immune Agents* (Dipankar Dasgupta)
- Methods that exploit the concept of *Danger Theory for Immunology* (Polly Matzinger).

This thesis will utilize the concept of Danger Theory to implement an effective Intrusion Detection System. Prior to reviewing each of the above techniques, as a first step, the working of an Ideal Intrusion Detection System is examined to understand its characteristics. This helps determine what is to be accomplished in the proposed model. Following this, the author devotes a separate chapter to let readers acquaint themselves with the functioning of the Human Immune System. This is done to enable the
metaphorical understanding of Human Immune System as Intrusion Detection System. A separate section is devoted to review and address the deficiencies of the existing immunological principles in Computer Security. Since the proposed model utilizes the concept of Danger Theory, the author explains this theory and examines its relevance to Intrusion Detection in the next chapter. The conceptual view and overall infrastructure of the proposed model is then presented.

The model is built on the strategy of Defense in Depth. The monitoring is done by mobile, attack resistant immune agents at four unique levels namely the network level, the application level, the host/system level and the protocol level. To enable exchange of critical Intrusion Detection information between components, peer-peer architecture is implemented. The agents act as communication interface and communicate with other agents and modules. They constantly listen for messages from peer agents, receive it, update their Knowledge Database and also act as a sender to transmit data and information to neighboring agents and components. These mobile agents are programmed to be self-learning so as to deal with novel threats. For reduction of false positives and effective correlation of analysis results and corresponding alerts, the concept of Danger Theory is adopted. The proposed model is designed to be accurate in its ability to distinguish between an event and an incident, scalable, flexible and adaptable.
2. Working of an Ideal Intrusion Detection System

Any attempt to compromise the Integrity, Confidentiality and Availability of a resource can be categorized as Intrusion. Although the majority of anomalous activities takes place from within the organization, the security policies that are put in place protect the inside network from outside forces. We need a technique that would eliminate both Internal and External penetrators. External penetrators are unauthorized users trying to gain access to the system, whereas internal penetrators are authorized users who misrepresent the privileges accorded to them [27]. An effective way of eliminating both internal and external penetration is by using the concept of Intrusion Detection. It is the process of gathering and analyzing information to identify, assess, and report the presence of intrusive activity in a system. [Fig 1] illustrates the approach of Defense in Depth or Elastic Defense. Adopting multiple layers of defense strengthens the overall security infrastructure as it seeks to mitigate the risk of one layer of defense being compromised.

The architecture is one of the most critical considerations in Intrusion Detection. An efficient architecture is one in which each host machine, its internal components and process performs its function in an effective and coordinated manner, resulting in efficient information processing, analysis and timely response. Three types of tiered architectures can be adopted for Intrusion Detection System design purposes namely, Single-tiered, Multi-tiered and Peer-Peer architectures [21].
As name suggests, in a Single-tiered architecture, the components of an Intrusion Detection System collect, process and analyze data themselves instead of passing the output they collect to another set of components [Fig 2]. Although the architecture is relatively simple and inexpensive, it is not efficient as the components are not aware of each other and operate independently. Any updates occurring at one component would not automatically update or affect other components.
Fig 2. Types of Intrusion Detection Architectures
In a Multi-tiered architecture however, Intrusion Detection System involves multiple components that pass information to each other [Fig 2]. Each component performs a specific task and then passes the output obtained to another component. Most of today’s Intrusion Detection Systems utilize Multi-tiered architecture due to its efficiency and depth of analysis. The main drawbacks of this architecture include increased cost and complexity involved in designing them.

As opposed to Multi-tiered architecture, which takes raw data, processes it and sends output to higher-order component, the Peer to Peer architecture usually involves exchanging critical Intrusion Detection information between peer components, each of which performs identical functions [Fig 2]. This architecture is simple to design although lacks sophistication due to absence of specialized components. It is therefore important to carefully choose the best architecture depending on the needs of the organization [21].

An Intrusion Detection System has three major tasks to perform. It has to first continuously monitor the system it is trying to protect. It should then try to detect if any deviation from the normal system behavior has taken place. If so, it has to immediately notify the system security officer that an attack is underway. After the notification, appropriate action is taken to abort the attack from causing irreversible damage to the system [2]. These tasks are accomplished by four components namely; agent, director, alert analyzer and system security officer [Fig 3]. A Multi-tiered Intrusion Detection architecture is employed. Agents are modules that reside in the protected computer system. They continuously monitor the system for the presence of malicious activity. They collect information from their source, segregate the information based on their relevance, preprocess the collected relevant information into a specific format, and finally send the information to the director. Ideally the agents would be flexible to the changing system and user behavior. The activities of the agent would take place in the background. At the same time it would not act like a black box. The operations and working of an agent would be examinable from the outside.
The director on receiving the information checks whether the information is enough to make its decision. If it needs more information, it requests the agent to collect further information and send it back. Also it can request the agent to preprocess the information that it received in a different format. After receiving additional information from the agent, it checks redundancy of information and eliminates unnecessary data. It then sends this information to the Intrusion Detection System Knowledge Database and the analysis engine.

Intrusion Detection System Knowledge Database has information about all possible attacks. The analysis engine compares the information that it received from the director with the information stored in the Intrusion Detection System Knowledge Database and analyses it. It uses multiple techniques to determine if a potential attack is underway. Rather than just looking for attacks, it has the capability of recognizing suspicious activities too. The comparison results are then sent to the director. In an ideal situation, the results of the analysis engine can not be modified to force false negatives to occur. In other words the analysis engine would resist subversion attacks. It would also be very difficult to fool even with the full knowledge of its working. Often when comparison takes place by the analysis engine, it produces false positives and false negatives. However, an ideal Intrusion Detection System is resistant to such attacks. The director after receiving information from the analysis engine sends an alert to alert storage unit.

The alert storage is a unit where all past alerts that has been issued is stored. The director then sends the information to the alert analyzer unit. This unit analyzes the attacks and categorizes them based on the threat level. It also looks for new worms and tries to spot violations. It now issues an alert to the system security officer. The alerts are issued based on the threat level and are issued in a timely manner. Attacks that could be potentially damaging are taken care of first. The system security officer then communicates with the attack response module. This module in turn communicates with the response policy
module. Counter measures are then taken based on the level of threat. If multiple attacks have same level of threat, the security officer would have the ability to handle that.

The Intrusion Detection System module also has a fault detection and isolation unit. This unit tries to look for faulty processors in the system. If it finds any, it would immediately isolate the faulty unit from the rest of the working modules so that Intrusion Detection System can continue monitoring without halting its operations. This is done so that the Intrusion Detection System Knowledge Database does not have to be rebuilt time and
again. For some reason if the system shuts down because of presence of a malicious activity or faulty unit, it would have the ability to provide graceful degradation of service. It should try to keep as much information as possible. Under ideal circumstances, the overall working of the Intrusion Detection System module would impose minimum overhead on the system it is trying to protect to avoid slowing the system down.
CHAPTER 3

3. Understanding Human Immune System

Our body is constantly under attack from various foreign organisms such as bacteria, viruses and infectious diseases. No human can live long without the Immune System that prevents pathogenic entry into our body. Human Immune System is a complex, intricate network of organs which produces cells and tissues that participate in the immune response. Our Immune System adopts the concept of defense in depth which is a multi-level protection that works at three unique levels [Fig 4]. At the first level, our body creates a surface barrier that prevents foreign pathogens from directly entering the body. This forms the highest level or the first line of defense. One good example of such external barrier would be our skin. Our skin is made of tightly packed cells that inhibit water entry and bacterial growth because of its slightly acidic nature. Once the pathogen successfully evades the external barrier and enters the body, the second layer of defense, the Innate Immune System takes over [9, 11].

Innate Immunity is inbuilt and offers basic defense against the infection. They prevent the entry of harmful microorganisms and bacteria into our tissues. However if they do successfully gain entry, they eliminate them before the occurrence of disease. Innate Immune System is non-specific (generic) in nature and does not exhibit memory of invaded pathogens or provide lasting protective immunity to the species. In most cases, it helps slow down bacterial growth and the spread of infection. However, if the Innate Immune System fails to protect the body against the threat the pathogen poses, it
immediately activates the third and the most powerful level of defense known as the Adaptive Immune System [9, 11].

Adaptive Immune System exhibits five fascinating properties such as immunological recognition, self / non-self discrimination, specialization, learning and memory. Unlike Innate Immune System, Adaptive Immune System is highly specific in nature. It has the capacity to sense and recognize an organism as foreign (non-self) as distinct from the body’s own self cells; respond to a pathogen by production of needed antibodies; remember the pathogen to which it responded to mount a much stronger and focused attack the next time the same pathogen is encountered; distinguish between pathogens and eliminate pathogens from our body [9, 11].

Fig 4. Immune System (Defense in Depth)
The major players in the Human Immune System are Lymphocytes, Antigens, Antibodies, Antigen Presenting Cells (APC) and Major Histocompatibility complex (MHC). Lymphocytes form the heart of the Immune System [Fig 5]. They are special class of white blood cells with the capacity to distinguish self cells from foreign cells [28]. Lymphocytes have specific binding areas or molecular structure on their surface known as receptors. The binding areas enable them to bind to a particular antigen. Antigens are molecular proteins that can be found on the surface of living substances such as foreign blood cells, bacteria, viruses and some non-living substances such as toxins, harmful chemicals, drugs etc. The lymphocyte receptors have complementary shapes to the localized region on the surface of antigens known as epitopes. Antigens are recognized by their epitopes binding to lymphocyte antibody receptors. Antibodies are a specific type of proteins produced by the lymphocytes in response to the invading foreign organisms. Each antibody is unique and defends our body against a specific kind of antigen [9].

Fig 5. Structure of a Lymphocyte [9]

There are two classes of lymphocytes that take part in the immune response. They are B cells and T cells respectively. B cells develop and mature in the bone marrow of our body. When the body gets invaded by foreign pathogens, B cells react in response by secreting antibodies [6]. Hence B cells are often referred to as ‘antibody secreting white
blood cells’. Antibodies are ‘Y’ shaped proteins with 2 ends. One end is called the variable end and the other is referred as the constant end. The variable end of the antibody always changes and is designed to sense and recognize matching foreign cells (antigens). The constant end of the antibody on the other hand remains the same and binds itself to the outside surface of the white blood cells to determine an appropriate mechanism to destroy a particular antigen. Antibodies secreted by each of the B cells has different variable end and are hence different for every invading pathogen. This is to ensure that our body produces enough antibodies to protect itself from a diverse range of antigens. So, any time our body encounters a foreign organism, the variable end of the antibody produced by the B cells gets itself bonded to the outside surface of the invading bacteria. Our body now knows that the invaded organism is dangerous and that it needs to be destroyed. B cells are capable of immunological memory. When a previously invaded pathogen enters our body for the second time, B cells that produce the right antibody are alert and are ready for it. This time however, they can react quickly and more efficiently than the first time. These memory cells remind the Immune System of all the antigen patterns that have been encountered by our body in the past [29].

T cells unlike B cells play a crucial role in cell mediated immunity. Cell mediated immunity is an immune response that does not involve secreting antibodies but rather involves eliminating infected self cells before it releases harmful toxins and viruses that can infect other cells. They can be easily distinguished from other lymphocyte types such as B cells by the presence of special binding area called T-cell receptors. T cells are made in the bone marrow just like B cells; however, the immature T cells travel to the thymus where they become fully mature. There are two classes of T cells, namely, Helper T cells and Killer T cells. Helper T cells as name suggests, help and stimulate B cells to produce antibodies. In addition to this, they also help Killer cells to develop and mature. Killer cells has been assigned the duty of destroying virus infected self cells to prevent virus reproduction and infection spread [30].
Only fully developed B and T cells have the capacity to participate in the immune response and fight infection. The main function of the Adaptive Immune System is to be able to accurately distinguish between self cells from non-self cells. During the birth of B and T cells it is possible that the developing B and T cells categorize proteins in the body as harmful antigens. When this happens the B and T cell receptors stick themselves to the self cell epitopes. Such cells must not circulate around the body as it will create an immune response against our own body cells and tissues. They are filtered by the process known as Negative Selection [Fig 6].

Negative Selection ensures that our body only produces lymphocytes that could bind themselves to the epitopes of the antigens not our own body cells (Self cells). During this phase all B and T cells that bind to self cells are killed instantly by a process known as Apoptosis instead of being released into the body. Programmed Cell Death or Apoptosis is a process which motivates such cells to commit suicide or kill themselves. As opposed to Apoptosis, Necrosis is a process where death of cell occurs due to cell injury or infection. Unlike Apoptosis, cells that die as a result of Necrosis, release harmful toxins into the body and infect other cells in the body [31, 32].
Once the B and T cells pass the negative selection test, they continuously circulate around the body looking for pathogens. The antibodies secreted by B cells needs to be activated either directly or indirectly to be able to recognize harmful antigens. The B cell antibody receptors chemically bind to the antigen epitopes with a certain affinity. The probability of bond forming increases as the affinity increases. In order to ensure that the Immune System can sense and recognize as many antigen patterns as possible, an exact match between the antigen epitopes and lymphocyte receptors is not necessary [Fig 7]. A match can still occur if the given number of contiguous features complements one another. In other words, if the set of contiguous features is long enough, it can make a match. The number of features required to bind before a match can be made is known as the Affinity Threshold. If a B cell antibody receptors bind themselves to an antigen epitope above a certain threshold, they get directly activated [10].
However, if a B cell antibody receptor binds to an antigen epitope with weak affinity, it seeks the help of T cells and MHC. MHC helps B cells by performing two important functions. It first binds itself to the hidden fragments of antigens that are not visible on the cell surface. It then transports the antigen infected cell to the surface of the B cell. In order for the T cell receptors to recognize the antigens, the antigens needs to first be processed. This is done by Antigen Presenting Cell usually B cells. APC can present the antigen in a form that T cell can recognize. The processed antigen is then stuck into the special molecule inside the MHC. Now the APC displays this foreign antigen complexed with MHC for T cell receptors to recognize. Now the T cells can easily bind to the MHC molecule on the surface of the B cell. When T cell receptor binds to the MHC molecule with a strong affinity, it immediately sends a chemical signal to the B cell which lets it get activated indirectly.
Once the B cells get activated, they divide into groups with same antigen binding properties. B cells which are unable to get activated by any antigen within a certain amount of time die off. In the Immune System, only the lymphocytes with high affinity for antigens are encouraged to reproduce. This helps to create a large number of lymphocytes that are specific a particular antigen pattern [10].

It is necessary to accurately distinguish self from non-self. Sometimes, it is possible that randomly generated receptors could bind themselves to self cells and initiate an autoimmune response (where the Immune System starts attacking the body). This is where the helper T cells come into the picture again. Most of the self cells circulate through the Thymus where all immature T cells mature. Hence, all maturing T cells get exposed to the self proteins in the Thymus region. During this time, if any of the maturing T cells binds itself to the self proteins circulating the Thymus, it will be aborted. Only the T cells that survive the maturing process will leave the Thymus. This helps the Immune System to ensure that the all mature T cells are tolerant to the self cells. This process is also known as centralized tolerance, as all immature T cells are tolerized in a central location (Thymus).

However, ensuring B cell tolerance is a harder. B cells get matured in the lymph nodes which are distributed all over the body unlike the Thymus. In this case the helper T cells provide a distributed censoring of B cells via a process known as co-stimulation [Fig 8]. In order to become activated B cells must receive co-stimulation in the form of 2 signals.
Signal one occurs when the B cell receptors bind to the antigen epitopes beyond the affinity threshold. Signal Two is provided to the B cell by the helper T cell. Helper T cell will only provide signal two if it recognizes the pathogen the B cell has captured. B cells ingest the pathogen and break them into protein fragments known as peptides. These peptides are then presented to the T helper cells. If the T cells do not recognize the pathogen, it concludes that what B cell has captured is a self protein and hence does not co-stimulate B cell by providing signal 2. If the B cell does not receive the signal 2 from the helper T cells, they die by Apoptosis (Programmed Cell Death). This helps B cells to adapt to a specific pathogen and eliminate infection [9, 11].
CHAPTER 4

4. Review of Existing Literature

This section gives a brief review of the existing immunological concepts in Intrusion Detection.

a. Negative Selection

The principle of Negative Selection introduced by Forrest et al. works on the basis of Adaptive Immunity which utilizes the law of discrimination to distinguish between self and non-self. In Negative Selection, all immune detectors that have the capacity to recognize self elements are eliminated. Only the detectors that can identify foreign elements are kept. ‘Self’ here refers to the normal behavior of the system. The self patterns are stored in the Knowledge Database. ‘Non-self’ refers to the anomalous system activities. The system generates random patterns and compares it with the self pattern in the database. If there is a match, the randomly generated pattern fails to become a detector and is hence discarded. If the randomly generated patterns do not match with the self, it becomes a detector and monitors system for intrusive activities. Negative Selection mechanism is robust and excellent in detecting changes to a system. However it falls short in other areas. Scalability is an important issue here. As the system that needs to be protected expands, it becomes a necessity to generate a large number of detectors so as to guarantee adequate coverage and accurate levels of detection. Such large number of detectors can become unmanageable. It is also prone to false alarms since a strict
distinction of self and non-self exists. Further, the algorithm fails to address the issue of attacks against the detector patterns [3, 4, 5].

b. Automated Computer Virus Detection

In Kephart’s approach known viruses are detected by the computer coded sequences and unknown viruses are detected by their unusual behavior. The approach uses the concept of decoy programs. These are programs that are placed in strategic locations in the memory in order to capture sample of viruses. The decoy programs are then examined from time to time for any changes that could have occurred. Modified decoy programs indicate the presence of virus. A process known as signature extractor is used to extract the signature of the virus captured by the decoy programs. The signature is then entered into the database for future reference. This approach effectively integrates the concept of Innate Immunity (detecting previously known viruses/misuse detection) and Adaptive Immunity (deriving a prescription by implementing decoy programs/anomaly detection). Further, this approach attempts to halt the viral spread in a networked environment by the use of ‘kill signals’. An infected machine sends signal to the neighboring machine letting it know that it has been infected along with signature and repair information that might be of use for detecting and eliminating infection. Once the neighbor gets this message, it immediately updates its database. If this neighbor also gets infected, it sends the signal to its neighbors. However, if it does not get infected it does not propagate the kill signal although it does update its database. One main issue that tend to prevail is that, it has not been determined under what circumstances should a kill signal be propagated to neighboring nodes and for how long. Erroneous kill signals could very well lead to false alarms and subversion attacks [33].
c. Mobile Agent based Intrusion Detection

Dasgupta proposed an alternative immune-based approach for Intrusion Detection. This approach employed mobile agents that monitored the system at multiple levels (process, system, packet and user levels) in a hierarchical manner. The warning signals produced by these intelligent agents are communicated in an effective manner to neighboring nodes. Decisions are made based on the collective warning signals received. Being a distributed system, compromising one node does not result in the take down of the whole system. The mobile agents exhibit excellent properties such as learning and memory with the capability of recognizing both known and unknown threats. However, this approach fails to provide any mechanism to ensure that the mobile agents themselves are corruption free. This fact rattles the whole foundation on which this architecture works as it is heavily dependent on these mobile agents and their decision-making process to ensure an attack free system [7].

Ensuring attack resistance capabilities in mobile agents is difficult. Mell and McLarnon proposed a methodology that provides attack resistance capabilities in mobile agents. The approach is not immune-based. According to the authors, attack resistance is achieved through five stages. In the first stage, the location of the mobile agents is randomized. This makes it hard for the attackers to pinpoint the exact location of the agent. The second stage involves eliminating single points of failure by removing centralized directory services. To achieve this, the authors propose a concept of ‘buddy’ relationship where an agent buddies up with another agent to constantly notify each other of their whereabouts. Following this stage, the buddies of agents that are killed by attackers will take evasive action to avoid the same fate since their location may also be known to the attacker. The fourth stage involves resurrecting killed agent. If an active agent gets killed, they will automatically be substituted by backup agents. The backup agents are constantly updated with state information of the active agent. In case of the
death of an active agent, the backup agents negotiate among themselves and pick a candidate for replacing the killed agent. In the final stage, the backup agents reestablish broken communication links so as to reconnect into the IDS hierarchy [23].

There is a growing concern that applying mobile agent technology to Intrusion Detection adds new vulnerabilities in the system. Ensuring security of agents that continuously circulate the system looking for anomalies has become very challenging and complex. This area is still very much in its infancy.

d. Emergence of Danger Theory

According to the self / non-self philosophy, ‘any entity that originates from the organism will not trigger an immune response where as an entity that originates outside of the organism will trigger an immune reaction’. Today many immunologists are debating the legitimacy of this concept. Renowned immunologist Polly Matzinger questions the principle both conceptually and experimentally. She argues that many ‘self’ components do induce immune reactions, and many ‘non-self’ components do not. She explains the falsity of the principle by stating that there is no immune reaction whatsoever to the foreign bacteria in the gut or in the food we eat although both originate from outside the organism. Further she points out that defining ‘self’ is difficult. As human body changes over time, ‘self’ does too. Once individuals attain puberty, their body changes drastically, but our Immune System does not reject that change [24]. This leaves a real gap in the self / non-self theory which states ‘non-self’ is learned during early fetal life. In addition, certain types of tumors fought against our Immune System, autoimmunity (where the body fails to recognize its own constituent parts as ‘self’ resulting in immune response against its own cells and tissues) and successful implants further strengthen Matzinger’s argument. After years of having trouble with the self/non-self view,
Matzinger finally abandoned it and introduced a new concept known as the Danger Theory.

According to the Danger Theory, immune response is triggered when diseased cells that die unnaturally induces alarm signals. Alarm or danger signals are actually harmful toxins released by cells in distress. The propagating signals create a danger zone around itself and only antibodies within the range begin immune reaction. This clearly lands us in a viewpoint where it is not the 'foreignness’ of an entity that triggers the immune response but the actual level of danger itself [16, 19].

Security experts are now trying to implement this fascinating philosophy in Intrusion Detection Systems. In the context of Intrusion Detection System, danger signals would be interpreted as unusual memory usages, access of unauthorized files, intruder presence, inappropriate disk activity and so forth. The system would be programmed in such a way that these signals would be induced after only limited infection to minimize the extent of damage to the system. The generated alarm signals would then be correlated with the Intrusion Detection System alerts. According to the Danger Theory concept, the alerts are classified into two main categories namely apoptotic (good/normal) and necrotic (bad/abnormal) respectively. Inappropriate threshold setting of the sensors is the root cause for false alarm production. It is believed that a proper balancing between the two types of alerts would result in an optimum sensor setting of the intrusion signature or anomaly threshold. The successful correlation of these alerts would then lead to the construction of intrusion scenario.

The Danger Theory based principle of danger zone establishment can also be applied to Intrusion Detection System. When the Intrusion Detection System has clear indications of an intrusive activity, it can activate the detection sensors that are spatially, temporally or logically near the original sensor emitting the danger signal. The propagation of danger signals would enable the system to immune itself from attacks [14, 15]. The concept of
Danger Theory and its relevance to Intrusion Detection Systems will be discussed in detail in the following sections.
5. The Danger Theory

Every cell in our body has a defined life cycle, a beginning and an end. There are two ways in which cells die. They either get killed (accidentally) by harmful pathogens or get triggered to commit suicide. Our body needs to get rid of unnecessary aged and damaged cells. This is natural and a part of development process and takes place in an orderly fashion. The process of deliberate life relinquishment of a cell is termed as *Programmed Cell Death or Apoptosis*. During the process of cell suicide, the cell that is activated to commit suicide has to be prevented from releasing its contents which could result in a local inflammatory reaction. In the case of Apoptosis, this is prevented as the cell that undergoes suicide sends out signals to nearby scavenger cells (*phagocytes*). These scavenger cells engulf and absorb the dying cell so that the cell membrane is still intact and not being leaky [Fig 9].
But not all cells die by Apoptosis. Sometimes cells die accidentally due to injury, infection, toxin exposure and inflammation. This death is not programmed or organized. The disorderly death does not send signals which inform the nearby phagocytes to engulf the injured cells. This makes it hard for the cleanup cells (phagocytes) to locate and digest the cells that die due to Necrosis. The injury received by the cells, compromises cell membrane which stores special digestive enzymes [Fig 10]. The membrane prevents the digestive enzymes inside from destroying the cell itself. Necrosis is almost always accompanied by release in these special digestive enzymes which accelerates unorganized chemical chain reaction. Unlike Apoptosis, Necrosis results in the release of harmful chemicals that damages their uninfected cells [34].
Danger Theory was built on the concept that proposed that the intracellular contents that were released by damaged cells were actually a form of danger signal that alerted the nearby Antigen Presenting Cells (APC) and activated them. However, only cells that die due to Necrosis would send out alarm signals, healthy cells and cells that die due to Programmed Cell Death should not. Rather than responding to foreignness in the case of self / non-self theory, the Immune System responds to actual danger. The danger model was built on existing immune signal model that utilizes the concept of self / non-self discrimination.

The immune signal model was first proposed by Burnet [35] [Fig 11]. B cell receptors found on the surface of B cells has the capacity to bind to antigens which triggers the production of antigen specific antibody. On the other hand, T killer cell receptors recognize and bind to infected cells that bear a specific foreign antigen. According to Burnet, this interaction of B and T killer cell receptors with antigens and infected cells initiates a stimulation that is adequate to activate the B and T killer cells and commence an immune response. This stimulation was termed as Signal One.
The One signal model was later extended by Bretscher and Cohn who included T-helper cell and signal two to create an associative recognition model or the Two Signal Model [Fig 12]. Since one signal model faced the issue of autoimmunity, Bretscher and Cohn suggested that if the initiation of immune response required the activation of two cells that could recognize specificities on the same antigen, the chances of commencing an immune response against self cells would be slim to none. T helper cell would double check to see whether the antigen that the B cell has captured is in fact foreign as opposed to self. When a B cell receives signal one, it passes this signal to T helper and awaits its response. If T helper cell also recognizes the antigen as foreign, it would send signal two confirming B cell to initiate an immune response. Absence of signal two (help signal) from T helper would lead to the death of B cell [18].
The two signal model version was later modified by Lafferty and Cunningham [Fig 13]. They proposed that T helper cells themselves need to be activated before they can send the help signal to B cells. They included Antigen Presenting Cell and a new signal termed as co-stimulation. Just like B cells, the recognition of antigens by T helper cells alone is not enough to fully activate them. Once the T helper receives signal one from B cells, they wait for signal two or co-stimulation from Antigen Presenting Cells. Antigen Presenting Cells can be B cells, Macrophages or dendritic cells. Since B cells are the least efficient Antigen Presenting Cells and do not express co-stimulatory molecules, either Macrophages or dendritic cells involve in the activation of T helper cells as they are professional and highly efficient. After the receipt of co-stimulatory signal by the Antigen Presenting Cell, the T helper cell gets fully activated. If the T cell receives signal one from B cell without the second signal from APC, the T cell is assumed to have recognized a self antigen and is forced to die by the process of Apoptosis [36].
Fig 13. Lafferty and Cunningham’s Model [36]

T helper cells receive the co-stimulation signal from the Antigen Presenting Cells which does not have the capability to distinguish self antigens from foreign antigens. The APCs simply presents any antigen it picks up. But the co-stimulation signal should only be issued to T helper cells for foreign antigens. If the immune response commences solely by the signals presented from the APCs, then autoimmunity should be a more common problem with the functioning of the Immune System. This fact led Charlie Janeway introduce the concept of ‘infectious non-self’ [Fig 14]. He suggested that the Immune System actually distinguishes between ‘infectious nonself’ and ‘noninfectious self’ [17]. He proposed that Antigen Presenting Cells has proteins on their surface known as the Pattern Recognition Receptors (PRR) that can identify patterns of molecules that are only associated with microbial pathogens like bacteria and cellular stress. Activation of APC takes place when the carbohydrates on the surface of the endosomal pathogens bind itself to the PRR (signal ‘0’). The APC then processes the pathogen (antigen) and then re-expresses it in a form that T cell can recognize. As described earlier, the APC expresses
these antigens as peptides in Major Histocompatibility Complex (MHC) molecules and regulate the co-stimulatory molecules to activate the T helper cells and commence an immune response that is now clearly targeted at foreign antigens. However, Janeway’s concept falls short when it comes to explaining immune response to tumors (attacks against self), successful transplants where the Immune System fails to attack non-self, absence of immune reaction to the food we eat or the bacteria in the gut and other dysfunctions that may lead to autoimmune diseases.

Matzinger’s danger model proposes that the immune signals that control the Immune System are endogenous or originate from within the organism (tissue or cell) and not exogenous in nature. They are actually alarm signals (signal ‘0’) that arise when cells either get infected or stressed out. The signals that these injured cells send out activate the nearby APCs. The model is similar to Janeway’s extended self/non-self discrimination model as it also assumes that the APCs themselves needs to be activated first by some

Fig 14. Janeway’s Model ( Priming of Antigen Presenting Cells) [17]
mechanism before they can activate T cells. The random signal one that the APC sends to T helper alone would not suffice. The APC needs to be fully activated and the key point is that healthy cells do not send alarm signals to the neighboring APCs. They send calming signals informing Antigen Presenting Cells of the absence of danger. However, damaged cells, the cells that die due to necrosis and stressed cells should send danger signals (signal ‘0’) and activate their local APCs. In contrast, Programmed Cell Death corpses would send ‘engulf and eat me’ signals to near by scavenger cells such as phagocytes. The fully activated APC would now activate the T helper cell by sending signal two. Matzinger also proposes that the efficacy of T helper cells can be extended by routing signal two through antigen presenting cells. So the antigen seen by the helper cell and the killer cell need not be the same. However, it is mandatory that both antigens need to be presented by the same APC therefore allowing T helper to prime many more T killer cells. Thus the danger model completely eliminates the idea of immune response to foreignness but rather introduces a fascinating concept of response to danger. The model looks for a source or an instance of risk rather than initiating an immune response as a result of a presence of foreign entity that does no harm to the body [15, 16, 19] [Fig 15].
Danger Theory operates by applying three laws of Lymphotics. According to law one, T cells require two signals to become fully activated. Signal one comes from two sources, B cell (T cell receptor binding to the MHC/peptide on the surface of B cell) and the random signal from antigen presenting cell. Signal two comes from fully activated APC (co-stimulation). So the T cells will be activated only if they receive signal one and two. If T cells receive signal one in the absence of signal two, they will be destroyed by PCD. T cells will ignore signal two in the absence of signal one. The second law of lymphotics states that, the co-stimulatory signals will only be provided to T cells by the professional APCs such as macrophages and dendritic cells for virgin and experienced T cells. B cells can also act as APCs; however, they can only re-stimulate experienced T cells but does not have the capacity to activate virgin T cells. The third law of lymphotics states that, once the T cells get activated, the activated effector stage only lasts for a short span of time. During this time, T cells do not require co-stimulation or signal two and can function to help activate B cells and kill antigens simply by receiving signal one.
However, once the effector stage becomes inactive, the T cells either die or revert back to resting stage. From the resting stage, they can be brought back to active stage only by the combination of both signal one and co-stimulation signal two [15].

In the following section, the application of Danger Theory to Intrusion Detection will be discussed.
6. Application of Danger Theory to Intrusion Detection

A Danger Theory based Intrusion Detection System would focus on accurate classification, correlation and balancing of alerts. The alerts are classified as either Apoptotic or Necrotic. Apoptotic alerts refer to ‘Programmed Cell Death’ in Human Immune System. With respect to Intrusion Detection System, they would correspond to low-level alerts which would sometimes result from authorized activities. Such alerts could also represent conditions and behavior that would have to exist for an intrusion to succeed (prerequisite for an attack). Necrotic alerts on the other hand refer to ‘untimely death of cell or group of cells as a result of injury, disease or other pathologic state’. With respect to Intrusion Detection System, they would correspond to alerts generated as a result of successful intrusions (consequence of a successful attack). Since the main function of a computer system is to provide information (flow of data from source to destination over a communication channel), the task of a security system is to restrict access to this information to legitimate users only. Necrotic alerts would be raised only when there is an evidence of interruption, interception, modification or fabrication of the normal information flow. Unlike Apoptotic alerts, Necrotic alerts would relate to severe system damage caused by a successful attack [14, 22].

The current Intrusion Detection Systems faces the problem of excessive false alarm production due to the inaccuracy in differentiating attacks from incidents. Danger Theory based Intrusion Detection approach would be based on the pre and post
conditions (prerequisites and consequences) of individual attacks to correlate related alerts together to develop an intrusion scenario. The key idea of DT based alert correlation is that consequences (Necrotic alerts) of certain attacks can be used as prerequisites (Apoptotic alerts) for other attacks. One alert can be linked to another alert when the consequence of the first alert is a necessary prerequisite for the second alert and this prerequisite was not satisfied before the first alert. The attack that corresponds to the first alert renders the attack that corresponds to the second alert possible. This enables the IDS to conclude that the first attack has been carried out so as to accommodate the second attack. The advantage of this methodology is that it is only necessary to specify properties such as prerequisites and consequences for individual attacks. From this, it is possible to automatically create complex intrusion scenarios by correlation analysis. In addition to this, it is also possible to deduce missing attacks. For example, two series of alerts are received by the system; the second series of alerts cannot be linked to the first one. This is because, the prerequisite required for the first attack of the second series of alerts is not any of the consequences of the last attack of the first series. Therefore, there is a missing alert that could effectively link the first and the second series of alerts. If attack signatures exists that could connect the first and the second series of alerts, the DT based IDS might decide to correlate the first series of alerts to the second series even though there is no corresponding match between prerequisite and consequence. Hence, it becomes entirely feasible to correlate missing alerts when key types of alerts are received [12, 14, 15].

DT based Intrusion Detection is promising in the sense that it would minimize the alerts arising due to false positives as it can quantify the degree of alert detection by appropriately tuning the intrusion signatures and anomaly thresholds. Striking an acceptable balance between Apoptotic and Necrotic alerts would enable the IDS to identify the most suitable intrusion signatures and anomaly threshold setting. Similar to memory cells that has the capability of remembering invaded disease causing organism in Human Immune System, intrusion signatures and thresholds are continuously redefined
as new attacks invade the system so as to familiarize the IDS of the latest threats. This significantly increases the detection and accuracy rate of the IDS [14, 15].

According to the concept of Danger Theory, danger signals (intracellular contents released by the damaged cells) propagate to nearby Antigen Presenting Cells and alerts them that a foreign pathogen has invaded the body. This alert propagation establishes a danger zone around the infected area. Similar theory can be implemented in Intrusion Detection Systems. When an intrusion detection sensor identifies the presence of unauthorized activity, it raises an alert. Danger alerts arising from one sensor can be transmitted to nearby sensors informing them of the intruder presence. The alerts can be either Apoptotic or Necrotic. After the proper alert correlation analysis, the IDS can effectively indicate the strength and the impact of the different intrusion scenarios possible. If the probability of an intrusion scenario is higher than the threshold set, the sensor can activate other sensors that are temporally or logically near this sensor emitting the danger signal. The activation of nearby sensors now establishes a danger zone around the area suspected to have the intruder presence. Although Danger Theory is a concept still in its infancy, researchers believe that it can be adopted to identify key types of Apoptotic and Necrotic alerts and understand the balance between them to effectively minimize false alarms [14, 15].
CHAPTER 7

7. Immune-based Event Incident Model for Intrusion Detection

a. Concept

The author’s concept is to create an Intrusion Detection System that in many aspects mimics what happens within the Human Immune System. Nature has gifted mankind with the Immune System that acts as a defense against microbes found in the environment surrounding us. Their function is to effectively and accurately differentiate our own cells from that of harmful foreign cells. When it comes to Intrusion Detection Systems however, characterizing the standard and non-standard behavior of a system can be very challenging. The normal working of the system is observed over a specific period of time to capture its properties. Anomalies can be detected if there are any deviations of the actual system activity from the categorized ‘safe’ behavior. Immune System has features of anomaly detection that allows it to define self and detect entities different from self (non-self).

However, a number of theories exist as to how exactly this distinction takes place within the Immune System. Recently, the one theory that has captured the attention of all immunologists is the Danger Theory. Danger Theory rightly pinpoints that all foreign substances that enter our body do not produce an immune response rather entities that deliberately causes harm or danger are the ones that effectively induces an immune response. Harmful toxins that get released during the process of unnatural cell death are
believed to produce ‘danger signals’. These signals are responsible for activating the leukocytes which are the main cells of the Immune System. Leukocytes are further classified as Phagocytes and Lymphocytes.

Phagocytes act as the first line of defense and help chew up invading harmful organisms. They are the part of Innate Immune System and are non-specific in nature. The non-specific detection and response of Innate Immune System can be closely compared to anomaly detection. In anomaly detection however, if intrusive activities are not exactly similar to the anomalous activities, we are left with two very interesting scenarios where anomalous activities that are not intrusive are categorized as intrusions (Events determined to be incidents) and intrusive activities that are not anomalous are categorized normal (Incidents determined to be events). The main issue is selecting an optimum threshold level so neither of the above two scenarios gets blown out of proportion. It is also equally important to properly select the features that need to be monitored. Again, the concept of Danger Theory can be utilized to set the appropriate threshold level. Any deviation from normal system activity that is positively determined as intrusion can be stored as a signature which in future could be used to identify an attack specific to that deviation.

Lymphocytes on the other hand help our body to remember and recognize organisms that have previously invaded our body. They are the part of Adaptive Immune System and are very specific in nature. Immune System also has features of misuse intrusion detection. The concept behind misuse intrusion detection is that all known attacks are stored as a signature pattern in the Knowledge Database. This way even minute variation of the same attack can be easily captured. Pattern matching, immunological memory of already invaded pathogens and specific and accurate recognition are some of the properties of Adaptive Immune System that can be compared to the properties of misuse detection. Effective synchronization of these two classes of cells protects our body against all types of diseases. The author’s Intrusion Detection System is built on this general philosophy.
This new methodology effectively considers the fact that the Human Immune System exhibits characteristics of both Anomaly and Misuse detection techniques. Danger Theory of immunology is used to address the issue of false alarms and alert correlation. The central pillar that holds the above concepts together is the use of attack resistant mobile agents also known as the **Intrusion Detection Squad** whose function closely resembles the function of leukocytes (Phagocytes and Lymphocytes) in Human Immune System. Leukocytes roam around our body looking for dangerous foreign organisms where as the mobile agents constantly look for anomalies and attacks within computer system. Without Leukocytes, Immune System can not exist and without the use of attack resistant mobile agents this methodology will not be feasible. Besides that, the proposed methodology has excellent learning and adaptive capability of the Human Immune System which makes it possible to effectively react to unknown attacks and also improve its response to recurring attacks.

**b. Architecture**

The proposed Intrusion Detection utilizes Peer-Peer architecture where critical Intrusion Detection information is shared between peer components performing identical functions. Monitoring takes place simultaneously by customized mobile agents at four levels namely the *Host level, Application level, Protocol level* and the *Network level*. At the host level, intrusions are identified by carefully monitoring and analyzing internals of the system such as state of the system (its stored information whether in Random Access Memory, in the file system or else where), system calls, application logs and other host based activities. At the application level, specialized agents are designed to monitor communication on application specific protocols. Since they are designed having specific application in mind, they are very accurate in detecting malicious activities for the application they try to protect. At the protocol level, intrusions are detected by monitoring and analyzing the dynamic behavior and state of the communication protocol between a connected device and the system it is protecting. Dynamic behavior is
monitored by enforcing the correct use of the protocol. At the network level, presence of malicious activity is identified by examining both the incoming and outgoing network traffic patterns between hosts.

The proposed *Event-Incident Model* utilizes an Intrusion Detection Squad to effectively carry out multi-level monitoring simultaneously [Fig 16]. The squad consists of 6 main players namely the **Assistant Patrol Agents**, **Incident Pattern Presenters**, **Correlator**, **Negotiator**, **Coordinator** and **Neutralizer**. The squad team works together with the **CIA** (*Confidentiality, Integrity and Availability*) **threshold unit**, **Peer information buffer**, **Knowledge Database** and **Anomaly Signature Converter** to accurately identify, categorize, memorize and protect itself from attacks. The Danger Theory based Event-Incident Model for Intrusion detection is a 6 phase process. Each phase denotes a distinct sequence of events which leads to the progression to the next phase.

1. **Recruitment: Phase One**

The **Coordinator** of the Intrusion Detection Squad is responsible for the recruitment of customized mobile agents. They generate two classes of agents for each of the four levels.

- **Assistant Patrol Agents (APA)**
- **Incident Pattern Presenters (IPP)**

**Assistant Patrol Agents** are carefully chosen by the process of **Positive Selection**. According to the immunological concept of Positive Selection, only T cells that have the ability to recognize self MHC are retained, the others die via apoptosis. This is because T cell receptors recognize antigens only in the context of MHC, hence the T cells must be tuned to recognize host MHC first. A similar approach is undertaken here. During Positive Selection, the Coordinator activates the **Random Profile Generator** unit to
produce random system profiles. Only Patrol agents that can recognize the system’s profiles are kept. However, agents that can not recognize the normal system profiles are deleted or rendered useless. The chosen Assistant Patrol Agents primarily roam in its home node within its specialized levels with particular intent and purpose. The agents check to see if the incoming traffic is normal by immediately activating the inbuilt Profile Deviation Detector unit which communicates with the Unit Two of the Knowledge Database. The Unit Two of the Knowledge Database has profiles of normal system behavior that have been observed over time. If the agent’s profile deviation detector finds significant deviation between the incoming activity and the stored profile, it concludes that the activity is abnormal and categorizes it as an incident. Assistant Patrol Agents only look for normal activity (event pattern) and any deviation is flagged suspicious.

Incident Pattern Presenters on the other hand are chosen by the process of Negative Selection. As discussed in chapter 3, the process of Negative Selection in immunology ensures that our body only produces lymphocytes (B and T cells) that could bind themselves to the epitopes of the antigens, not our own body cells. During this selection process, all B and T cells that bind to self-cells are eliminated by apoptosis and the others are retained. Incident Pattern Presenters are chosen utilizing a similar concept. During the negative selection, the Coordinator activates Random Pattern Generator unit to produce random incident patterns. Only agents that can successfully identify these patterns are kept. Agents that can not identify these patterns as incidents are deleted. IPP primarily roam in its home node within its specialized levels looking for known incident patterns. They check to see if an incoming traffic pattern is an incident or an event by activating the inbuilt Pattern Matcher unit which communicates with Unit Three of the Knowledge Database.
Fig 16. Architecture of Event-Incident Model
The unit three of the Knowledge Database has signature patterns of all known attacks. This unit is also constantly updated on discovery of new attack patterns. If the pattern matcher can successfully match the incoming traffic with pattern stored in the database it concludes that an attack is underway. Unlike APA, IPP only looks for abnormal or incident patterns.

2. Dispersal: Phase Two

Coordinator acts as the head of the Agent Team. Once the recruitment is completed, it is the Coordinator that decides when to send the agents for a neighborhood patrol. Agents are dispatched periodically to the neighboring nodes by the Coordinator. The Coordinator digitally signs the agent by encrypting it using a private key as a means to provide authentication of the associated agent. Following this, it sends it to a chosen neighbor. The neighboring node that receives the agent verifies to see if it is authentic by using the public key provided by the agent’s Coordinator or owner and decrypts the agent. This way the receiving node can protect its platform from corruption. After verifying the authenticity of the agents, the agents are allowed to migrate to its specific levels. The agents then quickly monitor the levels for suspicious behavior. Every agent’s home node has a Peer Information Buffer unit which is used to store critical attributes such as file size, permissions, modifications date of the node’s neighbors and perhaps also create checksum for the contents. Agents deliver their findings to the Coordinator (owner) which then processes the information and decides whether an action plan is required. While processing the agent information, it checks agent’s findings by communicating with the peer-information buffer to look for significant discrepancies. An action plan is taken based on the outcome of the check. It is the responsibility of the Coordinator to oversee all agent activity and inter-agent communication.
3. Alert Categorization and Intrusion Scenario Strength Analysis: Phase Three

The **Correlator/Negotiator Unit** is responsible for Alert Categorization and Intrusion Scenario Strength Analysis phase. Here, it is worth recalling Matzinger’s Danger Theory Model discussed in Chapter 5 [**Fig 15**].

![Danger Model](image)

**Fig 17. Comparison**

According to this model, the random signal ‘1’ (APC simply presents any antigen it picks up in the form of signal ‘1’) that the APC sends to T helper cells alone is not enough. The APC themselves need to be activated before they can send the co-simulation signal to T helper cells. Signals (signal ‘0’) send out by the cells in distress activate the nearby Antigen Presenting Cells. Following this the APC sends co-simulation signal ‘2’ to T helper cells and activates them. T helper would now send confirmation signal ‘2’ to B cells and activates them (B cell presents antigens to T helper and waits for confirmation.
signal ‘2’). Now T helper routes the signal ‘2’ through other APC’s which are responsible for propagating the danger signal. A similar methodology is used to implement the Event- Incident Model. A comparison between the two models is presented in [Fig 17]. The Assistant Patrol Agents sends signal ‘0’ to inform the Coordinator that it has discovered an abnormal activity in the system. Similarly the Incident Pattern Presenters sends signal ‘1’ informing the Coordinator of the existence of an incident pattern. The Coordinator will be activated if it receives either one or both the signals ‘0’ and ‘1’. The signals are a means of delivering the findings of the mobile agents to the Coordinator.

Although the Coordinator will get activated, it will not propagate danger signals just yet. The Coordinator will process the information received by checking with the Peer Information Buffer. After the processing, the Coordinator will send the processed information to the **Correlator/Negotiator Unit** and awaits the final confirmation signal ‘2’ which would positively indicate the presence of malicious activity in the system. The Correlator Negotiator unit receives the information from the Coordinator and then utilizes Danger Theory to analyze and categorize the alert information as either Necrotic or Apoptotic. This would help the Correlator determine the severity of the attack.

It is worth mentioning here that it is not the presence of any foreign activity that would raise an alarm but it is the actual measure of danger or abnormality that triggers it. For instance, system updates makes considerable changes to the system. Although the new updates are foreign to the system, they do not cause any potential harm to the system but rather improves the performance. The Correlator negotiator unit makes sure that the information provided by the Coordinator corresponds to actual incidents (potentially harmful system behavior) rather than events (normal behavior such as system updates etc). The Correlator does this by communicating with **Unit one of the Knowledge Database**. This unit consists of pre and post conditions of all known attacks. The Correlator unit checks to see if the current alert information matches any of the pre or post conditions in the database. This would help link one alert to the other when
consequences of the first alert become a necessary prerequisite for the other. Prerequisites can be closely linked to Apoptotic alerts which would correspond to low-level alerts that could sometimes result from authorized activities. Hence presence of an attack prerequisite does not necessarily mean the presence of intrusion. Consequences of an attack can be related to Necrotic alerts as they would indicate severe system damage or the presence of actual intrusion itself. By carefully correlating the Apoptotic and Necrotic alerts it is possible to build complex intrusion scenarios.

For instance, Code Red is an excellent example of a blended attack. This attack takes place through several stages (over time) and effectively combines multiple attacks. First it spreads itself through Buffer Overflow which is a very common attack although not serious. It then launches Denial of Service attack against a fixed Internet Protocol. So its prerequisites would be Buffer Overflow and Denial of Service attacks. Consequences of these prerequisites would result in ‘CODE RED’. The post condition of inadequate buffer checks leads to the precondition of Code Red (buffer overflow).

The Negotiator unit is responsible for setting the appropriate Confidentiality Integrity and Availability Threshold. The initial CIA threshold for peer nodes is set by communicating with the peer information buffer. However, once new alerts arrive and are segregated as Apoptotic and Necrotic, the Negotiator unit strikes a balance between the two alerts to determine the most suitable intrusion signature and anomaly threshold setting so as to reduce the production of false alarms. This automatically updates the CIA threshold unit. Once the Correlator/Negotiator unit positively determines the presence of harmful activity in the system, it sends a confirmation signal ‘2’ to the Coordinator.

4. Signal Propagation: Phase Four

Once the Coordinator receives the confirmation signal ‘2’ from Correlator / Negotiator unit, it immediately activates the Neutralizer Unit and starts to propagate the danger
signal to all its neighbors. It informs the neighbors as to which node is infected, how severe is the damage and what precautions to take.

5. Neutralization: Phase Five

Once the Neutralizer unit gets activated by the Coordinator, it takes necessary steps to calm the situation. Based on the type of the attack and the severity of the attack, action is taken to immunize itself from infected node.

6. Updating Knowledge Database: Phase Six

If the signature of the undergoing attack is not already saved in the Knowledge Database, the Coordinator feeds the detected anomaly into the Anomaly Signature Generator unit. This unit automatically converts the new attack into a signature pattern for future reference. Similar to the laws of Lymphotics discussed in chapter 5, Danger Theory based Event Incident Model has defined set of laws that it follows.

**Law 1:**

Coordinator becomes activated if

- Receives Signal ‘0’
- Receives Signal ‘1’
- Receives both Signal ‘0’ and Signal ‘1’
**Law 2:**

Coordinator Propagates Danger Signal if and only if

- Receives confirmation signal ‘2’ from Correlator/Negotiator Unit

Else

- Ignore Signal ‘0’ and signal ‘1’
- Become deactivated

**Law 3:**

After Propagation of Danger Signal

- Become deactivated
c. Attack Resistance

Fig 18. Attack Resistance

> Agents location completely randomized
> All agents are digitally signed
> Communication between agents through secure channel
> Back up copies of agents maintained
> Each agent has a co-operating agent associated with it
> Principle of Proof-carrying code utilized
One of the unique features of the Event-Incident Model is the ability of the Intrusion Detection Squad to protect itself and its platform from attacks. [Fig 18] illustrates how attack resistance is accomplished by the Model.

1. Randomized agent location

Once the mobile agents are recruited by the Coordinator, the Coordinator randomly assigns nodes that need to be monitored. The mobile agents are elusive in the sense that they continuously and randomly move around the network making it impossible to pinpoint their exact location. The list of neighbors that each agent visit and the order in which they visit keeps changing and the Coordinator refers to the Peer Information Buffer to assign this. The agents carefully analyses and determines the current state of the network and then sends this state information to the Coordinator which chooses an efficient and safe migration strategy for the agents.

2. Digital signatures

The Coordinator before sending the agents for a neighborhood patrol, digitally signs the agent code as a means of confirming the authenticity and the integrity of the code. It does this by utilizing Public Key Cryptography. The creator of the agent (Coordinator) uses a pair of cryptographic keys namely public key and private key. The owner (Coordinator) retains the private key as it is unique and specific to that owner, while the public key is made available to all. The Coordinator computes a hash of the agent program and produces a fixed length string as an output also known as digital fingerprint or digest. This agent is digitally signed by using the Coordinator’s unique private key. The agent and the signature are both encrypted by using the receiver's public key. Now the receiving node decrypts the agent and the signature by using its private key. This would
now expose the agent program with the signature. It computes hash from the program contents. In order to verify the signature it decrypts signature in the message back into the digest by utilizing the sender’s public key. It now compares it to newly computed digest to ensure it has not been tampered with.

3. Cooperating agents

Every agent has a cooperating peer agent to which it periodically informs its whereabouts. This cooperating agent would maintain an itinerary that contains the previous location, current location and the future location information of its peer. It checks the itinerary to see if any inconsistencies are present. If so, it immediately informs the owner of the agent that its peer has been corrupted.

4. Backup Copies Maintained

The Coordinator has all the necessary backup information for its agents as agents communicate its internal state with the Coordinator. In the event of a mobile agent failure or corruption, the Coordinator recruits a new agent to take over the place of the corrupted agent by copying the state information into the new agent. This enables it to easily resume the activities where it was left off.

5. Proof Carrying code

In a mobile agent environment every node’s Coordinator establishes a set of safety rules that would ensure the safe execution of the mobile code. The agent Coordinator, before sending mobile agents on patrol, attaches a formal safety proof that proves that the agent
indeed satisfies all of the safety protocols established by the receiver node. The receiving
node, upon the agent’s arrival, checks to see if the proof is valid. If so it concludes that it
is safe to execute the mobile agent application.

6. Every mobile agent maintains Path Histories

In order to avoid corrupted agents monitoring hosts, every mobile agent in the network
maintains an agent path history. This will help the node know as to which host was last
monitored by the agent. If the previously visited node is an infected node, the mobile
agent will not be allowed to monitor the current platform. The path history of the mobile
agents contains details such as the node previously visited, the node that is to be visited
next along with the complete path history. Whenever the agent visits a platform, the node
adds a signed entry to the path history. By merely looking at the signed entry, it is
possible to determine which node was last visited by the agent.

7. Encrypted Communication

Communication between nodes and agents take place via a secure encrypted channel.
This authenticates the sending and receiving party to each other.

d. Topology (Danger Zone Establishment)

[Fig 19] illustrates the topology of the Event-Incident Model. The topology assumption is
that two neighbors at least have one common neighbor. In [Fig 19], we have a network of
19 nodes. The Coordinator on node 8 recruits both Assistant Patrol Agents and Incident
Pattern Presenters. It then digitally signs them and sends them for a routine patrol to node
11. The patrol agents monitor node 11 and determine the presence of a suspicious
activity. They immediately communicate with their owner (node 8’s Coordinator) and inform their findings (signal ‘0’ and ‘1’). This activates the Coordinator.

The Coordinator then checks the validity of this information with the help of the Correlator/ Negotiator unit. Once it receives the confirmation signal ‘2’ from the Correlator/ Negotiator unit, it decides to propagate danger signal. Coordinator on node ‘8’ now sends the danger signal to all its neighbors (7, 14, 13, 9 10) excluding the infected neighbor 11. Each of the node that receives the danger signal checks to see if it is directly connected to the infected node 11. **Nodes that are not directly connected to the infected node (green nodes) will not propagate the danger signal. Only the nodes that are directly connected to the infected node (red nodes) propagate the danger signal.** Nodes 14, 13 and 9 will not propagate the danger signal but will immune itself from node 11 and update its peer information buffer.

**Fig 19. Topology**
Node 7 and 10 are directly connected to 11, hence they perform their own check by sending their mobile agents to node 11. This is done to confirm node 8’s results. If the results come back positive, they would propagate the danger signal to their neighbors. Node 10 would now propagate the signal to 9, 2 and 3. It will not send danger signal to its neighbor 8, since it received the signal from 8. Nodes 9 and 2 are not directly connected to the node 11 and hence will not propagate danger signal. However node 3 is at the close proximity of the infected node. Hence node 3 will propagate the danger signal to its neighbors 2, 4 and 18. Note that node 2 receives the signal twice. A node can receive the danger signal any number of times (from different nodes), however, it will not propagate the danger signal twice to a same node. When a node receives danger signal twice, it checks to see if both the signals pertains to the same infected node. If not, it repeats the process and propagates the second danger signal. Otherwise, the second signal is ignored. Nodes 2 and 18 are not directly connected to 11 however, node 4 is. Node 4 would now perform its own check and the propagate danger signals to 18, 5, 12 and 7. Since 7 is the only node that is directly connected to node 11, it propagates danger signal to its neighbors 8, 14, 15, 16, 6 and 12. Node 8 now receives the signal that it propagated. But since it has already immunized itself it would not propagate the signal again. Danger zone is established around all of the nodes that receive the signal. Nodes 1, 19 and 17 do not receive the danger signal and are not within the danger zone. If even one of the red nodes gets infected, its neighboring green nodes would become red nodes and start propagating danger signal. [Fig 20] illustrates the Event-Incident Model Intrusion Detection Scenario.
Fig 20. Scenario
CHAPTER 8

8. Conclusion

The following merits make the proposed Intrusion Detection System promising:

1. Recognition

Mobile agents of the Intrusion Detection Squad have the ability to accurately distinguish events from incidents. They do this by effectively communicating with the Knowledge Database Unit.

2. Learning and Memory

The first encounter with an unknown attack allows the mobile agents to learn, analyze and understand the nature of the attack. This new information is then fed into the Knowledge Database Unit so as to remember and respond quickly if hit by the same attack in future.

3. Diversity (4 levels)

Diverse set of mobile agents are created by the Coordinator to ensure security at four levels namely Host, Application, Protocol and Network level. Since the agents are customized, an accurate detection is made feasible.
4. **Scalable**

This network of Intrusion Detection Squad will scale very well for large networks. Since the danger signals are only propagated by the nodes that are in close proximity to the infected node, the area of danger zone is relatively small or restricted. Adding more nodes to the network will not affect the overall efficiency of the model. The mobile agents are also expendable and no restrictions are necessary on the number of agents created.

5. **No single point of failure**

By distributing relevant intrusion information among the group of peers, it is extremely hard for the whole system to be compromised. A failure of one node, will not affect the working of an entire network. As soon as a node receives danger signal from its peers, it immediately immunizes itself against the infected node. Since danger signals spreads very fast, it is hard to infect the whole neighborhood as by that time all of the nodes would have been immunized.

6. **Danger zone establishment through effective alert communication**

Danger signals sent out by the Coordinator does not simply inform that the node is infected. It informs the exact nature of the attack (whether necrotic or apoptotic), evidence of any relationship between the alerts, possible ongoing intrusion scenario and the strength of the intrusion. This effective communication of the alert information in the form of danger signal helps establish a danger zone.

7. **False alarm reduction by alert correlation**

Once new alerts arrive and are segregated as apoptotic and necrotic, the negotiator unit strikes a balance between the two alerts to determine the most suitable intrusion signature
and anomaly threshold setting. By choosing the appropriate threshold it makes it possible to reduce the production of false alarms.

8. Attack Resistance

Agents and agent platforms resist attack in this architecture in more ways than one. Randomizing agent location, use of digital signatures, adopting encrypted communication channel, utilizing cooperating agents and the process of proof carrying code makes this mobile agent architecture very secure.

By exploring the parallelism between immunology and computer security, the proposed IDS combines the above merits into hybrid architecture of Intrusion Detection and Response. It exhibits ability to automatically detect and respond to new anomalies and subsequently stores them in its Knowledge Database for improving its detection accuracy and efficiency in the future. Table 1 presents the analogies between the components of the proposed IDS and the Human Immune System.
Table 1. Analogies between Intrusion Detection Squad and Immune System

<table>
<thead>
<tr>
<th>INTRUSION DETECTION SQUAD</th>
<th>IMMUNE SYSTEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 INCIDENT PATTERN</td>
<td>BACTERIA / NON-SELF</td>
</tr>
<tr>
<td>2 EVENT PATTERN</td>
<td>SELF</td>
</tr>
<tr>
<td>3 NEUTRALISER</td>
<td>ANTIBODIES</td>
</tr>
<tr>
<td>4 APP-&gt; ASSISTANT PATROL AGENTS</td>
<td>T-HELPER LYMPHOCYTES</td>
</tr>
<tr>
<td>5 IPP-&gt; INCIDENT PATTERN PRESENTERS</td>
<td>B CELL LYMPHOCYTES OR ANTIGEN PRESENTING CELLS</td>
</tr>
<tr>
<td>6 PATTERN GENERATOR</td>
<td>GENERATES RANDOM PATTERN FOR NEGATIVE SELECTION</td>
</tr>
<tr>
<td>7 PATTERN MATCHER</td>
<td>RECEPTOR/EPITOPE BINDING</td>
</tr>
<tr>
<td>8 PROFILE GENERATOR</td>
<td>ABILITY OF PHAGOCYTES &amp; KILLER CELLS TO DIFFERENTIATE SELF/NON-SELF</td>
</tr>
<tr>
<td>9 PROFILE DEVIATION DETECTOR</td>
<td>PATHOGEN RECOGNITION RECEPTORS OF PHAGOCYTES FOR COMMON PATHOGEN ASSOCIATED MOLECULAR PATTERNS</td>
</tr>
<tr>
<td>10 CORRELATOR</td>
<td>INITIATES IMMUNE RESPONSE APOTOPSIS / NECROSIS</td>
</tr>
<tr>
<td>11 NEGOTIATOR</td>
<td>REDUCES / SUPPRESSES BODY'S IMMUNE RESPONSE</td>
</tr>
<tr>
<td>12 GIA THRESHOLD-&gt; CONFIDENTIALITY, INTEGRITY &amp; AVAILABILITY THRESHOLD</td>
<td>AFFINITY THRESHOLD</td>
</tr>
<tr>
<td>13 ANOMALY SIGNATURE GENERATOR</td>
<td>PRODUCTION OF MEMORY CELLS THROUGH PRIOR ENCOUNTER OR BY VACCINATION</td>
</tr>
<tr>
<td>14 COORDINATOR</td>
<td>DUAL FUNCTIONALITY</td>
</tr>
<tr>
<td></td>
<td>I POSITIVE SELECTION PROCESS / INNATE IMMUNE SYSTEM-&gt; NON-SPECIFIC</td>
</tr>
<tr>
<td></td>
<td>II NEGATIVE SELECTION PROCESS / DIRECT IMMUNE RESPONSE BY SIGNALING BETWEEN CELLS (LYMPHOKINES)-&gt;ADAPTIVE OR SPECIFIC RESPONSE</td>
</tr>
<tr>
<td>15 KNOWLEDGE DATABASE</td>
<td>SYMPTOMS OF VARIOUS DISEASES THAT THE BODY EXHIBITS</td>
</tr>
<tr>
<td></td>
<td>CHARACTERISTICS OF SELF MHC MOLECULES THAT IS USED TO UNIQUELY IDENTIFY NONSELF MOLECULES</td>
</tr>
<tr>
<td></td>
<td>IMMUNOLOGICAL MEMORY</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


agents”. Technical Report 95–008, Department of Computer Science, Purdue University, February 1995.


23. Mell, P., Mclarnon, M. “Mobile Agent Attack Resistant Distributed Hierarchical Intrusion Detection System”, In the Proceedings of RAID ’99, CERIAS, Purdue University, 1999


29. Online Encyclopedia. “B-Cells”


34. Online Encyclopedia. “Apoptosis and Necrosis”


