A Hybrid IDS Architecture Based on the Immune System

Marcelo Reis, Fabrício Paula, Diego Fernandes* and Paulo Geus
Computing Institute
State University of Campinas
Campinas, SP Brazil
{*marcelo.reis,fabricio,diego.fernandes,paulo*}*@ic.unicamp.br*

Abstract

The human immune system provides a rich source of inspiration for computer network security. Exploring this analogy the authors propose a hybrid intrusion detection architecture that has the same learning and adaptive capability of the human immune system. **Keywords:** IDS, computer immunology, adaptive detection, evolving reaction.

1 Introduction

From the viewpoint of traditional computer security, it is possible to guarantee the security of a computer system observing the following issues. It is necessary to correctly especify and implant a security policy, correctly design and implement the programs, and properly configure the system [6]. However, in practice, it is seen that security policies, programs' implementation, and systems' configuration might contain flaws that lead to an imperfect security [6].

A higher security level can be achieved by adopting additional resources and design models that very closely resembles the conditions in which most computer networks currently exist—a hostile and prone to flaws environment. It is possible to find in nature a defense model that has many features that are desirable for a security system: the human immune system.

Once it is able to guarantee the survival of an individual for almost 70 years, even though he/she encounters potentially deadly parasites, bacteria and viruses every day, the immune system has a very strong analogy with computer network security.

This analogy between computer security problems and biological processes was first recognized in 1987, when Adelman [2] introduced the term "computer virus". The connection between immunology and computer security initiated in 1994 with publications [3, 5], resulting in a series of other works.

*The authors are sponsored by FAPESP

The initial researches were concentrated on isolated mechanisms of the immune system and how they could be applied to improve the security of a system [6]. More recent work started to consider the overall framework of the immune system as a design model for a security system, based on a set of organizing principles of the human defense system [6]. However, most of the researches concentrate on development of anomaly intrusion detection systems. This approach, however, explores only a portion of the framework provided by the immune system.

In this paper, it is proposed a new approach on development of intrusion detection systems (IDS) based on the human immune system framework. This new approach considers the fact that the human immune system has many features of misuse detection (pattern matching, memory of known attacks, more specific recognition, for example) in addition to those features of anomaly detection that have been explored in past researches (knowledge of what is normal and detection of what is different from normal).

In this way, the authors propose a hybrid IDS model, based on the framework of the immune system, that is capable of detecting and identifing an attack, elaborating a specialized response measure, and recovering from the attack. Besides that, the proposed model has the same learning and adaptive capability of the human immune system, and so it is able to react to unknown attacks and to improve its response under subsequent exposures to the same attack.

This paper presents an improvement on the authors research presented in [1] and is organized as follows. Section 2 presents an overview of the human immune system. The proposed IDS is described in Section 3 and some analogies between the immune system and the proposed model are pointed in Section 4. Finally, Section 5 composes some conclusions about the work presented in this paper.

2 Immune System Overview

It is impossible to understand how the immune system can be used as a design model for a computer defense system without an overview of its framework. This section presents the immune system basic structures and explains the immune response. The material for this overview is largely based on [6, 8, 9, 10].

The immune system is composed by the innate and adaptive immune systems. The innate immune system represents the first defense line and it is distinguished by its innate feature, limitted capacity to differenciate an infectious agent from another (non specific detection) and its primary and non specific response (most often insufficient). Among its main components there are the physical and chemical barriers, such as the skin, and cells known as phagocytes that survey the body for foreign substances.

On the other hand, the adaptative immune system is able to identify a particular pathogen, allowing a more efficient response. Besides that, it is able to "memorize" an infectious agent and to respond more vigorously to new exposures to the same pathogen. It is composed of lymphocytes (T cells and B cells) and antibodies.

At the heart of the system is the ability to recognize and respond to substances called *antigens*. To do this, the immune system must perform pattern recognition tasks to distinguish molecules and cells of the body (called *self*) from foreign ones (called *nonself*). This pattern recognition is performed by the reaction between antigens and proteins (called *receptors*) on the surface of immune system cells. Antigens are the patterns to be matched, and receptors are a sort of complement of antigens. When an antigen binds to a receptor, a matching occurs and the immune response starts.

Phagocyte receptors can bind to a set of structurally related antigens and so its detection is not specific. B cell receptors, which are produced in soluble form as antibodies, can bind directly to free antigen. On the other hand, T cell receptors do not bind intact and free antigen, rather, they react with cell surface *major histocompatibility complex* (MHC) molecules that display antigens fragments, called *peptides*. B and T cell receptors perform specific recognition of antigens.

The ability to detect most pathogens is partly achieved by generating receptors through a random process. However, only the receptors that do not bind to self proteins are choosen through a process called *negative selection*. During this process, recently created receptors are exposed to most self proteins; if any receptor binds to these self proteins it is eliminated. Negative selection creates the knowledge of what is normal (self) to the immune system[1].

The immune system has the ability to make its protection more specific by learning and memory[2]. If the immune system detects a pathogen that it has not encountered before, it undergoes a primary response, during which it "learns" the structure of the specific pathogen, evolving a set of its cells with high affinity for that pathogen, through a process called *affinity maturation*. On subsequent encounters with the same antigen pattern the immune system mounts a secondary response, using high affinity evolved cells retained in immune memory, that is more precise and efficient.

All immune cells and products (such as antibodies) circulate in the bloodstream, tissues and lymphatic vessels, acting as sentries on the lookout for foreign antigens. When receptors bind to antigens, on a sufficient concentration, an matching occurs and a complex set of events, called immune response, takes place resulting in the destruction of the infectious agents. The immune response can be splitted into three phases, as follows.

Detection Phase

When phagocytes or lymphocytes find foreign antigens they engulf and destroy them. After that, they display the antigen fragments combined with MHC molecules on their surface. If the foreign antigen is already "known" to the immune system, specific antibodies may bind directly to the antigen, making microbes attractive to other immune cells.

Antigen Presentation and Lymphocytes Activation Phase

Phagocytes and B cells that display MHC molecules with antigen peptides attracts circulating, resting T cells. If a T cell recognizes the antigen-protein complex and binds to it, it becomes activated and stimulates the transformation of the B cells into antibody-secreting plasma cells. Activated T cells start to reproduce and B cells start to produce specific antibodies—an antigen specific army is raised.

[1]The knowledge of what is normal and detection of what is different from normal are anomaly intrusion detection features [7].

[2]The learning feature of the immune system relates to the learning of what is known to be "bad", and the immune memory is a sort of database of dangerous antigen signatures. These are features of misuse intrusion detection [7].

Antigen Elimination Phase

Specific antibodies bind to antigens marking them for destruction by phagocytes and cytotoxic T cells eliminate infected cells. As long as the concentration of foreign antigens decreases, all the chemical stimulus are gradualy contained, leading to the immune response end. At the end, high affinity lymphocytes are retained in immune memory for future responses.

3 Hybrid Immune Based IDS Model

All researches on computer immunology, such as [4, 5], have focused on random generation of receptors and the process of negative selection of receptors that do not bind to self proteins. This approach is used in the quoted researches for the development of anomaly intrusion detection techniques. Basically they produce a database of what is considered to be normal in the system, and randomly generate receptors that are tested against the database of normal behaviour. All receptors that fail to match any entry in that database is used to monitor the system, assuming that if it is activated, an abnormal situation has happened.

The new approach proposed in this paper is that the immune system also has some misuse intrusion detection features, and so represents a design model for a hybrid intrusion detection system. The immune memory is a database of signatures of known dangerous antigens, and antibodies and B cell receptors are signatures of specific antigens that the immune system has already encountered. These components of the immune system allow it to respond more efficiently to new exposures to a known invader. These are clearly misuse intrusion detection features, with an improvement—the immune system can autonomously change its misuse database (immune memory).

This section proposes an IDS model, based on the framework of the human immune system, that uses a hybrid architecture which applies both anomaly and misuse detection approaches [7]. Figure 1 illustrates this IDS model, presenting its components and the information flow between them. All its components are detailed as follows.

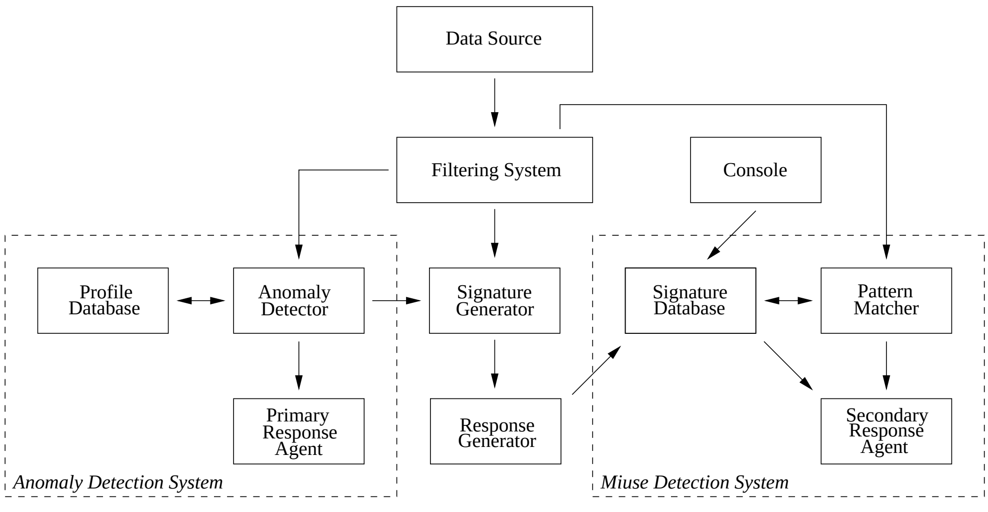

Figure 1: Hybrid immune based IDS model.

Data Source

The *data source* is responsible for collecting information and supplying a stream of event records to the *filtering system*. The nature of the information collected may vary according to the monitoring strategies adopted[3]: *host-based*, *network-based*, *aplication-based* or *target-based* [7]. The proposed IDS model is applicable to any of these strategies.

Filtering System

The *filtering system* provides audit reduction in order to identify and remove information that is redundant or irrelevant [7]. After filtering, the information stream is passed to the detection systems and, when required, to the *signature generator*.

Anomaly Detection System

Anomaly detection involves a process of establishing profiles of normal behaviors, comparing actual behavior to those profiles, and flagging deviations from the normal (assuming it indicates misuse of the system). This approach accommodates adaptations to changes in normal behavior over time, adding learning and adaptability to the IDS [7]. The components of the anomaly detection system are described as follows.

_ Profile Database

The *profile database* is responsible for storing the profiles that describe the behavior of the computer system. These profiles may be traced through quantitative analysis techniques, statistical measures, neural networks, genetic algorithms and immune system approaches. Moreover, profiles are periodically and automatically updated to provide adaptive detection [7].

_ Anomaly Detector

The *anomaly detector* receives the event stream from the *filtering system* and verifies if it represents anomalous behavior. To do so, it compares the informations received with the set of previously established profiles stored in the *profile database*. If any sign of abnormal behavior is detected, the *anomaly detector* activates the *primary response agent* and feeds the *signature generator* with the informations detected as abnormal.

_ Primary Response Agent

Once activated, the *primary response agent* initiates a series of contention measures to slow down or even block a probable attack. The *primary response agent* reaction is limited and general once the attack is not specifically identified yet. The main purpose of these primary response measures is to minimize damage until a specific and efficient response can be executed. Some examples of such primary responses are: priority level reduction or process blocking, remote login disabling, filesystem protection and alarms of intrusive activities.

[3] It is assumed that the anomaly detection system may use a different monitoring strategy from the one adopted by the misuse detection system.

Signature Generator

An innovating feature of the proposed IDS is the convertion of informations considered to be anomalous into a signature that specifically identifies the attack related to that abnormal behavior. This convertion introduces a learning capability, intrinsic to the anomaly detection, into the misuse detection system and provides a more efficient and precise detection of the attack in the future. In this way, the proposed IDS is able to automatically generate signatures of attacks that are unknown to the system. The *signature generator* is responsable for this convertion of anomalous informations into a signature of the attack. After the generation of the signature, the *signature generator* activates the *response generator*.

Response Generator

The *response generator* receives the signature of the attack and elaborates a set of countermeasures specific to that attack. Both signature and response produced are delivered to the *signature database*.

Misuse Detection System

Misuse intrusion detection comprehends the search for activity patterns that match a known attack or other violation of security policy. This approach has shown to be efficient and reliable, as a consequence it is used on most commercial IDS [7]. The components of the misuse detection system are described as follows.

_ Signature Database

The *signature database* responsible for storing the signatures of attacks, relating them to the respective response measures. The signatures are used by the *pattern matcher*, while the countermeasures are consulted by the *secondary response agent*. This way, the proposed IDS can specifically detect and respond to each manifestation of a known attack in the system. The introduction of new signatures and countermeasures into the *signature database* can be conducted in two ways:

1. Automatically by the *response generator*;
2. Or manually by the system administrator through the *console*.

_ Pattern Matcher

The *pattern matcher* receives the event stream from *filtering system* and matches it with the patterns stored in the *signature database*. If any pattern is found in the event stream, the *pattern matcher* activates the *secondary response agent*. The detection is conducted in real time and uses an approach based on state transition [7].

_ Secondary Response Agent

Once activated, the *secondary response agent* receives the pattern that was matched and queries the *signature database* for the specific countermeasures related to that pattern. So the *secondary response agent* executes the countermeasures.

Console

The interface between the proposed IDS and the system administrator is possible through the *console*. This interface allows the inclusion and removal of signatures and countermeasures in the *signature database*.

4 Analogies Between the Human Immune System and the IDS Model

As described in Section 2, the human immune system is divided into innate and adaptive systems. An analogy between these systems and the proposed IDS is illustrated on Figure 2.

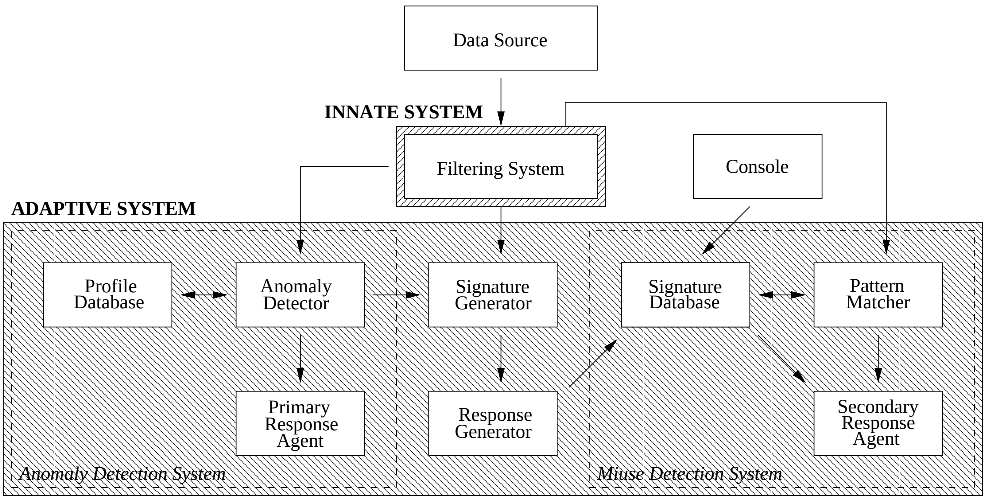

Figure 2: Analogy between innate, adaptive systems and the proposed IDS.

The innate system is partially represented by the *filtering system* whose function resembles the process of antigen presentation. Other main features of the innate system, non specific detection and response, are present in the anomaly detection system (*anomaly detector* and *primary response agent* respectively), which is properly modeled into the adaptive system considering the adaptive feature of the anomaly detection.

On the other hand, the adaptive system is represented by the components that implement learning and memory in the proposed IDS. Besides that, some of these components have other important features of the adaptive system, such as: accurate detection and efficient response.

Other analogies are presentated in Table 1, relating each component of the proposed IDS to the features of the human immune system.

5 Conclusion

The analogy between computer security and immunology represents a rich source of inspiration for development of new defense mechanisms, might it be algorithms and intrusion detection techniques, security policies aware of possible flaws or even entire

IDS Components	Immune System
Data Source	Source of self and nonself proteins
Filtering System	Antigen presentation process
Profile Database	Set of random generated receptors
Anomaly Detector	Fagocyte non specific detection
Primary Response Agent	Innate system primary response
Signature Generator	Production of memory cells
Response Generator	Specific antibodies production
Signature Database	Set of high affinity memory cells
Pattern Matcher	Detection though memory cells
Secondary Response Agent	Specific immune response
Console	Artificially acquired immunity through vaccines

Table 1: Analogies between the components of the proposed IDS and the immune system.

security systems. Exploring this analogy, the proposed IDS combines learning and specialization into a hybrid architecture of intrusion detection and response. This way, the proposed IDS is able of detect and respond to unknown attacks, improving its accuracy and efficiency on subsequent exploitations.

References

[1] M. A. Reis, D. M. Fernandes, F. S. Paula, and P. L. Geus. Modelagem de um Sistema de Segurança Imunológico. In *Proceedings of SSI'2001*, São José dos Campos, October 2001. (in Portuguese).

[2] F. Cohen. Computer Viruses. *Computers & Security*, 6:22-35, 1987.

[3] S. Forrest, A. S. Perelson, L. Allen, and R. Cherukuri. Self-nonself Discrimination in a Computer. In *Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy*, Los Alamos, CA, 1994. IEEE Computer Society Press.

[4] S. A. Hofmeyr, and S. Forrest. Architecture for an Artificial Immune System. *Evolutionary Computation*, 7(1):45-68, MIT, 1999.

[5] J. O. Kephart. A Biologically Inspired Immune System for Computers. In *Artificial Life IV*, 1994. MIT Press.

[6] A. Somayaji, S. A. Hofmeyr, and S. Forrest. Principles of a Computer Immune System. In *Proceedings of the Second New Security Paradigms Workshop*, 1997.

[7] R. G. Bace. *Intrusion Detection*. Macmillan Technical Publishing, 2000.

[8] I. Roitt, J. Brostoff, and D. Male. *Immunology*. Mosby, 1996. Fourth edition.

[9] *Understanding Vaccines*. U.S. NIH Publication No. 98-4219, January 1998.

[10] *Understanding Autoimmune Diseases*. U.S. NIH Publication No. 98-4273, May 1998.