Immuno-Inspired Autonomic System for Cyber Defense

Dipankar Dasgupta
Intelligent Security Systems Research Lab
The University of Memphis
Memphis, TN 38152
Email: dasgupta@memphis.edu

Abstract:

The biological immune system is an autonomic system for self-protection, which has evolved over millions of years probably through extensive redesigning, testing, tuning and optimization process. The powerful information processing capabilities of the immune system, such as feature extraction, pattern recognition, learning, memory, and its distributive nature provide rich metaphors for its artificial counterpart. Our study focuses on building an autonomic defense system, using some immunological metaphors for information gathering, analyzing, decision making and launching threat and attack responses. This on-going research effort is not to mimic the nature but to explore and learn valuable lessons useful for self-adaptive cyber defense systems.

1. Introduction

With the proliferation of complex Internet Computing, the tasks of integrating, deploying and managing computing resources have become very complicated. Moreover, protecting this heterogeneous computing base has become a more challenging issue (than ever before) for both the system administrator and the users. In order to protect this large cyber space, we need flexible, adaptable and robust cyber defense systems which can make intelligence decisions (in near real-time) for detecting wide variety of threats and attacks, including

- Active and passive attacks
- External attacks and internal misuses
- Known and unknown attacks
- Viruses and spam

There exist many software and hardware tools and techniques for cyber defense and each has its strengths and weaknesses. As intruders finding new ways to break in, security systems should be more flexible and intelligent enough to withstand both internal misuse and external intrusions. Influxes of new approaches are needed to enhance security measures. We are exploring immuno-inspired new security paradigm in order to address some of the threats and to build an autonomic cyber defense system.

The biological immune system is an adaptive defense system that is highly distributive in nature. It employs multi-level defense mechanisms to make rapid, highly specific and often very protective responses against wide variety of pathogenic microorganisms. It has evolved over millions of years probably through extensive redesigning, testing, tuning and optimization process. While many details of the immune mechanisms (innate and adaptive) and processes (humeral and cellular) are yet unknown (even to immunologists), it is, however, well-known that the immune system uses multilevel (and overlapping) defense both in parallel and sequential fashion. Depending on the type of the pathogen, and the way it gets into the body, the immune system uses different response mechanisms (differential pathways) either to neutralize the pathogenic effect or to destroy the infected cells. Though our body has continuously been exposed to various pathogens (known/unknown/harmful/benign), but handles most of them in an amazing delicacy with significant notice. Still this is not a full proof system; malaria, plague and other epidemics wiped out a large population at different times in history, and we are continually struggling to deal with new pathogenic challenges.

In next sections, some immunological features will be discussed along with their importance in building next generation cyber defense system.

2. Immune metaphor and cyber defense

From the information processing point of view, there are several immunological principles that makes the system very appealing, which include distributed processing, pathogenic pattern recognition, multi-layered protection, decentralized control, diversity, signaling, etc. These principles are also very important in developing next generation cyber defense system.

2.1 Distributed information processor

The immune system is a mobile agent system, where the agents circulate at various primary and secondary lymphoid organs of the body to perform immunogenetic functions. These differential migrations of lymphocyte subpopulations are carefully controlled to ensure that different variety of immune agents (specific to antigen) are present at every desired location. The lymph nodes and organs provide specialized local micro-environment (called *germinal center*) during pathogenic attack Such distributed processing centers are analogous to having mobile forensic laboratories for quick processing information (audit records) and devising strategies for deployment. We are developing an immunity-based architecture for security agents (with specific roles and functionalities). These agents can roam around the machines, nodes (or routers), and monitor the health of the network (i.e. look for changes such as malfunctions, faults, abnormalities, misuse, deviations, intrusions, etc.). The agents work in hierarchical fashion to recognize each other's activities and to take appropriate actions according to the underlying security policies [4-6]. Though the mobile agent technology is yet to mature, it can provide a platform for distributed attack detection and analysis.

2.2 Novel pattern recognizer.

The immune system can recognize and classify different novel-patterns (pathogenic patterns of interest) and generate selective responses. Self-nonself (or danger) discrimination may be one of the important tasks the immune system solves during the process of pathogenic recognition (see figure 1).

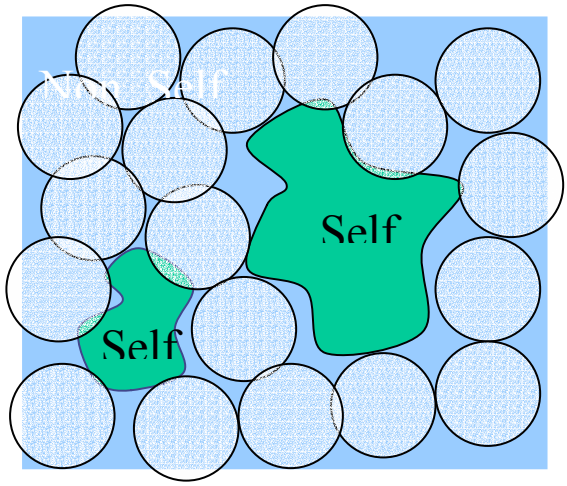


Figure 1: Conceptual view of self and nonself

This discrimination is achieved in part by T-cells, which have receptors on their surface that can detect foreign proteins (antigens). During the generation of T cells, receptors are made by a pseudo-random genetic rearrangement process. Then they undergo a censoring process, called negative selection, in the thymus where T cells that react against self-proteins are destroyed; so only those that do not bind to self-proteins are allowed to leave the thymus. These matured T cells then circulate throughout the body to perform immunological functions to protect against foreign antigens. Forrest et al. [9] proposed the negative selection (NS) algorithm based on self-nonself discrimination in the immune system. Different negative detector generation techniques are being studied, including real-valued, variable-size and fuzzy-rule detectors [8, 10]. The normal and the abnormal behaviors in networked computers are hard to predict, as the boundaries cannot be well defined. So the fuzzy logic can provide varying degree of normalcy in system behavior. The goal is, however, to evolve 'good' detector rules that cover the non-self space.

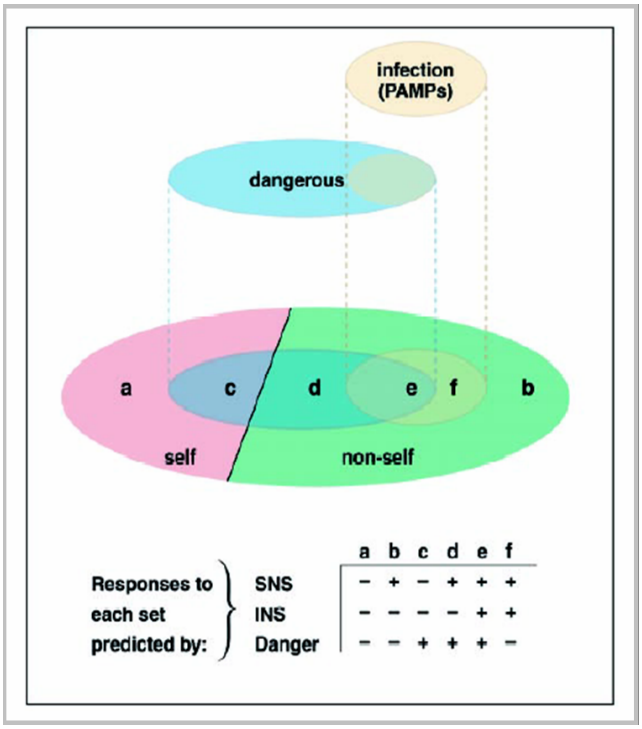
- 'good' rule:
  - It must not cover self space.
  - It has to be as general as possible: the larger the volume, the better.
- One rule may not enough, instead, a set of rules that solve collectively cover the non-self space with minimum overlap is necessary.

The aim is to find a small number of specialized detectors (as signature of known attack conditions) and other generalized detectors for unknown (or possible) attack conditions. Moreover, use of dynamic detector sets can adapt to a greater range of variations in system behavior. It is to be noted that NS approaches are logically different from the traditional methods for intrusion detection (attack signatures and/or normal profile are used). One of the advantages of using negative detectors is that the detectors can be distributed in different nodes (or hosts) with specific security coverage.

Based on new immune theory (Matzinger 1994, 2002), however, the immune system actually discriminates “some self from some non-self” (as shown in figure 2). Accordingly, some danger signals such as tissue damage triggers a myriad of immune reactions and responses.

The danger theory (DT) model appears to be more appropriate in cyber world as not all abnormal events

(nonself) represent attacks, rather a small percentage of such events are of real concern. It may be useful to develop computer security model based on DT, where some simple observations can trigger a chain of defense actions. But the challenge is clearly to define a suitable danger signal, a choice that might prove as critical.



Partition of the Universe of Antigens (Matzinger 1994, 2002)

SNS:
self and nonself (*a and b*)

INS:
noninfectious self (*a*) and infectious nonself (*f*)

Danger Theory (DT):
dangerous entities (*c, d, e*) and harmless ones

Figure 2: Partition of the Antigen Universe based on three models: SNS, INS, and DT (Matzinger 2002).

2.3 Multi-layered Defense System.

The immune system can be envisioned as a multilayer defense system with several mechanisms in each layer for protection against pathogens. The first layer is the anatomic barrier, composed of the skin and the surface of mucous membranes as physical barriers; physiological barriers such as destructive enzymes and stomach acids. In addition, it has two alternate defense pathways (innate and adaptive) may be considered as the second and third layers of defense. The *innate immunity*, also known as nonspecific immunity, is an unchanging (generic) mechanism that detects and destroys many invading organisms, whilst the adaptive immunity, also called acquired or specific immunity, represents the part of the immune system that is able to specifically recognize and selectively eliminate foreign microorganism and molecules. It responds to previously unknown foreign pathogens through on-line reinforcement learning and builds a response to them that can remain in the body for a long time. Adaptive immunity is characterized by learning, adaptability, and memory.

There are many security products (including COTS and GOTS) [12] that are available for computer systems to detect and report potential or existing computer system security irregularities at multiple layers. As illustrated in Figure 3, Firewall tries to block unwanted traffic and suspicious connections. Authentication and authorization tools examine the validity of users and try to prevent unauthorized usage. Access control mechanisms allow partitioning the information, program and data space based on certain criteria and user privileges. Many surveillance systems are developed so far, some of them perform active monitoring, and others are passive [1-2]. Reconnaissance tools, offline analyzers (and profilers) and trace back systems provide preventive and responsive measures. All these defense systems form a hierarchical cyber defense.

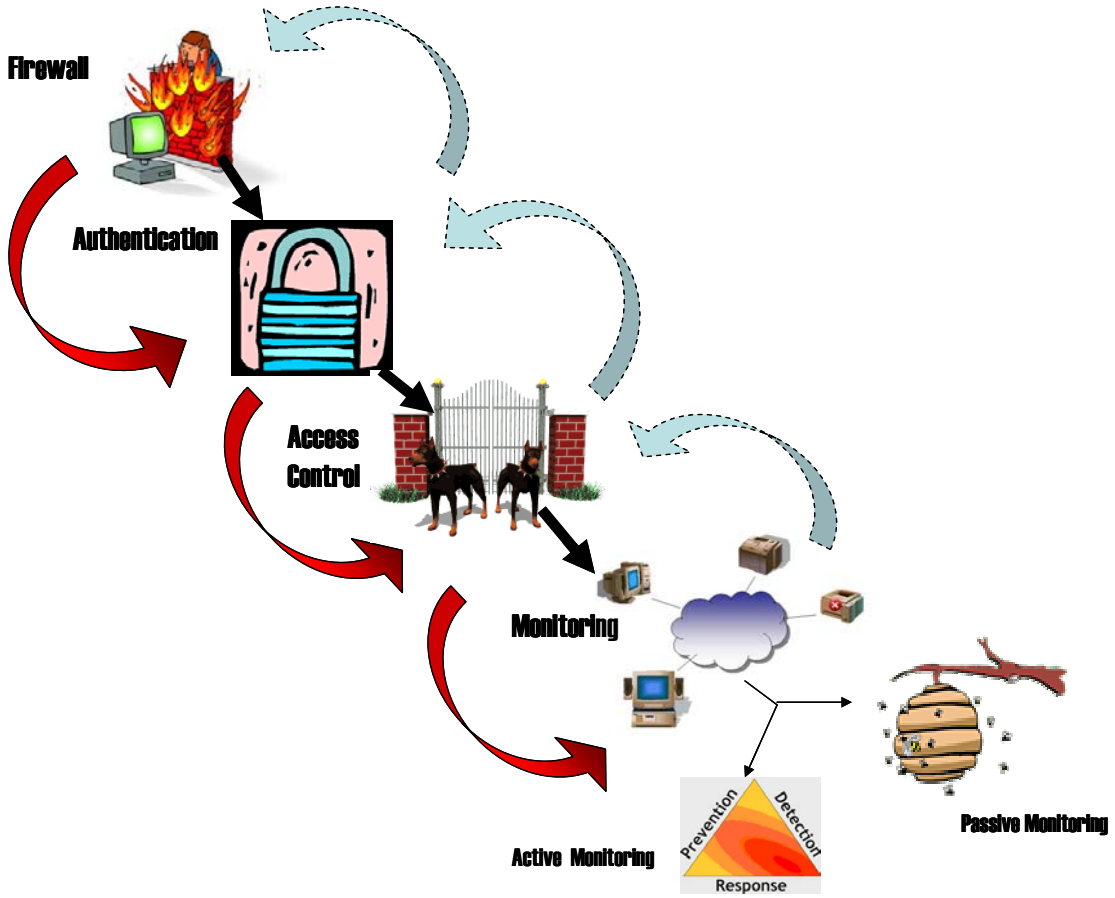


Figure 3: Multi-Layered Cyber Defense with feed forward and feedback signaling mechanism

2.3.1 Signaling and Message-passing

In the immune system, signal diffusion and dialogue are two kinds of communication schemes available. They take major role in sharing and passing information during immune response. In *immune diffusion*, the message is passed from one immuno-component to others without any feedback. Another scheme is called *immune dialogue,* where the immune system continuously exchanges molecular signals with its counter parts. The immune reactivity is determined by context, where self and foreign agents play upon each other. The body is under constant challenge to respond along a continuum of behavior and needs to adapt accordingly. These types of information sharing are missing in current security systems.

Though, there are a variety of tools (file integrity checkers, virus scanners, intrusion detectors, port scanners, etc.) available for multi-layered protection of cyber systems. Each tool has a specific purpose; some may overlap, while most don't. Most of these tools operate independently, without data exchange or consistent security policies. Each of them may have been developed by a different vendor, perhaps even competitors in the industry. Since there is no consistent data exchange between these tools, many attacks remain unnoticed. Moreover, security administrator intervention is usually required to analyze the acquired data and make decisions about what actions may need to be taken to prevent a compromise, or to recover from one. This may be a major bottleneck in building survivable autonomic system. The future cyber defense system should incorporate both feed forward and feedback mechanism as shown by arrows in Figure 3. It will help better understanding real attacks and their sources.

2.3.2 Co-stimulation

Making decisions based on multiple signals help to ensure tolerance and judge between dangerous and harmless invaders. This type of accompanying signal

helps in identifying an attack while minimizing false alarm and to generate decisive response in case of a real danger. Similar co-stimulation mechanism should be implemented among various tools that are specialized in detecting specific threats in an integrated cyber defense system in order to take accurate responses.

2.3.3 Decentralized Control

The immune system uses distributed control mechanism for learning, memory and associative retrieval to solve recognition and classification tasks. There is no single organ that controls the immune response; rather it handles the antigenic challenge through collaborative interaction. Use of such a defense mechanism may be hard to implement, but will help to avoid single point attack and make the system robust.

2.4 Multi-level Data fusion and correlation

In order to better understand diseases, experimental immunologists analyze biological processes at multiple levels such as molecular, cellular, protein and genetic for proper diagnose. In the same way, a cyber defense system should monitor networked computer's activities at different levels (such as application, user, system, process and packet levels) for correlating information among the observed parameters in order to determine intrusive activities (Figure 4). For example, such a system will, at user level -- search for an unusual user behavior pattern; at system level -- look at resource usage such as CPU, memory, I/O use etc.; at process level -- check for invalid or unauthenticated processes and priority violations; at packet level – monitor traffic (number, volume, and size of packets along with source and type of connections) information. It is to be noted that it is possible to extend monitoring sensors to firewall and router level as well. Moreover, different detection and representation schemes may be useful at different level for better detection of intrusions.

In the immune system, Antigen Presenting Cells (APCs) interpret the antigenic context and extract its unique features, by processing and presenting antigenic peptides on their surface. Each APC serves as a filter and a lens: a filter that destroys molecular noise, and a lens that focuses the attention of the lymphocyte receptors.

Sensory data fusion and information correlation [11] can provide better detection ability for wide range of attacks and threats at early stage of attacks. Moreover, the normalcy depends on correlations among different parameters. The independent values of two different parameters could be considered normal, but their combination could show abnormality. Accordingly, correlating information from multiple levels should be able to detect many attacks as illustrated in figure 5.

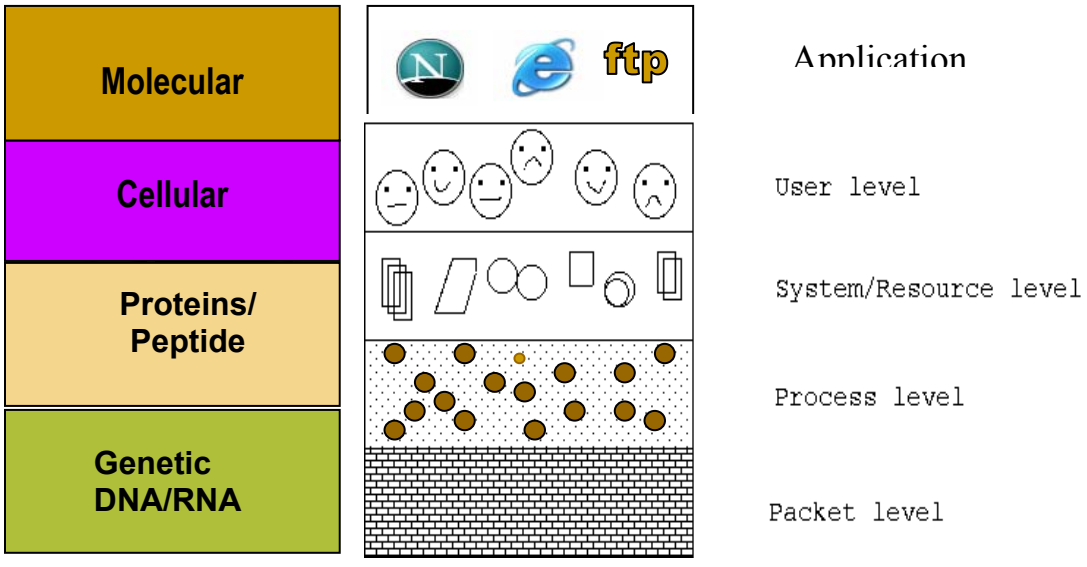


Figure 4: Multilevel monitoring and detection scheme

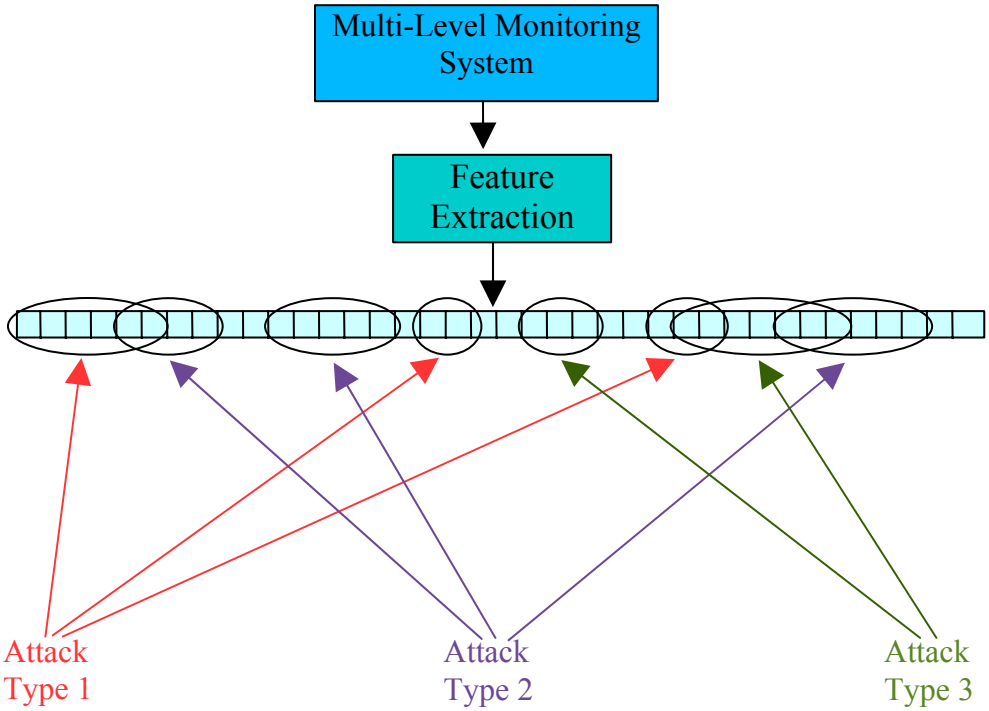


Figure 5: Shows the correlation of multi-level features in attack detection

3. Autonomic Cyber Defense System

An autonomic cyber defense system should be an integrated system of subsystems with many strategies and mechanism for overall protection. It should be flexible, scalable and adaptable which can provide a certain level of security assurance. It should ensure data availability, integrity, authentication, confidentiality, and nonrepudiation by incorporating protection, detection, and reaction capabilities. Such an integrated system should be able to identify irregularities that are linked to attempted or successful attacks, which may result in system failure or compromise. It should be able to detect security breaches that are internal, external, accidental, or intentional, in systematic fashion. Depending on the nature of intrusive activities, the system be able take automated responses (based polices and preferences of the organization). Such actions may include the following,

A1. Informing the system administrator via e-mail or other messaging system
A2. Change the priority of user processes
A3. Change access privileges of certain user
A4. Block a particular IP address or sender
A5. Disallow establishing a remote connection request
A6. Termination of existing network connection
A7. Restarting of a particular machine
A8. Logout user or close session

Figure 6 shows a conceptual model of the integrated cyber defense system, where the defense strategies are divided in three major areas. Under each defense strategy, several tools and techniques can be grouped together based on their functionalities. However, tools across different strategy should communicate with each other (as shown by arrows) and use uniform standard, protocol and policy in order to avoid conflict while sharing information. In may be possible that these tools run independently but work in concert based on some common security policy.

While building such a complex defense system may appear to be very difficult, but through a systematic approach of designing subcomponents and incremental integration, it is possible to develop such an autonomic system. More importantly, it may worth pursuing such efforts to combat sophisticated cyber terrorism. Further details of the proposed system will be provided in final version.

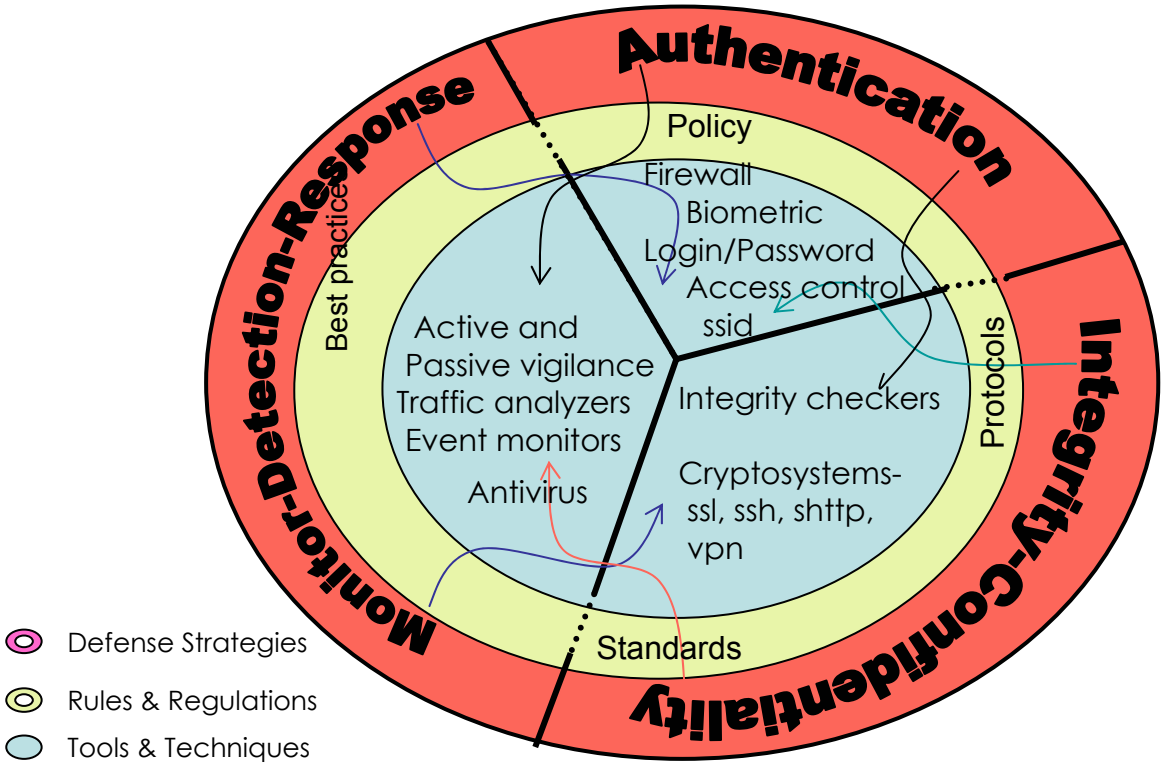


Figure 6: An integrated cyber defense system

4. Summary:

We can learn many lessons from the biological immune system in order to build robust and adaptive cyber defense system. The immune system employs a multilevel defense against invaders through nonspecific (innate) and specific (adaptive) immunity. Similar to the biological immune system, it is impractical to find and patch every security hole in a large network of computers. Thus, a multi-faceted and more comprehensive approach is most appropriate to cyber security. Modern medical sciences offer different treatment regiments to cure from diseases and also help to boost the immune system through vaccination. Vaccinations provide knowledge of known (and dangerous) pathogens and help the immune system to quickly handle viruses before they can cause damage to the body. In the same way, experts' knowledge about known attacks, viruses and worm be necessary (in order to avoid or minimize the training and learning process) to build the knowledge base (memory) of a computer immune system. As there is no single medicine available to cure all disease, there may not any single method to handle different type of cyber attacks. As intruders find new ways to break in, cyber security systems should be more flexible and intelligent enough to withstand both known and unknown attacks and the immuno-inspired approach may provide robust solution.

Our research focuses on exploring and exploiting immunological principles in developing a new paradigm in cyber security [3,5-7]. In particular, the goal is to develop a self-adaptive system that will perform real-time monitoring, analyzing, and generating appropriate response to wide variety of intrusive activities.

References

1. Bass, T. *Multisensor Data Fusion for Next Generation Distributed Intrusion Detection Systems*. Invited Paper, 1999 IRIS National Symposium on Sensor and Data Fusion, The Johns Hopkins University Applied Physics Laboratory, 24-27 May 1999.
2. Barrus, J and Rowe N C. *A Distributed Autonomous-Agent Network-Intrusion Detection*

and Response System. In Proceedings of the Command and Control Research and Technology Symposium, Monterey, CA, June 1998.

3. Brian, H and Dasgupta, D. *Mobile Security Agents for Network Traffic Analysis*. In the proceedings of the second DARPA Information Survivability Conference and Exposition II (DISCEX-II), Anaheim, California, June 13-14, 2001.
4. Chari, S. N and P. –C. Cheng. *BlueBox: A Policy-Driven Host-Based Intrusion Detection System.* In ACM Transactions on Information and System Security, Vol. 6, No. 2, pp 173-200, May 2003.
5. Dasgupta, D and Gonzalez F. *An Immunity-Based Technique to Characterize Intrusions in Computer Networks*. In the journal IEEE Transactions on Evolutionary Computation, Vol. 6, No. 3, June 2002.
6. Dunlap, G. T and Dasgupta, D. *An Administrative Tool for Distributed Security Task Scheduling*. Published in the proceedings of the Third Annual International Systems Security Engineering Association Conference, Orlando, March 13-15, 2002.
7. Dasgupta, D. *Immunity-Based Intrusion Detection Systems: A General Framework.* In the Proceedings of the 22nd National Information Systems Security Conference *(NISSC)*, October 18-21, 1999.
8. Dasgupta, D (Editor). *Artificial Immune Systems and Their Applications, ISBN 3-540-64390-7, Springer-Verlag, 1999.*
9. Forrest, S., Hofmeyr, S. A., Somayaji, A. and T. A. Longstaff. *A sense of self for unix processes.* In Proceedings of *IEEE Symposium on Research in Security and Privacy*, Oakland, CA, 1996.
10. Gomez, J and Dasgupta, D. *Evolving Fuzzy Classifiers for Intrusion Detection*. In the proceeding of 3rd Annual Information Assurance Workshop, June 17-19, 2002.
11. Debar H and Wespi, A. *Aggregation and Correlation of Intrusion-Detection Alerts*. In Proceedings of Recent Advances in Intrusion Detection (RAID), *W. Lee, L. Mé, A. Wespi (eds.)*, pp. 85-103, 2001.
12. Deswarte, Y., Abghour, N., Nicomette, V., Powell, D. *An Intrusion-Tolerant Authorization Scheme for Internet Applications.* In Sup. of the Proceedings of the 2002 International Conference on Dependable Systems and Networks (DSN), Washington, D.C. (USA), pp. C-1.1 - C-1.6, June 2002