International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:3, 2007

Danger Theory and Intelligent Data Processing

Anjum Iqbal and Mohd Aizaini Maarof

***Abstract*—**Artificial Immune System (AIS) is relatively naive paradigm for intelligent computations. The inspiration for AIS is derived from natural Immune System (IS). Classically it is believed that IS strives to discriminate between self and non-self. Most of the existing AIS research is based on this approach. Danger Theory (DT) argues this approach and proposes that IS fights against danger producing elements and tolerates others. We, the computational researchers, are not concerned with the arguments among immunologists but try to extract from it novel abstractions for intelligent computation. This paper aims to follow DT inspiration for intelligent data processing. The approach may introduce new avenue in intelligent processing. The data used is system calls data that is potentially significant in intrusion detection applications.

***Keywords*—**artificial immune system, danger theory, intelligent processing, system calls

I. Introduction

THE Artificial Immune Systems (AIS) are computational systems designed on the principles of natural Immune System (IS), which is highly distributed, adaptive and diverse system [1]. This is a novel and rapidly growing approach for solving complex computational problems intelligently. Artificial immune systems incorporate the ability to learn new information, recall previously learned information, and perform pattern matching in highly diversified manner. Researchers are making efforts to apply AIS paradigm to various domains. Examples include computing system intrusion detection, virus detection in files, information filtering, data clustering, process monitoring, equipment control, sensor fusion, image interpretation, detection of novel features in time series, and function optimization [2], [3].

Among immunologists, there are two distinct viewpoints about the main goal of IS, which provide the rationale for AIS design [4]. Thorough understanding of these viewpoints is of enormous importance for AIS researchers and designers.

The classical viewpoint about the main goal of IS is to discriminate between self and non-self, that is the discrimination of body cells and molecules from other invading cells and molecules. Currently this viewpoint is generally accepted by immunologist, and most of the AIS researchers are applying the same approach for their models. But this viewpoint has prominent question marks for it that may be answered by the other viewpoint called Danger Theory (DT) viewpoint. The DT states that immune system, instead of discriminating between self and non-self, looks for danger producing elements and events. Although DT is still controversial and requires more immunological research witnesses, it is believed that this viewpoint could help AIS researchers in improving their system designs [4], [5].

This paper aims to introduce a novel intelligent data processing approach inspired from DT. The derivation of biological abstractions for computational systems is a complex and creative process requiring knowledge of distinct fields [6]-[8]. We have tried here to convey the interdisciplinary research in simple and understandable manner. This may motivate compu-tational researchers to become part of AIS research community and contribute towards novel and potential field. The data used is system calls data available from the University of New Mexico website [9]. This data has significance for intrusion detection applications.

In the following sections, section II presents an overview of two distinct view points, self-nonself discrimination and danger cognition, about the main goal of immune system. Section III briefly describes a biological event that leads to danger. We have recently proposed a novel metaphor called DASTON [10], section IV introduces this metaphor. Section V depicts the presence of DASTONs in system call sequences. Finally, section VI concludes the effort. All these are arranged and supported with simple diagrams to provide comprehensive research base of an extremely complex field of study.

II. Overview of Distinct Viewpoints

A. Self Non-self Viewpoint

Classically it is believed for the immune system to function properly, it needs to be able to distinguish between the molecules of our own cells (self) and foreign molecules (nonself), which are a priori indistinguishable (Fig. 1). If the immune system is not capable of performing this distinction, then an immune response will be triggered against the self-antigens (substance that can trigger an immune response), causing autoimmune diseases. Not responding against a self-antigen is a phenomenon called self-tolerance, or simply tolerance. Understanding how this is accomplished by the immune system is called the self/non-self discrimination problem [2].

Manuscript received November 5, 2004. This work was supported in part by the Ministry of Science, Technology and Environment (MOSTE) Malaysia under vote no. 74022 and 74229.

A. Iqbal is with the Group on Artificial Immune Systems and Security (GAINS), Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia (phone: +6012-736-2147; fax: +607-553-2210; e-mail: anjum@siswa.utm.my).

M. A. Maarof is the head of Group on Artificial Immune Systems and Security (GAINS), Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia (e-mail: maarofma@fsksm.utm.my).

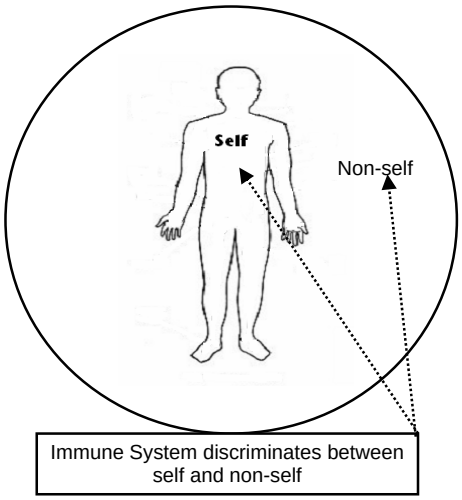

Fig. 1. Classical view point about the main goal of immune system that is the discrimination of self and nonself

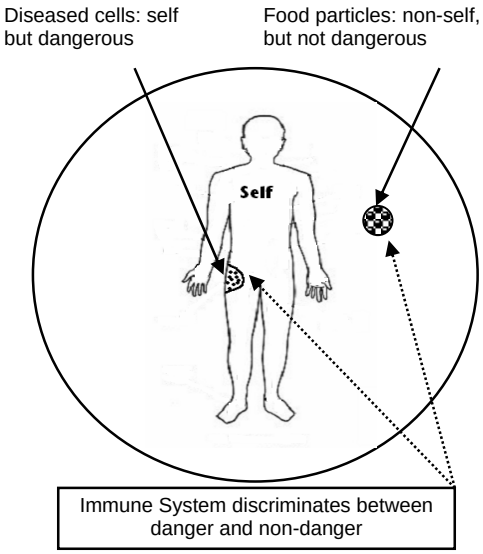

Fig. 2. Danger Theory viewpoint about the main goal of immune system that is looking for danger.

B. Danger Theory Viewpoint

Matzinger's Danger Theory debates the self-nonself point of view [11]-[15]. She points out that there must be discrimination happening that goes beyond the self-nonself distinction described above (Fig. 2). For instance [4]:

- There is no immune reaction to foreign bacteria in the gut or to the food we eat although both are foreign entities.
- Conversely, some auto-reactive processes are useful, for example against self molecules expressed by stressed cells, the cells undergoing dangerous interaction.
- The definition of self is problematic – realistically, self is confined to the subset actually seen by the lymphocytes (name of immune system cells) during maturation.
- The human body changes over its lifetime and thus self changes as well. Therefore, the question arises whether defenses against non-self learned early in life might be auto-reactive later.
- Other aspects that seem to be at odds with the traditional viewpoint are autoimmune diseases and certain types of tumors that are fought by the immune system (both attacks against self) and successful transplants (no attack against non-self).

III. Danger Producing Event

Genes are the codes of life encoding various structures and functions of living organisms. Pathogens are the micro-organisms that threat normal structures and functions of human (host) organs and systems. Genes of invading pathogens interplay with genes of host to produce poisonous or toxic products (Fig. 3), which convey presence of danger to immune system of the host. A small segment of the host gene may show susceptibility for pathogenic interplay. It is the segment which is contributing towards dangerous activities inside the host. Identification of these segments has immense significance in biology. We are not biologists and require inspiration only to apply this concept to our computational idea; therefore we are not going to explore biological details further. But the reader being attracted by the inspiration will have to go through the details of novel metaphor DASTON (DAnger Susceptible daTa codON), as described in [10].

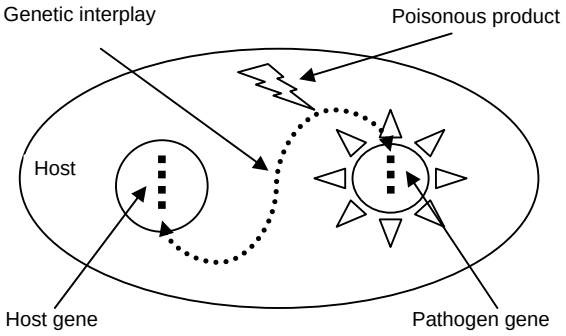

Fig. 3. The disease susceptible host genes interact with pathogen genes to produce poisonous products that cause danger.

IV. Introduction to DASTON

Based on the biological concept, briefly described in section III, we have proposed the presence of DASTONs in data [10]. These are the data chunks or points present in data heap that actively participate in data processing to retrieve specific information from that data when subjected to data or activity, which initiates the processing, as shown in Fig. 4. It is like presence of genetic segments in host that are

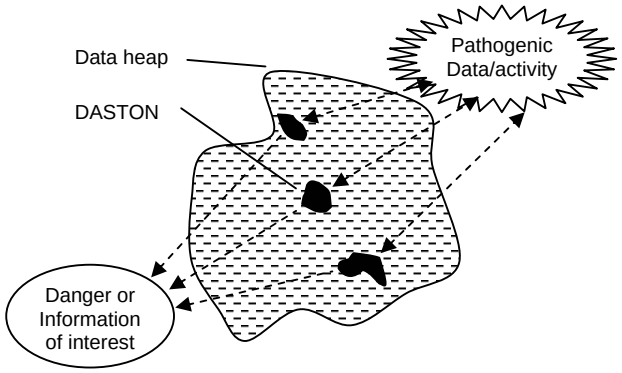

Fig. 4. DASTON present in data heap interact with incident data (named as pathogenic data) to produce required information (analogous to danger in danger theory)

susceptible to pathogenic interactions resulting in the production of toxic substances leading to danger. Real example could be that only specific fields in a database will interact with query fields to result required information. One may use his own creative analogy to implement this biologically inspired idea. The success of analogy depends on the degree of creativity and clarity in understanding the biological concept upon which it is based [6].

V. DASTON in System Calls Data

System calls data has enormous significance in intrusion detection applications. Most of the intrusions ultimately reflect through system call sequences of a process. Careful online or offline analysis of system call sequences may lead to efficient and reliable intrusion detection mechanisms. Processing of whole system call sequences will make the system inefficient. On the other hand, taking care of DASTONs only will reduce the amount of data to be processed, hence making the system efficient with improvement in its intelligence. By correlating system call sequences of normal and exploited processes, available for download from the University of New Mexico website [9], we have identified the presence of DASTONs. The result is shown in the plot of Fig. 5.

The size of data chunk, in this case, is two that is a pair of two adjacent system calls in a sequence. The numbers on horizontal and vertical axis are the numerical labels of system calls appearing at first and second position in a pair respectively. The crosses (X) are representing pairs appearing in both normal and exploit data, circles (O) represent pairs in normal sequence only, and dots (•) represent the pairs appearing in exploit sequence only. The pairs represented by "•" are those, which have had significant interplay with exploit script during the course of attack; like genetic segments susceptible to pathogenic interplay. These are the DASTONs present in exploited system calls sequence.

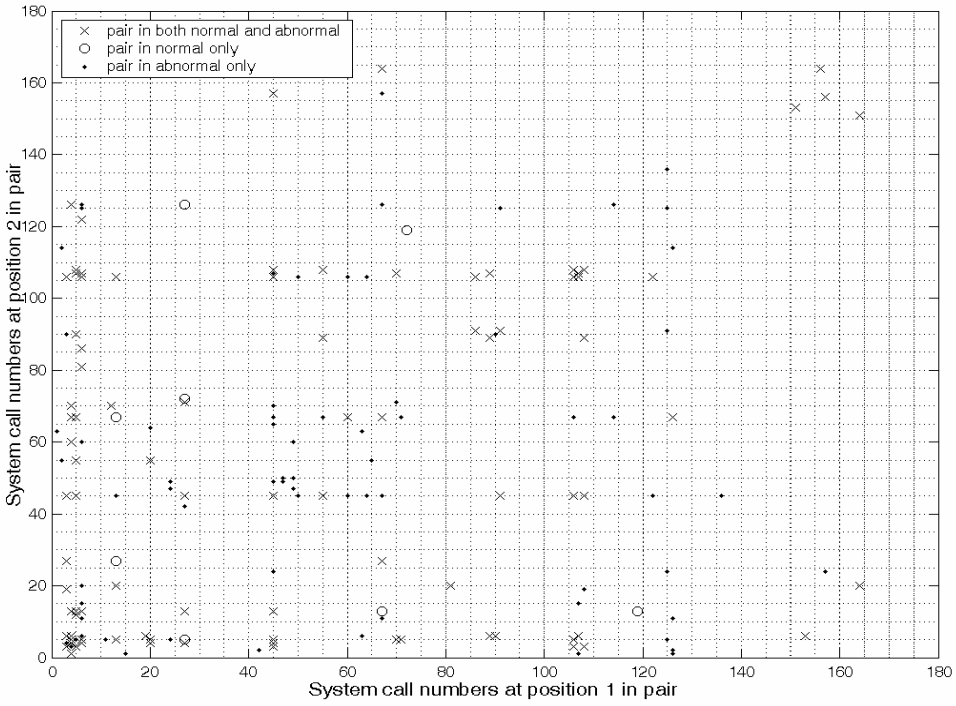

Fig. 5. Plot of normal and intrusion trace data for synthetic wu-ftpd; black dots are DASTONs those actively interacted with danger producing scripts (system call data from University of New Mexico).

VI. Conclusion

Artificial Immune Systems (AIS) is a relatively naive paradigm for intelligent computation. Danger Theory (DT) has shifted the classical viewpoint about the main goal of immune system. The DASTON is a novel metaphor in AIS field inspired from DT. Proper identification of DASTONS may lead to significant reduction in data to be processed and endow intelligence to the data processing. This effort has identified the presence of DASTONs in system calls data. Creative mappings and analogies may be used to identify DASTONs in various types of data. The mechanism of better identification and utilization of DASTONs will remain challenge in our future efforts. Anyhow, it has opened new avenue in biologically inspired intelligent computation that is Artificial Immune Systems.

Acknowledgment

We are grateful to academic network of ARTificial Immune SysTems (ARTIST) for their appreciations. We are also thankful to Dr. Polly Matezinger (the founder of Danger Theory) for her guidance.

References

[1] A. Somayaji, S. Hofmeyr, and S. Forrest, "Principles of a Computer Immune System," In *Proc. New Security Paradigms Workshop,* Charlottsville, 1998, pp. 75-82
[2] L. N. de Castro, and F. J. V. Zuben, "Artificial Immune Systems: A Survey of Applications," State University of Campinas, SP, Brazil, Tech. Rep. DCA – RT 01/99, 1999, Part I.
[3] L. N. de Castro, and F. J. V. Zuben, "Artificial Immune Systems: A Survey of Applications," State University of Campinas, SP, Brazil, Tech. Rep. DCA – RT 02/00, 2000, Part II.
[4] U. Aickelin, and S. Cayzer, "The Danger Theory and Its Application to Artificial Immune Systems," In *Proc. First International Conference on Artificial Immune Systems (ICARIS-2002)*, Canterbury, UK, 2002.
[5] A. Iqbal, and M. A. Maarof, "Distinct Viewpoints in Novel Biologically Inspired Computational Research: Artificial Immune Systems," In *Proc. Second International Conference on Artificial Intelligence Applications in Engineering and Technology*, Malaysia, 2004.
[6] S. A. Hofmeyr, and S. Forrest, "Architecture for an Artificial Immune System," *Evolutionary Computation Journal,* vol. 8 no. 4, pp. 443-473, 2000.
[7] J. Kim, "Computers are from Mars, Organisms are from Venus," *IEEE Computer,* vol. 35 no. 7, pp. 25-32, 2002.
[8] A. Iqbal, and M. A. Maarof, "Artificial Immune System and Immuno-informatics: Bridging Computation and Immunology," In *Proc. Kuala Lumpur International Conference on Biomedical Engineering (BioMed 2004)*, Malaysia, 2004.
[9] http://www.cs.unm.edu/~immsec/data-sets.htm
[10] A. Iqbal, and M. A. Maarof, "Towards Danger Theory based Artificial APC Model: Novel Metaphor for Danger Susceptible Data Codons," Italy: Springer-Verlag, 2004, ch. 13.
[11] P. Matzinger, "The Danger Model: A Renewed Sense of Self," *Science,* vol. 296, pp. 301-305, 2002.
[12] P. Matzinger, "The Danger Model in Its Historical Context," *Scand. J. Immunol,* vol. 54, pp. 4-9, 2001.
[13] S. Gallucci, M. Lolkema, and P. Matzinger (1999), Natural Adjuvants: Endogenous Activators of Dendritic Cells," *Nature Medicine*, vol. 5, no. 11, pp. 1249-1255, 1999.
[14] P. Matzinger, "The Real Function of the Immune System," Available: URL:http://cmmg.biosci.wayne.edu/asg/polly.html, 06-04-04.
[15] P. Matzinger, An Innate sense of danger, Seminars in Immunology, vol. 10, pp. 399-415, 1998.

A. Iqbal (M'04) became a Member (M) of IJIT in 2004. He was born at Jhelum (Pakistan) on June 2, 1963. He completed bachelor degree in 1986 from the University of Punjab Pakistan in mathematics and physics. A. Iqbal received Masters in electronics in 1989 from Quaid-i-Azam University Islamabad Pakistan. From May, 1990 to August, 1991 he received post graduate training in computer hardware and system software from Computer Training Center (CTC) Islamabad Pakistan.

He started his professional career in March 1989 as a COMPUTER HARDWARE ENGINEER in Computer Marketing Company (CMC) Pvt. Ltd. Lahore Pakistan. He served Informatics Complex Islamabad Pakistan as SCIENTIFIC OFFICER (1991-1995), SENIOR SCIENTIFIC OFFICER (1995-2002). Currently, he is PhD RESEARCH FELLOW in the Group on Artificial Immune Systems N Security (GAINS), Faculty of Computer Science and Information Systems (FSKSM), Universiti Teknologi Malaysia, Johor, Malaysia. His recent research interests include artificial immune systems, intrusion detection, immunoinformatics. Previously he experienced image processing, database design and development, and interfacing.

Mr. Iqbal is a member of international scientific committee for IJIT that conducts peer review of submitted articles.