A Systematic Literature Review Of Unknown Virus Detection And Artificial Immune System (AIS)

Irfan Iqbal

Abstract: Idea of un-known virus detection using Artificial Immune System was introduced at Fourth International Workshop on Synthesis and Simulation of Living Systems. This study is a Systematic Literature Review on "Un-known Virus Detection using Artificial Immune System". By detailed review of the relevant studies we have realized that combination of the techniques may lead to the better accuracy and efficiency.

Index Terms: artificial immune system, virus detection, Systematic Literature review, research methodology, Artificial Neural Network, Intrusions and anomaly detection, Artificial Intelligence.

1. INTRODUCTION

Various Artificial Intelligence techniques has been applied for Virus Detection, including Heuristics, Data Mining, Artificial Immune and Artificial Neural Networks [1]. Artificial Immune is a biologically inspired technique used by different researchers to detect the unknown viruses, intrusions and anomaly detection. In 1994 and 1996, IBM's Thomas J. Watson Research Center (in Yorktown Heights, New York), the Anti-Virus science and technology group initiated a work on automatic virus detection using an Artificial Immune System (AIS). Proposed system was able to detect and eradicate un-known viruses [2, 3]. Then after in 1997, the same group proposed an immune system for generation of un-known pathogen prescription. Staying with in the domain of the AIS, researchers used different intelligent techniques to implement their ideas. These techniques include: Negative Selection, Multi Agent, Chromosome based AIS, Danger Theory, Clonal Selection, Evolutionary Methods, Neural Networks, FSM Hidden Markov Models and Apoptosis [4-12]. Combination of multiple techniques has also been used to improve the detection rate and performance of the system [4]. Dynamic model for self/ non-self reduces the self set to half and evolutionary gene library generates efficient detectors which lead to the reduced detection time. It also reduces the False Positive and False Negative rates [13]. Another improvement was made by the introduction of GA based on RNA (Reverse Transcription Ribonucleic Acid) which provided memory such that during a complex search, algorithm can revert back and attempt to mutate in different direction in order to escape the local minima [5]. Latest technique proposed is Tri Tier Immune System which splits the AIS into three constituent parts i.e. Inherit, Adaptive and Parallel Immune Tier [14].Our Systematic Literature Review process is inspired by the guidelines proposed by Kitchenham [15].Methodology will be described in section 2. Results will be presented in section 3. Section 4 will discuss the answers to the research questions and section 5 is for the conclusions.

- *Irfan Iq bal is is currently working as research assistant and Docent in Computer Science Department in Qassim University, Saudi Arabia.*
  Email: e.eqbal@qu.edu.sa.

2 METHOD

In this section, we will describe the search process step by step.

2.1 Research Questions

The research questions addressed in this Systematic Literature review are:

RQ1. What do the previous studies say about how to detect "Un-known Virus using Artificial Immune System"?
RQ2. What is the accuracy and efficiency level of existing techniques to detect the un-known viruses?
RQ3. How accuracy and efficiency of the existing techniques can be improved?

Addressing to RQ1, there has been a lot of activity in this area since 1994. It started in 1994 when a publication was presented by IBM and they followed it in 1996 and 1997. There has been an increased interest in this area in the recent years due to increased virus complexity (Polymorphic Viruses) and threat rate. As conventional anti virus tools and techniques are unable to give some comprehensive solution [16].

While answering the RQ2 we shall focus on accuracy and efficiency issues of different techniques which have their background in Artificial Intelligence. RQ3 will be addressed in the conclusion remarks.

2.2 Search Process

We have conducted our search process using digital libraries/electronic resources provided by BTH (Inspec/Compendex, Libris, Ebrary, Google Scholar, ISI Web of Science and Scopus). Inspec/Compendex provides indirect links to the specific databases to which some article or paper actually belongs. So we searched our topic in selected journals, conference proceedings and Workshops as shown in Table 1. Then a team of two researchers had an insight of the results and established their relevance to the topic.

TABLE 1

SELECTED JOURNALS AND CONFERENCE PROCEEDINGS

Source	Acronym

IEEE Transactions on Evolutionary Computation	ITEC
Artificial Life and Robotics	ALAR
Journal of Computational Information Systems	JCIS
IEEE Parallel & Distributed Technology: Systems & Applications	IPDT
IEEE International Conference on Cybernetics and Intelligent Systems	CIS
IEEE Congress on Evolutionary Computation	ICEC
IEEE World Congress on Intelligent Control and Automation	IWCICA
Proceedings of the IEEE International Conference on Systems, Man and Cybernetics	IICOS
Proceedings of the International Symposium on Test and Measurement	ISTM
Proceedings of International Conference on Computational Science	ICCS
Proceedings of IEEE International Conference on Machine Learning and Cybernetics	IICML
Proceedings of International Conference Cryptology and Network Security	CANS
Proceedings of Genetic and Evolutionary Computation Conference	GECCO
Proceedings of International Conference on Artificial Immune Systems	ICARIS
Proceedings of Bio-Inspired Models of Network, Information and Computing Systems	BIONETICS
Workshop on Biologically Inspired Approaches to Advanced Information Technology	BioADIT
IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems	DAACS
Symposium on the Immune System and Cognition	AISB

2.3 Inclusion and Exclusion Criteria

We have devised our inclusion and exclusion criteria by answering the following questions:

- Is article found has discussed Artificial Immune System and virus detection?
- Is article not a duplicate version of already in the selected list?

2.4 Quality Assessment

Quality assessment criteria for our review based on following questions:

QA1. Is literature's proposed idea is purely based on Artificial Immune System and virus detection?
QA2. Is there any experiment included in research paper?
QA3. Does the author discuss the performance issues of the proposed techniques?

The quality questions are scored as follows:

QA1: Y (yes), the author explicitly discusses the AIS and virus detection, P(Partly), the author implicitly discusses the AIS and virus detection.
QA2: Y (yes), experiment is included, N(no), experiment is not included.
QA3: Y (yes), the author explicitly discusses the performance issues, P(Partly), the author implicitly discusses the performance issues between the lines, N(no), not discussed at all.

The scoring values are as under:

Y = 1, P = 0.5 and N = 0 as shown in Table 3.

We have assessed all the literature once again on the basis of quality questions and literature was reviewed by the participating researchers and they independently given their ratings of the papers. Some times it happened that they all did not reach an agreement.

3 RESULTS

3.1 Search Results

Queries based on different criteria produced the search results presented in Table 2. Selected relevant studies are presented in Table 4.

TABLE 4
SELECTED ARTICLES

ID	Title	Source	Ref
S1	Artificial immune system against viral attack	Computational Science - ICCS 2004	[17]
S2	An investigation of immune match, immune memory and complement operator in immune detection algorithm	Proceedings of the World Congress on Intelligent Control and Automation (WCICA)	[18]
S3	Multi-agents artificial immune system (MAAIS) inspired by danger theory for anomaly detection	Proceedings of the IEEE International Conference on Micro Electro Mechanical Systems (MEMS)	[6]
S4	Biologically inspired computer virus detection system	First International Workshop, BioADIT 2004	[19]
S5	Artificial immunity-inspired script-virus detection model	Journal of University of Electronic Science and Technology of China	[20]
S6	An antigen presenting cell modeling for danger model of Artificial Immune System	AISB 2004 Convention: Motion, Emotion and Cognition	[7]
S7	An artificial immune system-inspired multiobjective evolutionary algorithm with application to the detection of distributed computer network intrusions	Proceedings of GECCO 2007: Genetic and Evolutionary Computation Conference	[21]
S8	An artificial immune system architecture for computer security applications	IEEE Transactions on Evolutionary Computation	[22]
S9	An immunological approach to computer viruses detection	Computing and Information Systems	[9]
S10	Research of Trojan detection system based on artificial immune	Journal of University of Electronic Science and Technology of China	[23]
S11	Defending virus infection through extrinsic apoptosis	International Symposium on Information Technology	[10]
S12	A virus detection model based on immunology and SVDD	Journal of Computational Information Systems	[24]

S13	Neural networks for artificial immune systems: LVQ for detectors construction	4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS	[4]
S14	Immunity based virus detection with process call arguments and user feedback	Proceedings of the Bio-Inspired Models of Network, Information, and Computing Systems	[25]
S15	Tri-tier immune system in anti-virus and software fault diagnosis of mobile immune robot based on normal model	Journal of Intelligent and Robotic Systems: Theory and Applications	[14]
S16	A chromosome-based evaluation model for computer defense immune systems	Congress on Evolutionary Computation(IEEE Cat. No.03TH8674)	[12]
S17	An Artificial Immune Infrastructure for Network-attached Data Acquisition System	Proceedings of the International Symposium on Test and Measurement	[26]
S18	Combating computer viruses: IBM's new computer immune system	IEEE Parallel & Distributed Technology: Systems & Applications	[2]
S19	A retrovirus inspired algorithm for virus detection & optimization	GECCO 2006 - Genetic and Evolutionary Computation Conference	[5]
S20	Realization of a self-recognition algorithm based on the biological immune system	Artificial Life and Robotics	[8]
S21	A clonal selection algorithm by using learning operator	International Conference on Machine Learning and Cybernetics	[11]
S22	Immune system for cyberspace	Proceedings of the IEEE International Conference on Systems, Man and Cybernetics	[27]
S23	Evolution of self-definition	IEEE International Conference on Systems, Man, and Cybernetics	[28]

3.2 Quality Evaluation

Table 3 gives the quality score of each study based on the quality criteria defined in sub section 2.4.

TABLE 2
SEARCH RESULTS

Database	Search Query	No. of Results	Reflections
Inspec/ Compend ex	unknown AND virus detection	165	Irrelevant included
Inspec/ Compend ex	artificial immune system	3753	Too many records
Inspec/ Compend ex	artificial immune system AND virus detection	44	English 1996-2008

TABLE 3
QUALITY SCORE

Study	QA1	QA2	QA3	Total Score
S1	Y	Y	Y	3.0
S2	P	Y	Y	2.5
S3	Y	Y	P	2.5
S4	Y	N	Y	2.0
S5	Y	N	P	2.0
S6	Y	Y	Y	3.0
S7	P	N	P	1.0
S8	P	N	Y	1.5
S9	Y	Y	Y	3.0
S10	Y	N	N	1.0
S11	Y	Y	Y	3.0
S12	P	Y	Y	2.5
S13	P	Y	Y	2.5
S14	P	N	P	1.0
S15	Y	Y	Y	3.0
S16	Y	N	N	1.0
S17	Y	N	P	1.5
S18	Y	N	P	1.5
S19	Y	Y	Y	3.0
S20	Y	Y	Y	3.0
S21	Y	Y	Y	3.0
S22	Y	N	P	1.5
S23	Y	Y	Y	3.0

4 SYNTHESIS

Overall we have identified 23 relevant articles out of which 17 are purely discussing the topic under consideration while rest of the studies is discussing the topic partially. While going through the rigorous details of our final findings we have tried to aim our focus on the accuracy and efficiency of the techniques used by different researchers. Main techniques used by authors are: Negative Selection, Multi Agent, Chromosome based AIS, Danger Theory, Clonal Selection, Evolutionary Methods, Neural Networks, FSM Hidden Markov Models and Apoptosis. Combination of multiple techniques has also been used to improve the detection rate and efficiency of the proposed systems. Out of findings 14 articles are explicitly addressing the performance results. While in rest

of the studies results are available implicitly. As we have found evidences from the previous studies that combination of different techniques may help to device such a strategy which may lead to improvement in detection and efficiency.

5 CONCLUSION

It has been observed from the study that whenever some researcher has devised some combination of techniques he has realized good results as compared to the single technique. So we are of the opinion that one focus of the future research should be analyzing the diverse combination of different techniques.

6 LIMITATION OF SYSTEMATIC REVIEW

As we have explored the electronic resources provided by our institution, there may be a limitation in a way that we may not have access to some resources which may have articles relevant to our study area. Another limitation is that some of the articles are available in the provided resources but their full contents are not accessible. If all the reviewers do not reach an agreement about the quality of some article, there is no systematic process to handle this problem. This may result in exclusion of some really relevant article.

ACKNOWLEDGMENTS

My thanks and sincere appreciation go to Professor Dr. Yasir Muhammad Alghafeeli and professor Dr. Ali Hasan Husien Alahmadi, the finest teachers and mentors we could possibly want. Much thanks to Dr. Awwadh who provided continuous encouragement throughout my research work.

REFERENCES

[1] X.-B. Wang, et al., "Review on the application of artificial intelligence in antivirus detection system," Chengdu, China, 2008.

[2] S. Hedberg, "Combating computer viruses: IBM's new computer immune system," IEEE Parallel & Distributed Technology: Systems & Applications, vol. 4, pp. 9-11, 1996.

[3] J. O. Kephart, "A biologically inspired immune system for computers," Cambridge, MA, USA, 1994, pp. 130-9.

[4] S. Bezobrazov and V. Golovko, "Neural networks for artificial immune systems: LVQ for detectors construction," Dortmund, Germany, 2007, pp. 180-184.

[5] K. S. Edge, et al., "A retrovirus inspired algorithm for virus detection optimization," Seattle, WA, United states, 2006, pp. 103-110.

[6] H. Fu, et al., "Multi-agents artificial immune system (MAAIS) inspired by danger theory for anomaly detection," Harbin, Heilongjiang, China, 2007, pp. 570-573.

[7] A. Iqbal and M. A. Maarof, "An antigen presenting cell modeling for danger model of Artificial Immune System," Leeds, UK, 2004, pp. 43-4.

[8] S. Kwee-Bo, et al., "Realization of a self- recognition algorithm based on the biological immune system," Artificial Life and Robotics, vol. 7, pp. 32-9, 2003.

[9] M. Samy, et al., "An immunological approach to computer viruses detection," Computing and Information Systems, vol. 12, pp. 1-12, 2008.

[10] M. M. Saudi, et al., "Defending virus infection through extrinsic apoptosis," Piscataway, NJ, USA, 2008, p. 5 pp.

[11] Y. Ying and H. Chao-Zhen, "A clonal selection algorithm by using learning operator," Piscataway, NJ, USA, 2004, pp. 2924-9.

[12] W. Zejun, et al., "A chromosome-based evaluation model for computer defense immune systems," Piscataway, NJ, USA, 2003, pp. 1363-9.

[13] T. Li, et al., "An immune-based model for computer virus detection," Xiamen, China, 2005, pp. 59-71.

[14] T. Gong and Z. Cai, "Tri-tier immune system in anti-virus and software fault diagnosis of mobile immune robot based on normal model," Journal of Intelligent and Robotic Systems: Theory and Applications, vol. 51, pp. 187-201, 2008.

[15] B. Kitchenham, "Procedures for performing systematic reviews," Keele University, technical report2004.

[16] L. Hyungjoon, et al., "Biologically inspired computer virus detection system," Berlin, Germany, 2004, pp. 153-65.

[17] L. Hyungjoon, et al., "Artificial immune system against viral attack," Berlin, Germany, 2004, pp. 499-506.

[18] Y.-J. Zhang and H.-W. Leng, "An investigation of immune match, immune memory and complement operator in immune detection algorithm," Dalian, China, 2006, pp. 3618- 3622.

[19] J. O. Kephart, et al., "Biologically inspired defenses against computer viruses," San Mateo, CA, USA, 1995, pp. 985-96.

[20] C.-m. Liu, et al., "Artificial immunity-inspired script-virus detection model," Journal of University of Electronic Science and Technology of China, vol. 36, pp. 1219-22, 2007.

[21] C. R. Haag, et al., "An artificial immune system-inspired multiobjective evolutionary algorithm with application to the detection of distributed computer network intrusions," London, United kingdom, 2007, pp. 2717- 2724.

[22] P. K. Harmer, et al., "An artificial immune system architecture for computer security applications," IEEE Transactions on Evolutionary Computation, vol. 6, pp. 252-80, 2002.

[23] L.-t. Chen and L. Zhang, "Research of Trojan detection system based on artificial immune," Journal of University of Electronic Science and Technology of China, vol. 34, pp. 221-4, 2005.

[24] C. Liu, et al., "A virus detection model based on immunology and SVDD," Journal of Computational Information Systems, vol. 3, pp. 2043-2048, 2007.

[25] Z. Li, et al., "Immunity based virus detection with process call arguments and user feedback," Budapest, Hungary, 2007, pp. 57- 64.

[26] D. Yonggui, et al., "An Artificial Immune Infrastructure for Network-attached Data Acquisition System," Shenzhen, China, 2003, pp. 535-538.

[27] J. O. Kephart, et al., "Immune system for cyberspace," Orlando, FL, USA, 1997, pp. 879- 884.

[28] C. M. Kennedy, "Evolution of self-definition," New York, NY, USA, 1998, pp. 3810-15.