REVIEW

From Assoc.Prof. Veselina Gospodinova Zhecheva
Faculty of Computer Science and Engineering,
Burgas Free University

About PhD Dissertation
„MODELING OF NETWORK ATTACKS AND PROTECTION ALGORITHMS“

Of PETYA IVANOVA PETROVA,
For awarding the educational and scientific degree "Doctor"
In the professional field 4.6. "Informatics and Computer Science"

This review was prepared according to an Order of the Rector of Burgas Free University to open a procedure for the defense of the dissertation of Petya Ivanova Petrova, PhD student in self-study at the Centre for Computer and Technical Sciences of BFU. For this purpose, I have been provided with the relevant materials, including the dissertation, abstract, the cited order, minutes of doctoral exams, publications of the PhD student, etc. All of them individually and collectively fully meet the legal requirements and current practice of Bulgarian universities.
General presentation: the PhD student graduated from Burgas Free University, majoring in Informatics and Computer Science. She worked in the real sector, which helped her gain useful practical experience.

Subject of the PhD work. The presented text of the PhD dissertation is on a topic that is undoubtedly very relevant due to the fact that digital technologies have entered all spheres of public life. Especially during the crisis related to the global pandemic, online business services have become a major area of business development. This has led to an increase in security risks, although this issue has become increasingly important in recent decades. Offering the growing volume of online services leads to the expected high level of information and network security. The object of research in this dissertation - network security - corresponds to these trends in the IT business. All this determines the relevance of the topic of the dissertation. On this basis, I assess the topic as dissertable and very relevant for the IT sector needs.

Methodology. The applied methodology includes artificial intelligence methods, namely genetic algorithms and differential calculus.

Content. The dissertation has a volume of 114 pages and consists of an introduction, 3 chapters, a conclusion, a list of publications on the dissertation, a declaration of originality, an appendix and a list of references. A list of used figures and frequently used abbreviations is also included.

The purpose and tasks of the dissertation are formulated in the **Introduction**. An overview of the modeling of computer systems in a network environment, as well as modeling of security and its violation is made. Some methods for assessing the impact of malware on networks are described, as well as various types of unregulated impact on the system. Intrusion Detection Systems are described as a means of ensuring the system security and various algorithms with applications in the security field are considered. The need for research on improving security methods is justified. The **purpose** of the dissertation is mathematical modeling of processes in computer networks under the influence of malware and unregulated user behavior, construction and application of genetic algorithms for detecting intrusions into the computer network and data protection. According to this justification, one theoretical and three scientific and practical **tasks** are set, namely:

- Modeling of processes affected by a computer network with malware.

- Development and application of genetic algorithms for detecting intrusions into the computer network.
- Determining the fitness function in the genetic algorithm for detecting intrusions into computer networks.
- Implementation of network security by encryption with a genetic algorithm.

Chapter 1, entitled "Modeling Processes in a Computer Network Affected by Malware", contains research and development of a mathematical model of the behavior of the computer network and the dynamics of its components that are susceptible, exposed, infected and recovered after a malware attack. . To implement the mathematical model, tools from the field of systems of differential equations were used, which describe the state of the network. The SEIR model, used to describe the spread of biological viruses, is used to model the computer worm behavior in the network. The different states of the nodes are described, as well as the transitions between the states through the recovery functions and the dynamics of the node types. A solution of the system of differential equations is presented and the extreme values of the characteristics of the computer network in case of an attack with malicious software at equilibrium and nonequilibrium state of the system are determined. The results of numerical experiments at selected values of the considered parameters are presented and the dependences of some characteristics of the system (vulnerability, recoverability, exposure, etc.) on the time of equilibrium and imbalance of the system are examined. It is concluded that in endemic equilibrium, ie. in the absence of security software, the indicators characterizing the behavior of the computer network increase with time, while in imbalance, i.e. availability of security software, the parameters have a complex dependence on time.

Chapter 2 "Building a Genetic Algorithm for Detecting Intrusions into a Computer Network" describes a method for detecting violations based on a genetic algorithm. An overview of the classic intrusion and prevention systems has been made, and some basic types of attacks have been considered. The basic concepts of genetic algorithms and their application in network attacks are described. Predicate logic is used to describe the rules, and the states are represented by network characteristics (IP addresses of source and destination and port numbers, network protocols used, duration of communication, etc.). The used attributes of the rules with their types, ranges of values and description, as well as the chromosomes and the format of their codes are presented. Exemplary rules and their purpose are described. The main operations on chromosomes of a given generation are presented - selection, crossing and mutation and assessment of the evolution of the current generation of chromosomes through defined fitness functions. A pseudocode of the genetic algorithm is also included.

Chapter 3 Implementation of network security using an encrypted genetic algorithm is dedicated to an algorithm for data security during transmission over the network. An improvement of the classical RSA algorithm by selecting the initial parameters is presented. The application of the genetic algorithm is described. The algorithm is described in detail and its implementation in C # and the corresponding block diagram are presented.

Contribution. The reviewer accepts the defined by the PhD student scientific and scientific-applied results, namely:

1. SCIENTIFIC RESULTS

1.1. A model of the processes of susceptibility, exposure, infection and recovery of a computer network in case of malicious software, proposed with a system of differential equations for instantaneous and predictive assessment of the computer network - a system of equations.

1.2. An original solution of the system of differential equations in two cases - equilibrium for constant variables - analytical expression (37) - Chapter I and lack of equilibrium for time-dependent variables - analytical expression (65) - Chapter I, defining machine state classes in the computer network.

1.3. From the solution of the system of inhomogeneous differential equations are derived analytical expressions for calculating the network characteristics in case of susceptibility, exposure, infection and recovery (reconstruction) of machines in the computer network during an attack with malicious software.

2. SCIENTIFIC AND APPLIED RESULTS

2.1. A software tool has been developed to protect a computer network by encrypting the transmitted information using the operators of a genetic algorithm implemented with the C # programming language (Chapter III).

2.2. Software products implemented in the Matlab environment have been developed to illustrate the solutions of the system of differential equations for instantaneous and predictive estimation.

3. APPLIED RESULTS

3.1. The list of attributes, gene structure and codes of the network characteristics of the chromosome (rule) in the genetic algorithm has been expanded, expanding the scope of its action to detect network penetrations.

3.2. The structure of the fitness function of Firas Alabsi was evaluated and the basic operations on chromosomes of a given generation were illustrated with data obtained from a numerical experiment.

3.3. Experimental evaluations of the parameters A, AB of the fitness function of Firas Alabsi were made with data obtained from simulation of network communications of the type Normal, DoS, R2L, U2R, Probe, realized with randomly generated network characteristics of five chromosomal structures for each category .

These contributions build on existing analyzes and open new directions for research in the current and rapidly changing field of information security.

Publications and participation in scientific events. Six publications are presented, one of which is a report at a conference in Bulgaria and 5 are articles in scientific journals (two in SCOPUS indexed publications and two in the NACID list). Three of the publications are in English and three in Bulgarian. Three of the publications are indexed in SCOPUS, one with only author being the PhD supervisor (https://www.scilit.net/article/3a8dde1b1d44362e92cf2d2581715a50). The attached report shows that the candidate meets the minimum requirements for publications.

The reference list includes 114 sources, of which 98 articles and books and 16 - Internet resources. The bibliography is complete and up-to-date, most sources are in English, and several are in Bulgarian. We will note that the presented publications on the dissertation are also included in the list of used literature. It should also be noted that less than half of the articles and books (41 sources) are from the last 10 years, which is not much in a rapidly changing field such as information security.

The dissertation is diligently designed, all tables and figures are numbered. The summary is 36 pages long and correctly reflects the content of the dissertation.

In general, this dissertation meets the standard requirements for a doctoral dissertation, shows high academic erudition and practical skills of the doctoral student and deserves positive evaluation. In general, the structure of the dissertation is well developed, in a logical sequence of the individual sections with a relatively weak connection between them. The separate chapters are devoted to research on various aspects of security in relation to various threats. As strengths of the dissertation I define as follows:

- **Its relevance** - In the conditions of rapidly changing conditions in cyberspace, this study has significant practical / applied value.
- **The methodological approach and the used mathematical methodology** - The dissertation uses methods in the field of artificial intelligence (genetic algorithms) and differential calculus, which reveals good theoretical basis.

Related to the dissertation the following remarks, questions and comments can be made:

1. The review is more a place for a general introduction to the topic, justification and formulation of the purpose and objectives of the dissertation, rather than going into detail in the presentation of different approaches and algorithms.
2. There is a lack of systematic terminological review of the malicious means / types of attacks of modern systems, as well as a comparison of the various impacts and their frequency in the modern world. From there, a more in-depth justification of the selected models and studies could follow.
3. Chapter 1 does not clarify the impact of which security software is being tested, nor does it describe the units and intervals of change.
4. Point 2.9 states "As a future activity of the author… provides for the establishment of a database (rules, knowledge) for new unexplored structures of network characteristics of computer attacks, as well as methods for their response and prevention." In general, a database with signatures of known malicious means has been used for years. Can the PhD student explain more about the idea of building a base for new unknown means of attack? Database, database with rules and knowledge are different concepts, which one is envisaged?
5. The conclusion of Chapter 3 states that "the developed genetic algorithm has reliable and robust encryption with a high degree of protection of the transmitted message and can be applied in the exchange of data requiring a high level of security". However, no data from an experiment supporting this claim have been provided and no comparison with standard RSA has been made.

Conclusion

The above remarks and comments are not intended to downplay the work, which is a serious study in the field of information security through the methods of differential equations and genetic algorithms. The presented dissertation meets the requirements of the Regulations for the development of the academic staff at the Burgas Free University. Based on the above, I believe that the candidate Petya Ivanova Petrova meets all the criteria and requirements of the Law on the Development of Academic Staff in the Republic of Bulgaria, the Regulations for its implementation and the Regulations on the terms and conditions for obtaining scientific degrees and holding academic positions at the Burgas Free University and I propose to award her the educational and scientific degree "Doctor", professional field 4.6. "Informatics and Computer Science".

Signature:
/Assoc.Prof. V.Jecheva, PhD/

Burgas
20.05.2022 г.