Artificial Intelligence Methods for Cyber Threats Intelligence

ROUMEN TRIFONOV[1], SLAVCHO MANOLOV[1], RADOSLAV YOSHINOV[2],
GEORGI TSOCHEV[1], GALYA PAVLOVA[1]
[1]Technical University of Sofia,
8 Kl. Ohridski bul.,Sofia 1000,
[2]Telematics Laboratory
Bulgarian Academy of Sciences, Sofia
Akad.G.Bonchev St, bl. 8
BULGARIA
r_trifonov@tu-sofia.bg, slav1943@gmail.com, yoshinov@cc.bas.bg, gtsochev@tu-sofia.bg,
raicheva@tu-sofia.bg

Abstract: Following ENISA's findings on the two main trends in Cyber Defence development over the past few years - adopting the philosophy and methods of Military Intelligence and introducing Artificial Intelligence into technologies for counteraction of cyber attacks - the Faculty of Computer Systems and Technology at Technical University of Sofia undertook research on the application of intelligent methods for increasing the security in computer networks. While in the field of Tactical Cyber Threats Intelligence the research has already passed into the real-world prototyping phase, in the sphere of Operational Cyber Threats Intelligence (as in the international research community) the research is still at an early stage.

Key-Words: Cyber Threats Intelligence, Tactical, Operational, Artificial Intelligence, Multi-Agent Systems, Intrusion Detection and Prevention Systems, Behavioural Model, Machine Learning, Neural Networks, Reservoir Computing, Sequential Feature Selection

1 Introduction

The remarkable Cyber-Threat study conducted by European Network and Information Security Agency (ENISA) [1] is complemented by a series of conclusions and recommendations addressed to policy makers, business and research community. The first two research conclusions read as follows:

- definition of research roadmaps for Artificial Intelligence in Cyber Threat Intelligence. This could include (but not restricted to) attack pattern recognition and knowledge discovery and enrichment of cyber-threat context;
- development of security models based on agility/dynamics of Cyber Threats. This should also include the use of Cyber Threat Intelligence to assess efficiency and performance of implemented security controls.

These conclusions adequately reflect the radical changes over the past three-four years in the Landscape of the Cyber Threats Defense, expressed in two distinct trends.

The first one is concluded in the following: the conventional network defense tools such as intrusion detection systems and anti-virus focus on the vulnerability component of risk, and traditional incident response methodology became insufficient for certain actors because of the evolution in the goals and sophistication of computer network intrusions. A new class of threats, appropriately dubbed the “Advanced Persistent Threat" (APT), represents well-resourced and trained adversaries that conduct multi-year intrusion campaigns targeting highly sensitive economic, proprietary, or national security information. These adversaries accomplish their goals using advanced tools and techniques designed to defeat most conventional computer network defense mechanisms. Network defense techniques which leverage knowledge about these adversaries can create an intelligence feedback loop, enabling defenders to establish a state of information superiority which decreases the adversary's likelihood of success with each subsequent intrusion attempt. According to the vast majority of experts, the qualitative transition to new cyber defense tools must involve the widespread use of artificial intelligence methods to analyze information exchanged, network flows, sources of threats, and to plan effective impact measures, including proactive ones.

The other direction is the widespread use in Cyber Defense of the techniques and methods of traditional military science and military intelligence, including so-called “kill chains”. The term "kill chain" was originally used as a military concept related to the structure of the attack. The idea is to effectively counteract the opponent in the various phases of the attack or as a preventive action. The computer specialists at the Lockheed-Martin Corporation [2] are adapting this concept to information security, using it as a method of modeling penetration into a computer network. This model is gradually being adopted in the information security community for data protection by identifying cyber stages and corresponding countermeasures at each stage.

The "kill chain" model developed by Lockheed-Martin includes the following stages: Intelligence; Creation of the weapon; Delivery; Operation; Installation; Command and Control and Goal actions. Using a kill chain model to describe phases of intrusions, mapping adversary kill chain indicators to defender courses of action, identifying patterns that link individual intrusions into broader campaigns, and understanding the iterative nature of intelligence gathering form the basis of intelligence-driven computer network defense. Institutionalization of this approach reduces the likelihood of adversary success, informs network defense investment and resource prioritization, and yields relevant metrics of performance and effectiveness.

Following these trends, the Faculty of Computer Systems and Technology at Technical University of Sofia began research on the application of intelligent methods for increasing the security in computer networks. An essential section of this investigation is dedicated to the Cyber Threat Intelligence. The present article summarizes some results of a research done by the project team.

2 Basic Features of the Cyber Threats Intelligence Problem Formulation

The Cyber Intelligence or, more precisely, Cyber Threats Intelligence (CTI) has the following definition in the draft Bulgarian National Cyber Security Strategy [3]:

- establishing mechanisms and technical means to maintain an up-to-date picture of possible threats of different scale, sources and character, trends in geopolitical context development and relevant national cyber picture analysis and;
- development of capabilities to help identify attribution sources and take appropriate forms of protection and counteraction.

According to the documents of INSA (Intelligence and National Security Alliance) [4, 5, 6] the preparation of the intelligence in cyber operational environment is a systematic and continuous process of analyzing potential threats to detect a suspicious set of activities that may endanger systems, networks, information, employees, or customers by providing means to visualize and evaluate a number of specific penetration sensor inputs to bring up a particular threat. This process supports the organization's risk management strategy and decision-making in the area of information security. Its application identifies potential threats and assists security and risk managers selectively implement and maximize deep defense strategies by better understanding the critical points in time and space in the operating environment.

The Cyber Threats Intelligence Cycle [7], shown in Fig. 1, is a continuous process, through all stages of which a feedback and a steady evaluation by management are required.

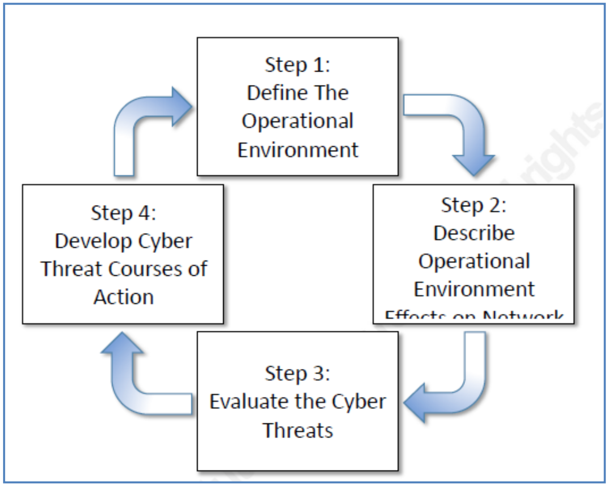

Fig. 1 *Cyber Threat Intelligence Cycle*

The development of a Threat Model is an important element of Cyber Threats Intelligence - in particular, identify the capabilities, intentions, and threat technologies that manage its behavior on the network. The Intelligence Team derives this analysis from information on current and previous threat operations. The knowledge of the possibilities of threat, intentions, technology, doctrine and tactics provides the basis for developing the Threat Model and detecting its vulnerabilities.

Cyber Intelligence Data [8] is the key to providing knowledge indispensable for proactive threat mitigation and protection of information systems data for theft and abuse. The ability to collect and analyze intelligence is realized through log file management tools, security event management, security information management, and file integrity monitoring.

Like its military analogue, the Cyber Threats Intelligence are developed at three levels: strategic, operational, and tactical. For the purposes of this study, the second two are considered:

- INSA defines [6] the operational level as: “The level at which campaigns and major operations are planned, conducted, and sustained to achieve strategic objectives within theaters or other operational areas. At this level, actors build the capabilities needed to support the tactical operations. They maneuver in cyberspace to position capability where they need to in order to be effective in their tactical missions. This is the level where a hacktivist group may plan both cyber and physical world activities to support their objectives;
- the definition of the tactical level is: “The level at which battles and engagements are planned and executed to achieve military objectives assigned to tactical units or task forces. Activities at this level focus on the ordered arrangement and maneuver of combat elements in relation to each other and to the enemy to achieve combat objectives”. The tactical level of the cyber domain is where malicious actors and network defenders maneuver against each other.

In 2016, ENISA developed so called Cyber-Threat Intelligence (CTI) “Big Picture” [1] (Fig. 2). It demonstrates all the elements involved in the attack with the relevant business processes, and shows to which of artifacts (components) the assets involved in the process are targeted.

The “Big Picture” demonstrates the relations with business processes and illustrates the context of different CTI components. It should be noted that issues of detailed knowledge of business processes are key to both attack planning and incident analysis. This contributes to identifying and illustrating the relationship between the various parts associated with CTI, and it is useful for business process analyzers to assess the specific threats to their organization.

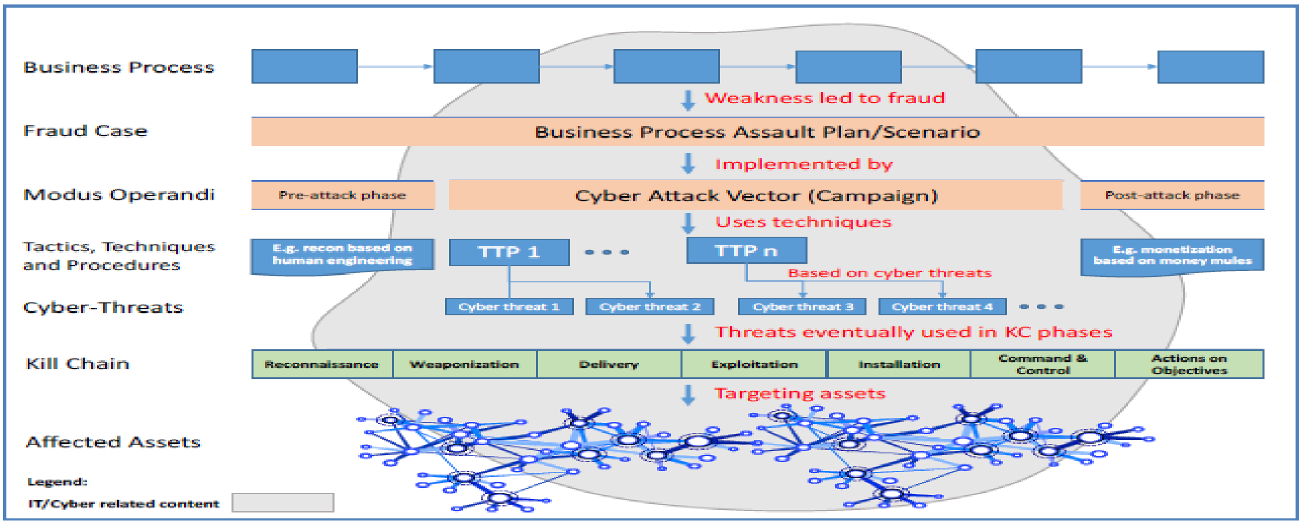

Fig. 2 *Cyber Threat Intelligence “Big Picture”*

3 Methods of Artificial Intelligence in Network and Information Security

As mentioned in the introduction to this article, world practice has already noted a significant number of various "Artificial Intelligence" applications in computer security. Without trying for a comprehensive classification, we could divide these methods into two main directions:

A. Conditionally named "distributed" or "network" methods:

A1. Multi-Agent Systems of Intelligent Agents;

A2. Neural Networks;

A3. Artificial Immune Systems and Genetic Algorithms, etc;

B. Conveniently named "compact" methods:

B1. Machine Learning Systems, including: associative methods, inductive logic programming, Bayes classification, etc.

B2. Pattern recognition algorithms;

B3. Expert Systems;

B4. Fuzzy logic, etc.

Having into account this variety of methods, it is of particular importance that adequate criteria are selected for the assessment and selection of a specific application for each specific solution. In the above mentioned project, the specification was carried out for two of the main sections of CTI.

4 Methods of Artificial Intelligence Suitable for Tactical Cyber Threats Intelligence

The Tactical Cyber Threats Intelligence [5] aims to detect immediate threats against the system and to provide an opportunity for their counter-action. Because of this, the elements of artificial intelligence interact directly with the devices for technical realization of the security policy: Firewalls, Intrusion Detection / Prevention Systems, Anti-Virus Software, Web Gateways and Network Snares.

The identification of attacks is a process of detecting pervasive events occurring during the operation of an information system. Similarly to high responsability process management systems, the requirement to recognize penetrating actions arises at the time of their occurrence and not after their implementation. Simultaneously with the detection of penetration attempts, it is necessary to start a mechanism for preventive actions that are related to the containment or isolation of the action of a source of attack and the activation of an active counteraction in order to block it and bring it into an incapacity.

The type of detection of attack depends on the nature of the threats (knowns, unknowns and combinations of the two types). A set of criteria have been developed for evaluation of the effectiveness of the discovery and the level of counter-performance. It is also extremely important to achieve the right balance between false positive results and false negatives. Incorrect positive results (so-called false alarms) may be no less harmful than false negative results.

During the development of the project, a comparative analysis of different methods of artificial intelligence in view of the above mentioned criteria was performed on bibliographic sources. It has been found that the methodology of abnormal tuned multi-agent systems [9, 10, 11, 12] outclass most traditional systems based on artificial intelligence in detecting attacks, particularly of unknown nature. The effectiveness of detecting hazards in multi-agent systems also outperforms traditional systems [13] (Fig. 3). The most important aspects of multi-agent-based Intrusion Detection and Prevention Systems (IDPS) systems are high precision, self-learning and sustainability [14, 15, 16].

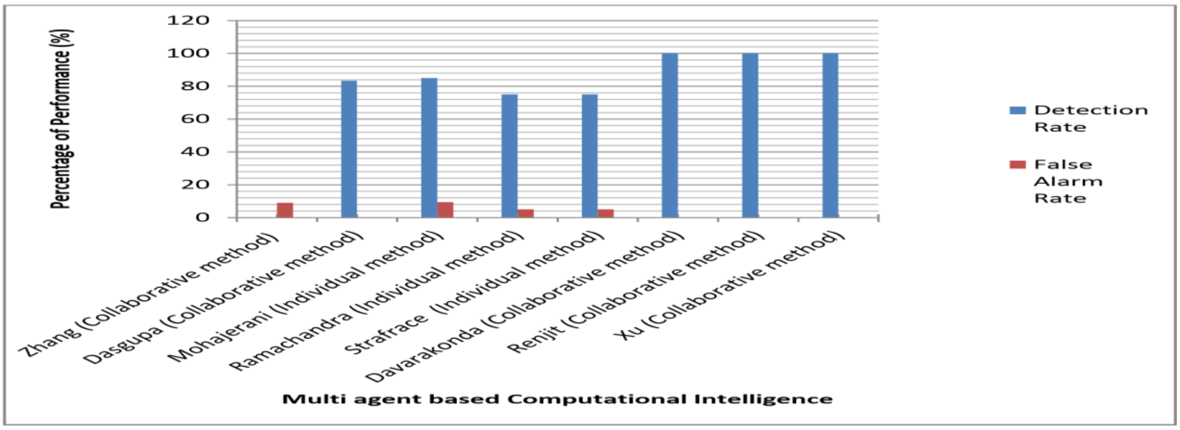

Fig.3 *Value of detection rate and false alarm rate of various multi-agent systems*

The practices described in the sources show that the results of the proper detection of threats using multi-agent-based systems are steadily increasing as the percentage of false alarms drastically decreases. Undoubtedly, multi-agent-based approaches can potentially reach increased flexibility, which will make them even more popular in the near future. Therefore, the experimental model created at the first stage of the project is a combination of multi-agent system and IDPS [17, 18, 19, 20].

Autonomous agents are computing systems that exist in a complex, dynamic environment, act independently in this environment, and thus realize a set of goals and tasks for which they are designed. For the purposes of the experimental model, agents from the Learning Agents class are used. Owing to their learning, they are capable to work independently in an originally unknown environment and become more competent than their initial knowledge. The learning agents consist of four conceptual elements: Learning Element, Performance Element, Critic and Problem Generator.

For our experiment a tentative system named as

Network Gateway Monitoring System (NGMS) (Fig. 4) was built. This is a multi-agent-based software framework, which consists of two parts - Network Prevention (NP) and Host Prevention (HP). The NP component works on the transport layer of the TCP / IP model, and checks the network traffic for detecting and preventing malicious packets and infiltration traces. The HP component works on the application layer of the TCP / IP model and the System Software layer of the operating system, and inspect the operation of the operating system and kernel activity to detect and prevent malicious code.

The experiments have been successfully conducted to verify and evaluate individual components and the entire platform. The proposed system succeeds in detecting attacks and malicious code that target the protected system with high accuracy and real-time. The NP component manages to characterize the normal behavior of the TCP \ IP protocol and to detect the attacks aiming to break the header of the packets. The HP component proved its high ability to protect against malicious code that affects Windows operating systems, no matter if the malicious code is in the kernel or focused on user activity.

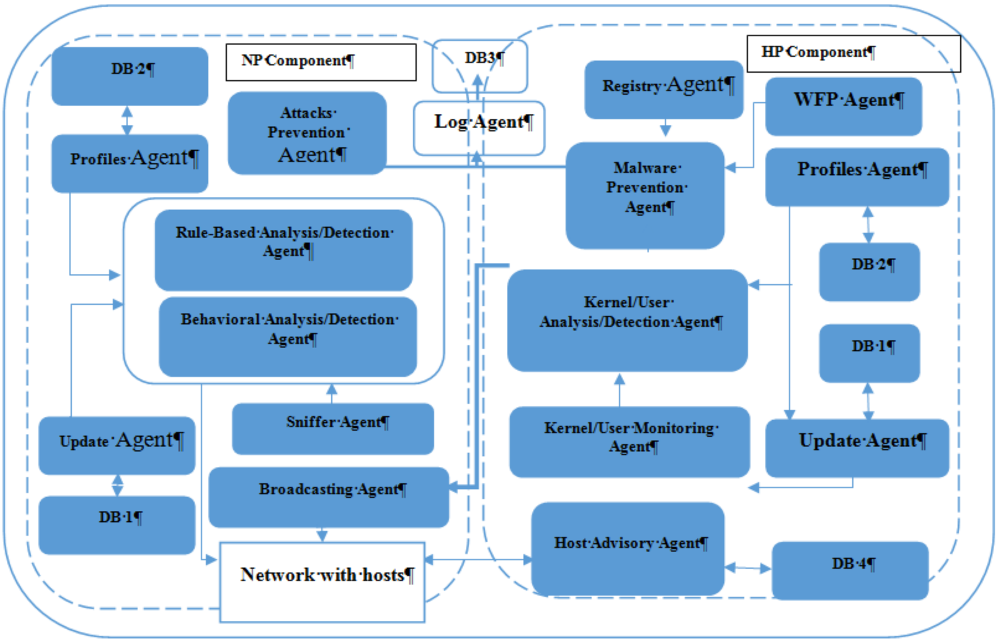

Fig. 4 *The experimental Network Gateway Monitoring System*

5. Methods of Artificial Intelligence Suitable for Operational Cyber Threats Intelligence

The ultimate goal of Operational Cyber Intelligence is to reduce risk to an organization’s critical mission and assets by: defining the operating environment; describing the impact of the operating environment; evaluating the adversary; and determining potential adversarial courses of action (COA). The Operational Cyber Intelligence provides a thread that links the probability and impact of a cyber attack with its strategic level implications by ensuring a coherent framework for analysis and prioritization of potential threats and vulnerabilities given the organization’s threat environment. Operational Intelligence is based on the Doctrine of Active Defense. Instead of searching for information regarding a specific attack against the organization, it focuses on analyzing the opponents' combat doctrines, weapon systems and attack and operational scenarios. This approach shifts the center of gravity to the ability to respond and block the outcome of the attack within the organizational environment or in its immediate vicinity.

Obviously, the basis for the automation of the Operational CTI must be the behavioral model of the likely adversary. It should be emphasized that the problem of using artificial intelligence methods in the Operational CTI is a completely new matter, and systematized literary sources have not yet been found. Only, there are reports concerning the use of behavioral analysis based on machine learning by the companies: Exabeam (USA), Darktrace (UK), CyberX (USA), Interset (Canada).

The TU-Sofia team concluded that the activity and the outgoing traffic in the network of the supposed adversary were to be the main source of information for building his behavioural model. This evokes analogies with the non-invasive brain -

computer interface whereby the physiological signals of the human brain (for example, through Electroencephalograms (EEGs)) can be used for human emotions evaluation [21]. Indeed, the streams of measured parameters received by different IP addresses of the monitored object using RFC 1757 Remote Network Monitoring methods [22] can be compared to EEG with n-number of channels.

If this analogy is applied in practice, first of all, on the order of the classification model of emotions [23], a basic classification of the behavior of the possible adversary, based on the needs of our research, must be constructed,. Currently, in the absence of references for such studies, it is assumed that this behavior can be divided for the present into two basic types: hostile and non-hostile.

In order to obtain the best possible performances, it is necessary to work with a smaller number of values which describe some relevant properties of the data retrieved from the network. These values are known as “features”. Features can be aggregated into a vector known as “feature vector”. Thus, feature extraction can be defined as an operation which transforms one or several signals into a feature vector. Identifying and extracting good features from signals is a crucial step, because otherwise the classification algorithm will have trouble identifying the class of these features, i.e., the behavioral state of the possible adversary. According to some researchers [24], it seems that the choice of a proper pre-processing and feature extraction method have more impact on the final performances than the selection of a good classification algorithm.

Therefore, following the analogy of the brain-computer interface, two basic tasks have to be solved:

- to find a suitable approach to selecting characteristics from which to derive features suitable for behavioral interpretation and validation. In doing so, the necessary inter-subject discrimination of the features for the subsequent classification must be ensured;
- to build and optimize an ensemble of classifiers based on trained models to be used to assess behavior.

Based on a study of literary sources, the Echo State Network (ESN) method was proposed as a mechanism for feature selection – this is a class of recurrent neural networks where the so-called “reservoir computing” approach for training is formulated [25]. It was found that using reservoir computing pre-training is beneficial for selecting the most relevant discriminative features and reaching good performance for behavior valence recognition. The main advantage of the ESN is the simplified training algorithm since only weights of the connections from the reservoir to the readout neurons are subject to training. Thus instead of gradient descent learning much faster least squares method can be used.

Exploring the feasibility of training cross-subject classifiers, we have settled on the Sequential Feature Selection (SFS) procedure [26] that reduces the inherent data variability and can lead to a high inter-subject behaviour status recognition accuracy. Starting from an empty set, SFS increments sequentially a new feature that best predicts the class at the current iteration. The process stops when there is no more improvement in the prediction. SFS is a very effective way to identify the dominant behavioral signatures across subjects. However, it is a computational heavy and time-consuming procedure, which was the main motivation to look for a computationally less intensive alternative.

As experiments are in their early stages, it is necessary to point out that the results are encouraging, but it is still too early to declare any definitive conclusions.

6 Conclusion

As can be seen from the above, the process of introducing Artificial Intelligence methods at the different levels of Cyber Threat Intelligence is at very different stages: while in Tactical Intelligence, it has long gone out of the phase of research and experiments and is used for building real effective systems, In the field of Operational Intelligence, these studies are in a very initial phase and require the commitment of substantial resources. Furthermore, the question arises as to the application of possible outcomes of Operational Iintelligence in the activity of Tactical Intelligence systems, which are intended to neutralize the immediate threats to computer systems and networks.

Acknowledgment

This research is conducted and funded in relation to the execution of a scientific-research project № H07/56 “Increasing the level of network and information security using intelligent methods” under the contract with National Science Fund in Bulgaria.

References:

[1] *ENISA Threats Landscape Report* 2016: 15 Top Cyber-Threats and Trends, January 2017

[2] [2] Eric M. Hutchins, Michael J. Cloppert, Rohan M. Amin *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion* Kill Chains Lockheed Martin Corporation 2015

[3] Republic of Bulgaria: *National Cyber Security Strategy "Cyber Resilient Bulgaria 2020"* 2016-03 NCSS Bulgaria final draft v 5 3

[4] *Cyber Intelligence: Setting the Landscape for an emerging Discipline, Intelligence and National Security Alliance* (INSA), 2011

[5] *Operational Level of Cyber Intelligence*, INSA, 2013

[6] *Operational Cyber Intelligence*, INSA, 2014

[7] Brian P. Kime *Threat Intelligence: Planning and Direction*, SANS Institute, 2015

[8] *Advanced cyber-security intelligence*, Quocirca, 2012

[9] Michael Luck, Peter McBurney, Christ Preist *Agent Technology: Next Generation Computing*, AgentLink II, January 2003

[10] S. D. Chi, J.S. Park, K.C. Jung and J.S. Lee Network Security Modeling and Cyber Attack Simulation Methodology, *Lecture Notes in Computer Science*, Vol. 2119, 2001

[11] V. Gorodetski, O. Karsayev, I. Kotenko, I. Khabalov Software Development Kit for Multi-Agent System Design and Implementation, *Lecture Notes in Artifical Intelligence*, Vol. 2296, Springer Verlag, 2002

[12] Molesini, A., Omicini, A., and Viroli, M. Environment in agent-oriented software engineering methodologies, *International Journal on Multiagent and Grid Systems*, 2007

[13] G. Gai, L. Rui, H. Wu, X. Hu An Improved Collaborative Method for Recommendation and Rating Prediction, IEEE International Conference on Data Mining Workshop, 2014

[14] Jai Sundar Balasubramaniyan, Jose Omar Garcia-Fernandez.*An Architecture for Intrusion Detection using Autonomous Agents*, Purdue University West Lafayette, 2007

[15] Taraka D. Peddireddy *Multiagent Network Security System using FIPA-OS*, University of South Carolina, 2011

[16] D. Dasgupta, F. Gonzalez, K. Yallapu, J. Gomez, R. Yarramsettii CIDS: *An agent-based intrusion detection system*, The University of Memphis, 2014

[17] Trifonov R., Manolov S. Tsochev G. Application of multi-agent systems for network and information protection, 28th International Conference on Information Technologies (Info-Tech 2014), Varna, Bulgaria

[18] Tsochev G, Trifonov R., Yoshinov R. Multi-agent framework for intelligent networks, 29th International Conference on Information Technologies (Info-Tech 2015), Varna, Bulgaria

[19] Tsochev G, Trifonov R., Naydenov G. Agent Communication Languages Comparison, 7th International Scientific Conference COMPUTER SCIENCE'2015, Durres, Albania

[20] Tsochev G, Trifonov R., Popov G. A Security Model based on Multi-agent systems, 30th International Conference on Information Technologies (Info-Tech 2016), Varna, Bulgaria

[21] Liu Y., Sourina O. and Nguyen M. K. Real-time EEG-based human emotion recognition and visualization, Proceedings of the Int. Conf. on Cyberworlds (CW '10), Singapore, 2010

[22] [RFC 1757 Remote Network Monitoring Management Information Base, Carnegie Mellon University, February 1995

[23] L. Bozhkov, P. Georgieva Classification models of emotional biosignals evoked while viewing affective pictures, International Conference on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH), Vienna, 2014

[24] Hammon P.S. and Sa V.R. de Preprocessing and meta-classification for brain-computer interfaces, *IEEE Transactions on Biomedical Engineering*, 54(3), 2007.

[25] Lukosevicius M. and Jaeger H. Reservoir computing approaches to recurrent neural network training, *Computer Science Review*, vol. 3, 2009

[26] Guyon I. and Elisseeff A. An Introduction to Variable and Feature Selection, *Journal of Machine Learning Research*, vol. 3, 2003