Assessment of Stability of Algorithms Based on Trust and Reputation Model

Ilya I. Viksnin, Radda A. Iureva, Igor I. Komarov, Anastasia L. Drannik
ITMO University
St. Petersburg, Russian Federation
wixnin@cit.ifmo.ru

Abstract—**Swarm robotic systems are actively developed and widely studied in the world scientific practice. It is expected that multiagent, distributed approach to creating artificial intelligence of autonomous systems will allow to solve a great number of complex problems in the areas of environmental protection, medicine, cleaning, patrolling, etc. Thereby the research of these systems (design and testing) in terms of information security becomes relevant. The key to a wide practical use of swarm robotic systems is the development of specific guidelines and algorithms for the organization of group actions. This research proposes the use of trust and reputation model for information security of swarm system. Swarm's agents generate trust levels to each other basing on the analysis of situation on the kth iteration step of the algorithm and using their sensor devices. On the calculated trust levels the collective recognition of saboteurs is carried out. To perform experiment software simulator was designed. It allows to vary the basic parameters of swarm robotic system (number of agents, number of targets, range of communication, number of saboteurs).**

I. INTRODUCTION

Research of swarm robotic system (SRS) information security (IS) is in the initial stage. There is speculation about the unity of the laws and regularities of interaction between "big" robots and agents with limited functionality in terms of group behavior. Thereafter expediency of research and development of SRS IS on the low-budget ranges and software simulations is obvious.

In accordance with the basic provisions of swarm robotics an object of study is the large group of miniature robots which form a system of decentralized control. The subject of research is the problem of SRS IS, which inherently affect two aspects of the functioning of the autonomous elementary agent: information interaction, which is realized by sending data and movement in space.

Trust is an effective mechanism for SRS decision- making process optimization. Such kind of algorithms is often considered in the literature as applied to P2P - systems related to e-commerce, social networks, etc.

The paper [1] comprehensively examines the general aspects of the usage of the categories of trust in relation to SRS. The basic approach is to enable agents to calculate the amount of trust they can place in their interaction partners. Further, the question is how the agent can gather such information about its counterparts' characteristics. This can be achieved, among other, through inferences drawn from the outcomes of multiple direct interactions with these partners or through indirect information provided by others.

Witkowski et al. [2] proposed a model whereby the trust is calculated based on the agent's behavior in previous actions. The paper refers to on-line trading and the agents select who they will trade with primarily on the basis of trust measure built on past experiences of trading with those individuals.

Papers [6], [7], [8] and [9] consider the various aspects of using trust for detecting deceptive agents in artificial societies. In [9], dealt with Consumer-to-Consumer Electronic Marketplaces, the problem of trust in online communities is studied. It is meant that there are agents in the system which intentionally distort information about their usefulness to community. Schillo et al [6] also addressed the problem of "lying witnesses". To optimize the performance of the social network, the authors proposed a mechanism for mutual authentication agents carried out on the basis of two criteria: the degree of honesty (1) and «the degree of altruism» (2), determining as being good to others at the expense of one's own utility. The issues were further developed in [7] and [8].

Cholez et al. [3] examined P2P networks in case of malicious node's presence and introduced an architecture based on agent's reputation. The idea is the every node can fetch the reputation information about group member, and therefore, judge how to treat user's requests. The reputation criteria are used: the way a peer contributes to the network, evaluate the quality of the shared content. Evolution of the reputation automatically updates related to peer contribution. Hereafter security algorithms for distributed networks, based on trust, are also considered in [5].

Despite the great interest in the SRS' trust problems, systems related with activity of buyers and sellers, as well as members of online communities fall into researchers' limelight first of all. Generally, we are talking about fairly intelligent agents that can be considered with respect to the trust issues

ISSN 2305-7254

rather complex cognitive and social aspects [12].

However, many proposed approaches and algorithms retain their efficacy when are used for communities of homogeneous agents with minimal cognitive and computing capabilities.

Messaging robots the false information from neighbors saboteurs is able to break:

- a proper functioning of the collective decision-making mechanism;
- targets of SRS,
- effective distribution oftargets.

It should be noted that the cryptography methods to protect communication channels in this work are not considered because of their contradiction with the basic principles of SRS (simplicity and homogeneity of devices). That is why we put the task of finding methods of protection of other kind.

This paper examines the usage of SRS trust mechanism. It is meant that SRS operates in an open environment, and actual threat for it is device's substitution or modification for a saboteur (attacker) which is described with reference to SR [11]. In this case cryptographic protection methods for mobile devices (proposed, in particular in [10]), are not directly considered. It is interesting to examine the own resources of SRS related to co-operatives, mass and mobility in terms of the possibility of their application for protection technologies.

In [12] trust and reputation security model are proposed. Using the algorithm, announced in this paper, as well as on general trends noted in the above-mentioned literature and applicable to the SRS, we made it our target to consider the main factors affecting the operation of the security trust mechanism for Swarm Robotics.

II. Problem statement and assumptions

With the growing interest in the SRS grows and the need to solve security problems. The development of SRS is the need to create mechanisms to ensure SRS IS. Unpredictable dynamics of the external environment, low computing power of individual robots, agents and lack of complete information on the status of the entire system make such systems particularly vulnerable to the threat of introduction of saboteurs, reducing the efficiency of the system.

The authors proposed the objective function, by which you can assess the impact of these parameters: N_{is}/N_s, where N_{is} is number of detected saboteurs, N_s is total number of agents.

It is clear that this objective function must be equal to 1 in the case when all saboteurs are identified, and 0 if none is detected.

Thus and so the purpose of this paper is to define the parameters under which the objective function will tend to 1, and to develop some strategies to achieve this level of definition of saboteurs.

Among previously unexplored parameters which can intuitively be offered as affecting the efficiency of achieving the target are the following:

- the dependence of trust algorithims' efficiency on the quantity of robotic system and the percentage of saboteurs in it;
- number of alternative targets in decision- making;
- robot communication range (number of neighbors which are considered in the calculation).

To achieve the stated purpose is required to develop software simulator to model the behavior of a mobile robotic system using the algorithm of trust and reputation factors.

III. Assessment of Algorithm's Stability

A. Trust and Reputation model

The basis of the developed model is the approach proposed in [12]. A class of so-called soft attacks which use interception of communications, formation and transmission of collective robots misinformation, as well as carrying out other actions that do not have identifiable symptoms invasion of saboteurs is SRS. To increase the measure of similarity (closeness) of objects belonging to the same category ("saboteur" or "legitimate agent") an algorithm for calculating the reputation of agents was formed as a measure of swarm's opinion about the quality of each agent. The realization of algorithm of distribution purposes in a team of robots was used to detect saboteurs.

The model assumes the following:

- robots have sensors, which can verify the distance between neighboring robots and target;
- robots-saboteurs can broadcast false information, not for all the neighbors, and not about all targets - i.e. model has at least a certain percentage of false data;
- robots evaluate information received from neighbors and what they "see" themselves, then make an opinion about the "reliability" of the neighbor and distribute it to other robots;
- saboteurs can also spread false evaluation information over other robots;

B. Simulation

Generated simulator should provide the following:

- generation of robot groups;
- creation of purposes;
- placement of robots and the targets on simulator's working space.

Generation of robot group involves creating a set of robots N={N1,...,Nn}. The group is characterized by the following parameters: DB - initial distance between robots, LG - a list of the targets of the swarm, n - the dimension of the swarm, d - number of saboteurs in the swarm. The robot is described by the vector of characteristics {i, E, CR, SR, T, L}, where i -

agent identifier, E - information about energy left, CR - range of communications, SR – range of sensors, T - type of robot, L – location.

Indicator of stored energy E is used to determine the "value" of the robot path to a target. Communication radius determines the number of robots that are in the current area of agent's information exchange; the radius of action of the sensors is used to determine the true "cost" of way for an interactive robot. This simulator has been assumed on the equality of the radii of communication and action sensors. This simplifies the process of determining the saboteurs because each robot that interacts with the others, will be able to check the information received from it.

Robot type determines its behavior. The robot, which has a type of "saboteurs", tries to mislead the remaining members of the swarm by giving them false information, thereby worsening the performance of the entire group. In turn, the robot with the type of "normal", is trying to distribute its task and to identify the saboteurs in the swarm in order to save efficiency.

The generated set of targets must have the following characteristics: i - id, L – target's localization. This set of characteristics for determining the stability of the algorithm trust coefficients in the simplest case, when all robots need to reach the target. Drawing an analogy with the real world, we can talk about the task of training the robots to their transportation (it is necessary to clarify that only robots are transported, they are not carrying anything), when all the robots have their own move to a certain area, from which they will transport swarm by some means or other.

Generally purposes' and robots' alignment in the work space of the simulator must be random.

Regarding the experiment, one must be able to fix the alignment parameters.

B. Performance of Experiment

The experiment starts with the distribution of the robots on the surface. To simplify this process we used the following algorithm:

- at the first stage one robot is randomly on the field, in a second step,
- at the second stage all other robots are randomly located within 150 units one from another and at least nearly 350 units from the first one.

The second stage is repeated till all robots are left used. Thus, the field, where the robots involved in the experiment are placed, has a radius of 350 units.

After this distribution are randomly selected agents which become saboteurs. The amount of saboteurs is determined before the experiment starts.

The next stage is the definition of neighbors. The distance at which a robot can "see» it neighbor is set for the entire swarm. This allows to determine how reliable are target's cost estimates, which are transmitted one robot to another during the value matrices' exchange.

The radius of visibility ranges from 40 to 350 units. If one of the robots has not neighbors, which it can observe with such radius value, the following visibility value range is selected (the step between radiuses' values is 10 units). This selection takes place before the moment when every robot has at least one neighbor observed.

This is followed by the stage of determining targets' distance. Each robot generates a cost estimate, which characterizes the cost of moving it from the current location to the target. Saboteurs are recorded of random values between 0 and 10 in their matrixes.

Then comes stage of neighbors' evaluation: each robot scans the array of neighboring agents targets' cost estimates and issues grades, based on the validity of the data: "1" - if the estimates are correct, "0" - if the agent is not in his line of sight and "-1" - if valuations seem false. Saboteurs are arranged all agents of the swarm "-1", does not matter do they see it or not, and "1" to each other. These grades are recorded in a matrix *V.*

Matrix *S* contains some results of matrix *V* analysis.

If robots *i* and *j* have issued each other positive grades, the value is increasing by 1 point, if negative it is reduced by 1 point. If these robots issue a similar grade to the robot *k,* values in the cell *S [i, j]* are incremented by 1 point, if the grades differ, the values are reduced by the same amount.

Basing on *S* the reputation of each agent *Q* is calculated. Reputation is the ratio of the sum of all positive grades to the sum of all grades.

Then matrix *H* is filled in such a way: if the robot *j* issued in the robot's *j* matrix *V* positive grade, then *H [i, j] = 1*, otherwise *H [i, j] = 0*; the matrix *G* is filled similar: if the robot *j* issued in the robot's *j* matrix *V* negative grade, then *G [i, j] = 1*, otherwise *G [i, j] = 0*. Further indicators are calculated *P* and *L*:

$$P[i] = \sum\nolimits_{j=0}^{N} H[i,j] \cdot Q[j] \tag{1}$$

$$L[i] = \sum\nolimits_{j=0}^{N} G[i,j] \cdot Q[j] \tag{2}$$

The nest stage is calculating index W[i]:

$$W_i = \frac{P[i]}{P[i] + L[i]} \tag{3}$$

If *W [i]* is more than the predetermined value (this value varies from 0.5 to 0.7), it is concluded that the robot is "good", otherwise - the saboteur.

D. Experiment

During the experiment:

1) The swarm of robots of given number is randomly generated (ranging from 50 to 1000);
2) The number of saboteurs is randomly set (ranging from 1% to 45%);
3) The number of informed agents is randomly set (ranging from 1% to 50%);
4) The local interaction radius (from 10 to 350) is set, defined by the number of neighbors for each robot that performs calculations. The minimum radius for the experiment is one neighbor;
5) In the first iteration step robots inform neighbors within a radius of communication about calculated cost of meeting the target and check the information received from neighbors. In the second step a model of reputation is formed;
6) At the final stage of the experiment opinion on each robot is collected and the results are averaged.

In the context of this model robots deceive all their neighbors. This model describes the behavior of robot-saboteur who has problems with the hardware, but does not deliberately seek to deceive their neighbors.

It appears that the success of the algorithm in determined conditions will facilitate identification of trends in more difficult conditions, since it will be performing some of the base case, which can be compared to several advanced models.

Developed software allows us to solve the problem of fixing the basic parameters of the system (size of groups, types of robots, radius of communication, etc.). This helps to debug its work and conduct experiments similar to the initial conditions to determine current trends.

E. Outcome of Experiment

Two main trends were revealed from experiments: noneffect of the number of indicator targets on the speed of determining the saboteurs in the network (1) and - (2) the effect of the number of neighbors of agent (i.e., the radius of a local communication) on the possibility of determining all the saboteurs in the system.

Trending performance in the absence of influence on the percentage of the identified saboteurs was carried out on several samples with different parameters.

The main parameters under consideration were:

- total number of robots,
- the number of targets,
- the range of communication.

Number of saboteurs' robots was fixed at 20% of the size of the swarm. This value was chosen after experimentation on a random sample since in this case, the definition of saboteurs occurs gradually with increasing radius of the interaction between agents.

The average values for groups of dimension 200, 500 and 1000 agents are presented in Fig. 1.

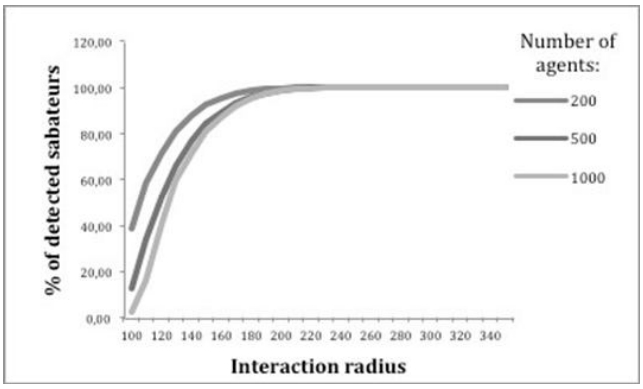

Fig. 1. Dependence of detected saboteurs' percent on the radius of communication

The averaged values are used as trend, in comparison with which the results of specific experiments can be made a conclusion about the existence or absence of the influence of the number of targets at the speed of determining the saboteurs in the system.

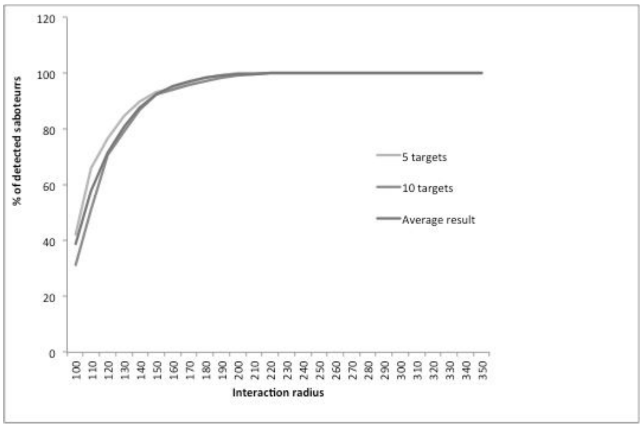

Fig. 2. Dependence of detected saboteurs' percent on the radius of communication for 5 and 10 targets

Fig. 2, 3, 5 and 4 show combined schedules of special cases of experiments for grouping dimension 200 agents and a graph, which is obtained from the mean values for the same group, but with the different number of targets.

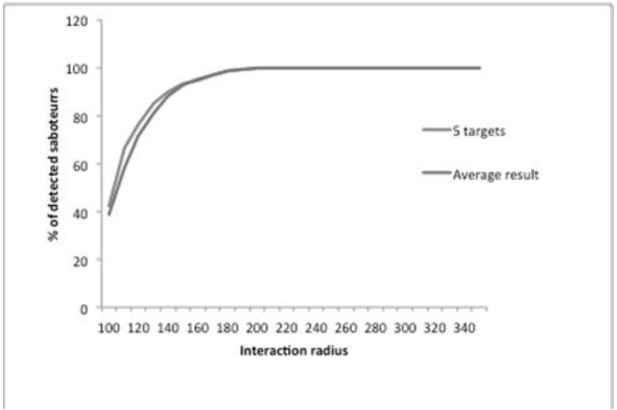

Fig. 3. Dependence of detected saboteurs' percent on the radius of communication for 5 targets

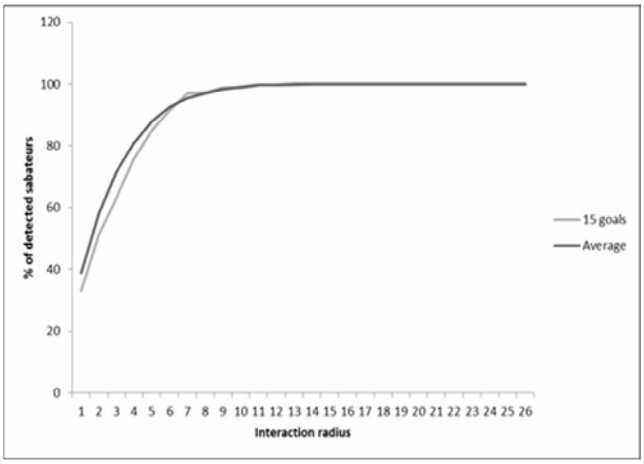

Fig. 4. Dependence of detected saboteurs' percent on the radius of communication for 15 targets

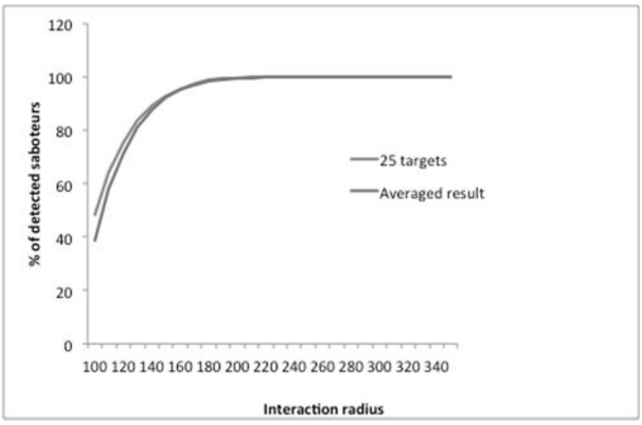

Fig. 5. Dependence of detected saboteurs' percent on the radius of communication for 25 targets

As a part of the study, every particular situation had no significant deviations from existing schedules. The data presented in figures leads to the conclusion about the absence of the influence of the number of targets on the result of the algorithm. Similar actions were carried out for groups of 500 and 1,000 agents.

Subsequently similar results were obtained, which allows us to talk about noneffect of targets on the percentage of detected saboteurs.

Follow-up study was conducted to assess the number of interconnections and the percentage of detected saboteurs. Previously it was stated that targets do not affect the results of the algorithm, so in subsequent experiments it was decided not to take this parameter into account.

Thus, only communication range and total number of agents and the number of saboteurs were taken into account.

Fig. 5 shows the final figure of required neighbors' number with a corresponding number of saboteurs.

IV. Discussion

Based on the schedule presented above, we can conclude that for the unambiguous detection of all saboteurs it is necessary that the average number of agents' neighbors exceeds this number at least twice.

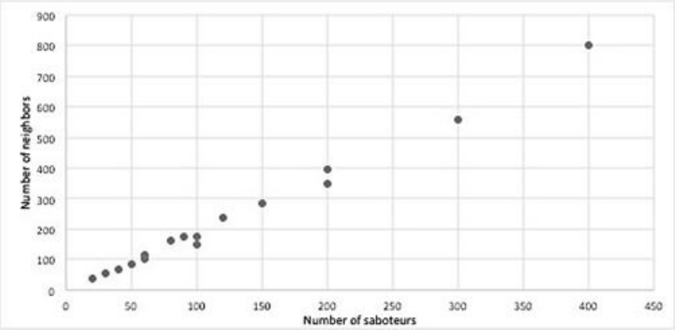

Fig. 5. Number of neighbors to determine all the saboteurs

Therefore, a condition in which the group becomes unhealthy, is the achievement of certain percentage of saboteurs.

When the number of saboteurs is over 40% of the number of the swarm, the definition of all saboteurs becomes impossible. If the number of saboteurs reaches 50% of the size of the swarm, swarm ceases to perform its tasks.

It should be noted that the resulting value of the number of neighbors varies from 150% to 200% of the saboteurs. There is some relationship between this number and the ratio of the number of saboteurs to the total number of robots.

Attempts to identify the greatest number of saboteurs, which allows a swarm instantly find all available saboteurs, and the smallest number of saboteurs, which is able to make swarm system inoperative, did not lead to an unequivocal result because there is a clear trend between these indicators and the average number of neighbors of each SRS agent.

With a large radius (increase of the number of neighbors), the percentage of detected saboteurs increases dramatically. Growth from 0% to 100% increases in radius which is placed at the values of 100 to 150 (depending on other parameters).

V. Conclusion

The key to wide practical use of SRS is the development of specific guidelines and algorithms for the organization of group performance. Increased risks of SRS IS cause urgent need to assess the well-known and new algorithms from the point of safety view. It should be noted that a common approach to ensuring SRS IS is not formed so far. New technologies often forget about IS to the latest stages of development, when it is undesirable (and sometimes expensive) to upgrade the whole technology. One of the problems of ensuring SRS IS is the contradiction between the development of algorithms for increasing efficiency of swarm and minimizing of number of agents which are informed about the target, on the one hand, and an increased risk of disinformation of swarm - on the other.

The software simulator, designed by authors to carry out experiments, allows to consider features of swarm intelligence and vary the basic characteristics of the system. Properly formulated simulation tasks allowe to neutralize the specifics of a particular implementation, and avoid unreasonably high requirements to the research polygon. Based on the results in the experiment, we can conclude the effectiveness of trust and reputation algorithm, provided that the number of saboteurs is at least half of the amount of swarm agents. Furthermore, it is revealed that the amount of SRS' tasks does not affect its efficiency. Thus, in future studies, this parameter can be offset.

REFERENCES

[1] C. Castelfranchi, R. Falcone, "Principles of trust for MAS: cognitive anatomy, social importance, and quantification", *in Proc. of the International Conference of Multi-Agent Systems (ICMAS '98*), 1998, pp. 72 – 79.

[2] M. Witkowski, A. Artikis, J. Pitt, "Experiments in building experiential trust in a society of objective-trust based agents", *in Falcone R., Singh M. & Tan, Y. -H. (eds.), Trust in Cyber-societies*, Berlin: Springer-Verlag, 2001, pp. 111–132.

[3] T. Cholez, I. Chrisment, O. Festor, "A Distributed and Adaptive Revocation Mechanism for P2P networks", *in Proc. of the Seventh International Conference on Networking*, 2008, pp. 290–295.

[4] S.D. Ramchurn, D. Huynh, N.R. Jennings, *Trust in multi-agent systems. The Knowledge Engineering Review*. New York: Cambridge University Press, vol.19, Iss. 1, Mar. 2004, pp. 1– 25.

[5] Garcia-Morchon O., Kuptsov D., Gurtov A., Wehrle K. "Cooperative security in distributed networks", *Computer Communications*, vol. 36, no 12, 2013, pp. 1284–1297.

[6] M. Schillo, P. Funk, M. Rovatsos, "Using trust for detecting deceptive agents in artificial societies", *Applied Artificial Intelligence*, Sp. Iss. on Trust, *Deception, and Fraud in Agent Societies,* 14 (8), pp. 825 – 848.

[7] S. Sen, P. S. Dutta, "The evolution and stability of cooperative traits", *in Castelfranchi, C. & Johnson, L. (eds.). Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*, vol. 3, pp. 1114 –1120.

[8] S. Sen, N. Sajja, "Robustness of reputation-based trust: Boolean case", *in Castelfranchi, C. & Johnson, L. (eds.), Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems*, vol.1, pp. 288 – 293.

[9] G. Zacharia, P. Maes, "Trust through reputation mechanisms", *Applied Artificial Intelligence,* 14, 2000, pp. 881 – 907.

[10] T. Sander, Ch. F Tschudin, "Protecting MobileAgents Against Malicious Hosts", *in Giovanni Vigna (ed.), MobileAgents and Security*, LNCS, Springer, 1998, pp. 44–60.

[11] F. Higgins, A. Tomlinson, K.M. Martin, "Survey on Security Challenges for Swarm Robotics", *in Proc. of the 2009 Fifth Int. Conf.on Autonomic and Autonomous Systems*, 2009, pp. 307–312.

[12] I. A. Zikratov, I. S. Lebedev, A. Gurtov, "Trust and Reputation Mechanisms for Multi-agent Robotic Systems", *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* vol. 8638, no. LNCS, 2014, pp. 106- 120.

[13] V. Gorodetski, I. Kotenko, O. Karsaev, "Multi-agent technologies for computer network security: Attack simulation, intrusion detection and intrusion detection learning", *Computer systems science and engineering,* no. 4, 2003, pp. 191–200.

[14] J. Carter, E. Bitting, A. A. Ghorbani, "Reputation formalization for an information-sharing multi-agent system", *Computational Intelligence*, vol. 18 (2), 2002, pp. 515-534.

[15] N. M. Karnik, A. R. Tripathi, "Security in the Ajanta mobile agent system", *Software - Practice and Experience,* vol. 31. no. 4, 2001, pp. 301–329.

[16] T. Sander, Ch. F. Tschudin, "Protecting mobile agents against malicious hosts", *in Giovanni Vigna (ed.) Mobile Agents and Security*, LNCS, Springer, 1998, pp. 44–60.

[17] J. Page, A. Zaslavsky, M. Indrawan, "Countering security vulnerabilities using a shared security buddy model schema in mobile agent communities", *in Proc. of the First Int. Workshop on Safety and Security in Multi-Agent Systems (SASEMAS 2004),* 2004, pp. 85–101.

[18] I. A. Zikratov, I. S. Lebedev, A. V. Gurtov, E. V. Kuzmich, "Securing swarm intellect robots with a police office model", *in Application of Information and Communication Technologies (AICT), IEEE 8th Int. Conf.*, Oct. 2014, pp. 1 – 5.