Threat and Intruder Models of Multi-agent Robotic System Using Police Office Model with Quantum Encryption*

Zakoldaev D. A.1, Vorobeva A. A.2

1PhD, associate Professor, Department Computer Systems Design and Security, St. Petersburg National Research University of Information Technologies, Mechanics and Optics, Russia, St. Petersburg
2PhD, associate Professor, Department Computer Systems Design and Security, St. Petersburg National Research University of Information Technologies, Mechanics and Optics, Russia, St. Petersburg
1d.zakoldaev@itmo.ru

Abstract. The subject of this study is to ensure the confidentiality of information in mobile robotic systems. To ensure the confidentiality of information, it is proposed to use quantum encryption methods. Existing developments in this area allow to ensure the correct generation of quantum keys between two elements. The paper considers the functioning of the multi-agent robotic system based on the Police Office Model, the general scheme of functioning and basic processes of information interaction are given. The authors consider the issue of the protection of the multi-agent robotic system by using quantum keys from various threats, offering a classification of threats. Based on the model of the multi-agent robotic system, an intruder model is proposed that includes the level of access of the intruder to the system and the period of interaction of the intruder on the system, based on the life-cycle model. As a result of the research, a model of generalized functioning of a mobile robotic system with encryption based on quantum keys, a threat model and an intruder model was proposed.

Keywords: multi-agent robotic system; quantum encryption; confidentiality; information security; threat; intruder

1. Introduction

Since the middle of the last century, we have seen a continuous process of optimization and automation of human activities. In the performance of various work, robotic devices begin to replace the human. To solve problems requiring complex actions, such as control of the terrain, search operations or cargo movement, special intelligent robotic systems are developed, called the multi-agent robotic systems (MARS). Examples of such systems can be found in (1-4). Such systems usually consist of a group of robots, each of which is simple and is capable of performing single simple tasks. The advantage of such systems over individual robots is that the use of a group of robots provides redundancy of the system. Also, using a group of robots to perform complex tasks instead of single robot can raise the efficiency of the system, speed up the process and make the existing solution cheaper (5,6).

As the complexity of the tasks increases, it becomes necessary to increase the complexity of the individual robot: for the solution of tasks of increased complexity, it is necessary to have additional devices, which makes the robotic device more expensive, increases the time and financial costs for its maintenance. In this case, the more elements present in the robotic device, the greater the probability of failure of a single element, which can affect the functioning of the robot as a whole. Using a group of simple and cheap robots increases the reliability and flexibility of the system: if one of the robots fails, it can be replaced by another (7,8).

Any modern information system that processes this or that information can be exposed to risks of violating the security features of such information by intruders. Cyber-physical systems (CPS) is no exception. The meaning of CPS is to connect physical processes that require the practical implementation of continuous control in real time with software-electronic systems (9). The situation is aggravated by the fact that the variety of CPS types and approaches to development of these systems make them vulnerable to different types of attacks (10).

2. Confidentiality assurance problem in Multi-agent robotic systems

Ensuring information security (IS) is one of the most important components of the proper operation of the MARS. The work of MARS is based on the interaction between agents of the system and on the interaction between the information and physical components of these agents. Such interaction is subject to various kinds of vulnerabilities, using which attackers can implement attacks on the system. To provide comprehensive protection of the system, it is necessary to ensure the basic properties of information security, which in the context of CPS can be interpreted as follows: (11)

- confidentiality - the ability to prevent disclosure of information to unauthorized users;
- integrity - the ability to withstand impacts aimed at unauthorized modification of information received and sent by sensors and elements of a robotic device;
- availability - the ability to remain operational and respond promptly.
It is customary to refer to the main mechanisms of attacks on MARS: (12)

- attacks on communication channels;
- difficulty in identifying and authenticating agents in the system;
- the physical introduction of robots-intruders which reduce the effectiveness of the system. Also, initially allied robots can become intruders in case of incorrect functioning.

To date, there are many approaches to ensuring IS and reducing the likelihood of the implementation of threats.

In (13), a model is considered that is based on changes of agent states, digital signatures, and the use of public key encryption. Agents can be in several protected states, depending on the type of activity.

A scheme for ensuring IS in a multi-agent system using police offices is described in (14). The essence of this method is to divide the territory of functioning of the system into areas and create nodes for management and control of the activities of agents. The paper (15) describes a modernized scheme based on police office model (POM) for ensuring IS in a group of robotic devices.

Using the methods of mobile cryptography, presented in (16), it is possible to ensure the integrity of the executable code by the agents. This method is based on obtaining information about the environment on the basis of calculating the value of some encrypted function.

In (17), the authors propose to use a “Buddy” security model. In such a system all agents are homogeneous. By interacting with each other and with the environment, agents monitor the events occurring in the system and are responsible for ensuring each other's security.

The authors of (18,19) propose an approach to the organization of secure communication using dedicated short-range communications technology, which is resistant to various IS breaches.

The authors of (20) propose using cryptographic methods of protection to ensure the integrity of information. The paper presents a method for protecting the integrity of information in multichannel unidirectional radio communication systems.

Thus, having considered a lot of existing methods of ensuring information security, the authors of current paper concluded that at the moment there are no universal approaches that can provide the necessary level of confidentiality in the MARS. Given the trends in the expansion of the use of MARS in various areas of our lives, from smart cities to medicine, the elements of such systems can transfer confidential data among themselves, the disclosure of which can bring various kinds of damage to those in whose interests it is to ensure the confidentiality of such data. Disclosure of confidential information can help an intruder attack the system and, given the growing prevalence of such systems in many areas of our lives, such attacks can have global consequences.

3. Police Office Model with quantum encryption

As a solution to the problem of ensuring confidentiality, the authors of this work suggest the use of a POM (14, 15) with the implementation of quantum encryption of information transmitted among agents. Classical encryption algorithms, whether DES, AES or RSA, these algorithms are based on mathematical algorithms and are not completely resistant to attacks. Quantum cryptographic algorithms use the principle of uncertainty in quantum physics (21). It is believed that if an attacker tries to eavesdrop a communication channel, this will change the state of the transmitted information and the attacker will be detected. The method of quantum key distribution between the two parties was first presented in (22). Examples of protocols for quantum key distribution are: BB84, B92, SARG04, E91, BBM92, etc. (22-26)

The authors propose the following scheme for ensuring the IS of MARS using autonomous unmanned robots capable of moving through a predetermined area. The system consists of two types of agents: motionless robot-bases, which constituting the set $B=\{b_1, \ldots, b_n\}$ and perform the role of police offices. Mobile robot-agents that make up the set $R=\{r_1, \ldots, r_m\}$ performing the task of moving from point A to point B. The bases have their own "coverage zone", the boundary of such a zone is determined by the limiting distance to which the technical communication devices of base are able to transmit information to agents.

The essence of the tasks for robots - moving from point A to point B. The tasks are constituting the set $T=\{t_1, \ldots, t_p\}$ and distributed by the bases to robots by a simple auction that allocates tasks via a sequence of first-price one-round auctions (27).
In the system there are two levels of information interaction (II): the lower and the upper. The upper level is the level of interaction between the bases. At this level, decentralized control is organized, i.e. the bases themselves make decisions about the distribution of tasks and the control of lower-level robots. On the top there is no central control element.

The lower level is the level of interaction between robots and bases, central control is organized at this level. Robots are dependent on the bases and are required to obtain permission for any movement. When performing a movement from the coverage area of the $b_i$-base and entering the coverage zone of the base $b_j$, the robot needs to go through the authorization procedure on the base $b_j$ to obtain the task from it. Such a procedure in the context of the POM is detailed in (15). The robot can only interact with the base in which “coverage zone” it is located. The information interaction of both levels is schematically shown in fig. 1. Each robot and each base has its own identification number, known in advance to all agents of the system. On the basis of this number, a verification is made of the existence of an agent.

![Figure 1: Top and low-level information interaction scheme](image)

When using quantum encryption algorithms in such a scheme, agents need to perform a positioning procedure before key generation. This procedure involves pointing the technical devices of the receiver and transmitter to each other to generate a quantum key. This action is necessary and is due to the needs of quantum encryption. Top-level agents-bases have a receiver and a transmitter for communicating at their level and a receiver and transmitter for communicating with the lower-level robots. Robots of the lower level have a receiver and a transmitter for communicating with the bases and do not have the opportunity to communicate with each other. This solution is difficult in terms of implementation, but this task is engineering and will eventually be solved.

In order to assess the level of information security, the authors propose a model of threats and the model of the intruder. In the future, it is appropriate to conduct a risk assessment, based on these models.

The function diagram on the figure 2 describes agents’ interaction using quantum encryption which includes the following stages:

1) on the first stage robot-agent $r_i$ sends communication request to a base-agent $b_i$, which control this area, $b_i$ checks if the agent exists;
2) agents $r_i$ and $b_i$ generate crypto key for encrypted channel;
3) after estimating of agents’ resources the base $b_j$ starts the tasks distribution between robot-agents. By the end of this stage, the robot $r_i$ gets suitable task;
4) $r_i$ moves from the area of the base $b_i$ to the area of the base $b_j$ which controls robot’s movement;
5) on the final stage $r_i$ sends communication request to the new controlling base-agent $b_j$. They generate new crypto-key and $r_i$ waits for the new task.
4. Threat model

In accordance with (28), the information security threat model is a physical, mathematical, descriptive representation of the properties or characteristics of information security threats.

As a threat model, there were identified sources of threats, factors that determine the possibility of implementing threats, methods for implementing threats and the consequences of implementing security threats.

A specific feature of the MARS is that the group consists of autonomous agents, and the achievement of the goal group is ensured by a set of parallel actions of agents with each other and with the environment. In addition to ensuring confidentiality, integrity and accessibility of information, an essential condition for the provision of IS is the lack of an energy security (ES) threats as lack of battery power and the threat of unavailability of the

Figure 2: Function diagram of system working using quantum message encryption
energy [29]. In the case of an attack, which affects the ES, the robots may be useless individually, or fail to reach the group goal, or do it spending much more resources, which will lead to the overall efficiency of the system.

In this model, a threat is a potentially possible event or an impact posing a threat of the breach of the IS of the MARS, which can lead to damage to the functioning of the MARS.

The main components of the threat are the source of the threat, the target object and the actual event (action, phenomenon), which entails damage to the functioning of the system. The sources of threats can be both external intruders, and internal ones, including employees serving the MARS and the MARS-agents themselves, whose software errors can lead to problems in the functioning of the system. The environment is also a source of threats, called natural (flood, earthquake, etc.). Target objects may be agents of MARS or communication channels of agents.

The IS threats of the MARS may be classified according to several basic characteristics. The first type is threats aimed at violating the confidentiality of information interaction, the implementation of unauthorized access (UA) to information transmitted through the MARS channels and used by its agents. Examples of such threats to the MARS are UA by an internal or external intruder to software agents, encrypted communication channels, etc.

Threats of integrity, unintentional data changes are attacks on the operation system, agent controllers, robot-agent sensors, unauthorized modification of data transmitted by agents.

The threats to accessibility are aimed at hampering access to MARS resources or making it impossible. The intruder creates conditions when access is impossible for some time, as an example, errors in the maintenance of MARS, the elimination of agents, the creation of obstacles for the moving robots.

Dividing by location of threat source. Within the framework of this classification, threats are divided into:
- internal ones, the sources of which are located inside the MARS. Internal attacks include sabotage, the introduction of enemy robotic agents into the system with the goal of destructive impact on the system, errors in the servicing of MARS by employees;
- external ones, whose sources are outside the MARS. External is the natural impact on the MARS, the harmful impact of software on MARS agents.

Dividing by the nature of the threat:
- threats related to natural ones are caused by the environmental impact on the MARS;
- artificial (subjective) threats are caused by human exposure to MARS accidentally or deliberately.

Most of the problems of ensuring the IS of the MARS are primarily related to deliberate threats. An example of an artificial deliberate threat is the brute force of the encryption key and access to information about the resources and tasks of agents and bases, with the possibility of modifying it and then transferring it to agents to reduce the effectiveness of the MARS targets proceeding.

To create a model of MARS threats and assess the implementation possibility, it is necessary to determine the initial security level of the MARS. For this purpose, the main characteristics of the MARS were identified and the values of the initial protection level of each component on a low-medium-high scale were determined, the evaluation is presented in table 1, the "+" sign indicates the appropriate level of security. None of the items evaluated was rated as "Low", so the table does not have this column. The initial level of security of the MARS is estimated as average (the index of the initial level of security $ Y_1 = 5 $ on a scale of 0-10). The results of probability and realization threats analysis demonstrated in table 2.

<table>
<thead>
<tr>
<th>MARS characteristics</th>
<th>Security level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Medium</td>
</tr>
<tr>
<td>MARS stability</td>
<td></td>
</tr>
<tr>
<td>Component integrity of MARS</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 1. Initial security level of the MARS.
At the next stage, the probability of realizing the threats identified by the authors was calculated. The value $Y$ was calculated using the formula $Y = \frac{Y_1 + Y_2}{20}$.

Table 2. The probability of threat realization

<table>
<thead>
<tr>
<th>MARS security threats types</th>
<th>Coefficient of threat realization</th>
<th>Possibility of threat realization</th>
<th>Probability of threat realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Threats related to the external environment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1. Natural disasters</td>
<td>0</td>
<td>Unlikely</td>
<td>Low (0.25)</td>
</tr>
<tr>
<td>1.2. Landscape changes</td>
<td>0</td>
<td>Unlikely</td>
<td>Low (0.25)</td>
</tr>
<tr>
<td>2. Threats implemented by an internal intruder</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1. UA to the software of base-agents</td>
<td>5</td>
<td>Medium</td>
<td>Medium (0.5)</td>
</tr>
<tr>
<td>2.2. UA to the software of robot-agents</td>
<td>5</td>
<td>Medium</td>
<td>Medium (0.5)</td>
</tr>
<tr>
<td>2.3. Errors in servicing of MARS</td>
<td>2</td>
<td>Low</td>
<td>Medium (0.35)</td>
</tr>
<tr>
<td>2.4. Sabotage</td>
<td>2</td>
<td>Low</td>
<td>Medium (0.35)</td>
</tr>
<tr>
<td>3. Threats related to the technical destruction of hardware</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1. Base-agent liquidation</td>
<td>0</td>
<td>Unlikely</td>
<td>Low (0.25)</td>
</tr>
<tr>
<td>3.2. Robot-agent liquidation</td>
<td>2</td>
<td>Low</td>
<td>Medium (0.35)</td>
</tr>
<tr>
<td>4. Threats that prevent the movement of robot agents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1. Creating static obstacles</td>
<td>0</td>
<td>Unlikely</td>
<td>Low (0.25)</td>
</tr>
<tr>
<td>4.2. Creating moving obstacles</td>
<td>2</td>
<td>Low</td>
<td>Medium (0.35)</td>
</tr>
<tr>
<td>5. Threats of unauthorized access using software, hardware</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1. Malware action</td>
<td>2</td>
<td>Low</td>
<td>Medium (0.35)</td>
</tr>
<tr>
<td>5.2. Open communication channel vulnerabilities</td>
<td>0</td>
<td>Low</td>
<td>Low (0.25)</td>
</tr>
<tr>
<td>5.3. Encrypted communication channel vulnerabilities</td>
<td>0</td>
<td>Low</td>
<td>Low (0.25)</td>
</tr>
<tr>
<td>5.4. Vulnerabilities of robot-agents sensors</td>
<td>2</td>
<td>Low</td>
<td>Medium (0.35)</td>
</tr>
</tbody>
</table>
5.5. Impact on software of the controllers

<table>
<thead>
<tr>
<th>2</th>
<th>Low</th>
<th>Medium (0,35)</th>
</tr>
</thead>
</table>

where the threat realizing coefficient $Y_2$ is:
- 0 - unlikely threat;
- 2 - low threat probability;
- 5 - medium threat probability;
- 10 - high threat probability.

The threat realizing coefficient $Y$ is calculated as follow:
- $0 < Y < 0.3$ - low;
- $0.3 < Y < 0.6$ - medium;
- $0.6 < Y < 0.8$ - high;
- $0.8 < Y < 1$ - very high.

5. Intruder model

Existing approaches to the developing of the intruder model are presented in the normative and methodological documentation and have a number of differences from the approaches presented in the scientific and technical literature. After the analysis of the recommendations for creating models of threats and the intruder (30), the authors of this work concluded that there are no universal approaches for compiling such models.

From the viewpoint of necessity for the presence of operator persons within the controlled area (CA), intruders were classified as follows:

• I category: persons who do not have the right to access the CA;
• II category: persons who have the right to access CA.

Also, all potential intruders are divided into:
- external (the attack implementation from outside the CA);
- internal (the attack implementation within the CA).

This division is schematically illustrated on the fig. 3:

![Diagram of intruders](image)

External intruders of the I category are:
- members of criminal organizations;
- outsiders who use spyware to conduct attacks on the system;
- former employees of the organization;
- competitors.

Such intruders can perform the following actions:
- attempts to destroy/damage the robot;
- attempts to intercept robot control;
- unauthorized access to information through special software influences.

Persons of category II can be divided into the following groups:
1. Staff performing maintenance and operation of MARS equipment:
   - have access to robots;
   - have data on the terrain on which MARS is deployed;
   - may cause equipment damage (intentionally/unintentionally).
2. Designers involved in the development of the robot:
   - have the ability to intentionally/unintentionally make changes to the robot model that are incompatible
     with the correct operation;

3. Engineers engaged in assembly robots:
   - have access to the robot and its components;
   - can intentionally/unintentionally at the time of assembly make changes in the design of the robot that
     are incompatible with the correct operation.

4. Programmers who develop software for the robots:
   - have access to the software of the robot;
   - have the ability to intentionally/unintentionally make changes to the software for incorrect work at a
     certain point in the task, such as authentication, distribution of targets, etc.

5. The MARS operator (the person who is responsible for monitoring the status and accomplishing the
   tasks assigned, manages the system):
   - have the ability to prioritize goals;
   - have the ability to deliberately/unintentionally ignore information about the incorrect operation of the
     system.

It is possible to use another approach to the intruder classification of the MARS IS, grouping them according
to the time of participation in the life cycle of the system. Such compliance of intruders and threats with respect

to the life cycle is shown in table 3. The MARS life cycle model includes the following stages:
- robotics system design;
- designing, assembling and tuning of MARS components;
- functioning of the MARS;
- MARS recycling.

In accordance with this classification, possible intruders can be divided into the following groups:
1) MARS developing:
   - MARS designers and developers.

2) Designing, assembling and tuning of components of MARS:
   - designers and engineers of MARS, technical stuff, software developers, MARS operators.

3) Functioning of MARS:
   - servicing stuff, MARS operators.

4) MARS recycling:
   - recycling service stuff who can use residual confidential data.

Fig. 4 illustrates this division. On each life cycle stage, potential intruders may be external intruders
described in the previous classification: criminal organizations members; outsiders; former employees of the
organization; competitors.

<table>
<thead>
<tr>
<th>Life cycle stages</th>
<th>Intruders</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARS developing</td>
<td>MARS designers and developers</td>
<td>Sabotage</td>
</tr>
<tr>
<td>Designing, assembling and tuning of</td>
<td>Designers and engineers of MARS, technical stuff, software developers, MARS</td>
<td>Sabotage, errors in servicing of MARS, impact on OS of controllers</td>
</tr>
</tbody>
</table>

Table 3. Conformity of threats and intruders with the life cycle of the system
6. Conclusion
A model of functioning of a multi-agent robotic system is proposed. The model uses an approach to the organization of functioning on the basis of the Police Office Model - the area of the model is divided into equal areas where the "police offices" are located, which are responsible for the development of optimal plans of agents' actions. Between the "police offices" is carried out information interaction, which allows to guarantee the development of a joint optimal plan. Low-level agents (elements of the mobile robotic system) carry out the action plans provided to them. The connection between agents and "police offices", as well as between "police offices" is carried out using quantum encryption keys.

Based on the proposed model of functioning, a threat model has been developed that includes a classification of threats, as well as an assessment of the likelihood of their implementation. The use of this model will allow us to assess the main threats, which facilitates the adoption of optimal solutions for the implementation of the countermeasures by the owner of the mobile robotic system.

In addition to the threat model in, the intruder model is presented, including an assessment of the possibility of the intruder appearing at various stages of the system life cycle. The main ways to intruder’s influence on the system are given.

The results obtained in this work can become the basis for further studies of the robustness of multi-agent robotic systems. Based on the proposed models of threats and the offender, it is planned to determine the methodology for assessing the risks of implementing threats to the information security of the multi-agent robotic system, and to develop a method for assessing the security of the system from various attacks.

7. Acknowledgement
This article was prepared with the Financial support of the Ministry of Science and Higher Education of the Russian Federation under the agreement No. 075-15-2019-1707 from 22.11.2019 (identifier RFMEFI60519X0189, internal number 05.605.21.0189).

References