Mathematical Modeling of User Perception in Information Security Systems

Mikhail A. Styugin*
Siberian State Aerospace University
Krasnoyarsky Rabochy, 31, Krasnoyarsk, 660014
Institute of Space and Information Technology
Russia

Alexey A. Kytmanov†
Institute of Space and Information Technology
Siberian Federal University
Kirenskogo, 26, Krasnoyarsk, 660041
Russia

*styugin@gmail.com
†aakytm@gmail.com
© Siberian Federal University. All rights reserved

The problem of the functional structures research is considered in this example of information systems. A feature of such research is that it is not always possible to ensure that the research results will match reality. This is a topic of current interest in the field of design and analysis of information security systems and software analysis for undeclared capabilities of systems in general. By undeclared capabilities, we refer to a functionality available in software that is invisible to users and can be used/exploited by an intruder. This paper presents a model of a researcher and of a functional object investigated by him. Based on this model, informational limitations of the researcher are shown. The mathematical model of the subjective structure of an investigated system is constructed. It is shown in which cases this structure is stable. This article answers the question of if the researcher can claim that his subjective functional structure corresponds to the actual structure of the investigated system. We provide examples of such approach on certain mathematical models of information security.

Keywords: mathematical models of information security, model of the researcher, information structure of the conflict, information flows, misinformation, black box model, graph theory.

DOI: 10.17516/1997-1397-2015-8-4-454-466

Introduction

Information management in the sense of the knowledge of structure of a system and its vulnerabilities is an important issue in the field of information security. Effectiveness of information security actions depends on the information that a security administrator has. More precisely, it depends on how his information coincides with the reality. On the other hand, effectiveness of information security depends on a potential intruder’s knowledge about the system: how close
his knowledge is to the real picture. There is a huge amount of information management tools which can be used in this case: logging systems, testing systems, security simulation systems, fake objects such as Honeypots, etc. However, at the present moment there are no modeling tools for such kind of systems.

For the modeling of such systems, it is necessary to split all the information on the system into an objective one and a subjective one. The level of similarity between a subjective image and the real structure of the system depends on the objects in the system to which the subject has an access and from which he can get a response. Under certain circumstances this level of similarity can dramatically affect the security of a system.

In the present article we propose one of the possible solutions to this problem.

1. Review of the relevant literature

Let us outline the main existing directions in the field of researching how of the attacker’s information is influenced and managed. The most well-known direction is obfuscation. The most commonly used concept of obfuscation refers to software source code. The concepts that are used less commonly refer to software architecture and hardware devices. We may refer the reader to the works [1–3] where the mentioned concepts are studied.

The main idea of such concepts is obfuscation of a source code or a system structure in such way that the decoding task becomes a non-trivial problem. The key difference between the obfuscation theory and the theory used in the present work is that in the obfuscation theory the attacker knows about unknown parts of the system and therefore he is able to formulate the problem of investigating the real system structure. In our situation, the attacker does not have any information about the real system structure so that he is unable to formulate the problem of obtaining such a structure.

As a basis for the present work we used the existing mathematical models of access control. The most common examples of such models would be the Harrison, Ruzzo, Ullman (HRU) Model [4], the Bell-LaPadula Model [5], the Model of Secure Information Flow [6], the Role-based Access Control Models [7]. The HRU Model can be considered as the most universal (but at the same time the least usable) since the current access status in any system can always be represented as a matrix. For improving usability, a few modifications of the HRU model were introduced, namely the Typed Access Matrix Model [8], the Dynamic-Typed Access Matrix Model [9], etc. However, these models do not take into account the structure of the information system participants’ awareness.

An exception here is the Quantitative Information Flow approach that became popular recently (see [10–12]). In these works, the concept of information flow is considered in terms of the Shannon entropy. Article [12] is the most interesting for us since the notion of information flow is expressed by considering the beliefs of individual users receiving this information flow. In other words, the information flows are considered in terms of the participants’ awareness of the system. The related concepts of misinformation and covert channels are considered as well. Despite the fact that in this work we study similar processes, our approach is completely different as we only use qualitative characteristics of information flows.

As a basis for our research, we take the HRU Model with added awareness of all participants. The approach described in the present work can be applied without loss of generality to various access control models including information flow modes.
2. **Functional system model**

Most information systems can be modeled using the black box model with input parameters which one can affect and output parameters which one can observe. We will treat a system structure as transformation laws that convert input parameters into output result. A black box structure can be represented as a set of elementary black boxes with inputs, outputs and connections between some of them. The output of each box functionally depends on its inputs. We call such a system of interacting black boxes a functional system.

One can make assumptions about the functional structure of the system. However, it is possible to control only a finite number of inputs and outputs. Parameter types (discrete, continuous, etc.) will also affect the result of the system research.

Consider the scheme given on Fig. 1.

![Fig. 1. Information flows structure](image)

In this case, researcher made an assumption about the black box structure. Namely he assumes the box consists of three objects, each of which represents a functional box, consisting of inputs and outputs. Object \( o_1 \) can be normalized by splitting it into the two independent objects with input \( x \) and outputs \( k \) and \( w \).

The functions of system objects are

\[
k(x), w(x), e(k, y), f(w, e).
\]

The resulting function of the system is

\[
f(w(x), e(k(x)), y).
\]

Here \( x \) and \( y \) are input parameters of the system. Parameter \( y \) is considered but not observed by the researcher.

Since the system has input parameters which are unobservable, the researcher cannot obtain an informative feedback from the investigated system. However, it could be still possible to learn something new about the system [13]. There could be some information on the topology of the system and on the properties of the information flows.

It is necessary to split the subjective and the objective view of the system for determining the conditions under which the informative feedback exists. To do this, one has to divide the view of the objects into subjective (which might not be real) and objective, and the information flows between those objects.

Thus, we get the subjective and the objective scheme of the information flows (Fig. 2).

In the scheme shown on Fig. 2, the object \( o_2 \) has an extra input \( y \) in the objective system while the subjective scheme does not have it. Will the researcher get an informative feedback from the system? It depends on the properties of the functions and on the topology of information flows.
flows network. Based on the properties of the researcher model described in [13], we can get an informative feedback from the system when we are able to introduce an equivalence relation on the set of values of the function, observed by the researcher based on unknown input parameters.

For the scheme mentioned above, these values will be

\[ F_{o_2} : K \times Y \to E_1 \times E_2, \]

so that the function may be written down as a direct product

\[
\begin{align*}
F_{o_2}^K : & K \to E_1, \\
F_{o_2}^Y : & Y \to E_2, \\
F_{o_2} : & = F_{o_2}^K \times F_{o_2}^Y.
\end{align*}
\]

One can introduce an equivalence relation on the range (set of values) of \( F_{o_2} \) by the formula

\[
F_{o_2}^\equiv : E_1 \times E_2 \equiv E_1 \times E_2, \\
\forall e_1 \in E_1, e_2 \in E_2, e_3 \in E_2 \Rightarrow (e_1, e_2) = (e_1, e_3).
\]

Furthermore, if the object \( o_3 \) performs similar mappings then it is possible to define an equivalence relation on the set of values of the function observed by the researcher. That is, the researcher gets the function with values which can be split into two sets. One of these sets can be indistinguishable i.e. not within the scope of the functional visibility of the subject (Fig. 3).
We call such information flows orthogonal since they have different domains of parameters and functions. It is always possible to separate orthogonal information flows by introducing duplicating objects. As a result we get a system with non-orthogonal flows (Fig. 4).

Fig. 4. A separation of orthogonal flows

In the resulting system, the researcher can get an accurate informative feedback on the top part of the information flows scheme. At the same time he cannot observe functional response of the bottom part. This is due to the proposition proven in [14]. It states that the subject can always compare the simpler model to the actual system and verify its validity by limiting the set of the functional visibility.

From now on we consider that all the flows in the system are non-orthogonal. Consider the situation described on Fig. 2. In this case, information flows can be iterative and noniterative. By a noniterative system, we mean the system in which the information flows exist continuously changing the values of the inputs of the black boxes. Any uncontrolled input in a noniterative system leads to the situation where we can never get informative feedback.

For an iterative system, we assume that the values of one or more inputs of an object stay the same for multiple experiments. Thus we can get an informative feedback from the other information flows.

Most technical and information systems are iterative. It is possible to investigate their structure even without having knowledge of the values of all the inputs. There are two key issues that are important from the practical point of view in the research of such structures:

- Which subjective schemes are stable for the researcher?
- What information can one obtain from the objective scheme based on the knowledge of the subjective scheme? How do these schemes correspond to each other?

To answer these questions, it is necessary to formalize the system in such a way that it becomes clear how the subject can adjust its assumptions on the real system structure based on the informative feedback of the system. Since the results of this study cannot be predicted, it is important for us to describe the maximal information of the real system structure that can be obtained from the incoming and outgoing information in the model.

3. The mathematical formulation of the problem

3.1. Definition of objective and subjective graph

A scheme with a given topology will be identified with a directed graph. A vertex of the graph corresponds to an object in the scheme. A directed edge corresponds to the direction of
information flow in the scheme. Note that we can consider edges with the same two vertices and opposite directions. We associate a graph with one vertex distinguished from all others with the scheme that has the subject (the researcher). This is the so-called selected vertex.

Thus, we consider directed graphs of the following form

\[ \Gamma = \{V, E\}, \]

where

\[ V = \{V_0, V_1, \ldots, V_k\} \]

is a set of vertices containing the selected vertex \( V_0 \), and

\[ E = \{E_1, \ldots, E_m\}, \quad E_i = E_{i_1i_2} = (V_{i_1}, V_{i_2}), \]

\[ i_1 \in \{0, 1, \ldots, k\}, \quad i_2 \in \{0, 1, \ldots, k\} \setminus \{i_1\} \]

is a set of ordered pairs of vertices from \( V \), called directed edges or just simply edges.

To the objective (real) scheme we assign the graph, which we call by objective graph or etalon graph. Graph that corresponds to the subjective scheme is called subjective graph.

### 3.2. Transformations of graphs

While investigating real structure of a scheme the researcher can remove certain information flows between objects that do not exist in reality (objective scheme) or add existing but previously unknown information flows. After a finite number of actions the researcher will come to the situation where he will be unable to remove or add anything new to a scheme (the scheme is stable). We call such a process of modification of a scheme when investigating its real structure a scheme reduction. The stable scheme that one obtains after the described process has been completed is called reduced subjective scheme or just reduced scheme.

In terms of mathematical language, we refer to the rules of addition and removal of information flows as graph transformations. Note that such transformations are irreversible.

By stable graph we refer to the graph that corresponds to a stable scheme. We refer to the reduced graph denoted by \( \widehat{\Gamma} \) as a stable graph obtained from graph \( \Gamma \) using graph transformations.

### 3.3. The formulation of the problem in a general form

Consider the objective graph \( \Gamma^o = \{V^o, E^o\} \) and the subjective graph \( \Gamma^s = \{V^s, E^s\} \) such that

\[ V_0^o = V_0^s = V_0, \quad V^o \cap V^s \neq \emptyset \quad \text{and} \quad E^o \cap E^s \neq \emptyset. \]

The question is how similar objective graph \( \Gamma^o \) and reduced (subjective) graph \( \widehat{\Gamma}^s \) can be?

### 4. Graph transformation rules

Note that in the case of a disconnected graph it makes sense to only consider the connected component containing the selected vertex \( V_0 \). Indeed if the researcher has no connection to some part of a scheme, then in no way can he get information about objects from that part. Therefore, without loss of generality, we only consider connected graphs.
**Definition 1.** By a selected subgraph of graph $\Gamma$ we call subgraph of $\Gamma$ whose set of vertices $V_I$ contains the selected vertex $V_0$ and whose set of edges contains all the edges of $\Gamma$ whose heads and tails belong simultaneously to $V_I$.

**Rule 1.** Any selected subgraph not containing cycles is stable.

This typically corresponds to the situation where the researcher either cannot act on the object, or he has no feedback from the object. In this case, he cannot receive any new information about the objects, and therefore, he cannot add or remove information flows.

**Definition 2.** By an oriented cycle $\Gamma_I$ of graph $\Gamma = \{V,E\}$ where $I = (i_1, \ldots, i_s)$, $i_k \in \{1, \ldots, m\}$ for $k = 1, \ldots, s$ we call the graph $\Gamma_I = \{V_I,E_I\}$ with

$$V_I = \{V_{i_1}, \ldots, V_{i_s}\} \subset V$$

and

$$E_I = \{E_{i_1i_2}, \ldots, E_{i_{s-1}i_s}, E_{i_si_1}\} \subset E.$$  

An oriented cycle containing selected vertex $V_0$ is called a selected cycle.

**Definition 3.** By a difference $\Gamma_1 \setminus \Gamma_2$ of graphs $\Gamma_1 = \{V_1,E_1\}$ and $\Gamma_2 = \{V_2,E_2\}$ we call the connected component of the graph $\Gamma = \{V_1,E_1 \setminus E_2\}$ which contains the selected vertex $V_0$.

**Definition 4.** By a feedback ring $\Gamma_0$ of graph $\Gamma$ we call the set of all the vertices and all the edges of $\Gamma$ contained in arbitrary oriented cycle of $\Gamma$ passing through the selected vertex $V_0$.

**Rule 2.** Suppose there is a vertex $V_k$ that belongs to an oriented cycle containing the selected vertex $V_0$. Suppose the mentioned cycle belongs to $\Gamma_0$ and $\Gamma^*$. If there is an edge $E_{ik} = (V_i,V_k)$ with the tail $V_k$ such that $E_{ik} \in E^*$ and $E_{ik} \notin E^*$ then the graph $\Gamma^*$ is unstable (Fig. 5).

This rule yields the restoration of hidden edges (i.e., existing in an objective graph, but not in a subjective graph) of a subjective graph $\Gamma^*$ that will be performed according to the step of the algorithm given in the next section.

**Rule 3.** If $\Gamma^* \setminus \Gamma_0 \neq \emptyset$ then $\Gamma^*$ is unstable (Fig. 6).

This rule yields the removal of non-existent edges of a subjective graph $\Gamma^*$ that will be performed according to the step of the algorithm given in the next section.

**Corollary 1.** A graph that contains no vertices strongly connected to the selected vertex $V_0$ is stable.
5. Graph reduction algorithm

Graph reduction algorithm consists of the two steps.

1. Addition of the edges missing in the subjective graph, but existing in the objective graph.

2. Removal of the edges existing in the subjective graph, but missing in the objective graph.

**Remark 1.** Note the following important points. First of all, each step, in general, may consist of a different number of operations of addition (removal) of edges performed sequentially. Secondly, the order of steps in the algorithm in general can affect the outcome (reduced graph). Therefore, we will implement the steps of the algorithm in a specific order. First, one fully performs the step of addition of all the possible edges and after that, the step of removal of edges. This will ensure one obtains the same result when the algorithm is finished.

We now describe in detail the steps of the algorithm.

Let $\Gamma_I$ be an oriented cycle containing the vertex $V_0$, $\Gamma_I \subset \Gamma^o$, $\Gamma_I \subset \Gamma^s$. By $\bar{E}_I$ we denote the set of all the edges of the graph $\Gamma$ which do not belong to the cycle $\Gamma_I$ and whose tails are in arbitrary vertices of $\Gamma_I$. By $V_I$ we denote the set of vertices which are heads of the edges of $\bar{E}_I$.

**Step 1 (addition).** If the graph is unstable according to the Rule 2 then by adding some edges it can be reduced to the graph

$$\widehat{\Gamma}^s = \{\bar{V}^s, \bar{E}^s\},$$

where

$$\bar{E}^s = E^s \cup \bigcup_I (\bar{E}_I \setminus \bar{E}^s_I), \quad \bar{V}^s = E^s \cup \bigcup_I \left(\bar{V}^o_I \setminus \bar{V}^s_I\right). \tag{1}$$

Here the union is taken over all multi-indices $I$ for which there exist oriented cycles $\Gamma_I$ containing $V_0$ and such that $\Gamma_I \subset \Gamma^o$, $\Gamma_I \subset \Gamma^s$.

**Remark 2.** If in (1) one uses $\bar{E}^o_I$ instead of $\bar{E}^o_I \setminus \bar{E}^s_I$ and $\bar{V}^o_I$ instead of $\bar{V}^o_I \setminus \bar{V}^s_I$, the result remains the same. However, the proposed form is convenient because the given sets are exactly the sets of new edges and vertices (which are not contained in the previous version of the modified subjective graph).

**Step 2 (removal).** If the graph is unstable according to the Rule 3, then by removing some edges it can be reduced to the graph

$$\widehat{\Gamma}^s = \Gamma^s \setminus \left(\bar{\Gamma}^s \setminus \bar{\Gamma}^o\right).$$

1. The main result

The main result of this paper answers the question, "under what conditions can a subjective graph be reduced to the objective one?"

**Theorem 5.1.** If the feedback ring of subjective graph contains all its vertices, and the set of vertices of subjective and objective graphs coincide, then the reduced subjective graph coincides with the objective graph. This can be written down as

If $V^o = V^s \subset \widehat{\Gamma}^s$, then $\widehat{\Gamma}^s = \Gamma^o$. 

- 461 -
6. Example of an interpretation of the model

Suppose there is a subject $S$ (in the system) which can initiate the process $p_1$ with parameter $x$.

The process $p_1$, in turn, starts the process $p_2$ with the parameter $y$ that is generated based on $x$.

**Process $p_1 (x: \text{par})$**

\[
\begin{align*}
&y = \text{func}(x); \\
&\text{Create Process } p_2 \ (y); \\
&\ldots \\
&\text{end.}
\end{align*}
\]

The process $p_2$ generates parameter $z$ based on $y$. After that, the file $f_1$ is open for writing and the process writes $z$ down to it.

**Process $p_2 (y: \text{par})$**

\[
\begin{align*}
&z = \text{func}(y); \\
&\text{Open } f_1 \text{ for write; } \\
&\text{Write } z \text{ to } f_1; \\
&\ldots \\
&\text{end.}
\end{align*}
\]

The subject $s$ can view the contents of the file $f_1$, i.e. he is able to run the command:

**Open $f_1$ for read;**

\[
\begin{align*}
&k = f_1;
\end{align*}
\]

As a result, we get the following information flows scheme (Fig. 7).

![Fig. 7. Information flows scheme](image)

This scheme can be represented as a graph (objective and/or subjective), see Fig. 8.

![Fig. 8. Graph representation of information flows](image)

Here we assume that the subjective graph coincides with the objective one, i.e. assumptions of the subject $S$ on the information flows structure in the system are true.
Now let us consider two graphs which are unstable according to the Rules 2 and 3 (Fig. 9 and 10, respectively).

![Unstable graph according to the Rule 2](image1)

![Unstable graph according to the Rule 3](image2)

In the first case, the process \( p_1 \) in fact does not initiate the process \( p_2 \). This leads to the situation when the subject cannot discover any feedback when affecting the file \( f_1 \). If there is no other feedback ring in the graph, he will lose all the objects and information flows in this chain.

In the second case, there is an extra process that has an access for writing to the file \( f_1 \). This fact will be discovered by the subject sooner or later so that the new edge will appear in the subjective graph.

### 7. Analysis of real systems

Application of the model described above can be considered for information security systems. Take for example the classical Harrison-Ruzzo-Ullman access matrix model [4] which is considered in information security models.

The model defines:
- \( O \) — the set of objects in the system;
- \( S \) — the set of subjects in the system (\( S \subseteq O \));
- \( R \) — the set of access rights types of subjects to objects;
- \( T \) — access matrix with rows corresponding to the subjects and columns corresponding to the objects (\( T[s,o] \subseteq R \)).

As a result of applying a primitive operator \( \alpha \), a transition from the state \( q = (S,O,T) \) to the state \( q' = (S',O',T') \) is performed. We denote this transition by \( q \rightarrow_{\alpha} q' \).

In our case, it is necessary to introduce the subjective states of the system for all the subjects of the system

\[
\forall s_i \in S \ \exists q_{s_i} = (S_{s_i}, O_{s_i}, T_{s_i}).
\]

Each subject in the graph defined by the state \( q_{s_i} = (S_{s_i}, O_{s_i}, T_{s_i}) \) has an area of misinformation (fake objects and accesses) and an invisible area (objects and accesses that he does not
see in the objective access matrix). Denote them as

\[ q_{\text{mis}}^{s_i} = (S, O_s, T, T) \]

Invisible area:

\[ q_{\text{inv}}^{s_i} = (S, O_s, T, T) \]

Harrison-Ruzzo-Ullman model (HRU) is used for analysis of safe-states of systems depending on the possibility of a transition to the state \( q' \) for which the forbidden permissions appear in the respective cells of the matrix.

If we introduce subjective states of the system, the HRU system cannot perform transitions if they do not exist in the subjective matrices of any of the subjects. Thus we can express initial state of the system with \( N \) subjects as follows:

\[ q_0 = (S_{s_1} \cup \ldots \cup S_{s_N} \cap S, O_{s_1} \cup \ldots \cup O_{s_N} \cap O, T_{s_1} \cup \ldots \cup T_{s_N} \cap T). \]

Obviously, when intersecting the sets, the system can switch from an unsecure state to a secure one and vice versa. Consequently, if the system in which only the objective states of the HRU automaton are being analyzed is unsecure, then it does not mean that we also get an insecure system starting from subjective states of the HRU automaton. This gives us an opportunity to obtain a secure system by adjusting subjective access matrices.

In order to reduce analysis of the system to analysis of the information flows, one has to define graph vertices as objects of the set \( O \). One also needs to introduce a function that transfers elements of the access matrix to the edges of the graph. For \( R = \{\text{read}, \text{write}, \text{execute}\} \), it can be done as follows:

\[ \Gamma = (V, E), \quad V = O, \quad E = \{(v_1, v_2) \mid \text{write} \in T[o_1, o_2] \text{ or execute} \in T[o_1, o_2] \text{ or read} \in T[o_1, o_2]\}. \]

We have to implement this construction both for objective and subjective graphs. As a result, we will get the objective and the subjective access graphs. This will allow us to answer the following questions according to the above described theory:

- How can misinformation of the subjects lead them to the stable subjective states?
- What information can be obtained by the subjects based on their access to the objects and entities of the system (i.e., what graphs can they obtain through the research)?
- In which case can the information security administrator claim that the subjective image of the system constructed by him corresponds to the objective structure (according to the Theorem 5.1)?

In order to bring the information flows graph back to the access matrix \( T' \) of the HRU system, one has to perform the following transformations:

\[ r \in T'[o_1, o_2] \iff r \in T[o_1, o_2] \quad \text{and} \quad (v_i, v_j) \in E, \]

where \( r \in \{\text{read}, \text{write}, \text{execute}\} \). The result is the modified access matrix of the system, in which the state can be secure. Thus, by controlling the user perception of the system structure, one can bring the system from an insecure state to a secure one.
8. Conclusion and applications

The model of a researcher considered in the present work allows us to make the system secure by means of user perception about the system and about misinformation to the users. Besides the Harrison-Ruzzo-Ullman model, one can consider and make similar additions to other information security models such as the Take-Grant model or mandate and role access control model [5]. The model can also be used for conflictive systems modeling. For example, it can be used for the construction of the initial payment matrices in theoretical-game-models.

Mathematical models developed in this work can also be used in the field of software testing since they allow one to compare the structure of the tested system to the tests with a stable positive feedback. Based on the developed functional systems research methodology, we can come to software design patterns, which would avoid uninformative feedback on the testing stage.

This work was supported by the Federal Grant-in-Aid Program under Governmental (Contract No. 14.574.21.0126 , unique identifier RFMEFI57414X0126, first author); the state order of the Ministry of Education and Science of the Russian Federation for Siberian Federal University (task 1.1462.2014/K, second author) and the Russian Foundation for Basic Research (grant 15-31-20008-mol_a_ved, second author).

References

Математическое моделирование информированности субъектов в системах информационной безопасности

Михаил А. Стюгин
Алексей А. Кытманов

В статье рассмотрена проблема исследования функциональных структур на примере информационных систем. Особенность такого исследования заключается в том, что не всегда возможно добиться того, что результат исследования будет соответствовать реальности. Это крайне актуальная проблема в области разработки и анализа систем информационной безопасности и анализа программного обеспечения на предмет недекларируемых возможностей. В статье дана модель исследователя и исследуемого им функционального объекта. На основании данной модели показаны информационные ограничения исследователя. Построена математическая модель субъективной структуры исследуемой системы, и показано, в каких случаях она является устойчивой. Дан также ответ на вопрос, в каком случае субъект может утверждать, что его субъективная функциональная структура объекта соответствует действительной. Приведены примеры реализации данного подхода на математических моделях информационной безопасности.

Ключевые слова: математические модели информационной безопасности, модель исследователя, информационная структура конфликта, информационные потоки, дезинформация, модель черного ящика, теория графов.