Conditions for creating perfectly secure systems *

M Styugin

Research fellow, Reshetnev Siberian State Aerospace University, Krasnoyarsk, Russia

E-mail: styugin@gmail.com

Abstract The present paper reviews a method for establishing secure information systems by complicating the possibility to research them for potential adversaries. A formalized model of a researcher and a definition of a research secure system are presented. A theorem for conditions required for creating a system perfectly secured from research. The Shannon’s theorem of absolute security of perfect secrecy ciphers in cryptography is an instance of the theorem presented in the paper.

1. Introduction

Complexity of information systems has been constantly increasing lately. Both hardware and software complexity is increasing. This also influences the development of information security technologies. As vulnerabilities in complex systems cannot be detected and eliminated completely at the development stage, therefore information on the system’s operating principles and algorithms has to be concealed. Technologies for such “concealment” are based on security methods for protecting information systems from research. In the last few years, there were numerous studies in this area and the number is growing every year.

In 2015 alone, there were over 150 articles on different methods of concealing data/processes in the field of information systems and malicious software. Some rather specific issues can be pointed out there, which are quite frequently reviewed in the researches including those presented at top-tier conferences.

One of the current trends is development of technologies, which complicate data gathering for adversaries by implementing Data Mining [1] and technologies for concealment of user actions. The most widely known research on that problem was conducted by Steven Alpern [2] and Shmuel Gal [3].

The technology of constant system modification to protect the system from research (Moving Target Defence - MTD) [4] and the technology for protection software code and algorithm should also be mentioned. Recently there were over 150 different MTD techniques [5] concerning such spheres as local area network security [6], protection from software code injection [7], protection from XSS attacks [8], protection from DDoS attacks [9], etc.

However, despite the avalanche-like growth of publications in that area there are still no exact mathematical methods for analysis and comparison of research secure systems. The existing mathematical models study only separate system properties to compare their effectiveness as in [10, 11]. Still there is no general model that would define ideal objects such as an absolutely research-secure system for example a definition of an absolutely

* The reported study was funded by RFBR according to the research project No. 16-29-09456 офи_м

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd
indistinguishable data transmission channel in [12] actually does not provide a formal definition of its indistinguishability.

In order to create such theory, perfect secrecy conditions shall be established, which are implemented in cryptography as applied to information systems in general and formulate a general kind of Shannon Secrecy Theorem, for which the Shannon absolute secrecy theorem would be an instance.

2. Formalization of Objective for Protection against Research

The key feature of the present study is formalization of the objective for protection against research represented as aggregate of inputs, outputs and a discrete internal state. That kind of formalization allows introducing ideal objects such as a Perfectly Secure System and demonstrating feasibility of implementing some of the solutions in information security.

Conditions, in which an intruder may operate relatively to the targeted item, have been formalized earlier [13]. Assume that the intruder’s targeted item has \( f(par) \) function, which connects the item’s inputs and outputs, where \( par \) is a set of all possible arguments of function, i.e. \( f(p_1, ..., p_n) \). Inputs of the item may also be parameters unknown, that is the current set of the item’s inputs is selected from a set \( <par> \in \{<par>\} \). Curly brackets will be used for indication of all possible values of a variable. In the latter case it denotes a set of all possible combinations of \( <par> \), which the intruder can consider as the target’s inputs. A set of input parameters may be divided into a \( \{par\}^v \) set of parameters visible for the intruder and a \( \{par\}^nv \) set of parameters invisible for the intruder.

\[
(\{par\}) = \{par\}^v \cup \{par\}^nv, \{par\}^v \cap \{par\}^nv = \emptyset.
\]

The set of the item’s possible outputs is similar to the previous set:

\[
(\{f\}) = \{f\}^v \cup \{f\}^nv, \{f\}^v \cap \{f\}^nv = \emptyset.
\]

Hence, the set of conditions in which an intruder may operate relatively to the targeted item can be represented with a tuple of thee values:

\[
<\{par\}^v = <par>, \{par\}^v, f \in \{f\}^v>.
\]

Where \( <par>^v \) is the number of variables, which the intruder selected as the item’s “input”, \( par \) is the current item’s input state, \( f \) is the current item’s output state. The above formula is slightly different from the one presented in [13]. Here it is represented in a structure, which is more convenient for conclusions below. The above tuple contains three conditions: an intruder may select the set of parameters for \( <par>^v \) item’s input incorrectly; the true set of parameters may not be a part of its \( \{par\}^v \) parameter visibility and the item’s output may be outside the set of its functional visibility \( \{f\}^v \). The “worst” state of a targeted system for an intruder may be represented by the tuple below:

\[
<\{par\}^v \neq <par>, \{par\}^v, f \notin \{f\}^v>.
\]

In that case the intruder will not even be able to set the objective for their activity as shown in [14].

Additional parameters need to be added to the tuple for further formalization:

\[
\{f^v(par') = f(par), f(par) \in \{f(par)\}^v, <par>^v = <par>, par \in \{par\}^v, <f^v = <f>, f \in \{f\}^v>.
\]

Equality of \( f^v(par') = f(par) \) indicates that an intruder establishes \( f^v(par') \) function of \( f(par) \) target item correctly. As well as equality of \( <f^v = <f> \) indicates that the intruder established the item’s output \( <f^v \) correctly. Condition of \( f(par) \in \{f(par)\}^v \) indicates that the item’s function is visible.

In addition it should be noted that from a practical perspective there is normally no point in differentiating between visibility of some of the variables and awareness of their values. In
case a variable is visible, then its value is known. Therefore, only the property of being one of
the visible elements shall remain in the tuple:
\[ \{ \text{par} \in \{ \text{par} \}^\times, f(\text{par}) \in \{ f(\text{par}) \}^\times, f \in \{ f \}^\times \}. \]

For the sake of notation convenience, compliance to the condition (membership) will be
designated as 1 and non-compliance to the condition (not a part of the set) will be designated
as 0. Therefore a system which is perfectly suitable for an intruder should be represented as
\((1, 1, 1)\). The above notation describes a system complying to conditions \( \{ \text{par} \in \{ \text{par} \}^\times, f(\text{par}) \in \{ f(\text{par}) \}^\times, f \in \{ f \}^\times \}\). First a definition for “intruder’s objective” shall be stipulated. Intruder’s objective is to bring a system to the state with values \((1, 1, 1)\). Hence, a secure
system is a system that cannot be induced into a state of \((1, 1, 1)\). It should also be noted that
in a real-life system when each two of the system parameters are visible it always implies
visibility of a third parameter. For example, an input variable and a system’s function are
known, hence an output value can be found.

Therefore, a secure system can exist only in states with one visible variable. Such states
have the structures shown below:
\[ (0, 0, 1) - \{ \text{par} \notin \{ \text{par} \}^\times, f(\text{par}) \notin \{ f(\text{par}) \}^\times, f \notin \{ f \}^\times \} \]
\[ (0, 1, 0) - \{ \text{par} \notin \{ \text{par} \}^\times, f(\text{par}) \in \{ f(\text{par}) \}^\times, f \notin \{ f \}^\times \} \]
\[ (1, 0, 0) - \{ \text{par} \in \{ \text{par} \}^\times, f(\text{par}) \notin \{ f(\text{par}) \}^\times, f \notin \{ f \}^\times \} \]

Definition 1. A secure system is a system with one known (visible) parameter (input value,
output value, input-output dependence function) that does not allow finding the other two
system parameters.

Absolutely Secure System

We can represent a Perfectly Secure System with the Shannon Secrecy encryption
principle [5] for each of the conditions.

Definition 2.1 A Perfectly Secure System in state \((0, 0, 1)\) is a system in which conditions
(1) are maintained:
\[ \Pr(f(\text{par}) = \{ f(\text{par}) \}) = \Pr(f(\text{par}) = \{ f(\text{par}) \}) \mid \text{par} = \{ \text{par} \} \). \]
\[ \Pr(f(\text{par}) = \{ f(\text{par}) \}) = \Pr(f(\text{par}) = \{ f(\text{par}) \}) \mid f = \{ f \} \). \]

Definition 2.2 A Perfectly Secure System in state \((0, 1, 0)\) is a system in which conditions
of (2) are maintained:
\[ \Pr(f(\text{par}) = \{ f(\text{par}) \}) = \Pr(f(\text{par}) = \{ f(\text{par}) \}) \mid \text{par} = \{ \text{par} \} \). \]
\[ \Pr(f(\text{par}) = \{ f(\text{par}) \}) = \Pr(f(\text{par}) = \{ f(\text{par}) \}) \mid f = \{ f \} \). \]

Definition 2.3 A Perfectly Secure System in state \((1, 0, 0)\) is a system in which conditions
of (3) are maintained:
\[ \Pr(\text{par} = \{ \text{par} \}) = \Pr(\text{par} = \{ \text{par} \}) \mid f = \{ f \}), \]
\[ \Pr(\text{par} = \{ \text{par} \}) = \Pr(\text{par} = \{ \text{par} \}) \mid f(\text{par}) = \{ f(\text{par}) \}). \]

Then it shall be defined in which conditions a system can exist in states \((1), (2)\) and \((3)\).
They shall be designated as \(S(1) = 1\) for fulfillment of conditions for \((1)\) and \(S(1) = 0\) for
failure to fulfill conditions. \(S(2)\) and \(S(3)\) shall be designated in the same manner.

Theorem 1.1 \(S(1) = 1 \Rightarrow |\{f(\text{par})\}| \geq \max(|\{f\}|, |\{\text{par}\}|)\)

Here the system can be in state \(S(1)\) when a set of different functions between system input
and output are not less than the maximum between sizes of input and output value sets.

Two conditions shall be reviewed for theorem proof.
1. When \(|\{\text{par}\}| < |\{f\}|\) and let us review non-fulfillment of the first part of the condition:
\[ |\{f(\text{par})\}| < |\{f\}| \]
Set \( \{ f_{\text{par}} \} \) shall be defined so that it would include all values of \( f \) from a specific \( \tilde{p} \text{ar} \) parameter by enumeration of all possible functions of \( \{ f(\text{par}) \} \). The following is obtained
\[
\{ f_{\text{par}} \} \equiv \{ f | f = f(\tilde{p}ar) \text{ for some } f(\text{par}) \in \{ f(\text{par}) \} \}.
\]

It is evident that \( |\{ f_{\text{par}} \}| \leq |\{ f(\text{par}) \}| \). Hence, it may be concluded that there is at least one \( f' \in \{ f \} \), which is such that \( f' \notin \{ f_{\text{par}} \} \). Thus,
\[
\text{Pr}(\{ f = \{ f' \} \}) = \text{Pr}(f = \{ f' \} | \text{par} = \{ \text{par} \}) = 0 \neq \text{Pr}(\{ f = f' \}).
\]

The obtained expression contradicts \( S(1) \) condition, which stipulates that
\[
\text{Pr}(f = \{ f \}) = \text{Pr}(f = \{ f \} | \text{par} = \{ \text{par} \}).
\]

2. Here is a case when \( |\{ \text{par} \}| > |\{ f \}| \) and the first part of the condition is not observed:
\[
|\{ f(\text{par}) \}| < |\{ \text{par} \}|
\]

The set of \( \{ \text{par} \} \) shall be defined so that it would include all the values of \( \text{par} \), which can be used for obtaining the result of \( \tilde{f} \) by enumerating all possible functions of \( \{ f(\text{par}) \} \). As a result the following is obtained:
\[
\{ \text{par} \} \equiv \{ \tilde{p}ar \} f(\tilde{p}ar) = \tilde{f} \text{ for some } f(\text{par}) \in \{ f(\text{par}) \}.
\]

It is evident that \( |\{ \text{par} \}| \leq |\{ f(\text{par}) \}| \). Hence, there is at least one \( \text{par}' \in \{ \text{par} \} \), which is so that \( \text{par}' \notin \{ \text{par} \} \). Thus,
\[
\text{Pr}(\{ \text{par} = \{ \text{par}' \} \}) = 0 \neq \text{Pr}(\{ \text{par} = \text{par}' \}).
\]

The obtained expression contradicts \( S(1) \) condition, which stipulates that
\[
\text{Pr}(f = \{ f \}) = \text{Pr}(f = \{ f \} | \text{par} = \{ \text{par} \}).
\]

which is equivalent to
\[
\text{Pr}(\text{par} = \{ \text{par} \}) = \text{Pr}(\text{par} = \{ \text{par} \} | f = \{ f \}).
\]

**Theorem 1.2** \( S(2) = 1 \Rightarrow |\{ \text{par} \}| \geq |\{ f \}| \).

This theorem assumes that a system can exist in state \( S(2) \) when size of the set with potential system’s inputs is greater or equal to the size of the system’s input set.

Proof is similar to Theorem 1.1

**Theorem 1.3** \( S(3) = 1 \Rightarrow |\{ f(\text{par}) \}| \geq \max(|\{ f \}|, |\{ \text{par} \}|) \)

This theorem assumes that the system in state \( S(3) \) shall have a size of possible outputs set that is not less than the maximum value of size of potential system’s input and output values.

Proof similar to Theorem 1.1

All the above conditions are necessary, but they are not enough for establishing a Perfectly Secure System.

According to the commonly used terminology a term “uniform” will be used for denoting an absolutely random function or a bit chain where each subsequent bit can be predicted with a probability not higher than \( \frac{1}{2} \) regardless of computational asymptotic complexity.

Here is a definition for the uniform-function
\[
f(\text{par}) = \text{uniform} \iff \forall \text{par}: \text{Pr}(f) = 1/|\{ f \}|.
\]

That means that a function is uniform when it is established in such a way that any output function value for an intruder is equally probable. That is the function itself is not probabilistic, but a determinate one. However, it should be randomly selected hence it does not provide any information to an intruder.

**Theorem 2.1**

1. When \( f(\text{par}) \) is uniform and \( |\{ f \}| \geq |\{ \text{par} \}| \), then \( S(1) = 1 \) then and only then when for any \( \text{par} \in \{ \text{par} \} \) there are \( f_1 ... f_n \), where for each \( f_i, i = 1 ... n \) there are \( f_i, i = 1 ... n \), in
which each \( f_i, i = 1 \ldots n \) there are \( c \), where \( m = \frac{|\{f(par)\}|}{|\{f\}|} \) so that \( f_i = f_k^i(par), i = 1 \ldots n, k = 1 \ldots m \).

2. When \( \{par\} \) is uniform and \( |\{f\}| \leq |\{par\}| \) then \( S(1) = 1 \), then and only then when for any \( f \in \{f\} \) there are \( \text{par}_1 \ldots \text{par}_n \), where for each \( \text{par}_i, i = 1 \ldots n \) there are \( f_1^i(\text{par}_i), \ldots, f_m^i(\text{par}_i) \), where \( m = \frac{|\{f(par)\}|}{|\{par\}|} \) so that \( f = f_k^i(\text{par}_i), i = 1 \ldots n, k = 1 \ldots m \).

The first part of the theorem asserts that for each \( \text{par} \) there are \( n \) function results and for each parameter and function result there are \( m \) functions, that can be matched with them. The second part of the theorem asserts that for each \( f \) output value there are \( n \) different parameters, which can be used for obtaining it by \( m \) different functions.

Formal proof of the above theorem is quite lengthy, so only the general principle of its development is presented here. Let’s review the first part of the theorem, when \( |\{f\}| \geq |\{par\}| \). In this case there shall be as many functions for transforming \( \text{par} \) into \( f \) to have the opportunity to obtain each value of \( f \) from each \( \text{par} \). Otherwise, it will result in non-fulfillment of \( S(1) \) condition. When the number of functions is sufficient to obtain each \( f \) from each \( \text{par} \) then they also shall be evenly distributed throughout all the pairs to make sure that probability of a pair’s occurrence is not higher than the occurrence probability for the other pairs, which also causes violation of \( S(1) \) condition. Hence, the number of functions \( f(par) \) shall be divisible by the number of output results designated by \( f \).

The second part of the theorem has a similar proof.

**Theorem 2.2**

When \( \text{par} \) is uniform then \( S(2) = 1 \), then and only then, when for each \( \text{par} \) and for each \( f \) there is a function \( f_1(\text{par}), \ldots, f_m(\text{par}) \), where \( m = \frac{|\{f(par)\}|}{|\{par\}|} \) which are such that \( f = f_k(\text{par}), k = 1 \ldots m \).

Proof is similar to Theorem 2.1.

**Theorem 2.3**

1. When \( \text{par} \) is uniform and \( |\{f\}| \geq |\{par\}| \) then \( S(3) = 1 \), then and only then, when for each \( \text{par} \in \{\text{par}\} \) there are \( f_1 \ldots f_n \), where for each \( f_i, i = 1 \ldots n \) there are \( f_1^i(\text{par}), \ldots, f_m^i(\text{par}) \), where \( m = \frac{|\{f(par)\}|}{|\{f\}|} \) are such that \( f_i = f_k^i(\text{par}), i = 1 \ldots n, k = 1 \ldots m \).

2. When \( \text{par} \) is uniform and \( |\{f\}| \leq |\{par\}| \) then \( S(3) = 1 \) then and only then, when for each \( f \in \{f\} \) there are \( \text{par}_1 \ldots \text{par}_n \), where for each \( \text{par}_i, i = 1 \ldots n \) there are \( f_1^i(\text{par}_i), \ldots, f_m^i(\text{par}_i) \), where \( m = \frac{|\{f(par)\}|}{|\{par\}|} \) are such that \( f = f_k^i(\text{par}_i), i = 1 \ldots n, k = 1 \ldots m \).

Proof is similar to Theorem 2.1.

**Conclusion**

The paper presented a formalization of an information system researcher including its limitations. A research secure system and an absolutely research secure system can be defined based on the researcher’s information limitations. A generalized theorem formulating the conditions required and sufficient for an absolutely research secure system was presented. The Shannon absolute secrecy theorem in cryptography is an instance of the presented theorem.
The created model confirms the result provided in [12]. The obtained theoretic results can be used for designing indistinguishable information systems when it is required to complicate research for an external agent.

References
[1] Cheng P, Lee I, Pan J-S and Lin C.-W., Roddick J F 2015 Hide association rules with fewer side effects. *IEICE Transactions on Information and Systems*. Volume E98D, Issue 10, Pages 1788-1798
[2] S. Alpern and Gal S 2002 Searching for an agent who may or may not want to be found. *Operations Research*, 50(2):311–323
[3] Gal S 2001 On the optimality of a simple strategy for searching graphs. *International Journal of Game Theory*, 29(4):533–542
[4] Jajodia S 2011 Moving Target Defense. Creating Asymmetric Uncertainty for Cyber Threats. Series: Advances in Information Securit. - 184 p
[5] Okhravi H, Hobson T, Bigelow D and Streilein W 2014 Finding Focus in the Blur of Moving-Target Techniques. *IEEE Security and Privacy*, vol. 12 pp 16-26
[6] Carvalho M. and Ford R. 2014. Moving-target defenses for computer networks. *IEEE Security and Privacy*. Vol. 12(2) pp 73-76
[7] Larsen P, Brunthaler S and Franz M 2015 Automatic Software Diversity. *IEEE Security and Privacy*. Vol. 13 (2) pp 30-37
[8] Portner J, Kerr J and Chu B 2015 Moving target defense against cross-site scripting attacks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 8930 pp 85-91
[9] Ma D, Xu Z and Lin D 2015 Defending blind DDoS attack on SDN based on moving target defense. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*. Vol. 152 pp 463-480
[10] Jin B Hong and Dong Seong Kim 2014 Scalable Security Models for Assessing Effectiveness of Moving Target Defenses. *Dependable Systems and Networks (DSN), 44th Annual IEEE/IFIP International Conference*. pp 515 - 526
[11] Jin Hong and Dong-Seong Kim 2012 HARMs: Hierarchical Attack Representation Models for Network Security Analysis. Originally published in the *Proceedings of the 10th Australian Information Security Management Conference*, Novotel Langley Hotel, Perth, Western Australia, 3rd-5th December, 2012
[12] Mikhail Styugin. 2015. Absolutely Indiscernible Data Transfer Channel. *Proceedings of The 14th European Conference on Cyber Warfare and Security (ECCWS-2015)*. pp 286-293
[13] Styugin M 2014 Protection against System Research. *Cybernetics and Systems: An International Journal*. Vol. 45 (4) pp 362-372
[14] Katz J and Lindell Y 2014 Introduction to Modern Cryptography, Second Edition. N.Y.: Chapman and Hall/CRC. 603 p