Malicious Hardware: characteristics, classification and formal models

Valeriy Gorbachov
Kharkiv National University of Radio Electronics
gorbachov@kture.kharkov.ua

Abstract

The paper addresses the threat to the security of using electronic systems, which may include malicious inclusions. The action classification of malicious hardware (MH) is given. The MH formal models, as well as formal model of unauthorized access, executed by MH, are based on the subject-object concept.

1. Introduction

Electronic Systems (ES) that contain embedded malicious hardware represent a serious threat, especially for government, aeronautic, financial and energy system applications. MHs can be implemented as hardware modifications to application specific ICs (ASICs), microprocessors, digital signal processors , or as IP core modifications for field programmable gate arrays (FPGA) [1].They are able to turn off the CPU, to send confidential information and bypass the software user authentication mechanisms. There are some important characteristics of this type of threat: standard testing methods, such as the common functional verification and Automatic Test Pattern Generation (ATPG) cannot always be used to solve the problem of detecting MH [2, 3]; identification of the threat sources without special tools is practically impossible; even in cases when an information security violation is detected, it is very difficult to prove that this action was performed by MH. These and other features make MHs very promising embedded devices for planning of electronic terrorism. Therefore, detecting and preventing approaches are in the attention centre of IT systems security investigation.

In section 2, we define the action classification of MHs based on the properties of information. We consider unauthorized access model and models of MHs in sections 3 and 4, respectively.

2. MH Action Classification

The development of effective methods designed to detect and block MHs depends on completeness of MH characteristics and classification schemes. In the frame of information security (IS) theory [4] the most important properties of information are confidentiality, integrity, and availability. Thereby MH actions violate these information properties. We propose to partition MHs into three groups: MHs which actions violate the confidentiality, integrity, and availability, respectively.

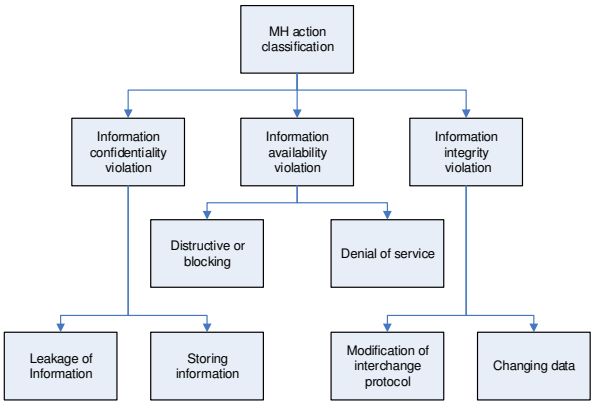

Figure 1. Malicious hardware action classification

Although, it is possible for MHs to be hybrids of this classification, e.g., have more than one action characteristic. We believe this MH action classification is more universal for the development of MH models.

In this paper, we follow and expand the MH action classification, proposed in [5, 6, 7].

3. Model of Unauthorized Access

In this paper, the MH model of unauthorized access is based on the subject-object approach [4]. This approach uses, firstly, the concepts: object, subject, and operation of access for the pair «subject-object»; secondly, the security theorem axiom: all problems of information security are described by accesses of subjects to objects. Aforementioned abstract concepts can be represented by the following physical representations in an electronic system:

Object (O_j) is a part of system resources, which is at time t in the passive state as for information and for other hardware components of the ES.

978-1-4799-7630-0/14/$31.00 ©2014 IEEE

Subject (S_i) is an electronic component, which is in the active state, in other words, able to access an object at time t.

Access is a category of subject-object model, which describes the process of impact performing by subjects on objects. Obviously, any access operation, including MH access, has to have the hardware support in the form of a channel.

Considering the MH, as a certain type of subject that is capable to realize the unauthorized access, it is necessary to note the following its features.

Firstly, MH, as a hardware - software resource of ES, in the passive state may be a part of an object.

Secondly, the information threat in an ES can come only from the active subject (MH), that currently possesses the resource management function.

Thirdly, the subjects can influence each other by means of objects that in turn can be changed by subjects. For example, an activation of MH.

In general, the pair (S_i, O_j) is connected by the set of allowed operations X_A. This set is defined by security policy (SP) and is a subset of the set X of all possible operations for each pair. At the same time, pair (S_i, O_j) may be connected by the set of not allowed operations X_{NA}. It is obvious that $X = X_A \bigcup X_{NA}$. The purpose of SP is to control and block operations from the set of X_{NA}.

The mechanism of access (including unauthorized access), consists of two stages: creating (activating) a subject and information exchange between objects.

Let us formalize the access operation of subject to objects with the aim of creating new subjects. This operation can be demonstrated by the following example. Suppose that a certain process running in a computer system needs to read some data from a hard disk (HD) at time t. For this purpose, the process sends the request to the hard disk controller. At this time, the process, which initiates the data acquisition is a subject, and the hard disk controller is an object. In doing so, the subject changes the contents of object's internal registers, thus changing its properties and creating a new subject in the hard disk controller. At the next time (t+1), the generated subject refers to the next object (directly to the controller, which manages the mechanics of a hard disk) and sends the corresponding requests to it, changing it, and thereby, creating a new subject, etc.

To formalize the mechanism of new subjects' creation, the following definitions will be used [4].

Definition 1. If an action of a subject S_j on an object O_i results in the creation of the subject S_k, the object O_i is called the source for the subject S_k. S_k is called the created subject.

Definition 2. The subject S_j, which creates a new subject S_k in the object O_i, in turn, is called activating or creating subject for the subject S_k,

Definition 3. The object O_i is associated with the subject S_j, if the subject S_j uses information of the object O_i.

Now let us introduce the operation:

$$\text{Create}(S_j, O_i, P') \rightarrow S_k, \quad (1)$$

which means, the S_j impacts the object O_i with the help of instruction stream P' in order to create the subject S_k. The operation (1) is called the subject creation operation.

If the subject S_j is an adversary and the subject S_k is a MH, then the activation process is described by the following [7]:

$$\text{Create}(S_j, O_i, P'_{unsp}) \rightarrow S_{MH}, \quad (2)$$

where P'_{unsp} is a unspecified instructions stream. The stream P'_{unsp} can be generated by either an external subject (adversary) or an internal subject.

If Create $(S_j, O_i, P') \rightarrow$ NULL (NULL is the empty set), the creation of a new subject in the object O_i with an activating impact of S_j is impossible.

Let us now formalize the mechanism of access in order to exchange information between objects. By definition, the objects of the system are passive. It is obvious, that in order to complete an exchange operation the existence of a data stream P'' between the objects is needed. This stream is always initiated and implemented by the subject (in case of unauthorized access - by the MH). This means that the operation of data stream creation is executed by the subject associated with the source object. Thus, the creation of the stream P'' between the object O_i (object - source) and the object O_j (object - recipient) should always be performed by two operations. First, the subject S_j creates the subject S_k in the object O_i (1). Then the subject S_k creates the data stream P'' from O_i to O_j . Thus the creation of the data stream P'' is described by the operation:

.

$$\text{Stream}(S_k, O_i, P'') \rightarrow O_j. \quad (3)$$

Thus, the access model for this case is represented by two operations (1) и (3):

$$\text{Create}(S_j, O_i, P') \rightarrow S_k,$$
$$\text{Stream}(S_k, O_i, P'') \rightarrow O_j. \quad (4)$$

Suppose that the subject S_j plays a role of an adversary and the subject S_k is a MH, then the model of unauthorized access realised by MH would be:

$$\text{Create } (S_j, O_j, P'_{unsp}) \rightarrow S_{MH},$$
$$\text{Stream } (S_{MH}, O_j, P''_{unsp}) \rightarrow O_m. \quad (5)$$

where P'_{unsp} and P''_{unsp} is unspecified instruction and data stream, respectively.

4. Formal Models of Malicious Hardware

Let us consider in more detail MHs violating the confidentiality, integrity and availability of information, develop their models for the purpose of their implementation in modeling SP.

4.1. MH that violates confidentiality of information

The main function of this type of malicious inclusions is information accumulation and transmission to a third party. The process of information accumulation consists of the activation and accumulation phase.

The operation of switching MH from standby to the active state for information accumulation is described using (2):

$$\text{Create } (S_j, O_j, P'_{unsp}) \rightarrow S'_{MH}, \quad (6)$$

where S'_{MH} is a MH in the accumulation state, the object O_j is always the source for the subject S'_{MH}.

The process of information accumulation is described using (3):

$$\text{Stream } (S'_{MH}, O_m, P''_{unsp}) \rightarrow O_{jмем}, \quad (7)$$

where O_m is a standard object of ES, that stores data; the object $O_{jмем}$ is the memory module of MH; P''_{unsp} is an unspecified data stream. The object O_m can be any object associated with the MH. Thus, the pair of operations (6) and (7) describes the process of information accumulation executed by MH.

The fundamental conclusion follows from the above-mentioned example: to ensure the security policy, it is necessary to control both the instruction stream P'_{unsp} and the data stream P''_{unsp}.

The process of information transmission by MH is implemented in two steps: activation for the transmission and transmission of information.

The activation of MH for the transmission of accumulated information is performed by the operation:

$$\text{Create } (S_j, O_j, P'_{unsp}) \rightarrow S''_{MH}. \quad (8)$$

where S''_{MH} is a MH in the state of information transmission. As in the previous case, the operation is activated by the instructions from the stream P'_{unsp}.

The transmission of accumulated information is performed by MH in the stream P''_{unsp}, which is described by the operation:

$$\text{Stream } (S''_{MH}, O_{jмем}, P''_{unsp}) \rightarrow O_m, \quad (9)$$

where P''_{unsp} is unspecified stream. O_m is an object-recipient.

Consider the process of switching MH from the active state to the standby. This procedure is a special case of the activation procedure of MH (8) and will differ by only instructions in the stream P'_{unsp}.

4.2. MH that violates the information availability

The main function of this type of MH is blocking/destruction of entire system or individual modules.

A formal model of a blocking/destruction describes the next stages of its functioning.

1. Switching MH from the standby to the active state is described by means of operation (6):

$$\text{Create } (S_j, O_j, P'_{unsp}) \rightarrow S''_{MH}, \quad (10)$$

where S''_{MH} is a MH in the blocking/destruction state; P'_{unsp} is an unspecified instruction stream.

In the active state MH of this type should create data stream P''_{unsp} in the object O_j (for example, the forbidden combination of signals that destructs or blocks the object O_m). That action model is:

$$\text{Stream } (S''_{MH}, O_j, P''_{unsp}) \rightarrow O_m. \quad (11)$$

In (10) the object O_j is a source for MH, in (11) the object O_j is always associated with MH.

As compared with the activation process, the switching destructive/blocking MH from the active state to the standby does not have any specific operations. It should be noticed that depending on functions implemented in this MH, it can lead to permanent or temporary disable the system. For this

reason, switching blocking MH to the standby mode will not be needed for all MHs.

4.3. MH that violates the information integrity

The main function of this type of MH is the destruction of data or modification of data transmission protocols.

In active state MH of this type manifest itself in appearing data streams aimed to violate the information integrity. MH may have several versions of the external manifestation: modification or substitution of destination addresses in various communication devices; modification of data for the purpose of integrity violation; changing system settings for the complete blocking of information security systems.

The formal model of the MH that violates the information integrity describes two steps of its behavior:

Step 1. The activation of MH:

$$\text{Create } (S_j, O_j, P'_{unsp}) \rightarrow S'''_{MH}, \qquad (12)$$

where S'''_{MH} is a MH in the state of violation of the information integrity.

Step 2. The modification of information:

$$\text{Stream } (S'''_{MH}, O_j, P''_{unsp}) \rightarrow O_m, \qquad (13)$$

where P''_{unsp} – data stream aimed to violate the information integrity of O_m.

In (12) the object O_j is a source for MH, in (13) the object O_j is always associated with MH.

The process of switching MH from the active state to the standby is similar to that described for MH violating the confidentiality of information.

5. Conclusion

The main conclusion of the work is:

1. The operation of MHs has common operation stages: standby, activation, active state, switching to the standby.
2. Any MH can be detected at one of the stages (excluding standby mode) by detecting unspecified streams (P'_{unsp}, P''_{unsp}).
3. In order to transfer the streams P'_{unsp} and P''_{unsp} channels are needed.
4. Models of MHs and unauthorized access can be used in order to develop effective methods for design protected electronic system architecture.

6. References

[1] D. Colins, "Trust in Integrated Circuits (TIC)", DARPA Solicitation BAA07-24, 2007 (www.darpa.mil/mto/solicitations/baa07-24/index.html).

[2] F. Wolff, C. Papachristou, S. Bhunia and R. S. Chakraborty, Towards Trojan-Free Trusted ICs: Problem Analysis and Detection Scheme, Design, Automation and Test in Europe, Munich, Germany, Mar 2008, pp. 1362 - 1365.

[3] Dan Wilt, "Trusted ICs Proposers Day Metrics Discussion", DARPA, 2007, (www.darpa.mil/mto/solicitations/baa07-24/Proposers_Day_Final.pdf).

[4] D.P. Zegzhda, Ivashko A.M. Fundamental of Information Systems Security - M: Telecom, 2000. - 452p.

[5] X. Wang, M. Tehranipoor, J. Plusquellic, Detecting malicious inclusions in secure hardware: Challenges and solutions // Proceedings of the 2008 IEEE International Workshop on Hardware-Oriented Security and Trust. – 2008. - Pages 15-19

[6] V.A. Gorbachov, I.N. Ivanisenko, Classification and formal models of hardware trojans, Applied Radio Electronics and Informatics, Sci. Journ., Kharkov: KhNURE.- vol. 6, 2007. № 2 pp. 306-310.

[7] Benjamin Sanno, Detecting Hardware Trojans, Ruhr-University Bochum, Germany July 22, 2009

[8] V.A. Gorbachov, Formal basis of malicious hardware blocking methods/ Applied Radio Electronics, Sci. Journ., Kharkov: KhNURE.- vol. 11, 2012. № 2 pp. 275-280.