Formal Specification of Security Properties using Z Notation

Shafiq Hussain, Peter Dunne and Ghulam Rasool
1Department of Computing, Engineering and Technology, University of Sunderland, UK
2Department of Computer Science, COMSATS Institute of IT, Lahore, Pakistan

Abstract: Software security is a challenging issue for distributed and open systems. Despite the importance of external protections of software systems, internal security has significant impact on the overall security of the software systems. In this study, internal security issues of software systems are addressed. Internal security of software systems is defined in terms of security properties: authentication, authorization, confidentiality, integrity, non-repudiation and resource availability. Internal security of software systems largely depends on the integration of these security properties into the software systems. A precise and unambiguous representation of these security properties is crucial for any successful secure system. Majority of the existing models are based on informal or semi-formal approaches to model these security properties. But no model is based on formal methods. Therefore, in this study, a formal specification of these security properties is presented in Z notation because formal methods are the only way to specify system properties unambiguously, completely and precisely. The resulting models are then analyzed by using Z/EVES theorem prover. The formal specifications of these security properties are analyzed only for syntax checking, type checking and automatic proofs of models.

Keywords: Availability, authentication, authorization, confidentiality, formal methods, integrity, non-repudiation

INTRODUCTION

Software security is a challenging and most demanding aspect of current software systems, particularly for web based systems. The emergence of sophisticated and complex distributed systems has made software security as an integral component of secure systems. Software security depends not only on the external protective measures but it also depends on the internal security aspects of software systems themselves. The internal security issues of software systems are more challenging and complicated as compared to external protections. This research is focused on the internal security properties of software systems. For internal security of software systems, security properties: confidentiality, integrity, non-repudiation, availability, authentication and authorization must be integrated into the systems. In authentication unique identification of a user or entity is established. Any user who wants to get access the system must have credentials to access the system by system must have predefined for that user. Authorization security property enables an authenticated user to access a particular resource in the system. Users, who want access, must have access rights for that resource. Confidentiality is the property which allows an authorized user to read the contents of a confidential resource. Confidentiality defines the read access policy of the system. Only those users can read confidential data who has rights to read that data. Integrity security policy allows an authorized user to write on a resource. Integrity policy defines write access for users of the system. Only those users, who have write permissions, can write data on a resource that has permissions for writing. Non-repudiation security property ensures auditing and logging mechanisms of software systems. Non-repudiation requires that every action by all the users of the system must be properly recorded and logged in the system. Availability security property enforces the availability of system resources to legitimate users. Every user of the system must get sufficient resources available in the system to complete his/her tasks. These security properties have been specified formally by using Z notation, a formal specification and modeling language and Z/EVES theorem prover has been used for analyzing the resulting models.

LITERATURE REVIEW

With the speedy developments in the area of distributed computing and web based systems, the needs to solve security issues become critical for smooth and successful operations of software systems (Subashini and Kavitha, 2011). For web based distributed systems, there is a need of secure communication mechanism because all communication is dependent on message passing in these systems (Savola and Abie, 2010). Security properties are the mechanisms used to stop an attack on software systems.
The security properties are required to be specified precisely at the requirements engineering phase of software development (Wolter et al., 2009). The security properties such as authentication, authorization, confidentiality, integrity, non-repudiation and availability are the main security properties and must be built into the software systems (Bertino et al., 2009; Menzel and Meinel, 2009). Authentication and authorization are the basic controls for software systems and all other security mechanisms such as confidentiality are dependent on these two properties (Jie et al., 2011; Welch et al., 2003; Thompson and Montague, 2011; Sandhu and Samarati, 1994). Authentication property makes it ensure for a user to be identified the system as a legitimate user of the system (López et al., 2007). Authorization is built on the authentication property and determines the access or denial of a user to access particular resources in the system (Joshi et al., 2005; Hu and Ahn, 2011; Ahn and Sandhu, 2000). Strict authentication and authorization mechanisms enable an enhanced and more secure environment for execution of business on the systems particularly electronic commerce and banking systems (Ferraiolo et al., 2001; Kolovski et al., 2007). Confidentiality property determines the read access to data in the system and grants read access to only those users who have read rights for that data (Bishop, 2003b). Integrity property determines the write access to data in the software system and grants write access to users who have write access for that data (Zhou et al., 2010; Xu and Liu, 2009; Thompson and Montague, 2011). Non-repudiation property ensures the auditing and logging mechanism in the software systems. It makes it mandatory to record any action performed by authorized users in the system (Feng et al., 2010; Onieva et al., 2009). Availability property manages the readiness of system resources for legitimate users. It ensures system resource must be available to legal users when demanded and no deadlock occurs (Milanovic and Milic, 2011; Bishop, 2003a; Arci et al., 2003). Formal methods are mathematics based techniques for developing mathematical models of software systems and enables to analyze software designs to remove flaws before implementation. Hence, reducing huge effort spent at the testing phase (Jacky, 1996; Xia and Tang, 2008; Zafar and Alhumaidan, 2011; Sidek and Ahmad, 2009). Formal specification helps us to specify, analyze and design computer systems ranging from commercial systems such as banking systems to critical systems such as air traffic control systems (Zafar, 2009; Coronato and De Pietro, 2011; Hussain et al., 2011). Software security is important area for the application of formal methods because security properties are impossible for exhaustive testing. Formal methods have been used successfully in many security systems such as SSAI and Correctness by Construction (Gilliam et al., 2001; Dimitrov, 2011; Woodcock et al., 2009).

The use of formal methods for software systems helps to analyze system properties by using powerful tools to produce formal proofs and making possible the production of more reliable and secure software systems (Almeida et al., 2011).

**FORMAL METHODS**

Formal methods are mathematics based tools and techniques used for developing mathematical models and designs of software systems. Formal methods can be applied at all the stages of software development life cycle from specification to implementation. Formal methods help to write formal specifications from the requirement documents in a precise way. In this way, common specification errors such as incompleteness and ambiguities can be removed. These specification errors are occurred due to the ambiguous nature of design methodologies being used for developing software applications. These design methodologies are based on informal and semi-formal techniques. The solution lies with the use of formal methods. By using formal methods, software systems are designed without ambiguities in formal specification language such as Z. The resulting models are then analyzed by using formal method tools such as theorem provers and model checkers. These models can be verified and validated by automatic formal methods tools. Formal proofs of desired properties can also be generated by using theorem provers such as Z/EVES and Isabelle. There are some areas of computer science such as security where exhaustive testing is not possible because security properties are very hard to prove before implementation. Therefore, formal methods can be used to validate security properties before implementation.

Theorem provers are normally used to generate formal proofs of software systems while model checkers are used to validate models and remove faults such as deadlocks. The use of formal method specification languages, model checkers and theorem provers depends on the nature of the application being developed. Formal methods can be categorized broadly into three types: formal specification languages such as Z, VDM-SL, VDM++, Object-Z, B-Method, Event-B and Alloy etc., model checkers such as SPIN, Alloy and Pro-B, theorem provers such as Z/EVES, Isabelle, Coq and RODIN. In this study, security properties have been specified in Z notation. Z/EVES theorem prover has been used for analysis of models produced in Z. Z is a model based language and built on set theory. Models in Z are a collection of state variables, invariants on state variables, set of operations and a set of pre and post conditions.

**INFORMAL MODEL**

In this study, six security properties have been considered. These security properties include
authentication, authorization, confidentiality, integrity, non-repudiation and availability. These security properties are defined informally as follows:

**Authentication:** Authentication is the mechanism in which unique identification of a user of the software system is determined. Authentication process ensures a login policy stored in the software system. Any user who wants to login into the system must provide unique credentials such as passwords defined in the login policy for that user.

**Authorization:** Authorization property grants or denies access to system resources for a particular user. If the user has access rights for intended resources, he or she is granted access to that resource otherwise, no access is granted. Access is granted to user depending upon the policy already defined for the users according to the roles of the users.

**Confidentiality:** Confidentiality property determines the read access to data in the system and grants read access to only those users who have read rights for that data. Confidential data can only be read by the users who are authorized to read that data. Some additional credentials are required to read confidential data.

**Integrity:** Integrity property determines the write access to data in the software system and grants write access to users who have write access for that data. Data can only be written by the users who are authorized for it. There must be a strict writing policy in the system to ensure integrity of data. The common issue of data integrity such as redundancy, inconsistency etc. must be considered.

**Non-repudiation:** Non-repudiation property ensures the auditing and logging mechanism in the software systems. It makes it mandatory to record any action performed by authorized users in the system. This policy sets limits on the legitimate users of the system and ensures that every action or activity of these users must be recorded. It helps for auditing at the later stages.

**Availability of resources:** Availability property manages the readiness of system resources for legitimate users. It ensures that system resource must be available to legal users when demanded and no deadlock occurs. The resource allocation algorithm should be defined in such a way that at least sufficient system resources remain available to legitimate users all the time. System resources must not be consumed by existing users of the system.

**FORMAL MODEL**

The formal models of security properties have been developed by using Z notation, a formal specification and modeling language. This formal model consists of static model and dynamic model. In static model, state variables are defined formally along with invariants on these variables. The invariants impose constraints on data hold by these variables. Furthermore, these invariants should remain true for all the time for smooth functioning of the system. In dynamic model, operations on the system are defined along with a set of pre and post conditions.

**Static model:** The static model of security properties starts with the definition of basic types for users, credentials, resources and actions on those resources. In the following specification USER represents set of users, CREDENTIAL represents set of credentials the user may have, RESOURCE represents set of resources in the system and ACTION represents the set of action allowed on system resources in the system:

\[ \text{[USER, CREDENTIAL, RESOURCE, ACTION]} \]

The following specification describes global types in the system. The global types are accessible in all the schemas on the system. In this model of security properties, only one global type OPERATIONS has been defined. This global type is used in the definition of confidentiality and integrity properties:

\[ \text{OPERATIONS::=READ|WRITE|APPEND|EXECUTE} \]

This schema describes the state of the system having state variables and invariants on them.

Access Control System

Users: P user
Resources: P resource
Action: P action
registered_users: User→P CREDENTIAL
user_resources: USER→P RESOURCE
resource_actions: RESOURCE→P ACTION
user_actions: User→P ACTION
authenticated_users: User→P CREDENTIAL
authorized_users: User→P CREDENTIAL
log: USER→P ACTION
owner: USER→P RESOURCE
dom registered_users ⊆ users
dom users_resources ⊆ users
∀u: USERS| u ∈ dom users_resources. User_resources u ⊆ resources
dom user_action ⊆ users
∀u: USER| u ∈ dom user_action . user_actions u ⊆ actions
dom resource-action ⊆ resources
∀r: RESOURCE| r ∈ dom resource_action
resource_action r ⊆ actions
dom authenticated_users ⊆ users
dom authorized_users ⊆ users
dom log ⊆ users
∀u: USERS| u ∈ dom log. log u ⊆ actions
This is the initialization schema and which ensures that at least one state of the system exists and system will work for at least one time.

\[
\text{init} = \emptyset
\]

Dynamic model: The following is the dynamic model of the system in which basic operations for the system have been defined as follows:

Basic operations: This operation adds users into the system. The pre-condition says that users to be added must not already be added into the system.

Add_Users

\[
\text{Add_Users} \triangleq \text{Access Control System}
\]

\[
\forall u : \text{USER} \mid u \in \text{users}. U \not\in \text{users}
\]

This operation adds users into the system. This operation requires that resources to be added must not already be added into the system.

Add_Resources

\[
\text{Add_Resources} \triangleq \text{Access Control System}
\]

\[
\forall r : \text{RESOURCE} \mid r \in \text{resources}. r \not\in \text{resources}
\]

This operation adds resources into the system and ensures only allowed actions on the system resources. The pre-condition says that actions to be added must not already be added in the system.

Add_Actions

\[
\text{Add_Actions} \triangleq \text{Access Control System}
\]

\[
\forall a : \text{ACTION} \mid a \in \text{actions}. a \not\in \text{actions}
\]

This operation adds registered users in the system. Only those users who have proper credentials can be registered in the system. They must be recognized in the system and must not already be registered in the system.

Add_Registered_User

\[
\text{Add_Registered_User} \triangleq \text{Access Control System}
\]

\[
\forall u : \text{USER} \mid u \in \text{dom owner}. u \not\in \text{resources}
\]

Formal model of security properties: Formal model of security properties developed in Z notation is presented in this subsection. All the six security properties: authentication, authorization, confidentiality, integrity, non-repudiation and availability of resources are specified in the following paragraphs.

The first schema below represents authentication security property. A user who wants to login into the system, enters user id and credentials. If the user supplied user id and credential are matched with those already stored in the system, the user is allowed to login into the system.

Authentication
Access Control System

\[ \text{u?} : \text{USER} \]
\[ \text{crdls?} : \mathbb{P} \text{ CREDENTIAL} \]
if \( \text{u?} \in \text{dom registered_users} \land \text{crdls?} = \text{registered_users u?} \)
then \( \text{authenticated_users}' = \text{authenticated_users} \cup \{(\text{u?} \leftrightarrow \text{crdls?})\} \)
elts \( \text{authenticated_users}' = \text{authenticated_users} \)

This schema represents the authorization security property. This schema determines the access to system resources for a particular user. A user enters user id, credentials, actions he or she wants to perform, resource on which actions are to be performed to the system. This schema checks whether the user is registered or not. If the user is a registered user then it checks whether the required actions are allowed on the requested resource and required actions are allowed to that user. If yes the user is allowed to access that resource.

Authorization

Access Control System
\[ \text{u?} : \text{USER} \]
\[ \text{crdls?} : \mathbb{P} \text{ CREDENTIAL} \]
\[ \text{acs?} : \mathbb{P} \text{ ACTION} \]
\[ \text{r?} : \text{RESOURCE} \]
\[ \text{u?} \in \text{dom authenticated_users} \]
if \( \text{acs?} \subseteq \text{Resource_action r?} \land \text{acs?} \subseteq \text{User_actions u?} \)
then \( \text{authorized_users}' = \text{authorized_users} \cup \{(\text{u?} \leftrightarrow \text{crdls?})\} \)
elts \( \text{authorized_users}' = \text{authorized_users} \)

This schema describes the formal model of confidentiality. Confidentiality refers to read access to system resources. A user, who wants to read confidential data, enters user id and system checks whether the user is an authenticated user and an authorized user. If the user is authenticated and authorized then system checks the operations defined in the system for that user. If the operation, defined for that user is read operation then read access is granted to that user.

Confidentiality

Access Control System
\[ \text{u?} : \text{USER} \]
\[ \text{op?} : \text{OPERATIONS} \]
\[ \text{u?} \in \text{dom authenticated_users} \]
\[ \text{u?} \in \text{dom authorized_users} \]
if \( \text{op} = \text{READ} \lor \text{op} = \text{APPENDREAD} \)
then \( \text{user_actions}' = \text{user_action} \cup \{(\text{u?} \leftrightarrow \text{acs?})\} \)
elts \( \text{user_action}' = \text{user_actions} \)

This schema describes non-repudiation security property. Non-repudiation ensures the logging of any event or activity performed by legitimate users. For actions to be recorded, a user must be an authenticated user and an authorized user.

Non Repudiation

Access Control System
\[ \text{u?} : \text{USER} \]
\[ \text{csn?} : \mathbb{P} \text{ CSN} \]
\[ \text{u?} \in \text{dom authenticated_users} \]
\[ \text{u?} \in \text{dom authorized_users} \]
if \( \text{csn?} \neq \emptyset \)
then \( \text{log}' = \text{log} \cup \{(\text{u?} \leftrightarrow \text{csn?})\} \)
elts \( \text{log}' = \text{log} \)

This schema defines the availability of resources security property. There are predefined and recognized set of resources in the system. Any authenticated and authorized user is allocated a set of resources to perform his or her actions. The free set represents resources which are not currently occupied. To ensure availability, the free set must never be empty.

Availability

Access Control System
\[ \text{Free} : \mathbb{P} \text{ RESOURCE} \]
\[ \text{u?} : \text{USER} \]
\[ \text{u?} \in \text{dom authenticated_users} \]
\[ \text{u?} \in \text{dom authorized_users} \]
if \( \text{owner u?} \neq \emptyset \)
then \( \text{free} = \text{resources} / \text{owner u?} \)
elts \( \text{free} \land \text{owner u?} = \emptyset \)

FORMAL ANALYSIS

In this study, formal analysis of models of security properties is done by using Z/EVES theorem prover. This analysis was done for syntax checking, type checking and automatic formal proofs checking of Z
Fig. 1: Formal analysis of security properties

paragraphs. Other types of analysis such as invariants preservation, domain checking, validation of certain properties and formal proofs are not in the scope of this study. For brevity, only one snapshot taken from Z/EVES theorem prover has been shown in Fig. 1. In this snapshot, we have three columns: syntax, proof and main window containing actual specification of properties. Here, “Y” in syntax column shows that the specification in the main window has correct syntax and correct type. The “Y” in proof column shows that the specified paragraphs have no pending proof and all the attached invariants, pre-conditions and post-conditions have been satisfied.

CONCLUSION AND FUTURE WORK

In this study, software security is discussed as external protections and internal mechanisms. Software security is mainly concerned with the internal security of the software system. For internal security, security properties such as authentication, authorization, confidentiality, integrity, non-repudiation and availability of resources has been defined and discussed. A review of relevant literature has been presented. Formal methods have been described very briefly as mechanisms to formally specify security properties. An informal description of security properties is also presented in this study. Then the formal model is developed along with very brief commentary on the formal schemas. At the end, brief analysis of formal models has been presented. In the future, these formal models will be refined to add more details in the models. Analysis of properties such as domain checking, invariants preservation and precondition calculation is also a future task. In the future, an investigation will be performed to generate formal proofs of these security properties to validate the models.

REFERENCES


