Research Article

Integrating Usage Control with SIP-Based Communications

E. Barka¹ and A. Lakas²

¹ Information Security Track, College of Information Technology, United Arab Emirates (UAE) University,
P.O. Box 15551, Al Ain, United Arab Emirates

² Networking Track, College of Information Technology, United Arab Emirates (UAE) University,
P.O. Box 15551, Al Ain, United Arab Emirates

Correspondence should be addressed to E. Barka, ebarka@uaeu.ac.ae

Received 25 September 2008; Accepted 8 October 2008

Recommended by Abderrahim Benslimane

The Session Initiation Protocol (SIP) is a signaling protocol used for establishing and maintaining communication sessions involving two or more participants. SIP was initially designed for voice over IP and multimedia conferencing, and then was extended to support other services such as instant messaging and presence management. Today, SIP is also adopted to be used with 3G wireless networks, thus it becomes an integral protocol for ubiquitous environment. SIP has various methods that support a variety of applications such as subscribing to a service, notification of an event, status update, and location and presence services. However, when it comes to security, the use of wireless and mobile communication technologies and the pervasive nature of this environment introduce higher risks to security than that of the old simple environment. In this paper, we introduce new architecture that implements a new type of access control called usage access control (UCON) to control the access to the SIP-based communication at preconnection, during connection, and postconnection. This will enable prescribers of SIP services to control who can identify their locations to approve or disapprove their subsequent connections, and to also set some parameters to determine whether a certain communication can continue or should terminate.

Copyright © 2008 E. Barka and A. Lakas. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Nowadays, our society is impacted by a revolutionary innovations in information technology that made communication around the globe seems like it is only a mile away. With these advances in technology, particularly in communications, we are also encountering a series of new problems on security and privacy issues.

Among these technologies are the protocols known as signaling protocols, which are used to carry call setup information. These protocols are used to set up, control, and maintain sessions for multimedia applications. Applications which require sessions to be set up include telephony, videoconferencing, and remote learning. SIP signaling protocols are used to set up sessions over packet-switched networks for IP-based multimedia applications such as voice-over IP (VOIP) and IP-based videoconferencing.

SIP has several methods that support a variety of applications such as subscribing to a service, notification of an event, status update, and location and presence services.

SIP provides flexible and real-time communications in a ubiquitous way which adds additional risks that result from the adoption of ubiquity, mobility, and heterogeneity.

Traditional access controls typically focus on the protection of data in closed environments, and the enforcement of control has been primarily based on identity and attributes of a known user. These types of access control lack a comprehensive, systematic approach to fulfill the security requirements of today's pervasive and ubiquitous nature.

To address these issues, we introduce in this paper a new architecture that implements a new type of access control called usage access control (UCON) to control the access to the SIP-based communication at preconnection, during connection, and postconnection. This will enable prescribers of SIP services to control who can identify their locations to approve or disapprove their subsequent connections, and to also set some parameters to determine whether a certain communication can continue or should terminate.

The remainder of this paper is organized as follows. In Section 2, we present the main motivation in using the model

of UCON for SIP-based communications. Section 3 provides some background on SIP and discusses its role in the pervasive environment from a communication perspective. We do this by first providing an overview of SIP applications and related security aspects. In Section 4, we introduce the next generation access control, depicted in the usage control model (UCON). Section 5 presents our architecture and explains the integration of SIP with UCON. Finally, Section 6 discusses future work and concludes our paper.

2. MOTIVATION

In this section, we present few factors that support the usage control over the traditional access control through the use of UCON model in the case of SIP applications. Traditional access control models that are currently used for SIP communications suffer from the lack of flexibility to specify security policies in the case of ubiquitous environments.

For example, in the case of pay-per-usage, there is a need to terminate the SIP communications when the caller is out of credit. Traditional access control provides mechanisms only to control access during the establishment of an SIP call session, but completely loses control over the session during the call (when user credentials may change), and after the session has terminated. In this case, no control is provided after the access is granted even when the caller runs out of credit. This problem has been partially solved by introducing the back-to-back UA (B2BUA) mechanism to SIP architecture by forcing all communications to go through the SIP server including the media connection. However, this approach created other issues including the creation of a bottleneck situation, and generated additional processing overhead at the server. UCON can solve this issue, by specifying policies that monitor the SIP communications before and during the call. Moreover, it mandates and enforces continuous compliance to access conditions. In the case of noncompliance to these conditions, UCON provides mechanisms for revoking the access. Monitoring and enforcement mechanisms are now collocated and distributed within the SIP UAs.

In overall, UCON benefits consist specifically in its ability to specify and enforce usage policies not only before access, but also during the access and after.

3. SIP OVERVIEW

SIP, which is one of the most popular IP signaling protocols, is an application layer protocol created by the Internet Engineering Task Force (IETF) [1] to allow entities to locate one another on a network and invite themselves to participate in a session. It is also responsible for maintaining and terminating a session. SIP is limited only to the session establishments and terminations. Once the participating parties negotiate the characteristics of their communication, the session is established. The participating parties use the Session Description Protocols (SDPs) to define the audio or the video bearer channel they would like to communicate through. SDP is embedded within the SIP messages.

FIGURE 1: SIP communication.

The architecture of SIP is based on a set of components which content varies based on the application deployed. The basic set of SIP components includes the following.

- (i) User agent (UA): works on behalf of users to set up calls and establishes a multimedia communication.
- (ii) Proxy servers (PSs): keeps track of location of UAs and facilitates the establishment of sessions between them.
- (iii) Registrar: UAs keep the registrar updated on their current locations by initially registering to it with their current location. UAs also update the registrar with their preferred reachability information.

Figure 1 illustrates the basic components of SIP.

3.1. SIP applications

SIP provides support for various text, voice, and video session-based applications. The most popular application is VoIP. SIP comes natively with its own telephony applications; however, it may be used for the control and the management of IP telephony of various systems. For instance, SIP has been adopted by the 3GPP initiative to control and manage the communication between 3GPP components for the 3G and 4G telephony systems [2].

Another application for SIP is the instant messaging and presence management [3]. Instant messaging (IM) is a text-based communication that has seen an increasing popularity on the Internet. IM is accompanied with the presence of functionality where any IM communication is assorted with information which includes status and location information such as that contained in “buddy” list. This extension includes functionalities such as

- (i) publishing and uploading of presence information;
- (ii) presence and event notification;
- (iii) delivering of instant messages.

3.2. SIP security

SIP provides flexible and real-time communications in a ubiquitous way. However, with doing so, it also adds extra

risks due to the new factors of ubiquity, mobility, and heterogeneity. Some of the threats that are inherent to the use of SIP are listed as follows.

(i) Registration hijacking: a registrar assesses the identity of a UA. The *From* header of an SIP request can be arbitrarily modified and hence open to malicious registration.

(ii) Impersonating a server: a UA contacts a proxy server to deliver requests. The server could be impersonated by an attacker. Mobility in SIP further complicates this scenario.

(iii) Tampering with message bodies.

(iv) Tearing down sessions—insert a BYE.

(v) Denial-of-service attacks—denial-of-service attacks focus on rendering a particular network element unavailable, usually by directing an excessive amount of network traffic at its interfaces. In much architecture SIP proxy servers face the public Internet in order to accept requests from worldwide IP endpoints. SIP creates a number of potential opportunities for distributed denial-of-service attacks that must be recognized and addressed by the implementers and operators of SIP systems.

Therefore, the security challenges facing SIP are to ensure the following.

(i) *Authentication*—SIP currently has the HTTP style digest mechanism, but it is not enough. A single sign-on authentication mechanism is needed.

(ii) *Authorization* using policy-based mechanisms—the read/write/execute controls that are embedded in file systems. Some people have recommended, and tried to implement, traditional access control models, but they are broadly categorized as discretionary access control (DAC) [4, 5] and mandatory access control (MAC) models [4–6]. Others have proposed new models such as role-based access control (RBAC) and task-based access control (TBAC) to address the security requirements [7].

The above mentioned solutions are not sufficient enough for providing security for pervasive environments that requires continuous control of the access “not just at the establishment of the connection” to the system resources by very heterogeneous types of users and devices; but also during and after a session is terminated.

Our approach introduces a new type of access control named usage access control (UCON) [8–10]. UCON is used here to control the access to the SIP-based communications: presession, during the session, and even at postsession. There are many benefits to UCON usage in SIP. This will enable subscribers to SIP services to control who can identify their locations, to approve or disapprove their subsequent connections, and to also set some parameters to determine whether a certain communication can continue or should terminate. UCON also contributes to solving some of the problems that are specific to some applications such as the necessity to have back-to-back UAs (B2BUA) [6]. The payer-usage application is an illustration of this problem. In this case, UCON may terminate the call when the caller runs out of credit.

The following section provides an overview of the usage control model (UCON).

FIGURE 2: UCON components.

4. USAGE CONTROL MODEL (UCON) OVERVIEW

In this section, we briefly review the general ideas of UCON and the core authorization models. The details of these models can be found in [8, 9].

The UCON model consists of six components, three of these components “subjects, objects, and rights” are considered core components and the other three “authorization rules, conditions, and obligations” are additional, and they are mainly involved in the authorization process (see Figure 2).

In UCON system, at least the authorization rules (specifically rights-related authorization rules) have to be included for authorization. Conditions and obligations can also be used in the authorization process. The following subsections describe UCON’s core components.

4.1. Subjects

Subjects are “Active” entities associated with attributes, and hold and exercise certain rights on objects. Attributes are properties of the subjects that can be used for the authorization process. Examples of attributes include identities, roles, credits, memberships, security levels, and so forth. A subject can be a user, a group, a role, or a process. A user is an individual entity that has certain rights on an object. A group is a set of users who holds same rights as a group. A role is a named collection of users and relevant permissions [6]. Groups and roles may have hierarchical relationships.

4.2. Objects

Objects are “Passive” entities that subjects hold rights on, whereby the subjects can access or use objects. Objects are also associated with attributes, either by themselves or together with rights. As for subjects, the attributes include certain properties that can be used for the authorization process. Examples of object attributes are security levels, ownerships, classes, and so forth. Object classes are used to categorize objects so authorization can be done based not only on individual objects but also on sets of objects that belong to same class [8, 9]. In some cases, objects or objects with attributes (i.e., classes) are associated with attributes together with rights. Examples of the attributes for objects

with rights are credits, roles, memberships, and so forth. The credits may be used to define how many credits are required to obtain a certain right on a specific object. For example, an e-book together with a read right may require \$10 or the book with an additional print right may require \$15.

4.3. Rights

Rights are privileges that a subject can hold on an object. Rights consist of a set of access/usage functions that enables a subject's access to objects. The authorizations of rights require associations with subjects and objects. Rights may or may not have a hierarchy.

UCON rights can be divided into many functional categories. The two most fundamental right categories might be a view and a modification. They are denoted as V and M , respectively, such that

$$R = \{V, M\}. \quad (1)$$

Modification includes change to an existing digital object and creation of a new object that reuses an original digital object. The range of V and M is denoted as

$$C = \{0, 1, \alpha\}, \quad (2)$$

where

- (i) "0" means closed to everybody (no one can access);
- (ii) "1" means open to everybody (everyone can access); and
- (iii) " α " means access approval is selective or controlled. The openness of the control or availability of object to public is expressed as $0 < \alpha < 1$ which means that 1 is most open to public and 0 is least open.

The following subsections describe the additional component of UCON.

4.4. Authorization rules

Authorization rules are a set of requirements that should be satisfied before allowing subjects' access to objects or use of objects.

There are mainly two kinds of authorization rules. The rights-related authorization rules (RARs), which are used to check if a subject has valid privilege to exercise certain rights on a digital object.

Examples of such rules may include identities or roles verification, properties checking, proof of payments, and so forth. Moreover, the obligation-related authorization rules (OARs), which are mainly used to check if a subject has agreed on the fulfillment of an obligation which has to be done after obtaining or exercising rights on a digital object. Examples of such rules may include metered payment agreement, usage log report agreement, and so forth.

The authorization rules are different from conditions in that the authorization rules are a set of decision factors used to check whether a subject is qualified for the use of certain

rights on an object, whereas the conditions are decision factors used to check whether existing limitations and status of usage rights on an object are valid and whether those limitations have to be updated.

4.5. Conditions

Conditions are a set of decision factors that the system should verify at authorization process along with authorization rules before allowing usage of rights on a digital object.

Conditions are of two types: dynamic conditions are conditions which include information that may have to be checked for updates at each time of usage. Examples of dynamic conditions are the number of usage times (e.g., can read 5 times, can print 2 times), and usage log (e.g., already read portion cannot be accessed again). On the other hand, static conditions include information that does not have to be checked for updates. Some examples of static conditions are accessible time period (e.g., business hours), accessible location (e.g., workplace), and allowed printer name.

4.6. Obligations

Obligations are mandatory requirements that a subject has to perform after obtaining or exercising rights on an object. In real-world implementation, however, this may have to be done by agreeing on the fulfillment of obligations before obtaining the rights and at the time obligation-related authorization rules are checked. For example, a consumer subject may have to accept metered payment agreements before obtaining the rights for the usage of certain digital information or should agree on providing usage log information to a provider subject before reading an e-book or listening a music file. Traditional access control has hardly recognized the obligation concept.

The most important properties that distinguish UCON from traditional access control models and trust management are the continuity of usage decisions and the mutability of attributes. Continuity means that a control policy may be enforced not only before an access, but also during the period of the access.

The control decision components are checked and enforced in the first two phases, named, predecisions and ongoing decisions, respectively. In the after-usage phase, we do not enforce any policy since there is no access control after a subject finishes a usage on an object.

Mutability refers to updates of the subject or the object's attribute that may occur as a result of the access. Along with the three phases, there are three kinds of updates: preupdates, ongoing updates, and postupdates. All these updates are performed and monitored by the system.

An update of subject or object attributes may result in a system action to permit or revoke an access. An update can affect not only the concurrent usage, but also other usages related to the same subject or object. An update on the current usage may generate.

5. SIP/UCON INTEGRATION

Integrating the UCON technology into ubiquitous SIP-based environment requires a careful mapping between the entities of UCON and those entities and components of the SIP. Following is a list of integrated components which require such mapping:

User/subjects

The concept of participants in SIP is represented as a user component in the UCON.

Permissions/rights

The concept of permissions in UCON will reflect all the privileges that an SIP participant needs to complete a task.

Objects

The objects in UCON are used to represent all entities in SIP that an SIP UA's seek to connect to.

Authorization rules

Authorization rules in UCON are the set of requirements that should be satisfied before any SIP UA and be permitted to establish any connection with any other SIIP entity.

Obligations

The concept of obligation in UCON can be represented in SIP as the set of actions that an SIP user is required to perform before and after the connection has been established.

Conditions

Conditions in UCON are represented in SIP by the set of environmental and system decision factors that must be continuously evaluated to make sure that their changes do not lead to changes in the connection status.

5.1. Architecture of the UCON/SIP

One of the most critical issues in using UCON for enforcing access into SIP environment is to use the concept of a reference monitor. The reference monitor (RM) has been introduced, and extensively discussed, by the access control community for years. The concept of a reference monitor was introduced and published by the ISO in a standard for access control framework [11]. The RM concept has been considered as the core control mechanism for access and usage of digital information. In classical access control, subjects can access digital objects only through the reference monitor, which is a process inside the trusted computer base that is always running and is a tamper proof.

FIGURE 3: Conceptual structure for UCON/SIP reference monitor.

The following section discusses our conceptual structure of UCON/SIP access control domains, based on the reference monitor.

5.2. UCON/SIP areas of control

The area of control in our architecture refers to the area of coverage where the rights to access the SIP objects is under the control of the reference monitor.

According to the standard [11], the reference monitor consists of two facilities: access control enforcement facility (AEF) and access decision facility (ADF). AEF and ADF interact such as every request to access an object is intercepted by AEF. The later asks the ADF for a decision on the request approval. ADF returns either “yes” or “no” as appropriate. The enforcement of this decision takes place at the AEF.

Our UCON/SIP reference monitor is similar but differs in the details from that of ISO. Figure 3 shows the conceptual structure of the UCON/SIP reference monitor.

As the figure shows, both the AEF and ADF include several functional modules. AEF contains

- (iv) *monitoring* module which is used to keep track of the changes of the attributes of the subjects and objects;
- (v) *updating* module which is used primarily to update the attributes of the subjects and objects.

The ADF is where all of the access granting decisions, the access maintenance, and the access termination take place. This facility includes three core modules that are utilized collectively in rendering access decisions as response to the (AEF) requests:

- (i) *authorization*;
- (ii) *condition*;
- (iii) *obligation*.

The authorization module uses subjects and objects attributes, and the access rules to check if a request is

FIGURE 4: Client-side reference monitor.

FIGURE 5: Server-side reference monitor.

allowed or not. The condition module uses the access rules and the contextual information to decide whether the conditional requirements, both system and environmental, for the authorized request are satisfied or not. Finally, the obligation module is responsible for handling decisions that are tied to actions that are required to be performed by the requestor before, during or after the access is granted. All existing obligations are monitored by the monitoring module and the outcome must be resolved by the update module in the AEF.

5.3. Areas of control architecture

To control the access to the SIP environment, our architecture considers two areas of controls, based on the location of the reference monitor. The server-side control domain (SCD) is the area where the reference monitor is located at the server environment and directly enforces the access policy to the system resources. The client-side control domain (CCD), on the other hand, is the area of control where the reference monitor is located at the client environment and enforces the access policy on behalf of the server. Figures 4 and 5 depict this architecture, respectively.

Figure 4 shows that the reference monitor does not reside within the area of the server, but rather at the client side. This setup provides for better usage control over system objects. In this case, because of the existence of the SRM, the system objects can be stored centrally or locally, in either case, the objects are under the control of the client instead of the server.

Figure 5 shows that, in the case of the server-side control, the control of subject's access to objects is done centrally. Moreover, in this setup, the subject can either be located

within the network or outside, and the SIP objects may or may not be stored in the client's storage, depending upon the criticality and sensitivity of the content of the object. If it is not that sensitive, then it can be allowed to reside outside of the server-side storage. However, if the content is very critical or very sensitive, the object must stay within the server-side storage.

Figure 6 illustrates the scenario of an SIP user agent client (UAC) initiating a call to a user agent server (UAS). This call is established through the SIP proxy. For this call to take place, the SIP proxy will verify the credentials of the caller and the callee. This scenario is executed through the following steps.

- (1) UAC initiates the call by sending an SIP INVITE to the SIP proxy.
- (2) The SIP proxy verifies the credentials of UAC and UAS by soliciting locally a decision from the ADF.
- (3) ADF looks up the access rules at the policy server, rules that are relevant to UAC and UAS, makes a decision on whether UAC is authorized to make call to UAS, and sends locally a notification to the AEF for the enforcement of such decision.
- (4) In the case of acceptance, the AEF notifies the proxy with the decision which will allow the proxy to proceed with the processing of the call establishment.
- (5) The proxy processes the call initiation and forwards the INVITE to UAS.
- (6) If UAS accepts the call, it will send an ACK to the proxy.
- (7) The proxy relays the UAS ACK to the UAC and includes within the message the details of UAS (contact address, call attributes, etc.).
- (8) At this point, the voice channel is established and UAC and UAS can now start communicating.

5.4. Application scenarios

Scenario 1. UCON allows SIP to continuously monitor and control a call session even after it is established. In the case of pay-per-usage, UCON may terminate the call when the caller runs out of credit. This is done through the UDF and UEF. After the caller's credit has been validated, the UDF at the SIP server sends the details of the credit to the UEF at the client side, embedded within the parameters of ACK message. The UEF uses the update module to keep track of the credit consumed. When the credit reached its limit, the UEF terminates the call.

Scenario 2. System abuse: UCON allows SIP clients to control who can call them and when. This is done through the use of authorization before process (described in Figure 7). This process relies on the UDF and UEF to decide on whether a UAC is permitted to contact a UAS at any given time. To accomplish this task, access rules can be dynamically updated on the policy server to allow the ADF to look up these

FIGURE 6: Authorization (before).

FIGURE 7: Authorization (during).

rules any time a call request is made and makes a decision on whether UAC is authorized to make call to UAS, and sends locally a notification to the AEF for the enforcement of such decision. In this case, the decision to allow a call to be made depends upon the positive acknowledgment that the proxy relays from the UAS to the UAC and includes within the message the details of UAS (contact address, call attributes, etc.).

Scenario 3. UCON allows SIP to set some obligations that users agree to perform prior to establishing a call. For

example, a user may agree to log certain type of information pertaining to the call that he/she is requesting. Through the obligation module, within the ADF, SIP can terminate the call if that obligation was not fulfilled.

6. CONCLUSION

In this paper, we introduced a new concept of access control for enhancing the security of the SIP-based communications. This concept is based on controlling usage of the various SIP components. We have integrated this new concept with

the SIP protocol to produce new architecture that helps in controlling the access to the SIP-based environment before, during, and after connections. Our architecture goes beyond what traditional access control models can provide. UCON benefits to SIP by providing comprehensive, systematic approach to fulfill the security requirements of today's pervasive and ubiquitous nature. In this work, we have presented an extended architecture which integrates UCON with SIP components in a manner that enables SIP to support applications in ubiquitous environment more securely. We have presented few scenarios where we showed the necessary mapping between UCON components and that of SIP. It is worth noting that this work provides mainly an architecture for applying UCON features to SIP-based communications in the context of usage and access control. However, this work does not pretend to provide a set of security mechanisms such as those for authorization and authentication algorithms.

This work opens several directions for further investigation. First of all, this architecture requires validation through a real-life implementation and deployment. Secondly, for this model to be complete a performance analysis of the integrated model is needed. We intend to extend this work in the future by providing an implementation of this architecture, and assess its performance with regards to its benefits.

REFERENCES

- [1] J. Rosenberg, H. Schulzrinne, G. Camarillo, et al., "SIP: Session Initiation Protocol," Internet Engineering Task Force, RFC3261, June 2002.
- [2] 3GPP, "TS 24.228: signaling flows for the IP multimedia call control based on SIP and SDP," September 2002.
- [3] IEEE std. 802.11, Wireless Medium Access Control and Physical Layer Specifications, 1999.
- [4] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models," *Computer*, vol. 29, no. 2, pp. 38–47, 1996.
- [5] W. Zhou and C. Meinel, "Implement Role-Based Access Control with Attribute Certificates," Universität Trier, Mathematik-Informatik, Forschungsbericht 03-03, 2003.
- [6] W. Stallings, *Network Security Essentials, Applications and Standards*, Prentice Hall, Upper Saddle River, NJ, USA, 2003.
- [7] R. Thomas and R. S. Sandhu, *Task-Based Authorization Control (TBAC): Models for Active and Enterprise-Oriented Authorization Management Database Security XI: Status and Prospects*, North-Holland, Amsterdam, The Netherlands, 1997.
- [8] J. Park and R. S. Sandhu, "The UCON_{ABC} usage control model," *ACM Transaction on Information and System Security*, vol. 7, no. 1, pp. 128–174, 2004.
- [9] J. Park and R. S. Sandhu, "Towards usage control models: beyond traditional access control," in *Proceedings of the 7th ACM Symposium on Access Control Models and Technologies (SACMAT '02)*, pp. 57–64, Monterey, Calif, USA, June 2002.
- [10] X. Zhang, F. Parisi-Presicce, J. Park, and R. S. Sandhu, "Logical model and specification of usage control," to appear in *ACM Transaction on Information and System Security*.
- [11] ISO/IEC 10181-3, Information Technology—Open System Interconnection—Security framework for open systems: Access control, <http://www.iso.org/>.