A Lightweight Secure SIP Model for End-to-End Communication

Weirong Jiang
Research Institute of Information Technology, Tsinghua University, Beijing, 100084, P.R.China
jwr2000@mails.tsinghua.edu.cn

Abstract

Session Initiation Protocol (SIP) is a signaling standard approved by IETF for real-time multimedia session establishment. Increasingly wide deployment brings much concern on SIP security. Current solutions for end-to-end signaling security either require user-side powerful performance support for heterogeneous security mechanisms, or assume that trust relationship is transitive and static. Yet no solution is suitable for weak terminals with inherent computational power limitations. It is necessary to consider a reasonable combination of security solutions could be provided by the end users and the network servers. This paper presents a hybrid security model, combining hop-by-hop and end-to-end security, that trusted neighbor servers help weak terminals to make end-to-end communication secure with lightweight overload. We believe it is also an engineering trade-off between capacity and security.

Keywords

SIP, security model, weak terminal, end-to-end, authentication, dynamic trust

1. Introduction

Session Initiation Protocol (SIP) [1] is an application-layer signaling and control protocol for creating, modifying, and terminating sessions including Internet telephone calls, multimedia distribution, and multimedia conferences. Flexible, extensible and open, SIP becomes the most promising candidate as the signaling protocol for IP telephony and it has been chosen by the Third-Generation Partnership Project (3GPP) as the protocol for multimedia application in 3G mobile networks. As the SIP-based service is getting popular, it is facing severe security threats. Significant research and development effort is devoted to the security enhancement to SIP [4].

SIP supports hop-by-hop security using Transport Layer Security (TLS) [6] and end-to-end security using Secure MIME (S/MIME) [7]. Hop-by-hop security assumes that a SIP UA (user agent) trusts all proxy servers along its request path to inspect the message bodies contained in the message while end-to-end security assumes that a SIP UA does not trust any proxy servers to check the message [2]. Hop-by-hop security cannot prevent attacks

from malicious intermediaries while end-to-end security provides higher degree of security and better level of performance.

Some of the existing security solutions [3] for end-to-end communication are designed based on hop-by-hop security with insufficient security. Others do not consider the inherent limitations of weak terminals such as mobile phones which appears a promising application of SIP. Motivated by avoiding imposing heavy load on end users, we've designed an appropriate lightweight end-to-end secure model for weak terminals.

The remainder of this paper is organized as follows. Section 2 introduces the security for end-to-end communication in multi-hop SIP network. Section 3 describes the limitations of weak terminals and security requirements. In section 4, we propose a hybrid security model for SIP. In section 5 we present an application case and evaluate our security model based on comparative work. Section 5 concludes the paper.

2. SIP Security for End-to-End Communication

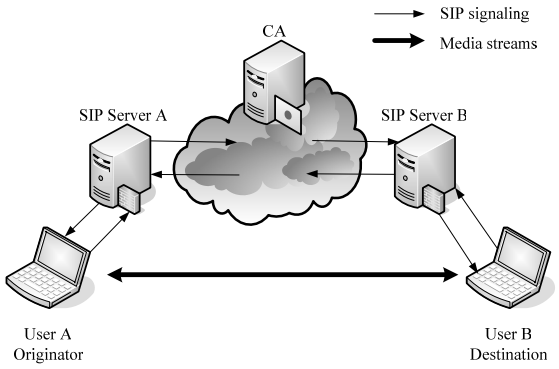

Figure 1. Multi-hop network for end-to-end SIP deployment

Figure 1 defines the secure multi-hop SIP network consisting of the following servers.

a) SIP servers [5]

Registrar servers provide Location services for mobile users. SIP allows abstract naming and dynamic registration and provides client-server security, namely mutual authentication, data integrity, and data confidentiality.

Proxy servers accept session requests made by a SIP UA, query the registrar server to acquire the recipient UA's addressing information and forward the session invitation directly to the recipient UA if it resides in the same domain or to a proxy server if in another domain.

Redirect servers realize the similar function as proxy servers but the only difference is that redirect servers generate redirection responses to requests, directing the client to contact an alternate set of URIs, rather than directly forward the session invitation to the recipients.

The three servers usually reside in a single SIP server, as SIP Server A or SIP Server B in Figure 1.

b) Certification authority (CA)

Certificate Authority is the core component of public key infrastructure (PKI) for distributing certificates. Admittedly, for both server to server or user to server, one can use all different kinds of key mechanisms, such as manual key, pre-shared key, signature, or

certificate. And certificate can be used as self signed or PKI. In our model, the certificate-based key management is preferred for its wide use and high scalability. The SIP servers have their own public key certificates obtained from CA, so servers can authenticate each other using PKI. For user authentication, the SIP server can either uses a pre-shared key with user or the user has a public key certificate from CA.

Figure 2 [3] shows four security steps for end-to-end communication within SIP signaling. SIP uses the existing security mechanisms, such as HTTP digest authentication, TLS, IPsec/IKE, S/MIME. There are no specific vulnerabilities in client-server security scheme for registration and in hop-by-hop security scheme (user-to-server and server-to-server security) for setup. The rest steps are end-to-end security measures for setup, which has some problems due to the specific limitations of weak terminals.

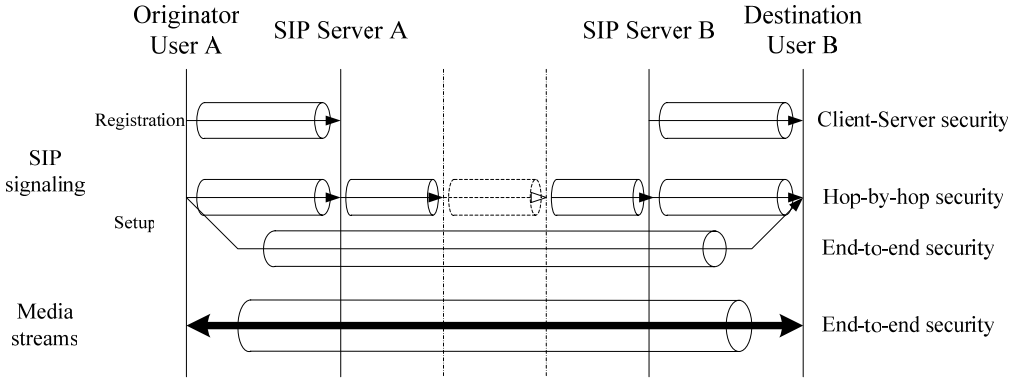

Figure 2. Steps of the end-to-end communication in SIP.

3. Limitations of Weak Terminals

Weak terminals such as mobile SIP phones have several fundamental limitations. They have very limited CPU power, memory size, and display area. It also requires less power consumption due to battery capacity, and their input device is harder to operate than the typical desktop computers [8]. Their supported security mechanisms may be scarce and heterogeneous. The networks they reside in may also have limitations including narrower bandwidth, more latency, poorer connection stability and less predictable availability. Moreover, some of them (e.g. wireless SIP nodes in ad hoc networks), roaming in a hostile environment with relatively poor physical protection, depending on relatively weak security schemes owing to their limited capacity, have non-negligible probability of being compromised [10]. And the secure channels also tend toward being revealed. Hence the trust relationships between nodes are prone to change dynamically. These limitations cause many difficulties for end-to-end security implementation, as discussed below.

Most security mechanisms process key management in SIP signaling takes four steps [3]:

1) Negotiate the cipher suite,
2) Perform mutual authentication using a long term key,
3) Set up a secure session to share a session key,
4) Exchange the session key in the secure session.

Correspondingly, limitations of weak terminals results in problems as follows.

A. Negotiation

End-to-end security requires the two parties negotiate and share the same authentication mechanism and cipher suite. But it is not easy for two weak terminals to accomplish an agreement once one of them does not support some security mechanisms.

B. Authentication

Mutual authentication between the two end points (users) needs long-term keys, such as pre-shared keys or certificates. But keeping every key or certificate of any unspecified individual is often costly for most weak terminals.

C. Setup delay

Key exchange for SIP signaling affects the post-dial delay and the worst case is multi-party session, such as conference call, with multiple media streams. Weak terminals with poor computational performance will just amplify these delays.

D. Static trust

The trust setup between two users will last until the next registration or until one of them revokes itself willingly. This static trust may bring some threats in vulnerable environment and the trustworthiness of peers must be taken into account.

In a word, the specific requirements for weak terminals in end-to-end security mainly include: lightweight security overhead and dynamic trust.

4. Hybrid Security Model

The new model proposed here is based on the assumption that a user trusts the first next-hop server and trusts an opposite-side user via transitive trust. We name the first next-hop server the neighbor servers which here act as security agents to share the security overhead for weak terminals. We also assume that a hierarchical CA system exist for intermediate servers. In most cases, servers are much stable than UAs so that it is easier for intermediaries to build a hierarchical CA system.

The proposed model meets the requirements identified in above section by leveraging the trusted neighbor server to share heavy load in fulfilling security schemes. Its security pattern is a combination of hop-by-hop (user-to-server) and end-to-end (server-to-server) security, shown in Figure 3.

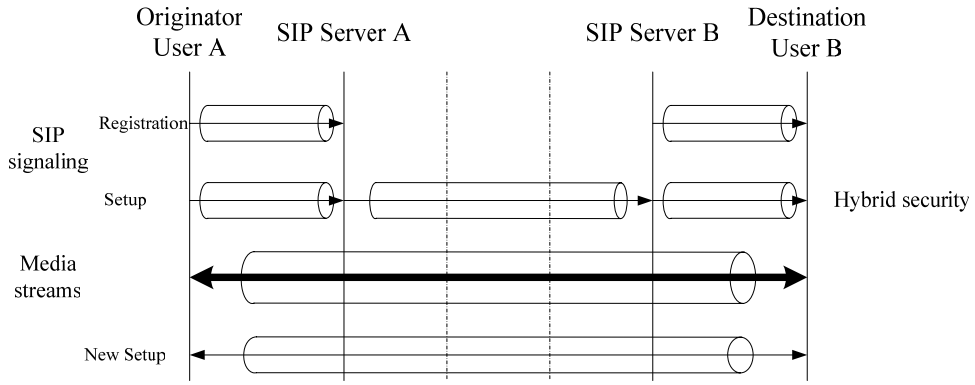

Figure 3. Steps of the hybrid model.

1) Security for signaling.

A user and a SIP server authenticate each other when the user registers his own location address. The user requests the server for the help on secure signaling due to its limited capacity. Also powerful users can choose not trusting the neighboring servers and initialize traditional end-to-end secure signaling by themselves. If requested, SIP servers neighboring to the users in the two ends authenticate each other when they begin to perform setup procedures initially. This allows weak users to omit the mutual authentication and establish a secure session for setup. The complicated cipher suite negotiation is executed by powerful neighbor servers. The originator only need to securely access the SIP server in its own domain. Figure 4 shows a prototype example of the flow.

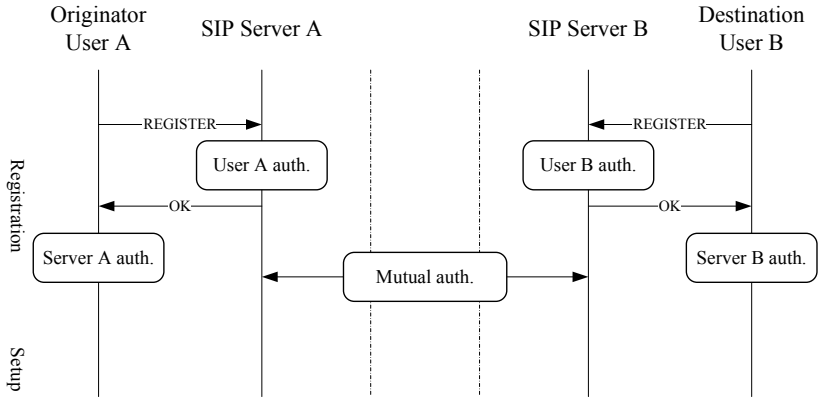

Figure 4. Example of secure signaling.

2) Security for media streams.

After the above step, the setup secure session has already been done at registration. The session key exchange can be executed during setup with extended SDP [11]. This allows the encryption of media streams at the application layer such as Secure RTP (SRTP) [12]. Figure 5 shows the prototype flow of key exchange and media encryption.

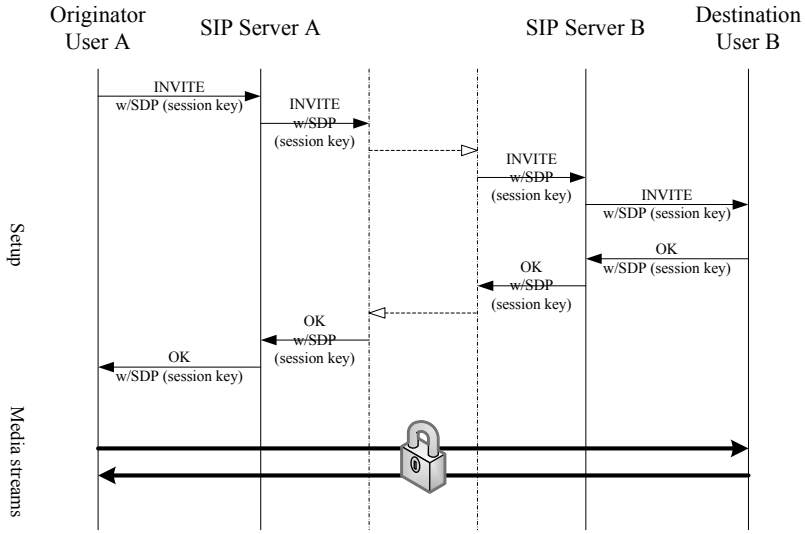

Figure 5. Example of secure media.

3) Security for dynamic trust.

Considering the dynamic trust, users should authenticate each other not only before the first time signaling but also before a new setup signaling. But still requesting the neighbor servers to carry out this work for users will cause much load and delay. In addition, the neighbor servers will easily become the objectives to be attacked if they are overburdened. To avoid imposing too much security load on neighbor servers, we use an additional pre-shared key named setup key in our model. Users use the setup keys to directly authenticate their peers in next setup signaling. The exchange of the setup key can be executed during registration at the first time and be executed during termination afterwards. Setup keys are

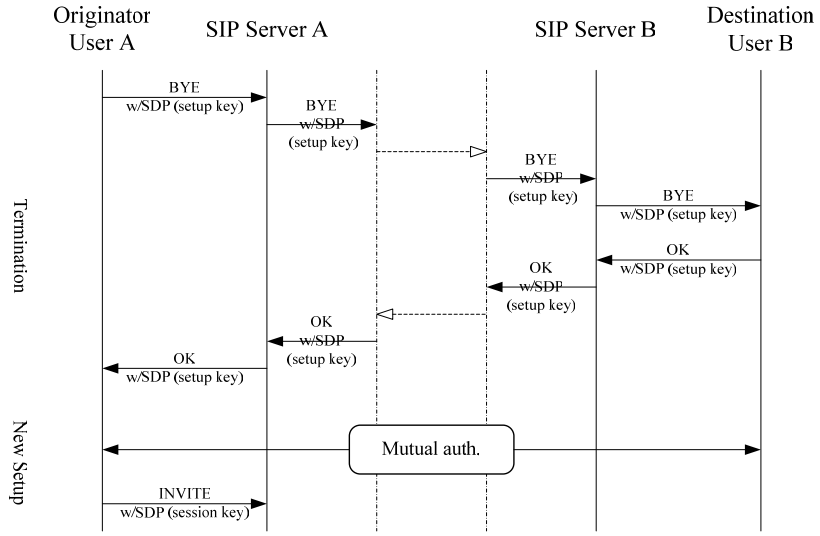

Figure 6. Example of dynamic trust.

also dynamic for their limited life time which is determined by the session participants. Figure 6 shows the prototype flow.

To sum up, we present the overall flow of our model in Figure 8.

5. Case Study and Evaluation

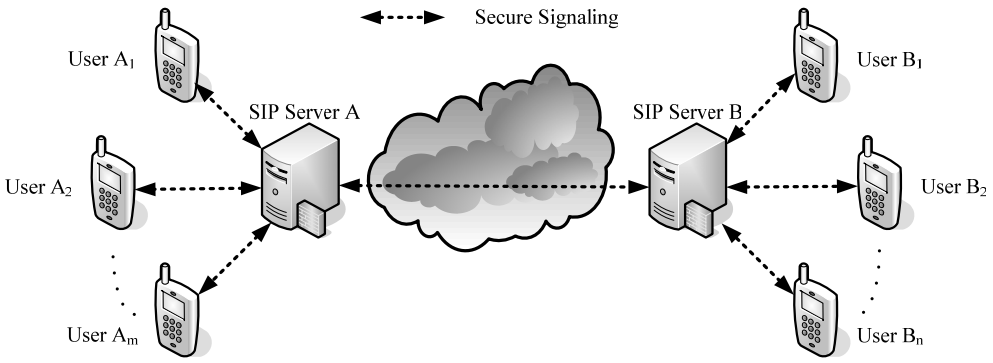

Figure 7. (*m*-to-*n*) Secure signaling

Figure 7 defines the application scenario. We use (*m* to *n*) to indicate that there are *m* originators in domain A and *n* destinations in domain B. In a direct peer-to-peer (i.e. user-to-user, user-to-server or server-to-server) signaling, caused by mutual agreement and

authentication, the additional load is L and the additional delay is D. We compare our hybrid security model with the conventional end-to-end security model in Table 1.

The hybrid model needs no direct user-to-user security mechanism agreement and no user-to-user authentication. If every originator in domain A goes to setup secure communication with every destination in domain B, they do not need to make secure signaling directly to peers one-by-one. They only need trust the neighbor servers.

Table 1. Comparison of the two security models.

Security model	**Total security load (*m* to *n*)**	**Max # of mechanisms a terminal need support**	**Setup delay (1 to *n*)**	**Dynamic trust**
End-to-end	$(m+n+mn+1)*L$	Max(m,n)	$n*D$	No
Hybrid	$(m+n+1)*L$	1	D	Yes

6. Conclusion and Future Work

Rethinking the end-to-end argument [9], it seems also an engineering trade-off between security and capacity. If the capacity is restricted by the limitations of ends, we have to consider lowering the security level to an appropriate extent. Trusting the one who is most worth trusting is a promising solution.

This paper discusses SIP security requirements for weak terminals in end-to-end communication and proposes a lightweight secure SIP model by combining the hop-by-hop security and end-to-end security. The trusted neighbor servers in our model provide crucial benefits in terms of key management and security load share. Dynamic trust is achieved by using setup keys to prevent malicious attack towards both weak terminals and neighbor servers.

Many security mechanisms, such as S/MIME, PGP, can be integrated into this model. It is a challenge to find an optimal practical solution for this model by simulations and verify the effectiveness of this model by comparative work in future. The life time of a setup key represents the level of trust relationship between two parties and how to establish an appropriate value of it, which should be discussed based on the specific application scenarios, is also an open research question.

7. Reference

[1] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, “SIP: Session Initiation Protocol,” IETF RFC 3261, June 2002.

[2] K. Ono and S. Tachimoto, “Requirements for End-to-Middle Security for the Session Initiation Protocol (SIP),” IETF draft-ietf-sipping-e2m-sec-reqs-06, March 2005.

[3] K. Ono and S. Tachimoto, “SIP signaling security for end-to-end communication,” Proc. of 9th Asia-Pacific Conference on Communications, APCC 2003, Sept. 2003.

[4] S. Salsano, L. Veltri, and D. Papalilo, “SIP Security Issues: The SIP Authentication Procedure and its Processing Load,” IEEE Network, 16(6): 38-44, Nov/Dec 2002.

[5] “Understanding SIP,” URL: http://www.sipcenter.com.

[6] T. Dierks and C. Allen, “The TLS Protocol Version 1.0,” IETF RFC 2246, Jan. 1999.

[7] B. Ramsdell, “Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification,” IETF RFC 3851, July 2004.
[8] T. V. Do, “WAP security: WTLS,”
URL: http://ece.gmu.edu/courses/ECE636/project/reports/TDo.pdf, May 2001.
[9] J. H. Saltzer, D. P. Reed, and D. D. Clark, “End-to-End Arguments in System Design,” ACM Transactions on Computer Systems, 2(4): 277-88, Nov. 1984.
[10] L. Zhou and Z. J. Hass, “Securing Ad Hoc Networks,” IEEE Network, 13(6): 24-30, Nov/Dec 1999.
[11] M. Handley and V. Jacobson, “SDP: Session Description Protocol,” IETF RFC 2327, April 1998.
[12] M. Baugher, D. McGrew, M. Naslund, E. Carrara and K. Norrman, “The Secure Real-time Transport Protocol (SRTP),” IETF RFC 3711, March 2004.

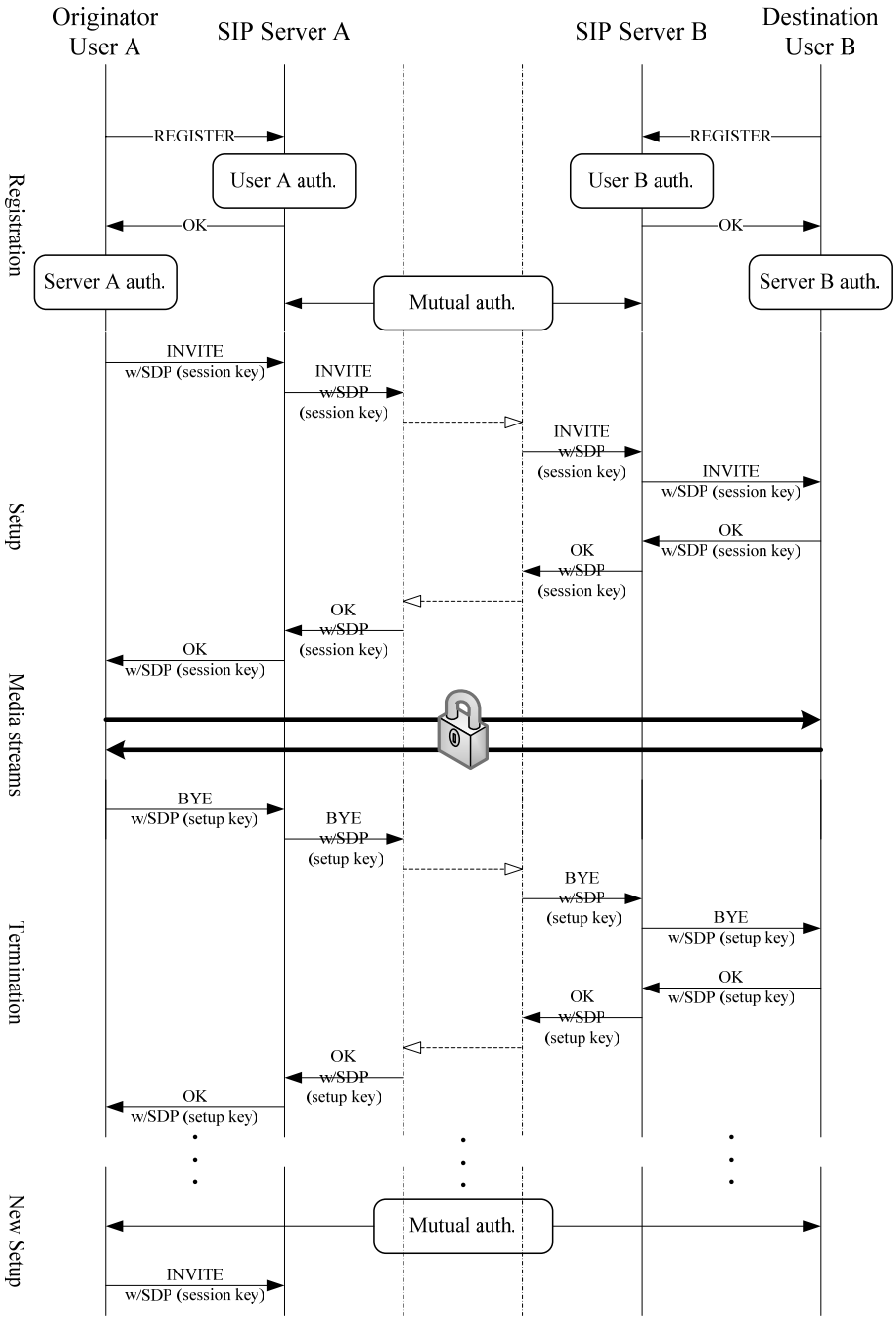

Figure 8. Example of overall flow