A Secure P2P Video Conference System for Enterprise Environments

Fuwen Liu, Hartmut Koenig

Brandenburg University of Technology Cottbus,
Department of Computer Science
PF 10 33 44, 03013 Cottbus, Germany
{lfw,koenig}@informatik.tu-cottbus.de

Abstract. Many emerging group oriented and collaborative applications such as audio/video conferences use the peer-to-peer (P2P) paradigm. Confidentiality is an often demanded feature for such applications, e.g. in business meetings, to provide group privacy. How to build a secure P2P video conference system is still an open issue. In this paper several possible solutions are discussed. We present a security architecture used for P2P video conferences that ensures confidential talks in an enterprise environment whose branches might be geographically dispersed.

1 Introduction

Video conference technology has been well studied and standardized by ITU-T in the H.323 recommendation [1]. H.323 based systems adopt the client-server communication model in which two centralized servers are applied to supporting group meetings: the *gatekeeper* for the group management and the MCU (*multipoint control unit*) for the distribution of the media streams. Although these systems are widely available, they possess several technical drawbacks. They are subject to a single point of failure and might become a performance bottleneck. Moreover, they are still pretty expensive what limits their wide deployment.

A P2P conference system as alternative approach is characterized by a distributed approach. All group management and media distribution functions of the system are assigned to the peers. The communication runs directly between the peers without passing a server. Thus, the drawbacks of H.323 based systems are basically eliminated. Peer-to-peer conference systems are well suited to setting up spontaneous conferences, because they do not depend on a certain infrastructure. So far only a few P2P video conference approaches have been reported like BRAVIS [2], DAVIKO [3], and the P2P-SIP architecture [4].

Security is of primary concern for such conference systems which are mainly deployed in *closed environments*, e.g. an enterprise whose branches are geographically dispersed, where they are usually used to discuss or to negotiate business topics. To protect H.323 based systems ITU-T released the recommendation H.235 [5] which specifies a security framework. This framework is not applicable to P2P conferences. Approaches dedicated to securing P2P conference systems have not emerged, yet. In

this paper, we propose such a security architecture and show how it can be incorporated into a P2P conference system using our video conference system BRAVIS as example. The remainder of the paper is organized as follows. After introducing the P2P video conference system BRAVIS in Section 2 we briefly describe the security requirements for a P2P video conference in Section 3. Next in Section 4, we discuss possible solutions for this issue. In Section 5 we introduce appropriate security architecture and show how it has been integrated in the BRAVIS system. Final remarks conclude the paper.

2 BRAVIS

BRAVIS [2] is a P2P multiparty video conference system designed for supporting collaborative groups in closed environments over the Internet. The essential technical features of BRAVIS system are following.

Hybrid P2P model

The P2P communication model distinguishes between pure and hybrid P2P models [6]. BRAVIS uses a hybrid P2P model. A SIP registrar is integrated in the system to allow peers registering their current IP address and retrieve the current IP addresses of the other peers for invitations**.** The hybrid model was chosen for two reasons. (1) No central authority is responsible for security related management functions like the identity management and the public key management in pure P2P systems what makes them prone to Sybil attacks [8], i.e. identity forgery. This allows an attacker to use different identities to attend conferences. To address this problem, a certificate authority (CA) was introduced in our system to centrally control the identities and the public keys. (2) Only one lookup operation is needed to locate a user when using the hybrid P2P model. In contrast, a pure P2P model like Chord [7] requires $O(logN)$ lookup operations for the same purpose.

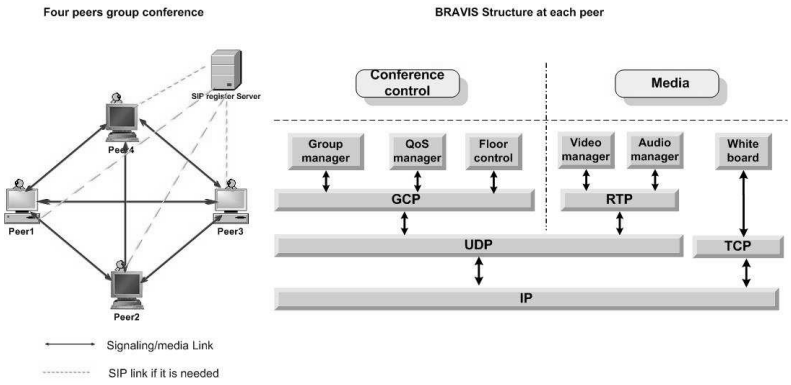

Fig. 1. Decentralized group management in the BRAVIS system

Figure 1 shows a four peer conference example and system structure at each peer. All peers of the group are assigned identical capabilities and properties. They use the

same system structure, i.e. all system control modules (group management module, floor control module, QoS module), and the media modules (video manager, audio manager, and whiteboard) are available at each peer. Thus each peer has the ability to supervise the composition of the group, to control the access to shared resources, and to tune QoS parameter without calling any additional server. Furthermore, media data transmissions take place among the peers involved in the current conference directly.

The system control modules run on top of the decentralized group communication protocol GCP [9] [10] which ensures the consistency of the conference control data among the peers. Based on this all peers possess the same view on the actual group state and can uniquely decide all group related issues by themselves, e.g. QoS parameter settings or floor assignments. GCP achieves this by providing virtual synchrony [11] to the upper layer modules. It assures that no data are lost, that data are delivered in the order as they are sent, and that all peers are updated equally.

3 Security requirements

Like other commonly used applications in the Internet, e.g. E-mail, a secure P2P conference has to support the well-known basic security features: *confidentiality, integrity, authentication*, and *access control*. Due to its decentralized structure and the real-time communication, a secure P2P conference system should meet some additional requirements beyond these basic demands:

End-to-end security

Usually two kinds of security services can be offered in an enterprise network: end-to-end security, or site-to-site and site-to-end security, respectively. The so-called end-to-end security means that messages are securely delivered from the sender' host to the receiver's host and that they are not accessible to any intermediate node or server along the transmission path. Site-to-site and site-to-end security mean that messages are merely protected during WAN transmission, while they are transmitted in the plaintext form within the site scope.

It is obvious that P2P conferences have to apply end-to-end security for several reasons: (1) Security threats occur not only during WAN transmission but also at local site as indicated in [12]. A significant number of threats originate from insiders. (2) In order to protect enterprise business secrets, enterprises demand that business information should be only accessible to group members but not to people outside the group, even if they belong to the same enterprise.

Group key management

In a secure P2P conference usually more than two participants are involved. A group key management protocol rather than a two-party key exchange protocol has to be applied to securing group communication. Two-party key exchange protocols are inefficient for group communication, because each member has to negotiate an individual key with the other group members. Each message sent to the group has to be separately encrypted with the respective keys of the group members, i.e. *n-1* encryptions are required. A group key needs only one encryption.

Flexible security policy enforcement
The security policy determines the desired protection level of a conference and specifies the security algorithms to be applied. The security policy of a P2P conference should be determined by the participants themselves rather than by a dedicated network administrator when running the conference, since a P2P conference is autonomous and consists of a transient group. The applied policy should be allowed to be attuned in the course of the conference to provide more flexibility for users.

Efficiency
Security always imposes additional processing burdens on the system. These burdens may pose a negative impact on the quality of service (QoS). For example, a secure conference incurs longer end-to-end communication delays due to message encryption/decryption. Therefore, the deployed algorithms and protocols should be efficient enough to meet the strict QoS requirements of real-time communication.

4 Overview of possible solutions

Nowadays virtual private networks (VPNs) are mostly applied for securing the communication across public networks. VPN functions can be introduced at different levels of the layered structure of the TCP/IP protocol stack. Correspondingly, there exist four kinds of VPNs: data link layer VPN, IPsec VPN, SSL VPN, and application layer VPN as shown in Figure 2.

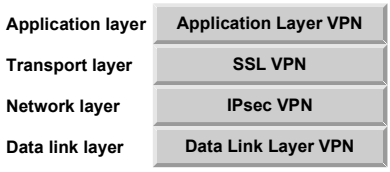

Fig. 2. Kinds of VPNs

Data link layer VPNs
Data link layer VPNs could be constructed using one of three protocols: Point-to-Point Tunneling Protocol (PPTP) [13], Layer 2 Forwarding (L2F) [14], and Layer 2 Tunneling Protocol (L2TP) [15]. They were commonly applied to dial-up communications between a mobile user and the gateway of its enterprise network to provide site-to-end security rather than end-to-end security when they are used with IPsec together.

IPsec VPNs
IPsec VPNs are enterprise networks which are deployed on a shared infrastructure using IPsec technology. The most important advantage of IPsec is that it is transparent to applications. Any IP based applications without modifications can get total protection when it is deployed. However, several disadvantages inherently exist when it is used for a P2P conference.

- *Inflexible security policy enforcement*
  Prior to the deployment of an IPsec VPN, the associated security policies must be manually configured in the related IP nodes. This specific task is usually only

allowed for the network administrator but not for general users, because IPsec is implemented in the kernel [16].

- *Difficulty to offer end-to-end security*
  IPsec VPNs operate at the network layer which is the lowest layer to provide end-to-end security in theory, but in practice IPsec VPNs rarely adopt a host-to-host architecture to provide end-to-end security for data transmission. This is because the configurations (e.g. security policy enforcement) on each host have to be manually carried out by the network administrator. This is an unbearable burden for the system administrator, especially for a large number of users [17].
- *Inefficient group communication*
  Currently IPsec does not support a group key management but only a two-party key management.

SSL VPNs

SSL VPNs are based on the commonly used protocol SSL (*Secure Socket Layer*) for secure data transmissions at the transport level. It was standardized by IETF where it is called TLS [18]. SSL VPNs are extensively used in HTTP-based applications to provide end-to-end protection between client and server. Like IPsec VPNs, there are problems for their use in P2P conferences:

- *Inefficient group communication*
  This is simply because the handshake protocol of SSL deals with the key management only for two parties rather than for the whole group members.
- *Merely supporting TCP-based applications*
  SSL merely supports TCP based applications, since its design assumes that the underlying layer offers a reliable transport. If SSL is applied in connection with UDP based applications, packet losses are viewed as security breaks that force to release the communication [18].

Application layer VPNs

Application layer VPNs use the security functions embedded in the respective applications. Due to its embedded implementation it can provide a more tailored protection compared to the underlying layer VPN technologies. Moreover, appropriate security algorithms and protocols such as a group key management protocol could be readily integrated into the system to meet the security requirements mentioned in Section 3. The major drawback of application layer VPNs is that some modifications have to be made in the applications to add these security functions. The designed security architecture is solely available for the designed application.

Table 1. Comparisons of VPNs

Security requirements	Date link layer VPN	IPsec VPN	SSL VPN	Application layer VPN
Basic security services	Yes	Yes	Yes	Yes
End-to-end security	No	Difficult	Yes	Yes
Group key management	No	No	No	Yes
Flexible security policy enforcement	No	No	Difficult	Yes
Supporting TCP and UDP-based applications simultaneously	Yes	Yes	TCP only	Yes
Transparent to applications	Yes	Yes	Yes	No

Summary

Table 1 shows that Data link layer VPNs and SSL VPNs are inappropriate for P2P conferences, because the first one does not provide end-to-end security and the latter one does not support UDP based applications. A straightforward solution would be the direct use of existing IPsec VPN infrastructure to support a P2P conference. Unfortunately, IPsec VPN is scarcely used in a host-to-host fashion to support the end-to-end security. Moreover, missing group key management and inflexible security policy enforcement make it difficult for IPsec VPNs to supporting P2P meetings in a dynamic and efficient manner. To fully meet the security requirements of P2P conferences the design of dedicated security architecture seems the most appropriate way, even it is more costly.

5 Secure BRAVIS system

In this section we introduce the security architecture designed for our P2P conference system BRAVIS. Our aim is to ensuring confidential P2P meetings on an end-to-end basis. For this purpose, a security layer has been inserted in the BRAVIS architecture presented in Section 2. It has been placed between the application and the communication layer. The resulting structure of the secure BRAVIS system is depicted in Fig. 3.

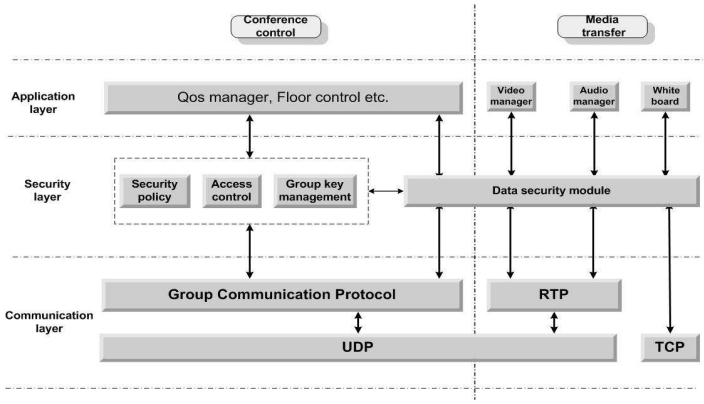

Fig. 3. Secure BRAVIS system

The security layer is composed of several modules. Each module fulfills a dedicated security function. They are shortly explained in the sequel.

Security policy module

This module decides which security level and what kind of security algorithm are enforced for the conference. Four security levels are distinguished. A *level zero* conference corresponds to a normal conference, where no special security function is applied. In a *level one* conference the joining of the group involves a mutual authentication, but the data exchange is not further protected. A *level two* conference besides the mutual authentication encrypts all signaling data and one or two media streams (video, audio, or whiteboard). *Level three* conferences are the most secure ones. All

exchanged data are protected. The entrance into the conference is only allowed after the successful mutual authentication. In addition, two different operation modes for managing the security policy were introduced: moderation and voting. In the moderation mode, one participant is designated as moderator who solely decides all security demands. When the moderator leaves the conference, he/she can hand over the moderation right to one of the remaining members. In the voting mode all group members share the same right to decide about the security policy. The security policy used in the conference is determined by voting.

Access control module

In BRAVIS the entrance into the meeting is by invitation. Each participant in the meeting can invite a new partner based on a social agreement with the other partners. No constraint is imposed on the callers for their calling activities, but an access control is applied to the invitee. Each participant on its own decides who can invite it. This is achieved by the use of an access control list (ACL) which is maintained by each participant. When a participant receives an invitation message, the required mutual authentication procedure is invoked. If this authentication is successful, the participant will check its ACL to examine whether the inviter has the right to call him/her. If true, it may accept this call.

Group key management module

A decentralized group key exchange protocol used by group members to manage the group key themselves should be deployed to match the P2P communication model. Several protocols are available for this purpose such as TGDH [19], the protocol proposed by Rodeh et al. [20], and others. However, they still possess shortages in the respect of security and efficiency. To overcome these shortages we designed and implemented an efficient and secure decentralized group distribution protocol for our system, called VTKD (*virtual token based key distribution)* [21]. VTKD consists of two parts: a mutual authentication of the partners and a secure key renewal. The latter is triggered when the group composition changes, i.e. when members join or leave the group. The public key signatures based mutual authentication between the inviting group member and the invitee is invoked when a new member joins. This ensures that the group key is only delivered to an authenticated member, while the new member can be sure that the received key is in fact shared with the inviting parties.

Data security module

The data security module is used to ensure the data confidentiality and integrity during a conference. The participant can separately select different security algorithms for the protection of the four kinds of data (video, audio, whiteboard, signaling). Standard encryption algorithms are used to process audio, whiteboard, and signaling data in real-time due to their small data size. For the real-time video transmission, a specific encryption algorithm is needed to meet the stringent QoS requirements and to handle the large amounts of video data (the bit rate of a MPEG2 video stream typically ranges between 4 and 9 Mbps [22]). The end systems in a multiparty P2P video conference have to simultaneously compress/decompress and encryption/decryption the outgoing video stream and all incoming video streams in real-time. This imposes a high processing burden. Therefore, we developed a novel video encryption algorithm [23] which is fast enough to meet real-time demands with a sufficient security.

6 Final remarks

P2P conference systems represent a new trend in the development of video conferencee systems. They provide a couple of interesting advantages compared to the traditional server based H.323 systems. Security is of primary concern for these systems to ensure the confidentiality of the talks, especially when applied in enterprise environments. In this paper we analyzed the feasibility of VPN technologies to secure P2P conferences. We showed that lower layer VPN technologies are not flexible enough for P2P conference applications. We argued that a security architecture especially designed for a P2P conference system (i.e. an application layer VPN) is a more appropriate solution to meet the stringent security and efficiency requirements. As example for such a security architecture, we presented the security solution applied in our P2P video conference system BRAVIS [2].

References

1. ITU-T: Recommendation H.323 v5--Packet based multimedia communication systems. July 2003.
2. The BRAVIS peer-to-peer video conference system. http://www.bravis.tu-cottbus.de.
3. The DAVIKO system. http://www.daviko.com.
4. D. A. Bryan and B. B. Lowekamp: Standards-Based P2P Communications Systems, Proceedings of the 2005 Virginia Space Grant Consortium Research Conference, April 2005
5. ITU-T: Recommendation H.235v3-- Security and encryption for H-series (H.323 and other H.245-based) multimedia terminals. May 2003.
6. D. S. Milojicic, V. Kalogerali, R. Lukose, K. Nagaraja, J. Pruyne, B. Richard, S. Rollins, Z. C. Xu: Peer-to-Peer Computing. HP white paper HPL-2002-57, March, 2002.
7. I. Stocia, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan: Chord: A scalable peer-to-peer lookup service for internet applications. In Proc. of ACM SIGCOMM 2001, pp 149-160, 2001.
8. J. R. Douceur: The Sybil Attack. IPTPS'02, March 2002.
9. E. C. Popovici, R. Mahlo, M. Zuehlke, and H. Koenig: Consistency Support for a Decentralized Management in Closed Multiparty Conferences Using SIP. In Proc. of IEEE ICON 2003, pp. 295 – 300.
10. M. Zuehlke and H. Koenig: A Signaling Protocol for Small Closed Dynamic Multi-peer Groups. In Z. Mammeri and P. Lorenz (eds.): HSNMC 2004, Springer LNCS 3079, pp. 973 – 984, 2004.
11. G. V. Chockler, I. Keidar, and R. Vitenberg: Group communication specifications: A comprehensive study. ACM Computing Suryes 4 (2001) 427-469.
12. ITU-T manual: Security in Telecommunications and Information Technology. December 2003.
13. K. Hamzeh, G. Pall, W. Verthein, J. Taarud, W. Little and G. Zorn: Point-to-Point Tunneling Protocol (PPTP). RFC 2637, July 1999.
14. A. Valencia and T. Kolar: Cisco Layer Two Forwarding (Protocol) "L2F", RFC 2341. May 1998.
15. W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zorn and B. Palter: Layer Two Tunneling Protocol "L2TP", RFC 2661. August 1999.
16. R. Perlman and C. Kaufman: Key Exchange in IPsec: Analysis of IKE. IEEE Internet Computing 2000
17. S. Frankel, K. Kent, R. Lewkowski, A. D. Orebaugh, R. W. Ritchey and S. R. Shama: Guide to IPsec VPNs. NIST Special Publication 800-77, January 2005.
18. T. Dierks and C. Allen: The TLS protocol Version 1.0. RFC 2246, January 1999.
19. Y. Kim, A. Perrig, and G. Tsudik: Simple and fault-tolerant key agreement for dynamic collaborative groups. ACM CCS 2000, pp. 235–244.
20. O. Rodeh, K. P. Birman, D. Dolev: Optimized Group Rekey for Group Communication Systems. In Proc. NDSS 2000, pp. 39-48.
21. F. Liu and H. Koenig: An efficient key distribution protocol for small closed peer groups. GI/ITG-workshop on peer-to-peer systems and applications. LNI Proceedings V.P-61, pp 163-167, 2005.
22. B. G. Haskell, A. Puri, and A. N. Netravali: Digital Video: An Introduction to MPEG-2. Kluwer Academic.
23. F. Liu and H. Koenig: A Novel Encryption Algorithm for High Resolution Video. In Proceeding of ACM NOSSDAV'05, Stephenson, WA, USA, June 2005.