基于SIP的多方视频会议模型设计与实现

柴二建,冯子亮
(四川大学 计算机学院,四川 成都 610065)

摘 要:用SIP协议构建视频会议系统已成为一个趋势,但在设计和实现过程中也存在一些缺陷和问题,如无成熟标准、安全性差。为解决这些问题,提出一个基于SIP的多方视频会议系统模型。该模型采用分层设计思想,利用成熟的网络协议和加密技术,保障了系统的实用性和安全性,是一个可扩展到广域网的适应性很强的多方视频会议解决方案。通过系统设计和实现验证了该方案的可行性和实用性。

关键词:视频会议系统;H.323;SIP;公钥加密
中图分类号:TN948.63 **文献标识码**:A **文章编号**:1673-629X(2008)05-0047-04

Design and Implementation SIP-Based Multiple Parties Video Conference System

CHAI Er-jian, FENG Zi-liang
(College of Computer Science, Sichuan University, Chengdu 610065, China)

Abstract: It has been a trend to use the SIP protocol in the construction of the video conference systems. However, there are some drawbacks and problems in the system design and implementation, such as no mature architecture, no security countermeasures against network attacks. Presents a new model for SIP-based mutiple parties video conference system. This model is designed based on the hierarchical design theory. The mature network protocols and encrypt techniques are used in this model which guarantees the practicability and security of the viedo conference. Furthermore, it can be easily applied in the WAN environment by changing the configuration. The design and implementation validates the feasibility and usability of this model.

Key words: video conference system; H.323; SIP; public-key encryption

0 引 言

随着我国宽带建设的进一步普及,以及电子政务和信息化建设的不断推进,视频会议产品已经越来越广泛地应用到了各个领域。现在市场上的视频产品大多是基于H.323协议的,但H.323标准过于严格,而且实现起来较难,系统的自由度较小,而且硬件设施比较昂贵。现在只有一些特殊的领域和一些特大企业,如政府机构、军队、一些特大的跨国公司等采用视频会议产品。广大的中小企业、学校同样有着视频会议、语音会议等多方面的潜在需求,但由于成本过高和网络环境不理想,使得视频会议系统很难在这些用户中普及。因此寻找一种成本低、灵活性高,在现有网络环境下可以进行视频通信的技术很有必要。而利用SIP灵活、简单,易实现,独立与硬件的特点,可以在现有的网路环境中很好地实现会议功能。正好解决了当前视频系统推广的瓶颈问题,有利于视频系统的进一步普及,也将提高我国的信息化水平。

收稿日期:2007-08-25

作者简介:柴二建(1984-),男,江苏连云港人,硕士研究生,研究方向为SIP多媒体传输、图像处理与图像信息检索;冯子亮,副教授,研究方向为图形图像、空中交通管理等。

1 SIP介绍

SIP是IETF(Internet工程任务组)提出的、不同于H.323的一种支持多媒体会话的IP网络信令控制协议[1]。SIP是应用层的协议,独立于传输层,用于描述生成、修改和终结多个参与者之间的会话。

它采用文本表示,简单、易于实现;功能扩充性和网络扩展性良好。SIP消息可以携带任何格式的消息体;通过携带不同的消息体可完成不同的数据业务,通过定义新的方法和消息头域可丰富SIP自身的呼叫控制。SIP的主要功能包括:资源定位、加入服务会话、会话参数协商等。

SIP作为多媒体通信的应用层控制(信令)协议,能够建立、改变和终止多媒体会话。按逻辑功能分,包括及一个用户代理客户机(UAC)、一个SIP代理服务

器、一个用户代理服务器(UAS)。UAC 发起一个请求,SIP 代理服务器作为终端用户的位置发现代理,UAS 接受或响应这个呼叫。SIP 可用于发起新的会话,也可以用于邀请成员加入已经建立的会话[2]。

SIP 只提供了会话过程中的消息呼叫和通信的流程及消息格式等基本的功能,并未提供包括会议控制、媒体数据传输在内的其他服务,因此还必须借助其他的协议如 RTP,才能构建完整的网络会议系统[3]。因此,到目前为止,SIP 在视频会议上还未有一个成熟的方案,普及度还较低[4]。SIP 使用的是类似于 HTTP 的文本协议,在安全方面也存在一定的问题。

2 系统模型设计

利用 SIP 实现灵活、简单的特点,结合局域网良好的网络传输条件和现有的各种成熟的网络传输协议,设计了一个 SIP 视频会议模型。该模型包括了完整的 SIP 结构和功能,能进行局域网内多方视频会议。包括了 SIP 用户注册、管理,会议的创建、管理,以及会议过程中的多方控制。在实现过程中,考虑到广域网中会议特点和 SIP 的强大定位功能,加入虚拟的 DNS 等定位服务器,来实现 SIP 消息的转发和终端定位功能。而且加入了身份认证机制,并利用公钥机制对消息体进行加密,提高会议安全性。改进后的系统模型如图 1 所示。

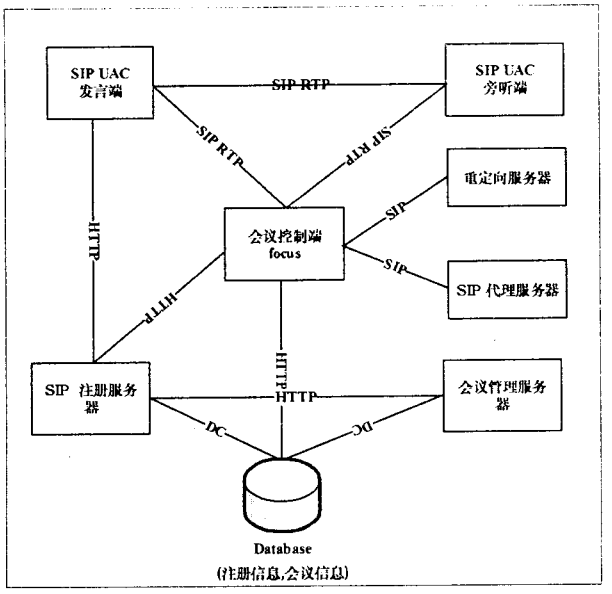

图 1 基于 SIP 的视频会议模型

由图 1 知,整个系统模型由三类实体组成,服务器、SIP 终端以及数据信息库。

SIP 终端包括 SIP UAC 和会议控制端,它们是用户直接可控部分,向用户提供视频会议的功能。如会议参与者可以通过 UAC 加入会议,提交发言要求,主持人可以通过会议控制端对整个会议进行监控如会议开始、结束、发言转换,验证终端用户身份,应答会议参与者的请求。

服务器包括 SIP 注册服务器、会议管理服务器,SIP 代理服务器和重定向服务器,它们主要完成三个方面的工作:一是为会议进行提供信息支持,会议参与者通过 Web 浏览器的方式向 SIP 注册服务器注册自己的信息,会议创建者可登陆会议管理服务器选择会议的参与者并发布会议信息。重定向服务器和 SIP 代理服务器为会议控制端提供 UAC 的位置信息,控制端将该信息转发到经认证后的 SIP 客户端;二是管理会议信息,会议参与者可以通过会议管理服务器创建并发布会议信息,删除会议,还可以保存和查询历时会议信息;三是提供身份认证和加密措施,确保系统的安全性。SIP 注册服务器要求每个用户注册时输入一个公钥信息,在发布会议信息之前将信息用该用户的公钥使用 RSA 算法进行加密,这样就完成了对用户的身份的认证和信息的加密功能。在会议过程中,控制端也通过此密钥使用 RSA 加密控制信息来确保系统的安全性能。

数据库主要用于存储用户的注册信息和会议信息,必须确保信息的一致性、完整性、安全性。任何对数据库的操作都要经过身份的验证,只有会议管理员才能登陆会议管理服务器来对数据库信息进行操作。从图 1 可看出,在设计中将数据库和 Web 服务器隔离,降低了攻击的概率。

模型充分依靠现有条件先成熟的协议如 HTTP,RTP 来实现功能上的需求,降低了实现的技术难度,而且增加了模型的推广性。

3 工作原理

3.1 会议流程

如图 2 所示,进行一次多方会议由 8 个流程完成:

①会议参与者向 SIP 注册服务器注册信息,如 SIPID,密码,参与者的公钥,Email 等信息。

②会议发起者登陆管理服务器创建一个会议。

③启动会议。发起者启动控制端,并通过 Email 的形式向参与者发布会议信息。包括控制端的 SIPURI,会议组地址(用于广播)和其它的一些信息。

④参与者启动 SIP UAC,加入会议组等待消息。

⑤控制端广播发布会议通告,提示用户登录。

⑥用户输入认证信息(SIPID,密码)登陆。经过控制端身份确认后,进入会议。此时,控制端获得 UAC 的 SIP URI 并更新列表。

⑦UAC 向控制端发布 SIP BYE 信息,退出会议。

⑧控制端广播BYE请求,结束会议。

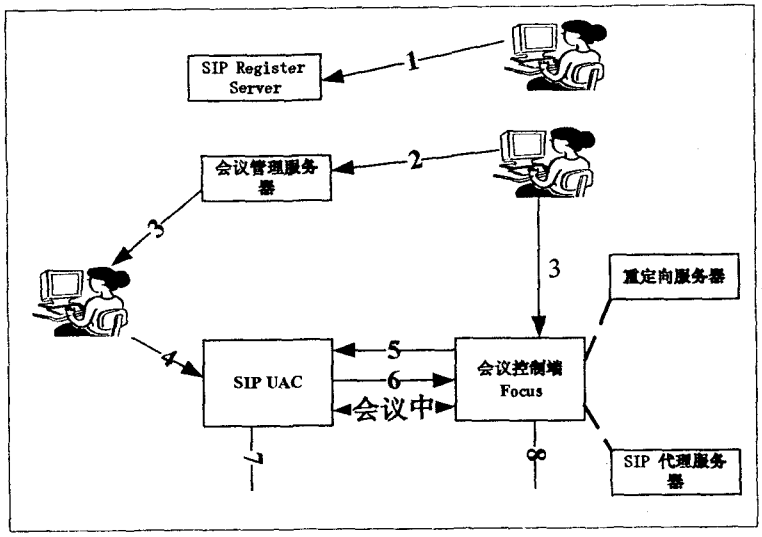

图2　会议流程图

3.2　工作机制

(1)SIP UAC向控制端请求视频传输。

SIP终端登陆控制端,并得到控制端的身份验证后,向控制端要求加入会议。假设SIP UAC为A,控制端为终端C。

①终端A向控制端发送INVITE请求。

②终端C从下层获得INVITE消息,解析得知是一个INVITE消息,解析header中的"to"得知是发向自己的消息,查看自己的"role"的值为"invited",即自己正忙于和另一个终端建立连接,如B。因此,它构建了一个"Trying"的response作为回复。解析INVITE的Contact header获得回复地址,利用UDP向终端A发去以通知A等待,C处于繁忙状态。

③终端C和其他终端的连接建立完成后,准备和终端A建立连接。它向A发送一个"Ringing" response,通知A做好准备。

④终端C向A发送成功的OK response,并携带了自己的媒体描述SDPC用于进行连接建立前的媒体协商。

⑤终端A从底层获得OK response消息,解析"From"header知,该消息是对自己发出的请求的响应,比较自己的媒体信息后,直接向"To"发送ACK。终端C收到来自A的ACK后,便根据双发协商的媒体信息来建立多媒体传输连接。

⑥终端A想退出已建立的多媒体连接,向C发送BYE消息。

⑦终端C接到BYE消息后,向A发送ACK。双方的连接断开。

(2)SIP UAC向其他的SIP UAC请求点对点的多媒体通信。

为了支持终端的移动性和会议安全性,每个终端内只具有控制端的地址信息("控制端信息表")和一个"用户名册",其它用户信息由控制端向SIP终端转发。因此,当终端之间建立点对点的多媒体通信时,它们的SIP消息需经过代理服务器(本例中用控制端来模拟)转发[5]。假设终端A向终端B请求建立点对点会话,消息流程如图3所示。

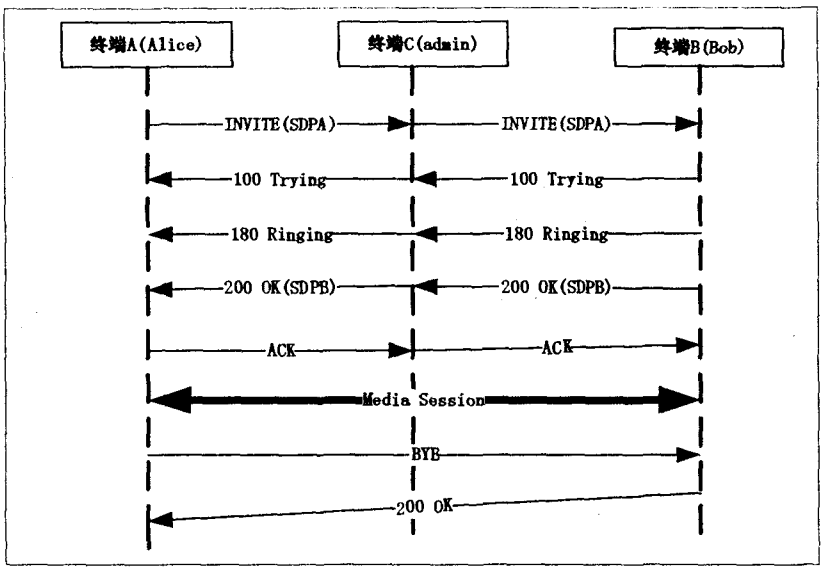

图3　终端A和B建立点对点通信

①终端A在"用户名册"中选择BOB与其进行点对点通信。首先A构建和终端B建立连接请求的SIP消息。终端A解析来自上层的SIP消息后,解析知是一个INVITE的request,然后继续解析"to"header,得知是发向BOB的请求。查询"控制端信息表",该目的地址非控制端,则发送至域名主机(此时由控制段C模拟)查询SIP URI信息。

②控制端C收到SIP消息后,解析得知是一个INVITE类型的request,解析"to"header,获得目的终端的URI为:<sip:bob@session.com>,为自己域名范围内的地址,查询"参与者信息表",获得目的地址,构建新的request,发送到终端B。

③终端B收到消息②后,解析得知是一个INVITE的request,然后解析"to"得到目的SIP URI:<

sip:bob@session.com>,是发向本客户端的请求,解析"contact"header 获得回复地址。因此根据自身状态,构建 response 消息,发向需回复的地址。

消息(3)~(11)的过程与第一种情况类似,不再详细说明。

(3)会议管理和控制。

各终端加入会议后,均和控制端建立了多媒体连接,控制端获得所有参与者终端的多媒体信息。能够对每个终端进行浏览,断开连接,要求发言等操作。

终端浏览,从安全角度考虑,只有控制端能浏览各终端状态,每个终端只能看到控制端(或发言端)的状态,而每个终端之间的状态是封闭的。

断开连接,控制端可以断开和任何终端的连接,向该终端发送 BYE 请求,该终端用 ACK 响应后,断开两者之间的多媒体连接。

选择发言端,控制端选择要发言的用户,向该终端发送 NOTIFY 消息,该终端用 ACK 回复后,控制端停止向多媒体分组中发送数据,发言终端向该分组发送自己的多媒体数据。

3.3 身份验证和消息加密

(1)身份验证。

会议参与者身份验证,SIP UAC 登录会议时要输入一对用户名/密码,通过验证后才可以进入会议流程。控制端身份验证,为了防止第三方窃听 SIP 消息,冒充会议控制端。利用用户的密钥对对控制端的身份进行验证。会议控制端用私钥对广播通告进行加密,会议工程中,SIP 控制端用自己的私钥对包信息中的主体信息 SDP 进行加密,SIP UAC 利用控制端的公钥进行解密,如果成功说明控制端的身份合法。

(2)消息加密。

会议发起者向参与者发送会议信息的 EMAIL 内容用参与者的公钥进行加密,防止消息窃取;会议过程中,所有终端都用下一跳的终端公钥进行加密,目的终端解密后进行正常的消息交互。

4 系统实现

模型的软件实现借鉴网络协议中的分层思想,下一层为上一层的操作提供"服务"[6]。从下至上分为三个层次:SIP 协议栈、会议流程控制和策略控制以及用户界面。

4.1 SIP 协议栈

SIP 协议栈主要为上层提供 SIP 消息的构建、发送和解析功能。同时提供一些基本的多媒体处理功能。包括多媒体数据的 RTP 传播,音频视频数据的采集、编码、播放等。SIP 消息管理层的结构如图 4 所示(由各个类组成)。

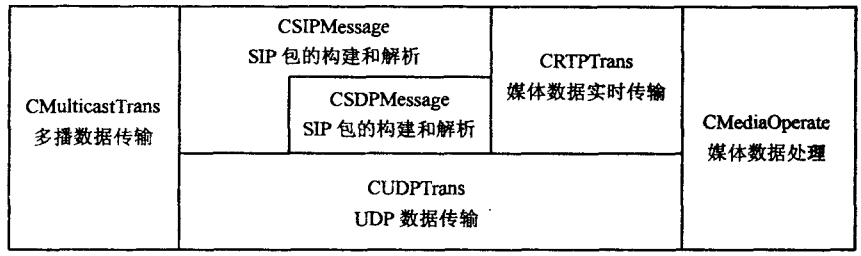

图 4 SIP 协议栈结构

4.2 会议流程控制层

SIP 本身不支持会议管理和控制,实现了一个会议控制层,它由一系列的会议控制消息和用户状态组成,根据会议控制消息来控制状态的转移[7]。流程控制层能够根据用户的功能性要求来调用 SIP 管理层的相关服务来达到用户的要求,是承上启下的一个层次。根据本系统的功能要求,用户的功能调用分为会议流程及策略控制两个方面。另外还有点对点的多媒体通信的控制。该层次由三个类组成:会议流程控制 CProcessControl,会议策略控制类 CStrategyControl 和对点对点多媒体通信的控制 CPToPConnect。

4.3 用户界面层

用户界面为用户使用该系统的功能提供接口。用户界面主要包括会议控制端的界面和 SIP 终端的界面(包括 SIP 发言端和 SIP 旁听端)。

5 结束语

探讨了基于 SIP 协议的视频会议系统,提出了一个完整的可扩展到广域网的多方会议模型。利用 SIP 灵活、简单、成本低的特性解决了现阶段 H.323 视频系统难于推广的问题。在模型中使用成熟的网络协议,如 HTTP,RTP 等,具有很好实用性和推广型;利用通行证和 RSA 公钥策略对终端身份进行认证和 SIP 消息进行加密,保证了会议过程的安全性;在软件上分层实现,使得该模型具有一定的重用性和较好扩展性。

多方会议模型仅仅实现了视频会议基本的一些功能,不能支持移动终端的访问。可以预测,SIP 必将在无线网的多媒体传输中扮演重要的角色,因此,此模型在移动终端的支持性能上还要加以完善。

参考文献:

[1] Rosenberg J, Schulzrinne H, Camarillo G, et al. Session Initiation Protocol[S]. RFC3261.2002.

[2] 傅志辉,梁荣华. 基于 SIP 协议远程教学系统模型[J]. 浙

(下转第 53 页)

找候选项集可以从事务数据库中项目数最多的事务开始,将其对应的项集作为候选项集,引出算法。

算法1:

(1) 初始时,设事务数据库 D 中有 m 个项目 $I_j(j=1,2,\cdots,m)$,D 对应的0-1矩阵为 $A(D)$。在 D 中找出项目数最多的事务(对应于 $A(D)$ 的含1最多的行向量,可能不止一个),将它们对应的项集作为候选项集,这些候选项集的集合记为 T。

(2) 记 T 中每个候选项集所含项目数为 d,若 $d=0$,说明找不到频繁项集,退出程序;否则,当 $d>0$ 时,执行(3)。

(3) 利用结论1计算出 T 中每个候选项集的支持度计数。找出所有满足支持度计数大于等于 min_sup 的项集,若存在,它(们)就是频繁项集,算法到此结束;否则执行(4)。

(4) 将 T 中每个候选项集分成 C_d^{d-1} 个子项集(这里每个子项集所含项目数为 $d-1$),作为候选项集;再找出 D 中所含项目数为 $d-1$ 的事务对应的项集(相应于 $A(D)$ 的含 $d-1$ 个1的行向量),也作为候选项集。将 T 更新为这些候选项集构成的集合,返回(2)。

3 实例分析

针对图1的事务数据库 D,它的0-1矩阵 $A(D)$ 见图2,假设最小支持度计数 min_sup = 2。根据算法1,矩阵 A 的行向量中1个数最多的为第4、8行(有4个1),对应于事务 T_4 和 T_8,相应项集为 $\{I_1,I_2,I_4,I_5\}$ 和 $\{I_1,I_2,I_3,I_5\}$,即 $T=\{\{I_1,I_2,I_4,I_5\},\{I_1,I_2,I_3,I_5\}\}$,$T$ 中每个候选项集含 $d=4$ 个项目。根据结论1,$\{I_1,I_2,I_4,I_5\}$ 和 $\{I_1,I_2,I_3,I_5\}$ 的支持度计数均为1 < min_sup,因此它们均不是频繁项集。将项集 $\{I_1,I_2,I_4,I_5\}$ 和 $\{I_1,I_2,I_3,I_5\}$ 分成含3个项目的子项集:$\{I_1,I_2,I_4\}$,$\{I_1,I_2,I_5\}$,$\{I_2,I_4,I_5\}$,$\{I_1,I_4,I_5\}$,$\{I_1,I_2,I_3\}$,$\{I_1,I_3,I_5\}$,$\{I_2,I_3,I_5\}$,作为候选项集;此外,D 中含3个项目的事务对应的项集有 $\{I_1,I_2,I_5\}$,$\{I_1,I_2,I_3\}$,也作为候选项集。根据结论1,这些候选项集的支持度计数如图4所示。

项集	支持度计数
I_1, I_2, I_4	1
I_1, I_2, I_5	3
I_2, I_4, I_5	1
I_1, I_4, I_5	1
I_1, I_2, I_3	2
I_1, I_3, I_5	1
I_2, I_3, I_5	1

图4　候选项集的支持度计数

由图4可以看出,只有两个候选项集的支持度计数满足 min_sup 条件,分别为 $\{I_1,I_2,I_5\}$ 和 $\{I_1,I_2,I_3\}$,则此事务数据库的频繁项集为:$\{I_1,I_2,I_3\}$ 和 $\{I_1,I_2,I_5\}$。

4 结束语

关联规则挖掘是当前数据挖掘领域的主要研究课题[6]。文中主要在Apriori算法的基础上,改进了计算支持度计数和寻找频繁项集的方法。改进后的算法较原Apriori算法在时间和空间上有了明显的提高。

参考文献:

[1] Han Jiawei, Kamber M. 数据挖掘概念与技术[M]. 范明,孟小峰译. 北京:机械工业出版社,2007.

[2] 刘以安,羊　斌. 关联规则挖掘中对Apriori算法的一种改进研究[J]. 计算机应用,2007,27(2):418-420.

[3] 王柏懿,刘寒冰,靳　书,等. 基于矩阵的关联规则挖掘算法[J]. 微计算机信息,2007,24(5-3):143-145.

[4] Agrawal R, Limielinski T, Swami A N. Mining Association Rules Between Sets of Items in Large Database[C]// Proceedings of ACM SIGOD Conference on Management of Data. Washinton DC:[s.n.],1993:207-216.

[5] 张梅峰,张建伟. 基于Apriori的有效关联规则算法的研究[J]. 计算机工程与应用,2003,39(19):196-198.

[6] Agrawal R, Srikant R. Fast Algorithms for Mining Association Rules in Large Database[C]// Proceeding of the 20th International Conference on Very Large Databases. Santiago, Chile:[s.n.],1994:487-499.

(上接第50页)

江工业大学学报,2007,35(2):159-162.

[3] 吴志军,马　兰,沈笑云. Visual C++ 视频会议开发技术与实例[M]. 北京:人民邮电出版社,2006.

[4] 计算机基础教程网. SIPSAP及SDP协议组合应用的研究[J/OL]. 2006-10-20. http://www.itwen.com/03office/06file/file20061020/66563.html.

[5] 陈华林,盛翊智. SIP协议中的媒体协商[J]. 技术交流,2005,25(4):71-79.

[6] 曾庆珩,胡瑞敏,边学工. 基于SIP的集中式会议控制模型及实现[J]. 计算机工程,2005,31(3):198-200.

[7] 张友波,张焕强,孙利民. 基于SIP的视频会议系统设计与实现[J]. Computer Engineering,2005,31(21):167-169.