

基于 P2P 的 SIP 网络研究

李桂林, 李建华

(中南大学 信息科学与工程学院, 长沙 410075)

(rainbow101@163.com)

摘 要:在现有 P2P 应用和 SIP 协议特点的基础上,提出了分层 P2P-SIP 网络设计方案,详细分析了 P2P-SIP 节点的实现机制。该设计采用 P2P 技术提高传统 SIP 网络的可靠性、自组织性,并解决了异构 P2P-SIP 网络互通问题。

关键词:P2P; 会话初始协议; 分层 P2P-SIP 网络; 互通

中图分类号:TP393.02 **文献标志码:**A

Research of SIP Internet based on P2P

LI Gui-lin, LI Jian-hua

(School of Information Science and Engineering, Central South University, Changsha Hunan 410075, China)

Abstract: A hierarchical P2P-SIP architecture was proposed based on the features of the existing Peer-to-Peer (P2P) applications and Session Initiation Protocol (SIP). Implementation mechanism of the node was introduced in details. The traditional P2P technology was used to improve the reliability and self-organization of SIP network, and the interworking problem among the heterogeneous overlays was resolved.

Key words: Peer-to-Peer (P2P); Session Initiation Protocol (SIP); hierarchical P2P-SIP overlay; interworking

会话初始协议(Session Initiation Protocol, SIP)^[1]作为下一代网络和软交换技术的一个重要发展方向,已成为目前研究和应用的热点。Peer-to-Peer(P2P)^[2]系统拥有高扩展性、高容错性、健壮性的特点,这种特点得益于 P2P 不依赖中心节点而依赖网络节点的自组织对等协作的方式发现资源、共享资源。P2P-SIP 网络的关键技术是如何采用 P2P 分布式算法进行资源定位以及解决异构 P2P-SIP 网络互通问题。由于使用 SIP 标准通信协议,P2P-SIP 更容易与其他软交换设备或 PSTN 互通。

1 背景和相关研究

Chord^[3]是一种基于 DHT 的结构化 P2P 算法,具有负载均衡性和查询结果返回的确定性。它有一个环形逻辑拓扑结构,每个节点维护 $\log(N)$ 条关于其他节点的信息(N 是 P2P 网络里的节点数目)。P2P 网络的查询以增加呼叫建立延迟为代价,只需通过查询 $O(\log(N))$ 个其他节点就能找到目标节点。Chord 的迭代和递归查询可直接映射为 SIP 的重定向和代理特性。

Skype^[4,5]是 Kazaa 体系结构的 VOIP 应用系统。它采用私有协议,存在中央认证服务器。在某种程度上,Skype 体系结构与传统 SIP 基本相同,除了 GI(Global Index)技术为新加入节点搜索超级节点 SN 的在线用户。这些超级节点类似于 SIP 注册服务器、代理服务器等,为其他用户提供路由查找。但是 Skype 采用泛洪查询,严重影响了网络带宽。它最大的优点是在节点的应用程序中配置等价于 STUN 和 TURN 服务器功能模块解决 NAT 问题^[6],不像现有的 SIP 应用需额外配置外在的服务器。

传统 SIP 网络采用 Internet 的客户机/服务器工作方式。

SIP 网络包含两类实体:用户代理和网络服务器。SIP 协议是一种有状态的协议,在不可靠 IP 网络上控制会话的建立、修改和终止。传统 SIP 网络也可以看成预设一些固定超级节点(SIP 服务器)的 P2P 系统,但它基于 DNS 查找而不是哈希值。采用纯 P2P 框架可以提高现有 SIP 网络的可靠性、自组性,能动态适应节点的单点失效。

关于 P2P 技术和 SIP 协议有效结合的研究主要有两个方面:P2P-over-SIP^[7]和 SIP-over-P2P^[8,9]。前者是 P2P 的查找中采用 SIP 信令,如 SIPshare^[7]是一种采用了 SIP 的 SUBSCRIBE 和 NOTIFY 消息、非结构化的 P2P 文件传输系统。后者是现有的 P2P 技术代替传统 SIP 网络的 DNS 和 SIP 服务器,但是这样却存在两个问题:1)在一个 P2P 网络(如 Chord)的节点很难通过独立的 P2P 节点定位服务路由 SIP 消息到达另外一个 P2P 网络(如 Pastry);2)当采用 P2P 技术提供节点定位服务时,怎样继承传统 SIP 的有状态代理服务器的特征也是需要考虑的。

2 P2P-SIP 网络系统架构

考虑异构 P2P-SIP 网络的互通性以及继承传统 SIP 有状态特征,本文提出如图 1 所示的分层 P2P-SIP 系统架构。多个 P2P-SIP 网络(如 Pastry, CAN 等称为本地 P2P-SIP 网络)采用各自的定位路由算法实现分布式 SIP 机制。为了连接底层异构 P2P-SIP 网络,在其上存在一个公共 P2P-SIP 网络。每个本地 P2P-SIP 网络选择一个或者多个性能高的节点成为公共 P2P-SIP 网络的一员充当网关节点,路由异构 P2P-SIP 网络的消息,称之为 P2P 服务器。如图 1 所示的节点 X 同是本地 CAN P2P-SIP 网络和公共 P2P-SIP 网络的节点,是 P2P 服务器,实现类似网关的功能。

收稿日期:2007-02-13;修回日期:2007-05-28。

作者简介:李桂林(1982-),女,湖南长沙人,硕士研究生,主要研究方向:计算机网络与通信、软件工程;李建华(1963-),男,湖南长沙人,教授,博士研究生,主要研究方向: workflow、企业信息集成、软件工程。

而且,每个 P2P-SIP 网中的某些节点是可以设置、有状态的,传统的有状态代理服务器的可控特征就能很好继承。P2P 服务器是成为这样有状态节点的合适的候选者。

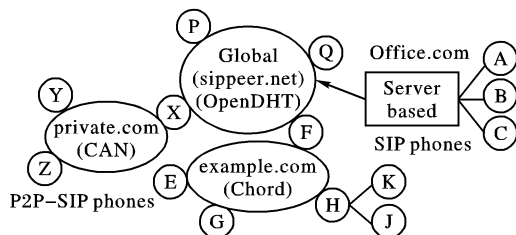


图1 分层 P2P-SIP 系统架构

3 P2P-SIP 节点框架

图2是一个 P2P-SIP 节点的逻辑框图(PO 是 P2P-SIP 网络),存在 Overlay Layer、Protocol Layer 和 NAT Traversal Layer 三个逻辑层和 PO UserAgent、PO Core Logic、Bootstrap 和 Media 逻辑块以及应用程序编程接口 API。

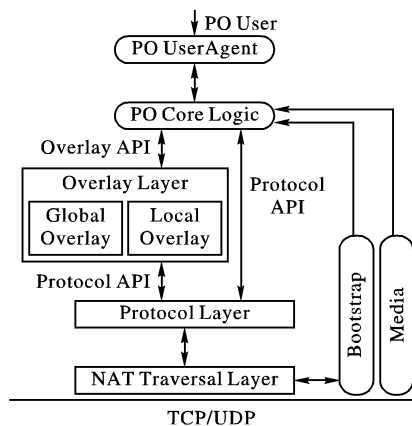


图2 P2P-SIP 节点逻辑框图

Overlay Layer:为了能支持异构 P2P-SIP 网络的节点通信,这一层包括 Local 和 Global Overlay 两部分。当一个节点成为本地 P2P-SIP 网络 SN 时,启动 Local;而在成为 P2P 服务器时启动 Global Overlay。它主要执行三种抽象操作:加入、离开和查找,基于本地和公共 P2P-SIP 网络的具体分布式算法,如 DHT 或者直接的泛洪算法。这一层的另外一个功能是提供证书管理和授权,所以,它同样处理节点和用户资源注册。Overlay Layer 主要包含两类数据类型,一个分布式算法使用和维持的路由表,如某个节点的前导节点、指针路由表和后继节点列表等。另外一种数据类型用来存放和处理 P2P-SIP 网络的用户记录。

Protocol Layer:它完成与信令相关的处理,这一层的信令协议可以是 SIP、扩展支持本地和公共 P2P-SIP 网络的 SIP 协议或者其他的信令协议如 H. 323 等。当节点挂靠完成以后,Protocol Layer 开始启用,执行各节点和资源的注册登记、会话建立和发送即时消息等。

NAT Traversal Layer:它使 Protocol Layer 的信令消息能穿越 NAT 和防火墙。它基于现有的 NAT 穿越机制,如 ICE^[6] 等。

PO User Agent:它包含一个用户接口,一个 P2P-SIP 节点可以包含多个此逻辑块。例如,一个节点是用户代理角色时,这个逻辑块可以提供为好友列表,若它是代理服务器角色时,可以是一个配置工具。

PO Core Logic:它提供具体节点的逻辑控制,因为它置于 PO User Agent 和更低层之间,所以可看作是一种中间件组件。P2P-SIP 节点的不同角色可以是用户代理、代理服务器、重定向服务器、注册服务器和网关。

Bootstrap:当一个新节点想要加入 P2P-SIP 网络时,它负责采用多种方式挂靠一个 PO 节点。

Media:它是独立的一部分,在用户定位、呼叫建立完成以后负责媒体流的传送。

应用编程接口 API:Overlay API 存在于 Overlay Layer 和 PO Core Logic 之间,它的主要目的是提供高逻辑层接入、控制和与 Overlay Layer 通信。Protocol API 交互 Protocol Layer 和 PO Core Logic 以及 Over Layer 之间的信息。它的主要目的是提供 P2P-SIP 网络的编码以及信令消息与信令协议的互相转换。

4 P2P-SIP 消息和节点操作分析

为了描述的方便,下面主要针对信誉度高的某本地 P2P-SIP 网络设计,借鉴 Skype 的设计思想,具有公网 IP 地址、有较高硬件配置和较大网络带宽的主机成为 SN 来履行 ON 代理网关及 SIP 代理服务器的职责,而其他主机则成为 ON。分布式路由算法选择 Chord,仅由 SN 组成,动态分配。而其他本地 P2P-SIP 网络可采用其他的分布式路由算法,公共 P2P-SIP 网络实现机制与本地 P2P-SIP 网络类似。

4.1 SIP 的 P2P 扩展

如图2所示的 P2P-SIP 网络中 protocol layer 所有的消息包括维持 DHT 的消息,查询消息等都使用 SIP 消息。为了能够完成这些操作,根据 SIP 扩展的原则^[10]需要增加部分头域和参数:

1)“P2P”标签:在所有的 P2P 相关的交互信息中的 Require 和 Support 头域需包含这个标签以标志是否对 P2P 操作的支持;

2)Node-ID 和 Resource-ID:分别标识节点和用户 ID 的哈希值,哈希函数采用 SHA-1;

3)overlay:用于区别各个 P2P-SIP 网络拓扑环,例如 global overlay, private.com overlay 等;

4)DHT-NodeID:这个头域包含的是消息发送方的相关信息,DHT-NodeID: <NodeID> <IP> <port>;

5)DHT-Link:这是一个非常重要的头域,扩展它用于传递前导节点、路由表、后继节点的信息,它的格式为:DHT-Link: <type> <depth> <Node-ID> <IP> <Port>;其中, type 取值分别为: P 代表前导节点, S 代表后继节点, F 代表 Finger Table; depth 说明在 Link 代表的相对位置或者深度;

6)alg 参数:说明 ID 是用什么 Hash 算法得到,它用于确保某 P2P-SIP 网络拓扑环的所用 Hash 算法是一致的,这样 ID 才会同一个 Hash 空间;同时它也给选择不同的 Hash 算法留下了空间;

7)user = node; user 参数是 SIP 的 RFC 3261 规范里已有的参数,这里给它加多一个可选的值“node”,这个值用于在 URI 中区分目的节点是 node 还是 phone 等(phone 和用户相关)。

节点和用户标识符用 SIP URIs 表示。URI 的端口部分是根据 RFC 3261 的规定组成,而且必须包含 alg 和 overlay 参

数。对于节点 URI, URIs 的用户信息部分必须是 Node-ID, 是由其 IP 地址和端口哈希的数值; 需包含 user = node URI 参数。而对于用户 URI, 它的用户部分必须是没有经过哈希算法的用户名, 可以包含用户参数例如 user = phone; 必须包含 resource-ID URI 参数。如一个用户名是 bob, 它的 IP 地址是 192. 168. 0. 7, 所在本地 P2P-SIP 网络名是 private. com overlay, 节点和用户经过 rsa-sha1 哈希算法得到 Node-ID 和 resource-ID 分别是 86ff438a22 和 723fedaabl (为了简短都只取了前十位), 那么它们的节点 URI 和用户 URI 分别是 SIP: 86ff438a32@192. 168. 0. 7; user = node; overlay = private. com overlay; alg = rsa-sha1; 和 SIP: bob@192. 168. 0. 7; user = phone; overlay = private. com overlay; alg = rsa-sha1; resource - ID = 723fedaabl。

P2P-SIP 网络的消息实现两种功能: 1) 与传统 SIP 系统中 SIP 消息类似, REGISTER 用来注册用户, INVITE 建立会话等; 2) 在 P2P-SIP 网络中 REGISTER 消息执行 DHT 的操作、维护 DHT 的状态。

4.2 认证

当用户第一次登录 P2P-SIP 系统时, 系统应当确定用户是使用自己的用户 ID 登录, 而不是使用非法取得的其他用户的 ID。由于在 P2P 系统中不存在一直在线的服务器, 可以由负责响应的节点生成一个临时的口令, 并将口令发送到用户的电子邮件信箱中。这样要求用户 ID 是一个合法的电子邮件地址, 用户就可以从邮箱中取到这个 ID 的口令。例如: Alice 使用的用户 ID 为 alice@domain. com, 并且以这个身份发出 REGISTER 消息, 响应节点收到 REGISTER 后, Alice 节点和响应节点之间就可以进行标准的 SIP Digest 验证注册过程。可以设定一定的口令过期时间, 比如一个月。

4.3 注册处理

在 P2P-SIP 网络中, SN 既是它下挂 ON 的代理服务器, 同时还作为 Chord 环中的 SN 负责某些其他节点的寻址 (即存储它们的地址信息)。某个 ON 挂靠的代理服务器和在 Chord 环中为其负责的节点一般不是同一个 SN, SN 为哪些节点负责取决于节点的标识符以及在 Chord 环中相应的哈希映射关系, 某个 SN 所负责的节点信息存放在该 SN 的关键字段中。SN 本身也可以作为用户代理。

节点注册的操作主要有两步, 即 SN 挂靠 (如果节点本身是 SN, 则不需要此步) 和 chord 注册。

挂靠的方式主要有如下三种:

1) 局部范围广播: 考虑到 SIP 定义 224. 0. 1. 75 作为 IPv4 的广播注册地址^[1], 本文考虑较小 TTL 的广播实现局部范围内的节点访问, 从而查找更多的 SN 信息。由于是在局部范围内查找, 所以 SN 的查找不能完全依赖于广播方法;

2) SN 地址缓存, 即在本地保存上次查找的 SN 列表, 如果列表中存在可用的 SN (即 SN 的环境没有任何变化), 将会大大改善注册的效率;

3) 预设固定 SN, 可以在应用程序中 (如 Skype 实现机制) 预设一些固定 SN。如果节点在初始化时使用前面方法都无法找到合适的 SN, 则使用这种方法确保最终找到可用的 SN。新节点在选择合适的 SN 进行挂靠后, 通过该 SN, 按照已有的哈希空间和查找算法找到 Chord 环中为其负责的 SN, 并向它发送 REGISTER 消息进行 Chord 注册。REGISTER 消息头部

的 To 字段被填充为用户标识符, 如 sip: alice@home. com, Request-URI 字段填充为 SN 的 IP 地址及端口号, 如 sip: 192. 2. 1. 2: 5060。

挂靠及注册完成以后, ON 会定期发送 REGISTER 消息给其挂靠的 SN。同时, 所有 SN 也会相互之间以及向它们下挂的 ON 定期发送 REGISTER 消息, 从而监控网络各节点的活动状态。如果 ON 满足 SN 的条件, 还会进行 ON 到 SN 的转变。为了可靠性, 每个 SN 都会有一个后继节点列表, 长度为 $\log(N)$ 。每个 ON 的注册信息都会复制给为其负责 SN 的后继节点列表中的 SN 以备份。

Chord 环中的每一个 SN 都有一个关键字字段, 保存该 SN 所负责 ON 的信息, 它与 ON 的 IP 地址等关联。如图 3 所示的 Chord 环中, 设定 $m = 5$, 即为节点标识符和关键字的位数。节点的数字均是由用户标识符通过 overlay protocol layer 处理后得到的数值。根据 Chord 节点路由表的计算和关键字分配原则, 得到节点 1 和 15 的路由表, 其中 i 表示路由表的第 i 项, start 对应指针表项的起始。succ. 代表第 i 项的后继节点, keys 为该节点的关键字。图 3 中 SN1 为 ON31 负责, 所以它的关键字有 31。

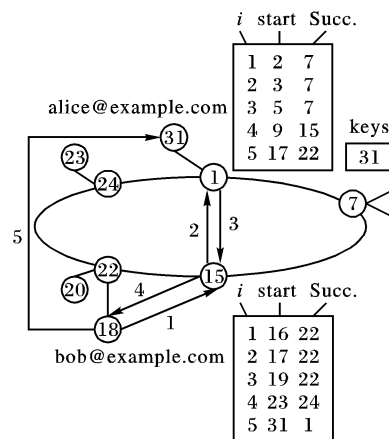


图3 节点指针表及通话建立

4.4 用户查找及连接建立

对于给定的用户标识符, 经过哈希函数运算就可以得到 Chord 环标识符, 即关键字, 根据这个关键字可以找到该用户负责的 SN, 获得该用户的 IP 地址等信息, 然后通过 SIP 协议直接和它建立连接从而建立会话。通话建立以后, 语音数据通过相关的媒体传输协议进行传输。

在图示 3 中, 以用户 bob 查找 Alice 为例, 例如 bob 和 Alice 所在的节点分别是 ON18 和 ON31, 在 Chord 环中, 为它们负责的节点分别是 SN22 和 SN1, 而为它们提供挂靠的节点分别是 SN15 和 SN24。具体查找及连接建立的流程如下:

1) bob 的节点 ON18 将邀请 Alice 进行通话的 SIP INVITE 消息发送给其挂靠的节点 SN15, 该消息中带有哈希 Alice 标识符的关键字 31 等信息;

2) SN15 在 Chord 环中找到为关键字 31 负责的节点 SN1, 并向其发出请求;

3) SN1 返回 ON31 的信息以及其挂靠的 SN24 的信息给 SN15, 其中包括 IP 地址和 NAT 及防火墙类型等信息;

4) SN15 将得到的信息返回给 SN18;

5) ON18 根据这些信息直接和 ON31 建立连接, 进行通信。

4.5 节点退出

ON 退出网络时,首先要向其挂靠的 SN 发送注销消息 (REGISTER 消息设置 expires = 0), 然后其挂靠的 SN 会把这个消息转发给 Chord 环中为该 ON 负责的 SN, 为其负责的 SN 将该 ON 的信息从指针表中删除并更新 Chord 消息, 最后其挂靠的 SN 对该节点进行注销。ON 的退出不会影响整个网络系统。

SN 退出网络时需要更新其下挂节点以及 Chord 环中所有相关 SN 的信息,需完成以下三个工作:

1) 将其下挂的所有 ON 转发给其他 SN 来代理,从而保证这些 ON 的正常通信。SN 只需为其下挂的 ON 向 Chord 环中其他 SN 发送 REGISTER 消息即可。一旦某个 SN 接受了 ON, SN 和 ON 之间将会定期相互发送 REGISTER 消息更新节点状态;

2) 将其负责的所有节点信息转给相关 SN, 即关键字的转移;

3) 更新所有与该 SN 相关联的其他 SN 信息,即将所有指针表包含有指向该 SN 的指针的节点都必须把该 SN 替换成 SN 的后继节点。

4.6 P2P-SIP 网络与传统 SIP 网络的互通

参考文献[1]的外拨代理服务器,传统 SIP 网络的一个 SIP 用户可用公共 P2P-SIP 网络的某个 P2P-SIP 节点作为外拨代理服务器,参与到 P2P-SIP 网络中;也可以根据文献[11],在不改变现有 SIP 系统的条件下,利用 SIPpeer^[11]转发 P2P-SIP 和传统 SIP 系统的请求给 request-URI 而不是进行 DHT 查找。

4.7 P2P-SIP 系统性能分析

由于本框架是基于 Chord 而建立的,因此本框架的性能取决于 Chord 的性能以及 SN 的个数和处理能力。

如果定义 Chord 环中 SN 的个数为 N , 系统的注册用户个数为 n , 节点标识符和关键字的位数为 m , 每个 SN 负责最多 $k = n/N$ 个节点,即存储最多 $k = n/N$ 个关键字。考虑到 Chord 环中查找一个节点平均需要遍历 $O(\log(N))$ ^[1] 个节点,因此可以得出本模型中 Chord 环更新、用户注册和通话建立等操作的复杂度为 $O(\log(N))$, 而节点的加入或退出等操作的复杂度为 $O(\log^2(N))$ 。

假设 SN 定期发送更新 Chord 节点状态信息的频率为 γ , 注册用户定期发送的注册信息的间隔时间为 t , 通话频率服从平均值为 c 的泊松分布, 节点加入或退出的概率服从平均值为 λ 泊松分布, 则每个 SN 平均发送消息的速率 M 就是 Chord 更新、用户注册、通话建立以及节点加入或退出所需要发送消息的平均速率之和, 即:

$$M = \gamma \log(N) + c \log(N) + k \log(N)/t + \lambda (\log^2 N)/N$$

SN 消息发送速率还受到主机资源的制约,如果 SN 每秒可以处理 R 条消息,考虑到经常加入或退出的节点一般不是 SN, 所以 λ 一定会比较小,因此对于等式 $R = M$, 可以粗略算出 $N_{\max} = 2^{R(\gamma+c)}$ 。如果 SN 每秒可以处理 $R = 10$ 条消息, 节点更新频率以及通话频率为每 60 s 一次, 即 $\gamma = c = 1/60$, 则网络中节点个数最大约为 $2^{10 \times 30}$, 所以本框架具有足够的扩展性。

基于 Chord 的 P2P 技术在 SIP 网络中的应用增加了网络的可伸缩性和可靠性,同时也增加了通话建立的延时,这是由

于 P2P 特性及 Chord 机制本身而造成的。在由 10 000 个节点组成的 Chord 网络中,平均查找路径长度为 6^[3], 因此, P2P 网络中由于用户查找的时间大约是传统 C/S 网络所需时间的 6 倍。

实践证明, Skype 应用中, 在网络状态良好的情况下, 从用户拨号到响铃的时间间隔大约是 2 s, 最差情况下是 3 ~ 8 s^[4]。而本模型采用 DHT 方式进行用户查找, 这个比 Skype 采用的泛洪法更加高效。因此, P2P-SIP 网络中 DHT 查找所造成的延时是可以接受的。

5 结语

提出了一个分层 P2P-SIP 系统架构和 P2P-SIP 节点框架。该系统在 SIP 协议通用性的基础上, 继承了 P2P 系统的可扩展性和可靠性, 而且保证了异构 P2P-SIP 网络的互通性, 另外它可以和传统 SIP 系统协同工作, 这些优点以增加呼叫延迟为代价。用 Chord 的 DHT 作为 P2P-SIP 的底层, 扩展 SIP 消息支持 P2P 的操作, 分析了此 P2P-SIP 节点在注册认证、资源定位等方面的操作。该系统架构对于 P2P-SIP 的标准化工作具有一定的参考价值 and 一定的应用前景。

参考文献:

- [1] ROSENBERG J, SSCHULZRINNE H, CAMANILO G, *et al.* SIP: session initiation protocol, RFC 3261[S]. Internet Engineering Task Force, 2002.
- [2] SCHOLLMEIER R. A definition of Peer-to-Peer networking for the classification of Peer-to-Peer architectures and applications[C]// Proceedings of the First International Conference on Peer-to-Peer Computing (P2P'01). Washington: IEEE Computer Society, 2001: 101 - 102.
- [3] STOICA I, MORRIS R, KARGER D, *et al.* Chord: a scalable Peer-to-Peer lookup service for Internet applications[C]// SIGCOMM. New York: ACM Press, 2001: 149 - 160.
- [4] Skype: Free Internet telephony that just works[EB/OL]. [2007 - 01 - 12]. <http://www.skype.com>.
- [5] BASET S, SCHULZRINNE H. An analysis of the skype Peer-to-Peer Internet telephony protocol, CUCS-039-04[R]. New York: Columbia University, 2004.
- [6] ROSENBERG J. Interactive connectivity establishment (ICE): a methodology for network address translator (NAT) traversal for the session initiation protocol (SIP) [Z]. draft-ietf-mmusic-ice-00 (work in progress), 2003.
- [7] SIPshare: SIP beyond voice and video[EB/OL]. (2004 - 09 - 16) [2007 - 01 - 12]. <http://www.research.earthlink.net/p2p/>.
- [8] BRYAN D, LOWERKAMP B, JENNINGS C. A P2P approach to SIP registration and resource location[Z]. draft-bryan-sipping-p2p-02, IETF, 2006.
- [9] BRYAN D A, LOWERKAMP B B, JENNINGS C. SOSIMPLE: a serverless, standards-based, P2P SIP communication system[C]// First International Workshop on Advanced Architectures and Algorithms for Internet Delivery and Applications (AAA-IDEA). [S. l.]: IEEE Press, 2005: 42 - 49.
- [10] CAMARILLO G. SIP 解密[M]. 北京: 人民邮电出版社, 2003: 117 - 138.
- [11] SINGH K, SCHULZRINNE H. SIPpeer: a session initiation protocol (SIP)-based Peer-to-Peer Internet telephony client adaptor[R]. New York: Columbia University, 2005.