

IMS 中 SIP 会话建立时延的研究

贾 宏, 万晓榆, 樊自甫

(重庆邮电大学 下一代网络应用技术研究, 重庆 400065)

摘 要 :为了在 3G 核心网中支持多媒体会话, 3GPP 在 R5 中引入了 IP 多媒体子系统(IMS), 并选择 SIP 作为 IMS 中的呼叫/会话的控制信令。IMS 可以提供多媒体业务和语音业务, 对 QoS 提出了更高的要求。主要讨论 IMS QoS 的一个重要参数-SIP 会话建立时延。分析了 SIP 会话建立时延模型和相关因素并从 3 方面提出了优化措施, SIP 的传输协议选择, 信令重传机制的改进以及对 SIP 信令的压缩, 最后进行了性能仿真。

关键词 :IP 多媒体子系统; 会话初始化协议; 服务质量; 会话建立时延; 定时器; 信令压缩

中图分类号 :TN915 文献标识码 :A 文章编号 :1000-7024 (2008) 07-1853-02

Research of SIP call setup delay in IMS

JIA Hong, WAN Xiao-yu, FAN Zi-fu

(Next Generation Network Application Technology Institute, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract : For supporting multimedia call in 3G core network, IMS is introduced into R5, and SIP is chosen as the call/session control signaling by 3GPP. IMS can provide multimedia and audio services, its QoS should be better. An important QoS parameter of IMS-SIP call setup delay is researched, the model of call setup delay and the correlative factors are analyzed. Measures of optimization in three ways is put forward, including transmission protocol, retransmission timer and signaling compression. At last, performance simulation is made.

Key words : IMS; SIP; QoS; call setup delay; timer; signaling compression

0 引 言

IMS(IP multimedia subsystem)是 3GPP 在 R5 版本的核心网中引入的一个多媒体子系统, 3GPP2 完全采用了 IMS 作为 CDMA2000 实现 IP 多媒体业务的网络框架, 而 ETSI TISPAN 在其 NGN 网络中也重用了 IMS 架构。IMS 可以向用户提供综合的语音、数据和多媒体等业务, 因此需要提供一定的 QoS 保障机制。呼叫双方信令交互的时延是 QoS 的一个重要参数。SIP 作为 IMS 的呼叫控制信令, 由于其自身的特点以及无线信道带宽有限, 使得 IMS 中基于 SIP 的会话建立的时延太大, 因此有必要进行优化。

1 IMS 中 SIP 会话建立

根据文献[4], 在一个 IMS 域中建立一个 SIP 会话, 至少需要 3 次请求响应, 共 6 条消息交互。如图 1 所示。

用户代理客户端(UAC)通过发送 INVITE 消息来呼叫用户代理服务器端(UAS), 双方通过几条 SIP 消息交互后完成媒体的协商, 随后进行资源预留。UAC 资源预留成功后, 发送

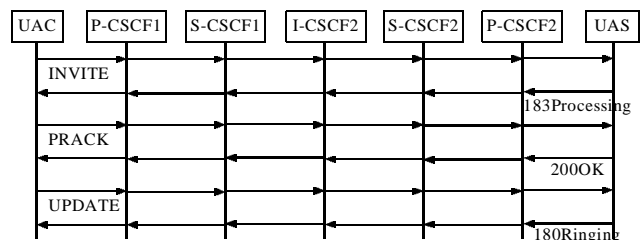


图 1 IMS 中 SIP 会话建立

UPDATE 消息通知 UAS, 如果这时 UAS 也完成资源预留, 则返回 180 振铃消息, 一个 SIP 会话就建立起来。

2 SIP 会话建立时延模型

本文讨论的是 SIP 会话建立时延, 指从 UAC 发起呼叫开始, 到收到 UAS 回铃音为止的一段时间。时延主要分两部分: 网络传输时延, 包括传输和重传时延; 网络中节点处理时延, 包括排队时延, HSS 和 DNS 查询时延等。其中最重要的是网络传输时延, 受底层的传输协议和纠错机制等影响。

收稿日期: 2007-04-10 E-mail: jhmaomao_@163.com

基金项目: 重庆市教委科学技术研究基金项目 (KJ070515、KJ070512)。

作者简介: 贾宏 (1979 -), 女, 四川南充人, 硕士研究生, 研究方向为下一代网络技术; 万晓榆 (1963 -), 男, 博士, 教授, 研究方向为下一代网络技术; 樊自甫 (1977 -), 男, 硕士, 讲师, 研究方向为下一代网络技术。

2.1 网络传输时延

根据文献[1], SIP网络传输时延分为:无线侧的发送时延,端到端的传输时延,数据帧丢失而导致的重传时延。传输时延模型主要建立在以下几个前提的基础上:

- (1)和空中信道的不可靠性相比,忽略INTERNTE的误码率;
- (2) SIP消息重传概率和无线信道 FER成线性关系,并且不考虑应答消息出错导致的SIP重传;
- (3)采用无线接入方式,所用的无线信道参数参考CDMA接入网。

如果一条SIP消息包含 k 个帧,则该消息的传输时延为

$$T = D + (k-1)\tau$$

式中: D ——INTERNTE上的SIP消息传输时延, τ ——一个无线帧的周期。

由于无线信道的不可靠性,数据帧丢失会导致SIP消息的重传。根据文献[5],第 i 次重传的SIP定时器大小 $Tr(i)$ 和定时器的初值 $Tr(1)$ 是2的指数关系

$$Tr(i) = 2^{i-1} * Tr(1)$$

式中: $Tr(1)$ 通常是对RTT(round trip time)的一个估计值,默认是500 ms^[5]。则第 i 次重传的时延计算如下

$$Tr(i) = T + Tr(i)$$

假设 P 是SIP消息重传的概率, FER 是无线信道的误帧率,则

$$P = k * FER$$

根据以上分析,一个SIP消息的平均传输时延为

$$T_{transm} = T + \sum_{i=1}^m P^i * Tr(i)$$

根据文献[5],SIP重传次数 i 最大为7。

2.2 SIP节点处理时延

文献[2]、[3]讨论了SIP节点处理时延。用户终端UE和CSCF服务器都可看成是一个SIP节点。节点处理时延包括排队时延,HSS查询时延,DNS查询时延等,主要跟SIP节点的排队算法,以及DNS和数据库所处的位置等有关系。相对于传输时延,SIP节点处理时延对整个会话建立时延的影响较小。

2.3 SIP会话建立时延模型

根据以上分析,IMS中建立一个SIP会话的总时延如下

$$T_{session} = \sum_{j=1}^n T_{transm}(j) + \sum_{j=1}^n T_{processing}(j)$$

式中: n ——IMS中建立一个SIP会话所需要的消息条数, $T_{processing}$ ——一条SIP消息在传送过程中的节点处理时延。

3 SIP会话建立时延的分析和优化

通过前面模型的分析可以看出SIP会话建立时延随着FER增长以指数规律增长,交换的消息的数量、消息长度以及重传机制对于会话建立的时延有很大的影响,这些方面的优化将可以缩短会话建立时延。通过分析,本文提出了以下措施来减小时延。

3.1 采用UDP作为SIP的承载层协议

SIP可以采用TCP或UDP作为底层的传输协议^[5]。如果采用TCP承载,则建立一个SIP会话首先要进行TCP三次握手,至少增加3条TCP消息的交互时间,时延模型中 n 值将增大,影响SIP会话建立的网络传输时延。如果采用UDP承载,

UDP是无连接的协议,所以不用增加额外开销。因此采用UDP协议能够减少交换的消息的数量,优化SIP会话建立时间。

3.2 采用自适应的重传定时器

从上述的数学模型可以看出当FER较差时,重传时延是整个SIP会话建立时延的重要组成部分,因此考虑消息重传机制的改进以减少平均会话建立时延。定时器初值 $Tr(1)$ 是个很关键的参数,如果太短,那么反馈的应答消息还在半路上,消息就开始重传了;如果太长,那么当丢失出现的时候,会增加不必要的会话建立时间。目前SIP中的 $Tr(1)$ 采用的是固定的时间间隔,无法根据当前的消息的传输时间做出适当的调节。为了减少重传,本文不采用固定定时器,而是将SIP呼叫与无线信道的性能结合起来,采用一种自适应的重传定时器。UAC的 $Tr(1)$ 定义如下

$$Tr(1) = D + (k_1 - 1) * \tau + D + (k_2 - 1) * \tau + T_{processing}$$

采用和UAC相同的方法定义UAS的 $Tr(1)$ 如下

$$Tr(1) = D + (k_2 - 1) * \tau + D + (k_3 - 1) * \tau + T_{processing}$$

式中: k_1 ——INVITE请求消息包含的无线帧数, k_2 ——183应答消息包含的帧数。 k_3 ——PRACK消息包含的帧数。

3.3 采用SIP压缩机制

采用SIP压缩机制主要是通过改变SIP消息的长度来降低时延。典型的SIP消息的大小由几百到几千字节,为了适合在窄带无线信道上传输,IMS对SIP进行了扩展,支持SIP消息的压缩^[6]。当无线信道一定时,一条SIP消息所含帧数 k 仅取决于消息大小。从时延模型可以看出, k 不仅影响SIP消息传输时延,还影响SIP重传的概率,对自适应的定时器来说, k 还成了影响定时器初值的关键因素。

从时延模型看出,当无线信道FER增长时,SIP重传概率也随之增加。而重传定时是成指数增长,增加的重传时延将严重影响整个传输时延。因此考虑采用SIP信令压缩算法来减小消息尺寸,这样既可以减小SIP重传概率,又可以减小定时器时钟。经分析,文本压缩算法LZSS适合SIP压缩,为了提高压缩比,再对其输出进行HUFFMAN编码。

4 性能仿真

本文采用UDP承载以及自适应定时器,验证了当FER增大时,SIP压缩对会话建立时延的影响。根据文献[2], D 和 τ 分别取值100 ms和20 ms,考虑了对DNS和HSS时延的优化措施后, $T_{processing}$ 取平均值210 ms^[3]。表1给出了压缩前后UDP消息长度以及所包含的无线帧数,图2给出了仿真结果。

仿真结果表明:采用SIP压缩机制能有效的降低会话建

表1 消息长度以及帧的数量

SIP消息	消息长度/ bytes	帧的数量 (19.2kbps)	压缩后消息 长度/bytes	压缩后帧的数 量(19.2kbps)
INVITE	953	24	523	14
183 Processing	1 148	29	197	5
PRACK	1 333	34	276	7
200OK	926	24	55	2
UPDATE	1 114	29	82	3
180 Ringing	490	13	69	2

(下转第1858页)

窗口、属性窗口、工作空间管理窗口及鹰眼窗口等组成。系统提供了空间数据和属性数据的管理、显示、查询、分析、更新和信息共享等功能,可满足校园规划和管理的基本需求,基本上实现了校园管理的科学化、网络化和智能化。

3.4 系统网络发布

本系统使用 ArcIMS 建立发布站点^[7-8],选择 Html Viewer 二次开发定制方式,提供 WebGIS 服务。通过创建地图服务,组织需要发布的数据及设置地图的符号、注释、比例尺等数据表现形式。在 Dreamweaver 中设计网站,提供图层显示、地图浏览、距离测量及图形数据和属性数据的多种方式查询等功能,以矢量图片的方式将校园地图发送到客户端,允许用户在客户端使用高级功能改变地图的表现形式,如为地图要素加标注,选择空间要素等。客户端浏览界面如图 5 所示。

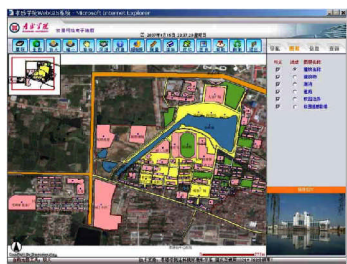


图 5 系统网络发布界面

在资源共享的同时,网络安全问题至关重要,为保证数据安全,系统设有管理员负责数据维护 and 用户授权,对一定范围内的用户开放系统。同时对数据进行分类管理,一般用户限定浏览校园平面数据;对于规划、管线等数据,用户只有申请注册,经审查确认身份并授权后,才能登录系统,在权限许可范围内浏览相关信息。

(上接第 1854 页)

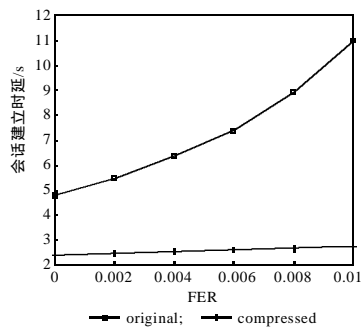


图 2 SIP 压缩前后的会话建立时延比较

立时延。如果不采用压缩,当 FER 增大时,会话建立时延将随着 FER 增加而快速增加。采用 SIP 压缩算法后,明显的削弱了 FER 对时延的影响。

5 结束语

本文根据 IMS 中的端到端 SIP 会话建立时延的数学模型,分析了影响时延的重要因素,并提出了优化时延的措施。仿真结果表明 SIP 压缩可以明显改善 SIP 会话建立的时延,但是其压缩/解压对无线信道误码率的影响还有待进一步的研究。为了

4 结束语

通过对校园地理信息系统的开发以及应用,可得出如下结论:利用组件技术开发地理信息系统,既可利用 MapObjects 强大的地图功能,又可充分发挥可视化开发语言 VB 简单、高效的特点。它不仅缩短了开发周期、降低了开发成本,而且所开发的系统具有可靠性好、便于维护等优点,是一种比较成功的开发模式。今后,系统应进一步开发与完善在校园规划、专业功能等方面的应用,同时还应在数据的更新、备份及维护等方面进行更为深入的研究,使系统的空间数据和属性数据具有时效性,使系统更加适应实际管理的需要。

参考文献:

[1] 唐伟奇.校园地理信息系统的开发与实现[J].科学技术与工程,2006(8):1102-1105.
[2] 游晓明,刘升.基于 GIS 的数字校园教务管理可视化方法研究[J].微电子学与计算机,2004,21(8):36-38.
[3] 游晓明,帅典勋,刘升.基于 GIS 的数字校园教学管理系统的设计与集成[J].计算机工程与设计,2005,26(11):2895-2897.
[4] 姜真杰,方陆明,寿韬.GIS 技术在数字校园建设中的应用[J].计算机工程与设计,2003,24(8):60-62.
[5] 李永红,邓红艳,赵敬东,等.组件式 GIS 开发的实践[J].计算机工程与设计,2005,26(4):1090-1092.
[6] 薛伟.MapObjects:地理信息系统程序设计[M].北京:国防工业出版社,2004:212-213.
[7] 赵朋,刘刚,张喜杰,等.一种基于 ArcIMS 开发 WebGIS 的方法及应用[J].计算机应用研究,2005(5):258-260.
[8] 唐达,杨宇曦,李璞.综合城市环保信息系统的设计与实现[J].计算机工程与设计,2006,27(23):4575-4578.

降低时延,还可以采用信道编码等纠错机制来减少 SIP 的重传,或者采用无线帧的重传而不是 SIP 重传来减少重传的信息量。

参考文献:

[1] Dai Bin, Wang Fufong, Ke Jinshui. Performance analysis of signaling using SigComp scheme in narrowband system[J]. IEEE CCNC, 2006, 1:376-379.
[2] Hanane Fathi, Shyam S Chakraborty, Ramjee Prasad. Optimization of SIP session setup delay for VoIP in 3G wireless networks[J]. Mobile Computing, 2006, 5(9):1121-1132.
[3] Kist A A, Harris R J. SIP signalling delay in 3GPP[C]. Fremantle WA: Sixth International Symposium on Communications Interworking of IFIP - Interworking, 2002:211-222.
[4] 3GPP TS 24.228, v.5.3.0, Signalling flows for the IP multimedia call control based on SIP and SDP - Stage 3[S].
[5] RFC 3261, SIP: Session initiation protocol[S].
[6] RFC 3320, Signaling compression (SigComp)[S].
[7] 3GPP TS 23.228 v.5.4.1, IP multimedia subsystem(IMS)-stage2[S].
[8] 周大刚,阮露,龙昭华.基于 SIP 的移动性管理[J].计算机工程与设计,2005,26(11):2937-2941.