文章编号:1000－2367(2015)06－0141－05 DOI:10.16366/j.cnki.1000－2367.2015.06.026

一种基于应用层的安全保密 IP 电话 QoS 保障方案及实现

王 杰,郑小东,薄树奎

(郑州航空工业管理学院 计算机科学与应用系,郑州 450015)

摘 要:提出了一种基于应用层的安全保密 IP 电话 QoS 保障方案并实现.在语音加解密过程中,通过多缓冲区的合理设置和加密模式的设计,降低了时延和带宽.针对网络丢包问题,提出了一种改进的错误隐藏技术,并与前向纠错技术集成应用,较好地完成了丢失包的恢复.采用自适应播放延迟方法,解决了时延抖动所导致的语音播放不流畅问题.在模拟发生包丢失和包错序的网络环境中进行了测试,IP 电话在满足安全保密通信的前提下,时延、带宽和丢包率等指标都达到了商用电话的标准,具有较强的实用性.

关键词:安全保密;IP 电话;QoS;错误隐藏;前向纠错;时延;带宽;丢包率

中图分类号:TP393 **文献标志码**:A

IP 电话也称 VoIP(Voice over IP),它利用 IP 网络传递语音数据信息,可以实现点对点的语音通信,还能以较低代价实现电话会议功能,极大地满足了人们即时通话的愿望[1].但 VoIP 采用开放性的 IP 网络进行数据传输,因此在安全通信和服务质量(QoS,Quality of Service)方面难以得到可靠保证[2].

现阶段保证 IP 电话语音质量的方法大致可分为两类:基于 IP 层的 QoS 机制和基于应用层的各类 QoS 技术.基于 IP 层的 QoS 机制可以提供更可靠的语音质量保证,但由于目前网络环境十分复杂,基于 IP 层的 QoS 机制还没有形成统一的标准,且实现需要很高的费用.而随着网络技术的发展,基于应用层的各类 QoS 保证技术的性能也不断提高,逐步能满足 IP 电话的通话质量要求[3].由于 IP 电话的 2 个主要协议 SIP 协议和 H.363 协议已经设计了 VoIP 框架和安全机制,较多研究工作围绕着这两个协议来展开[4].但在实现 VoIP 安全机制的同时对服务质量保障方面研究较少.

本文通过对 IP 电话保密通信过程中时延、带宽、丢包率等影响语音质量的因素分析,通过选择合理的加密模式和设置合理的缓冲区,来实现语音加解密过程中时延、带宽等服务质量控制.提出了一种改进的错误隐藏技术(EC,Error concealment),并结合一种基于校验的前向纠错技术(FEC,Forward Error Correction)较好地解决了网络丢包率过多的问题,保证了 IP 电话安全保密通信中的服务质量.

1 常用 QoS 保障技术

1.1 错误隐藏技术

网络协议通常可以准确提供包丢失位置,接收端的缓冲区也存储有一定量的丢失位置前后的语音数据.EC 技术则利用这些数据来构造丢失包语音数据的替代数据.常用的错误隐藏技术可分为三类:插入技术、插值技术和重建技术[5].

1.2 前向纠错技术

前向纠错技术是一种冗余编码技术,它依靠额外传送的冗余编码数据来修复或弥补因网络包丢失给接收端带来的语音数据损失[6].前向纠错编码技术在当前发送的包中携带更早些的包中的信息,当一个包丢失

收稿日期:2014-11-02;**修回日期**:2015-06-11.

基金项目:国家自然科学基金(41001235);河南省教育厅科学技术研究重点项目(12B520061);郑州市科技局科技发展计划项目(20120432).

第 1 作者简介(通信作者):王 杰(1978－),男,河南潢川人,郑州航空工业管理学院副教授,主要从事计算机网络、信号处理方面的研究,E-mail:wangjiew@126.com.

时,可以从随后的包中利用冗余数据重构丢失的包中语音数据[7].这种方法没有考虑语音信号的特性,相对增加了传输的数据量,不过这一方案对突发的包丢失是很有用的.FEC 可以吸收大部分的丢包率,保持高音质.但是采用 FEC 技术需要占用一定的网络带宽.

2 基于应用层的安全保密 IP 电话总体结构

2.1 安全保密 IP 电话通信流程

安全保密 IP 电话的通信流程如图 1 所示.在发送端采用静音检测技术,分别生成语音帧和静音帧,并采用 FEC 技术,生成校验数据,然后将语音数据和校验数据一起封装,生成数据包发送.在接收端设置两个足够大的缓冲区队列,分别存储数据包中的语音数据和校验数据.接收端首先采用 FEC 技术,利用收到的包中的语音数据和校验数据,尽量恢复被判决为丢失包中的语音数据.对于 FEC 技术无法恢复的数据,则采用错误隐藏技术.在相应的缓冲区中生成丢失的包中语音数据的替代数据,尽量降低语音数据的丢失对语音总体质量的影响.最后接收端以稳定平滑的速率,依次将语音数据从缓冲区队列中取出,进行语音解码或者静音恢复后播放.

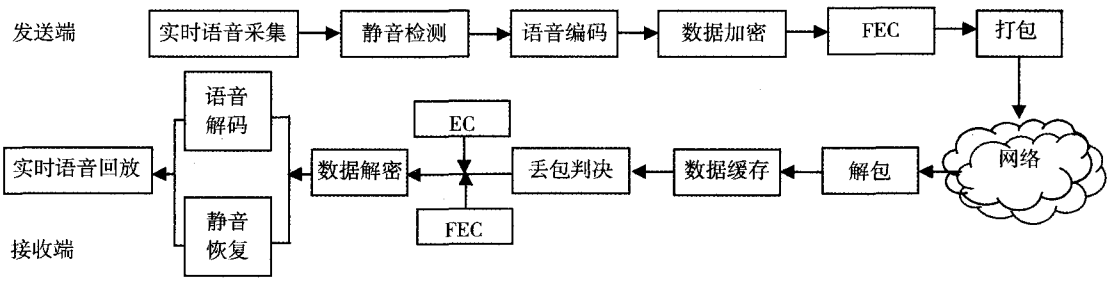

图1 安全保密IP电话通信流程

2.2 QoS 控制涉及的主要问题

2.2.1 语音质量实时检测

网络电话的语音数据传输是基于不可靠的、无连接传输协议 UDP,它不提供包丢失的重传机制,因此不能解决包丢失问题.实时传输协议 RTP 的引入便是基于这样的考虑.其能够为网络电话等实时应用提供了在一定范围内调整业务服务质量的依据.

2.2.2 包丢失的合理判决

对于网络电话等实时应用来说,因拥塞而被网络丢弃或没有按时到达接收端缓冲区的包,都被视为丢失.由于网络电话对实时性要求比较严格,因此采用时延特性好的数据恢复技术:前向纠错编码技术和错误隐藏技术.

2.2.3 接收端缓冲区的设置

不同的语音包在网络传输过程中时延的大小不同,包到达终点很难达到匀速,顺序也有可能颠倒,甚至发生丢包现象.如果时延抖动较大会造成语音播放的断续及部分失真.有效解决时延抖动的方法是在接收端设立数据缓冲区.将语音数据在缓冲区暂存一段时间,随后以稳定平滑的速率将语音数据从缓冲区中取出播放.

3 基于应用层的保密 IP 电话 QoS 控制

3.1 语音加解密过程的 QoS 控制

如果对采集到的每帧数据采用同种方式进行压缩处理,会大量增加带宽的占用.因此,对所采集到的语音帧首先进行静音检测,对静音的语音帧采用低比特的静音压缩,从而减少带宽的占用.

采集到的语音数据首先存放在缓冲区中,然后取出进行加解密操作.为了增加处理速度,减少时延,采用了多线程技术.把密钥序列存放在多个缓冲区中依次取出,再采用不同线程分别处理语音数据的更新、密钥

序列的更新和加解密,从而达到减少时延的目的.

分组密码的工作模式常见的是计数器模式和 ECB 模式. 8 bit 计数器模式解密不会出现错误扩散,而在 ECB 模式下,1 bit 传输错误会造成 1 个分组解密错误. 因此,ECB 模式适用于在低错误传输率的网络中提供服务,8 bit 计数器模式适用于在传输错误率较高的网络中服务.

3.2　错误隐藏技术在 QoS 控制中的应用

传统的错误隐藏算法中判决是否发生了丢包的依据是:对于一个语音包 P,如果其后续包中有 K 个语音包提前到达,那么接收端将判定 P 丢失. 这样虽然能有效处理少量丢包和包错序问题,但是当网络丢包率较高,连续丢失多个包发生概率较大时,语音播放的均匀性、连续性就会变的相对比较低[8].

对丢包判决条件做出如下改进:接收端设定一个时钟 T_c,并设播放延迟为一个时限 T. 设 t_i 为第 i 个 RTP 包 P_i 的发送时刻,a_i 为第 i 个 RTP 包到达接收端的时刻,T_s 为发送间隔,基本上等于发送端单包语音数据的生成时间. 如果接收端收到的第一个 RTP 包是 P_n,由此可知第一个 RTP 包 P_1 的发送时刻为 $t_n-(n-1)T_s$.

当包 P_n 到达时,接收端就把该包中的语音数据写入其相应的缓冲区 $B'_N(N'=n\%N)$. 时钟设为 $T_c=t_1=t_n-(n-1)T_s$ 并开始计时. 如果时钟到达时刻 t_1+T,则对缓冲区队列进行循环播放,查看第一个缓冲区是否有需要播放的数据,有则播放. 如果没有,则采用错误隐藏技术进行处理,生成丢失语音数据的替代数据播放. 然后查看第 2 个缓冲区是否已经存入,如果有则播放. 如果没有则采用错误隐藏技术生成替代数据播放,然后查看第 3 个缓冲区,依次类推.

3.3　前向纠错技术与错误隐藏技术的集成应用

在恶劣网络条件下,丢包率往往很高,连续丢包的可能性也相应增大. 如果只采用错误隐藏技术,则无法有效降低语音数据的丢失对总体语音质量造成的影响. 因此使用错误隐藏技术与前向纠错技术,使它们共同对发生了丢包的网络环境中的语音质量进行保障[9].

本文采用了基于校验码的 FEC 技术,因此在接收端设置两种缓冲区:校验数据缓冲区和语音数据缓冲区. 校验数据缓冲区用 B'_{N1} 表示,用来存放校验数据,B'_{N1} 的数目与语音数据缓冲区 B'_N 相等.

每个用于存储 RTP 包语音数据的缓冲区对应一个校验数据缓冲区. 接收端收到一个 RTP 包 P_i,解包后把语音数据 P_i 存入相应的语音数据缓冲区 B'_N,同时把校验数据 P'_{i-1} 存入相应的校验数据缓冲区 $B_{N'}$,其中 $N'=i\%N, N_1=(i-1)\%N, N'$ 为接收端设置的每种数据缓冲区的个数.

假设接收端判决单个 RTP 包 P_k 丢失,利用后续的 RTP 包 P_{k+1} 中校验数据 P'_K,与前一个 RTP 包 P_{k-1} 中的语音数据恢复 P_k 中的语音数据. 假设接收端判决连续两个 RTP 包 P_k 和 P_{k-1} 丢失,则利用后续两个 RTP 包 P_{k+2}, P_{k+3} 中的数据恢复 P_k, P_{k+1}. 假设接收端判决连续 3 个 RTP 包 P_k, P_{k+1}, P_{k+2} 丢失,则首先采用重复技术用 P_{k-1} 作为 P_k 的替代数据,然后利用后续的 P_{k+3}, P_{k+4} 对前面的 P_{k+1}, P_{k+2} 中语音数据进行恢复. 对于无法用 FEC 技术恢复的数据,则采用重复技术生成该数据的替代数据.

由于要用FEC技术处理RTP包 P_k 和 P_{k+1} 的丢失事件,就必须至少设置4个语音数据缓冲区来存储 P_k, $P_{k+1}, P_{k+2}, P_{k+3}$,相应的校验数据缓冲区也必须至少有 4 个. 在这里两种缓冲区各设置了 8 个.

接收端恢复语音数据流程如图 2 所示. 图中阴影框表示包 $P_{k-1}, P_{k+1}, P_{k+2}, P_{k+3}$ 被判决为丢失,其相应的缓冲区没有新数据写入. 接收端采用 FEC 技术恢复丢失包中的语音数据 $P_{k-1}, P_{k+2}, P_{k+3}$,而对于其无法恢复的数据 P_{k+1} 则采用重复技术,用 P_k 作为 P_{k+1} 的替代数据.

3.4　自适应播放延迟的实现

网络中时延抖动是普遍存在的. 接收端通过设置缓冲区,将收到的语音数据暂存一段时间以后播放,以便抑制时延抖动对播放语音音质的影响. 播放延迟指语音数据的采集与播放之间的时延[10].

当在播放端采用与发送端同样的包间隔时,同一个语音段 k 的所有 n_k 个语音包将具有同样的播放延迟,假设其为 $p_k(A)$,这样对同一个语音段的数据包,其播放时间为:

$$p_k^i(A)=t_k^i+p_k(A).$$

播放延迟调整算法的目标就是寻找合适的算法 A 估计 p_k,使得相应的 $p_k(A)$ 满足适当的折衷. 由于实时应用的需要,只能从语音段 $k-1$ 的网络时延数据来估计语音段 k 所需要的 $p_k(A)$.

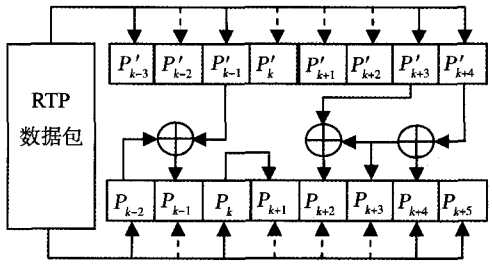

图2 接收端恢复语音流程

其估计式如下：

$$p_k = u_{k-1}^{n_{k-1}} + \beta * v_{k-1}^{n_{k-1}},$$

其中 u_k^i 为语音段 k 的第 i 个语音包的时延估计. v_k^i 为语音段 k 的第 i 个语音包的时延抖动估计. β 可以控制时延和丢包之间的折中.

语音段 k 的第 i 个语音包的传输时延 d_k^i 定义为：

$$d_k^i = a_k^i - t_k^i.$$

由传输时延 d_k^i 来估计 u_k^i 和 v_k^i. 指数加权平均算法的 u_k^i 和 v_k^i 的估计式如下

$$u_k^i = \alpha * u_k^{i-1} + (1-\alpha) d_k^i,$$

$$v_k^i = \alpha * v_k^{i-1} + (1-\alpha) \mid u_k^i - d_k^i \mid,$$

其中：$u_k^0 = u_{k-1}^{n_{k-1}}$，$u_1^1 = d_1^1$.

此算法对平均延迟和时延抖动都采取了指数加权平均，典型的 $\alpha = 0.997\ 800\ 1$.

为了实现语音的连续性，降低延迟抖动对语音质量的影响，在接收端设置了 8 块 360 字节语音数据缓冲区. 这相当于引入了 3 个 RTP 包中的语音帧长共 135 ms 的播放延迟.

4 测试结果与分析

在校园网上对安全保密 IP 电话进行了测试，两端的主机上分别装有安全保密 IP 电话的终端软件和 sniffer 抓包工具.

安全保密 IP 电话的带宽受 RTP 包的大小和加密模式的影响. 一个 IP 包有 90 ms 的语音信息，为了恢复丢包而采用错误隐藏技术需要附加冗余信息，另外为了实现语音加密，在计数器模式和 ECB 模式下的 RTP 包大小不一样，再加上 UDP 包头和 IP 包头. 因此在 IP 网中的实际带宽，ECB 模式为 7.6 kb/s，计数器模式为 7.16 kb/s. 目前商用网络电话的带宽仅语音编码最低为 5.3 kb/s，加上其他打包和协议的开销，在 IP 网中至少需要 8.89 kb/s 的带宽，与商用网络电话相比，采用 ECB 模式时，节省带宽为 14.6%，采用计数器模式时，节省带宽为 19.5%.

实验时启动安全保密 IP 电话终端软件一段时间，计算语音分组的时延和丢包率，用 sniffer 软件模拟双向流量，调整发包频率，得出使用 QoS 控制后的统计如表 1 所示.

表 1 QoS 控制的参数统计

发包频率/(ms^{-1})	4	6	8	10	12
丢包率/%	4.3	3.8	3.2	2.1	0.5
时延/ms	320	290	278	269	168

商用网络电话的丢包率要求控制在 5%范围以内，时延控制在 450 ms 以内，从表 1 中测试结果可以看出，两项性能指标均超过了商用电话的要求，提高了系统的实用性.

所设计的自适应播放延迟效果明显，在选择发包间隔为 6 ms，采用 ECB 模式加密时，用专业的语音 MOS 计算器测得改进后的值为 4.3，用计数器模式加密时，测得改进的值为 3.9，语音质量达到了商用电话标准.

5　结束语

本文在分析安全保密 IP 电话通信过程中，影响语音质量的因素进行了分析，并提出了一种基于应用层的 QoS 保障方案并实现，与其他 IP 电话相比，创新点如下：1）在应用层上解决安全保密 IP 电话的 QoS 控制问题；2）在语音加解密过程中，支持两种加密模式，能支持对不同速率和错误率的网络应用；3）对错误隐藏算法进行改进，并结合前向纠错技术，较好地控制了语音丢包率；4）通过自适应播放延迟的实现，保证语音播放的流畅性，提高了 IP 电话的实用性.

所实现的安全保密 IP 电话在校园网络中试运行，其各项性能指标均超过了商用 IP 电话的标准，适用于对语音通信有安全保密需求的应用场合.

参　考　文　献

[1] Nisar K, Hasbullah H, Said A M, et al. Internet Call Delay on Peer to Peer and Phone to Phone VoIP Network[C]. Computer Engineering and Technology ICCET 2009, Singapore, 2009.

[2] Suhansa R, Jerry K. Mead Curriculum and Laboratory Development for IP Telephone: A Survey of Electronics/Networking Technology Program[C]. 2010 3rd IEEE International Conference on Computer Science and Information Technology, Cheng Du, 2010.

[3] Jongkuk L, Kidong N, Daeyoung K. Effect of Network factors on VoIP[C]. 13th International Conference on Advanced Communication Technology, Gangwon-Do, 2011.

[4] 李印清，王　杰，郭秋萍. 端到端 IP 电话的安全保密系统设计与实现[J]. 河南师范大学学报(自然科学版)，2008，36(3)：32-34.

[5] 王　锦，梁　科，李国峰. 基于进程的 SIP 终端的设计[J]. 南开大学学报(自然科学版)，2014(5)：60-64.

[6] 石　婕，仲伟波，葛秀梅. 动态混沌加解密及其在 VoIP 中的应用[J]. 计算机科学，2014，41：268-271.

[7] 杜水明，张　聪，王　赞，等. 一种基于 VoIP 的改进 WSOLA 丢包隐藏算法[J]. 计算机应用与软件，2014(10)：266-268.

[8] 闻英友，罗　铭，赵　宏. VoIP 网络基于签密的安全机制的研究与实现[J]. 通信学报，2010，31(4)：8-14.

[9] 杨明烽. 峰火通信 PTN 全业务 QoS 解决方案浅析[J]. 电信科学，2010，26(9)：57-59.

[10] 崔红霞. IP 网络中 QoS 机制的研究[J]. 网络安全技术与应用，2007，7(4)：13-15.

A Secure IP Phone QoS Guarantee Scheme and Implementation Based on Application Layer

WANG Jie, ZHENG Xiaodong, BO Shukui

(Department of Computer Science and Application, Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015, China)

Abstract: Based on application layer this paper presents a secure IP phone QoS guarantee scheme and implementation. In the voice encryption and decryption process, the design of multi buffer reasonably and encryption mode, reduce the time of delay and bandwidth. In view of the network packet loss problem, an improved EC and FEC are integrated to acomplish the packet loss recovery. The adaptive playout delay method solve nonfluency of playback caused by the voice delay jitter. Secure IP phone is tested in the environment of the network simulated occurrence of packet loss and packet disordering. The results of test show that time delay, bandwidth comsumption and the lost probability of system can satisfy the demand of secure IP phone QoS, and which has a better use.

Keywords: security; IP phone; QoS; EC; FEC; delay; bandwidth; packet loss rate