

嵌入式 Ad-Hoc 网络下的实时语音通信*

杨大力, 谢维波, 杜智敏

(1. 华侨大学 计算机科学与技术学院, 福建 厦门 361021;

2. 华侨大学 厦门软件园嵌入式技术开放实验室, 福建 厦门 361008)

摘要: 综合 Ad-Hoc 网络技术、多媒体开发技术、嵌入式技术和 RTP 实时传输技术, 设计实现了一套在嵌入式 Ad-Hoc 网络环境下的全双工实时语音通信系统。语音编码、抖动处理、静音检测有效地提高了系统的语音传输质量。实验表明, 该系统性能稳定, 适应性强, 有较广泛的应用。

关键词: 嵌入式; Ad-Hoc; 语音通信; RTP

中图分类号: TP391.41

文献标识码: A

文章编号: 1674-7720(2011)23-0065-04

Real-time voice communication for embedded Ad-Hoc environment

Yang Dali, Xie Weibo, Du Zhimin

(1. College of Computer Science & Technology, Huaqiao University, Xiamen 361021, China;

2. Open Laboratory of Embedded Technology, Xiamen Software Park, Huaqiao University, Xiamen 361008, China)

Abstract: Based on the synthesis of Ad-Hoc network technology, multimedia technology, embedded technology and RTP technology, the paper designs and implements a set of a full-duplex real-time voice communication system for embedded Ad-Hoc network environment. Audio coding, eliminating jitter and voice activation detection effectively improve the quality of voice communication. Experimental results show that the system has stable performance, strong adaptability, and has a wide range of applications.

Key words: embedded; Ad-Hoc; voice communication; RTP

无线自组织网络(Ad-Hoc)最早起源于 20 世纪 70 年代美国国防部高级研究署(DARPA)资助的分组无线网络 PRNET(Packet Radio Network), 其最初的动机之一是满足战场生存的军事需求。Ad-Hoc 是一种无中心、自组织、对等式、节点可移动和多跳路由的移动通信网络, 这种网络结构不需要中心节点, 具有良好的自组织能力、抗毁性能、机动性和鲁棒性^[1-2]。无线自组织网络主要应用在军事、民用、治安、救灾等复杂而环境恶劣的场所中。在这些恶劣环境中, 没有铺设通信基础设施或者通信设施遭到了破坏, 迫切需求 Ad-Hoc 网络的投产使用。而实时语音通信作为人与人之间一种最直接的沟通手段, 应用在军事、治安、救灾等领域是必然的需求。本文正是基于这种需求, 设计并实现了一套在嵌入式 Ad-Hoc 网络环境下的实时语音通信系统。本设计涉及 Ad-Hoc 网络、Linux 无线网卡驱动、实时流媒体传输等技术。首

先, Ad-Hoc 网络采用了按需路由中的 AODV 协议^[3-4]来搭建实验的网络环境, 在嵌入式平台(MagicARM2410)上进行了 USB 无线网卡驱动 RT73 的移植, 对网络节点的分布做了合理的布局和连通性测试; 其次, 结合多媒体编程技术和流媒体实时传输技术^[5], 实现了语音采集和实时通信, 对语音通信质量做了分析和改善。与传统的语音通信比较, 本系统具有独特的优势: 不需要铺设基础通信设施, 无中心, 自组织, 可移动, 多跳全双工语音通信, 可以应用在治安、矿井、救灾等领域。

1 整体方案设计

本语音通信系统的移动终端是由广州致远电子有限公司开发的教学实验开发平台 MagicARM2410, 该实验平台采用 ARM920T 内核的 S3C2410A 微处理器, 工作频率高达 203 MHz, 扩展有充足的存储资源, 具有 I²S 数字音频接口和 8 英寸真彩 TFT 液晶屏, 满足系统的需求。网卡选用的是华硕的 USB 无线网卡 WL-167g, 该网卡的芯片是 RT73 芯片, 拥有官网开发的 Linux 驱动, 还

* 基金项目: 厦门市重点产学研项目(厦经技[2009]233-03); 福建省自然科学基金项目(2010J01334)

网络与通信 Network and Communication

有开源社区提供的驱动, 有较多的文档资料可供查询。开发平台的操作系统是 Linux-2.6.14.1, 在该操作系统下对 AODV 路由协议、UDA1341TS 声卡驱动、USB 驱动、RT73 无线网卡驱动进行移植。

语音通信系统终端主要分为音频采集、音频编码、传输、接收、音频解码、音频播放 6 个模块, 如图 1 所示。系统采用多线程技术, 分为音频采集、编码传输、接收解码和音频播放 4 个线程, 收发缓冲使用环形缓冲区, 其作为临界资源, 采集与编码传输竞争一个环形缓冲区, 播放与接收解码竞争一个环形缓冲区^[6], 如图 2 所示。

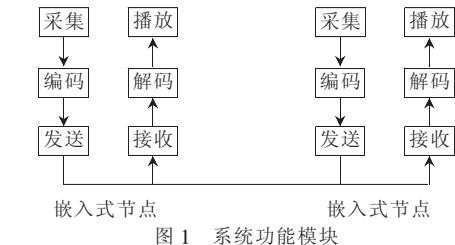


图 1 系统功能模块

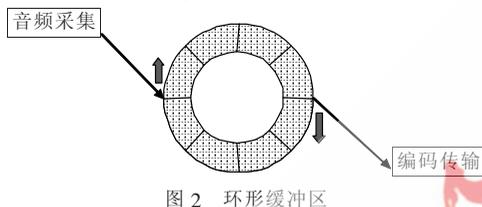


图 2 环形缓冲区

2 系统实现

2.1 搭建 Ad-Hoc 网络环境

本系统使用的路由协议源码是 aodv-uu-0.9.5, 这是一款成熟的 Linux 平台下的 AODV 路由协议实现, 移植到本系统只需要针对操作系统内核做一些相应的修改和参数设置。使用交叉编译工具 arm-linux-gcc3.4.1 对 AODV 源码进行交叉编译, 得到适合 Linux-2.6.14.1 内核的路由协议内核态模块 kaodv 和用户态模块 aodv。为了对网络节点进行合理的拓扑分布, 在室内使用了屏蔽的方案, 如图 3 所示, 进行 4 个节点的路由线路全双工语音通信时, Node1 分别与 Node3、Node4 互相设置 MAC 地址屏蔽, Node2 与 Node4 互相设置 MAC 地址屏蔽, 这样就能保证按照本文设计的路由线路方式进行语音传输。

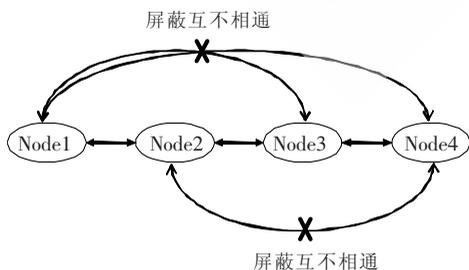


图 3 语音通信方案设计

2.2 音频参数设置

音频信号是一种连续变化的模拟信号, 计算机只能

记录和处理二进制的数字, 自然音源的模拟音频信号必须经过量化成为数字音频信号之后, 才能被计算机进一步处理。采样和量化是音频编程程序员理解声音数字化的两个关键步骤, 进行音频编程时经常用到的技术指标有采用频率、量化位数和声道数。常用的采样频率有 8 kHz、16 kHz、44.1 kHz、48 kHz 等; 常用的量化位数有 8 bit、12 bit、16 bit; 声道有单声道和双声道。本系统使用的是 8 kHz、16 bit、单声道。

//设置采样频率

Arg=采样频率;

Status=ioctl(dsp_fd, SOUND_PCM_WRITE_RATE, &Arg);

//设置量化位数

Arg=量化位数;

Status=ioctl(dsp_fd, SOUND_PCM_WRITE_BITS, &Arg);

//设置声道数

Arg=声道数;

Status=ioctl(dsp_fd, SOUND_PCM_WRITE_CHANNELS, &Arg);

2.3 音频采集

系统使用 read() 函数来从声卡读取数据, 其函数原型为: int read(int fd, char*buf, size_t count)。其中, fd 是音频设备文件的标识符, buf 是存储音频数据的缓冲区, count 是要采集的音频数据字节数。单次采集所用时间为: (采样字节数 count × 8 × 1000) / (采样频率 rate × 量化位数 bits × 声道数 channels) ms。因此, count 不能设置太大, 会增加系统的时延, 影响语音的实时性, 本系统单次采样 128 B, 采样时间为 8 ms。程序伪代码为:

```
while (conditions)
```

```
{
```

```
//访问临界资源环形缓冲区
```

```
Pthread_mutex_lock();
```

```
采集音频数据;
```

```
写入环形缓冲区;
```

```
//增加信号量值
```

```
Sem_Post();
```

```
//退出临界区
```

```
Pthread_mutex_unlock();
```

```
}
```

2.4 编码传输

Ad-Hoc 网络的带宽是有限的, 大量的数据传输会增加网络的负荷, 影响通信质量, 对语音进行适当的编码是很有必要的, 这可以有效地减轻网络的负担, 更利于语音通信。复杂的语音编码算法拥有高效的压缩比, 但由于计算量大, 会增加语音传输的延迟。本系统选用了著名的波形编码算法 G711 编码, G711 编码算法比较简单, 计算量小, 可减少一半的网络语音数据负荷^[7]。语音传输使用了 RTP 实时传输协议, RTP 协议常用于流媒体系统, 可有效监控语音传输质量, 这使它成为 IP 电话产业的技术基础。程序伪代码为:

《微型机与应用》2011 年第 30 卷第 23 期

```

while(conditions)
{
//等待信号量唤醒
Sem_wait();
//访问临界资源环形缓冲区
Pthread_mutex_lock();
从环形缓冲区提取音频数据;
G711 编码语音数据;
发送 RTP 语音数据包;
//退出临界区
Pthread_mutex_unlock();
}

```

2.5 接收解码

系统终端接收到语音数据包时,提取 payload 语音数据,并解码出 PCM 音频码,存放在环形缓冲区。程序伪代码如下:

```

while(conditions)
{
//访问临界资源环形缓冲区
提取 RTP 语音数据包;
解码语音数据;
将语音数据放入环形缓冲区;
//增加信号量值
Sem_post();
//退出临界区
Pthread_mutex_unlock();
}

```

2.6 音频播放

播放线程与接收解码线程竞争同一个环形缓冲区,从环形缓冲区提取音频数据,并写入声卡。程序伪代码如下:

```

While(conditions)
{
//等待信号量唤醒
Sem_wait();
//访问临界资源环形缓冲区
Pthread_mutex_lock();
从环形缓冲区提取音频数据;
播放音频数据;
//退出临界区
Pthread_mutex_unlock();
}

```

3 改善语音通信质量

本设计针对节省网络带宽和消除抖动两方面对语音的通信质量进行了优化。通过实验测试优化前后的语音通信质量,结果证明优化后确实具有更好的效果。下面分别介绍这两种优化方案。

(1)静音检测节省网络带宽。大量的研究表明,在一路上全双工电话交谈中,只有 36%~40%的语音信号是活动或有效的^[8]。电话交谈过程中某一时候往往只有一方

在讲,另一方在听,而且讲话时有大量停顿。如果能通过某种手段检测到通话过程中的静默时段,停止发送语音数据包,则可以节省大量的网络带宽。本设计在静默检测时使用 VAD(Voice Activation Detection)技术,分为短时能量检测和短时过零率检测两部分。短时能量检测可以很好地区分出浊音和静音,而短时过零率可以很好地区分出静音和清音,因此,将两者结合起来,就可以检测出静音段。下面是算法描述:

- ①对输入语音信号进行高通滤波和加窗处理;
- ②计算一帧的能量及其平均值;
- ③该帧平均能量值是否大于能量阈值,是则判定为语音帧并返回,否则转步骤④;
- ④前 3 帧是否有语音帧,有则判定为平滑语音帧并返回,否则转至步骤⑤;
- ⑤计算该帧过零率;
- ⑥该帧过零率是否在过零率阈值范围内,是则判定为语音帧并返回,否则判定为静音帧。

(2)消除抖动。所谓抖动,是指语音数据包没有按规律的时间间隔到达接收端,造成语音播放时声音断断续续。严重的抖动会导致语音数据包不按顺序到达接收端或者丢失,导致声音混乱无序^[9]。抖动是经常发生的,分组到达出现少量时延很常见。为了消除抖动,本文采用了延迟回放技术。如图 4 所示,接收端在收到第一个数据时并不马上开始回放,而是将传入数据放入缓存中,该缓存称为抖动缓存(Jitter Buffer)。在接收端设置一个适当大小的抖动缓存,按照 FIFO 队列进行操作,当语音数据在缓存中积累到一定程度(K 个数据包)时开始按时序播放语音,这样声音就比较平滑连续。

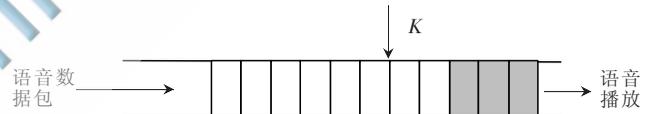


图 4 抖动缓冲区

4 系统测试及结论

本文对搭建的嵌入式 Ad-Hoc 网络环境进行了连通性测试,并在该基础上进行了实时语音通信的实验。网络连通性测试确保各个节点之间能互相通信,保证数据的路由转发路径按照设计的方案进行。实时语音通信实验监测了 2 个节点路由线路、3 个节点路由线路、4 个节点路由线路的语音通信质量,以及使用静音检测技术前后发送的语音数据包占总采样数据包的比例。

(1) 本文在搭建 Ad-Hoc 网络环境中设计了部分节点之间相互屏蔽 MAC 地址以保证路由转发按照设计线路进行的实验方案。下面使用 ping 命令对 4 个嵌入式节点线路的网络进行了数据跟踪测试,如图 5 所示,数据包确实按照所设计的方案进行路由转发。

(2) 分别在 4 个网络节点上移植了本实时语音通信系统,进行全双工语音通信,实验对语音数据包的收发

```

root@ubuntu:~/home/ubuntu# ping 192.168.0.5 -n
PING 192.168.0.5 (192.168.0.5) 56(134) bytes of data:
64 bytes from 192.168.0.5: icmp_seq=0 ttl=62 time=4.17 ms
RR:
  192.168.0.2
  192.168.0.3
  192.168.0.4
  192.168.0.5
  192.168.0.4
  192.168.0.3
  192.168.0.2

```

图5 路由路径测试

情况做了统计,统计结果如表1所示。数据表明,本嵌入式实时语音通信系统在4个节点线路以内的Ad-Hoc网络环境中进行全双工语音通信有较好的效果。

表1 语音数据包丢包率

测试时间/s	发包总数	收包总数	丢包总数	丢包率/%	
2个节点	≈162	20 236	20 084	152	0.75
3个节点	≈180	21 592	21 287	305	1.41
4个节点	≈164	20 416	18 868	1 548	7.58

(3)针对使用静音检测技术前后进行了对比。在系统终端对采集的音频数据包进行静音检测,在一段正常对话时间内分别统计被判定为语音数据包的总数和被判定为静音数据包的总数(如表2所示)。测试数据表明,通过使用静音检测技术确实能节省将近40%的网络带宽。

表2 静音检测技术性能分析

测试时间/s	采样包总数	发送包总数	发送百分比/%	
不使用静音检测技术	60	7 492	7 492	100
使用静音检测技术	60	7 494	4 752	63.41
使用静音检测技术	120	14 986	9 034	60.28
使用静音检测技术	180	22 476	13 422	59.72

本文综合了Ad-Hoc网络技术、Linux系统语音编程、静音检测技术和RTP实时传输技术设计实现了一套应用于嵌入式Ad-Hoc网络的实时语音通信系统。本系统性能良好,有较高的实际应用价值,可以应用在治

安、矿井、救灾等领域中。本文的下一步工作是针对节点的移动性、多节点多路由路径作进一步的改进,不断提高本系统的性能。

参考文献

- [1] 藏婉瑜,于勤,谢立,等.按需式ad hoc移动网络路由协议的研究进展[J].计算机学报,2003,25(10):1009-1017.
- [2] Zhan Yimin, Lin Baocheng. VoIP technology and it's application in broadband IP network[J]. CATV Technology, 2002,9(15):12-18.
- [3] 王雅辉,刘家康,赵胜辉.基于OPNET仿真平台的AODV协议的性能研究[J].光电技术应用,2004(6):60-64.
- [4] 王忠恒,张曦煌.移动Ad Hoc网络AODV路由协议的改进[J].计算机应用,2010(2):333-336.
- [5] 杨静文,卢益民,胡浩.Linux下实时音频传输的实现[J].电声技术,2005(2):63-65.
- [6] 杨水清,张剑,施云飞.ARM嵌入式Linux系统开发技术详解[M].北京:电子工业出版社,2008.
- [7] 焦传斌,于保华,李治柱.新型宽带语音编解码器AMR-WB的研究[J].计算机仿真,2005,22(1):150-159.
- [8] 于洪涛,徐静波,冉崇森.语音噪声压缩新算法[J].计算机工程与科学,2006,28(10):71-13.
- [9] COMER D E, STEVENS D L. Internetworking with TCP/IP volume 3 client-server programming and applications Linux/POSIX sockets version[M]. USA:Pearson Education,2001.

(收稿日期:2011-06-15)

作者简介:

杨大力,男,1986年生,硕士研究生,主要研究方向:嵌入式技术,网络技术。

谢维波,男,1964年生,教授,博士,硕士生导师,主要研究方向:嵌入式技术、数字信号处理。

杜智敏,男,1986年生,硕士研究生,主要研究方向:嵌入式技术,网络技术。