

基于 WiFi 嗅探的感知系统研究

赵飞飞^{1,2} 金彦亮¹ 熊勇²

(1. 上海大学 通信与信息工程学院 上海 200444; 2. 中国科学院上海微系统与

信息技术研究所 无线传感网与通信重点实验室 上海 200050)

摘要: 为了获取某一区域内开启 WiFi 功能的移动终端的数量,设计了一种基于 WiFi 芯片 MT7620A 和嵌入式操作系统 OpenWrt 的移动终端感知系统。在深入地研究了移动终端加入 WiFi 网络流程的基础上,通过选择合适的硬件平台和软件系统,在监听模式下利用抓包工具抓取移动终端发出的嗅探帧,并解析出嗅探帧中的媒体访问控制地址等数据。重点阐述了感知系统的软件实现机制和关键技术,关键技术包括在 U 盘上外扩交换空间以及系统抓取并保存移动终端数据的算法。实验结果表明系统具有使用方便、能持续稳定抓包等特点,具有一定的商业价值。

关键词: 移动终端检测; WiFi; 嗅探帧

中图分类号: TN99 **文献标识码:** A **国家标准学科分类代码:** 510.99

Sensing system based on WiFi probing

Zhao Feifei^{1,2} Jin Yanliang¹ Xiong Yong²

(1. School of Communications and Information Engineering, Shanghai University, Shanghai 200444, China;

2. Key Laboratory of Wireless Sensor Network & Communication, Shanghai Institute of
Microsystem and Information Technology, Shanghai 200050, China)

Abstract: In order to obtain the number of mobile devices WiFi on in some region, a WiFi device sensing system based on MT7620A chipset and embedded operation system OpenWrt was designed. The principle how a mobile device was studied, and then suitable hardware and software platform were chosen. After that, the system was set to monitor mode. With the help of sniff pro, probe requests from mobile devices were captured and media access control addresses were parsed. Implementation mechanism and key technology of the sensing system were elaborated, which include the creation of swap partition on a U disk and the algorithm of how related data captured and saved. The experiment result shows that our system can capture Probe Request stably and is ease to use, and it has some commercial value.

Keywords: mobile device detection; WiFi; probe request

1 引言

随着互联网时代的到来,各种智能终端迅速占领手机市场,从而 WiFi 也就成为了生活中必不可少网络连接工具^[1]。由于在现实生活中,移动终端是人们随身携带的通讯娱乐设备,因此检测某一区域的移动终端数目,可以估计出某一区域的人员密度,这些数据作为互联网时代大数据的一种^[2],在安全预警和疏散机制、商场分析顾客行为和智能家居等方面具有重要的理论意义和使用价值。

通常,为了获得某一区域的移动终端数目,采用的方法主要有蓝牙检测^[3]、移动终端信令检测^[4-5]。蓝牙检测是通过检测终端发出的蓝牙信号来检测移动终端存在与否的,但是蓝牙检测距离比较短,并且人们使用蓝牙功能的频率

越来越低。基于移动终端信令数据进行监测是一种新兴的动态探测技术,它通过分析加密的用户移动终端信令数据来获取某一基站覆盖区域的移动终端数目,但难以区分一个移动终端是处于室内和还是室外。

在 IEEE 802.11 系列协议中,开启了 WiFi 功能的移动终端会定期发送 Probe Request 帧,以主动发现附近的访问接入点,并告知网络自己的存在^[6],该帧中含有终端唯一的媒体访问控制(media access control, MAC)地址。因此,如果能捕获移动终端广播的 Probe Request 帧,并从中解析出它们的 MAC 地址,然后对某一时段内不同 MAC 地址的数目进行统计,即可实现移动终端数目的实时检测。由于嗅探帧含有与用户有关的信息,近年来越来越多的学者也开始从各种角度对其应用进行研究。文献^[7]从收集到的

Probe Request 帧推测人们的社会关系和所用语言等;文献[8]研究了如何通过该帧对移动终端用户进行位置跟踪;还有的文献从隐私的角度研究,提醒人们保护WiFi隐私^[6,9]。这些文献虽然提及了Probe Request 帧的采集过程,但是并没有深入讲解该技术的实现方式。而且它们多是通过装有无线网卡的笔记本电脑,并配合Wireshark 此类数据分析软件来实现嗅探帧的捕获,由于设备昂贵和携带不便,并不适合推广商用。

本文设计并实现了一个WiFi移动终端感知系统,系统硬件部分选择集成了高性能芯片MT7620A的WiFi开发板,降低了系统的研发周期,还保证了系统硬件的质量。

系统软件使用了开源的OpenWrt系统,在编译该系统固件的过程中,可以灵活选用需要集成的组件或驱动,降低设计成本、提高设计效率的同时,也可以实现系统功能的定制开发。

论文剩余部分的组织结构如下:第2部分介绍了系统的嗅探原理;第3部分介绍了基于WiFi的移动终端感知系统的总体架构,并从硬件平台的选型和程序的设计方面做了详细描述;第4部分介绍了系统在办公环境下的实验测试情况;第5部分介绍了系统的应用前景;最后是结论部分。

2 嗅探原理

根据IEEE 802.11协议,无线移动终端要和因特网建立数据链接,就必须接入WiFi网络。WiFi网络的接入过程如图1所示,主要有3个步骤:

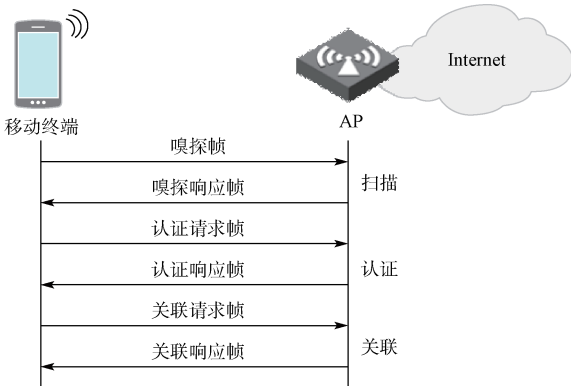


图1 WiFi移动终端接入过程

步骤1:扫描。WiFi移动终端依次在每个信道上广播嗅探帧,侦听到此帧的AP以单播的形式发送嗅探响应帧做出回应,这样移动终端就知道附近存在有哪些AP了。

步骤2:认证。移动终端会根据网络的信号强度等因素,决定加入其中的某一个网络,在成功加入该网络之前,需要进行身份认证。移动终端通过向AP发送认证请求帧申请认证,而AP以认证响应帧做出响应,该响应帧中会说明认证是否成功,并给出原因代码。

步骤3:关联。一旦认证成功,移动终端和AP就开始进行关联,关联是通过关联请求帧和关联响应帧来实现的。只有移动终端和AP关联成功之后,它才可以与AP建立数据链接而加入网络^[10]。

文中只关注WiFi接入过程的扫描阶段,因为支持WiFi的移动终端在开启WiFi功能后,不管它们是否已经接入WiFi网络,都将会继续搜索附近有无WiFi网络存在。由图1知,扫描主要是通过终端广播一种名为“Probe Request”的嗅探帧来实现的,在WiFi终端与AP通信的过程中,存在着3种类型的帧,即管理帧、控制帧和数据帧,Probe Request是一种管理帧。如图2所示,每一个Probe Request帧中都含有移动终端的MAC地址(图2中的Address 2,即嗅探帧的源地址),并且该帧没有加密。因此,可以设计一个感知系统来提捕获该帧,然后从中解析出移动终端的MAC地址等信息,在某一时段,统计出不同移动终端MAC地址的数目,即可得到该区域某一时段的移动终端数目。

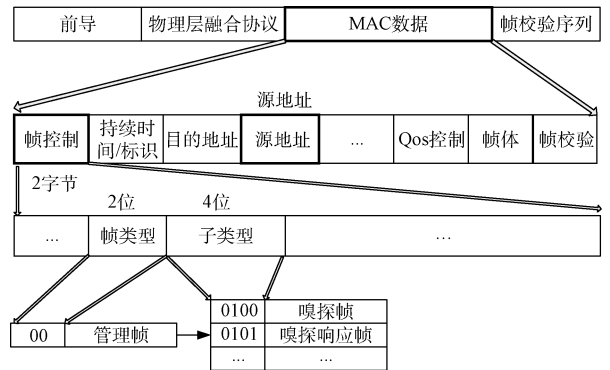


图2 802.11帧格式

MAC地址,又称物理地址和硬件地址。它用十六进制数表示,共6个字节。其中,前3个字节为组织唯一标识符(organizationally unique identifier,OUI),由IEEE负责分配给无线网卡厂商,后3个字节称为扩展标识符,由各厂商自行分配^[11]。如MAC地址为9C:C1:72:CF:84:84,其中前3个字节是IEEE分配给华为公司的,所以MAC地址前3个字节为9C:C1:72的移动终端,均为华为公司产品。

3 移动终端感知系统架构

综上所述,只要通过某种技术手段,捕获Probe Request帧,并解析出该帧中所需的MAC地址,进行统计处理,就可以得到区域内实时移动终端数目。如果每个人携带一部移动终端,甚至可以估算出该区域的人员密度。基于此原理,本文设计了一个基于WiFi的移动终端感知系统,用来捕获移动终端发送的Probe Request帧,然后从中解析出移动终端的MAC地址,对不同的MAC地址进行统计处理,即可实现移动终端数目的实时监测。如图3所示,感知系统主要有3部分组成。

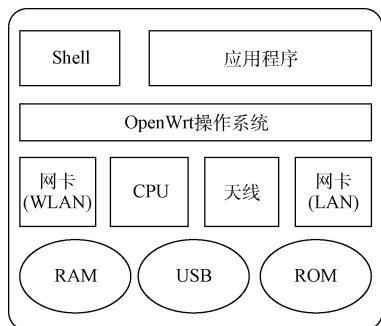


图 3 感知系统架构

1) 硬件平台, 即 WiFi 开发主板, 主板上集成了 CPU、RAM、FLASH ROM、NIC (network interface card, 包含 WLAN 和 LAN) 以及天线等功能模块。CPU 负责数据的处理; RAM 为程序的运行提供内存空间; Flash ROM 由于具有掉电不易失性, 可用于存储代码和保存系统采集的数据; NIC 有两种, WLAN NIC 负责将主板接入以太网, LAN NIC 用于提供 WiFi 网络或有线局域网; 天线负责 WiFi 空口数据包的侦听与接收。

2) 操作系统 OpenWrt。OpenWrt 是精简版的 Linux 系统, 具有资源丰富、移植方便、源代码开放、免费等优点^[12]。它用途广泛, 能运行在很多品牌的 WiFi SoC 上, 比如高通、博通、联发科以及英特尔等。并且, OpenWrt 提供了一个完全可写的文件系统及软件包管理工具 Opkg, 这样就可以从本地软件仓库或互联网软件仓库上下载并安装 OpenWrt 支持的软件包, 提高了系统软件开发的效率。

3) shell 和其它应用软件。shell 是一个系统工具, 它是系统跟硬件交互时的中间介质, 通过它可以把指令传给操作系统, 然后内核再去支配硬件去执行各种操作。此外, 感知系统功能的实现还依赖于其它一些软件包, 比如设备驱动, 语言工具、抓包工具等。

3.1 硬件平台选型

硬件平台是支撑感知系统运行的核心基础设施, 选择合适的硬件对于系统功能的良好实现具有至关重要的作用。选择硬件平台最关键的是选择 WiFi 芯片, 为了降低设计周期, 文中选择厂家现成的 MT7620 开发板作为硬件平台, 系统模型如图 4 所示。该开发板主要由 WiFi 模块核心板和底板组成, 其中核心板集成了 MIPS 架构处理器 MT7620A、RAM、天线等模块, 底板集成了电源、Flash ROM、网卡以及 USB 接口等。该开发板尺寸小, 层次清楚, 外围接口丰富, 非常适合二次开发, 它主要具有如下优点:

1) 拥有 580 MHz MPIS 架构的高性能 CPU core, 并且缓存处理机制出色。这样既能实现快速有效抓包, 也能使采集的数据得到及时有效地处理。

2) 可以运行多种版本的 Linux 系统, 并可根据需要进行编程。感知系统需要烧录 OpenWrt 操作系统, 并在此基

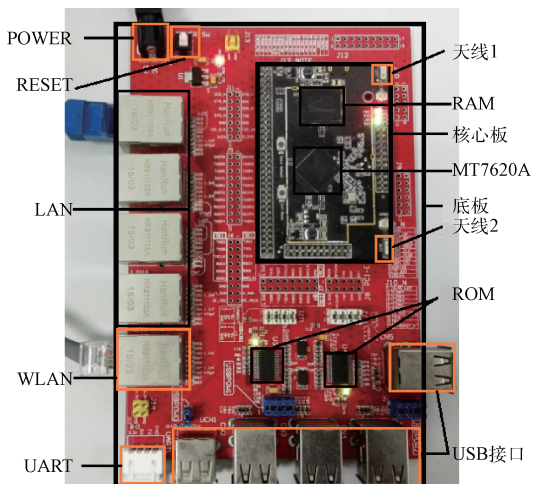


图 4 硬件系统模型

础上编写运行程序。

3) 运行内存大, 高达 128M RAM。由于检测程序运行时要占用一定的系统资源, 这有利于程序的稳定运行。

4) 完美支持 32 M SPI FLASH, 存储空间大。感知系统采集的数据要暂时存储在开发板 ROM 中, 大的 ROM 空间可以存储更多的数据和程序。

5) 支持 USB 2.0, 可扩展存储空间。USB 端口可以用来外扩存储空间, 以实现数据的随时读取, 从而增强了系统的健壮性。

6) 和同类产品比较, 价格低廉, 降低了开发成本。

3.2 软件设计

软件部分主要包括嵌入式操作系统 OpenWrt 的编译固化、外扩 U 盘并建立 Swap (交换分区)、检测程序等开发和设计, 以下是各部分的具体描述。

3.2.1 OpenWrt 的编译固化开发

OpenWrt 是开源系统, 从 OpenWrt 官网下载源码包, 并在 PC Linux 系统环境 (Ubuntu 14.04 版本) 下进行配置编译, 开发代码如图 5 所示。运行配置命令后, 系统弹出一个配置界面, 用于选择芯片的类型型号, 并根据应用需求自

```
# 更新系统软件
zff@ubuntu:~$sudo apt-get update
# 安装编译需要的工具包
zff@ubuntu:~$ apt-get install gcc g++ binutils patch bzip2 flex zlib1g autoconf
gawk make gettext bintils patch libz-dev asciidoc subversion
# 下载OpenWrt源码包
zff@ubuntu:~$git clone git://git.openwrt.org/openwrt.git
# 进入源码文件夹
zff@ubuntu:~$cd openwrt
# 更新扩展软件包
zff@ubuntu:~/openwrt $./scripts/feeds update -a
# 安装扩展软件包
zff@ubuntu:~/openwrt $./scripts/feeds install -a
# 进入配置界面
zff@ubuntu:~/openwrt $make menuconfig
# 开始编译
zff@ubuntu:~/openwrt $make V=99
```

图 5 OpenWrt 固件编译代码

主添加需要集成的组件,这里添加了用于支持USB存储拓展的kmod-sus-storage、kmod-usb-storage-extras和kmod-usb等软件包。编译成功后,可在源码文件夹的子目录里生成适合MT7620A的OpenWrt固件。需要注意的是,由于编译过程中需要从网站上下载各种源码包,所以编译过程中一直要让PC处于联网状态,否则编译会失败。将OpenWrt固件成功烧录到WiFi开发板后,即可利用SecureCRT软件通过SSH协议进入感知系统编程。

3.2.2 开发Swap分区

文中所用的嵌入式开发板物理内存有128 M,但相对于感知系统所运行的程序来说,内存仍然不够。为使感知系统能长期稳定运行,通过USB端口外扩了一个2 G大小的金士顿U盘,并在该U盘上建立了Swap分区。由于Swap分区大小一般设置为物理内存的2~2.5倍,所以文中将Swap分区设置为300 M,实现步骤如下:

Step 1:在PC的Ubuntu系统环境下将U盘格式化成两个分区:约1 700 M的EXT4分区和300 M的SWAP分区。

Step 2:在开发板上安装支持Ext4文件系统的软件包。将U盘插入到开发板的USB端口后,系统就能检测到U盘上的分区,代码如下:

```
root@OpenWrt:~# opkg update
root@OpenWrt:~# opkg install block-mount
kmod-fs-ext4
```

Step 3:将U盘插入开发板,并在U盘上建立文件系统,将系统映像移动到U盘的Ext4分区,以获得更多的空间用于存放各种文件,代码如下:

```
root@OpenWrt:~# mkdir /mnt/sda1
root@OpenWrt:~# mount /dev/sda1 /mnt/sda1
root@OpenWrt:~# tar -C /overlay -cvf - | tar -C
/mnt/sda1 -xf -
```

Step 4:确定U盘之前创立的Swap有效,并激活Swap分区。代码如下:

```
root@OpenWrt:~# mkswap /dev/sda2
root@OpenWrt:~# swapon /dev/sda2
```

Step 5:编写Swap分区启动脚本,使系统启动后自动完成Swap分区的激活。

3.2.3 检测程序设计

本文用Python语言和爬虫框架Scapy编写了系统的检测程序,在检测程序运行前,要将系统工作模式设置为监听模式,设置代码如图6所示。监听模式是WiFi终端7种工作模式中的一种^[13],该模式下,天线只能侦听无线信号,而不会发送信号。将感知系统设置为侦听模式,在该模式下系统不参与AP的竞争机制,不干扰其它AP的正常运行,而且侦听只涉及到WiFi终端的扫描阶段,不涉及终端的认证和关联,只捕获解析Probe Request帧,不涉及终端发送的数据帧,保护了用户的隐私。感知系统采集的数据

既可通过以太网传送到PC数据库,也可保存到系统ROM中,文中采取的是在ROM中新建文件phonemac.csv保存数据的方法。MT7620A可以在13个信道上侦听,为了简化设计,文中将探针所侦听的信道设置为常用的信道1。

```
root@OpenWrt:~# vim /etc/config/wireless

config wifi-device radio0
    option type mac 80211
    option channel 11
    option hwmode 11g
    option path '10180000.wmac'
    option htmode HT20
config wifi-iface
    option device radio 0
    option network lan
    option mode monitor
    option ssid OpenWrt
    option encryption none
    option hidden 1
```

图6 监听模式设置代码

为实现移动终端MAC地址等信息的获取与保存,程序需要设计2个线程来执行,即WiFi空口数据侦听线程和移动终端MAC解析存储线程,算法流程如图7所示。其中,如果一个移动终端的MAC地址在10 s内多次检测到,则认为该移动终端尚未离开系统覆盖区域。这样如果10 s内检测到重复的MAC地址,将更新该MAC地址的解析时间TIME_LAST_SEEN,否则将在文件phonemac.csv的末尾新建一个条目。

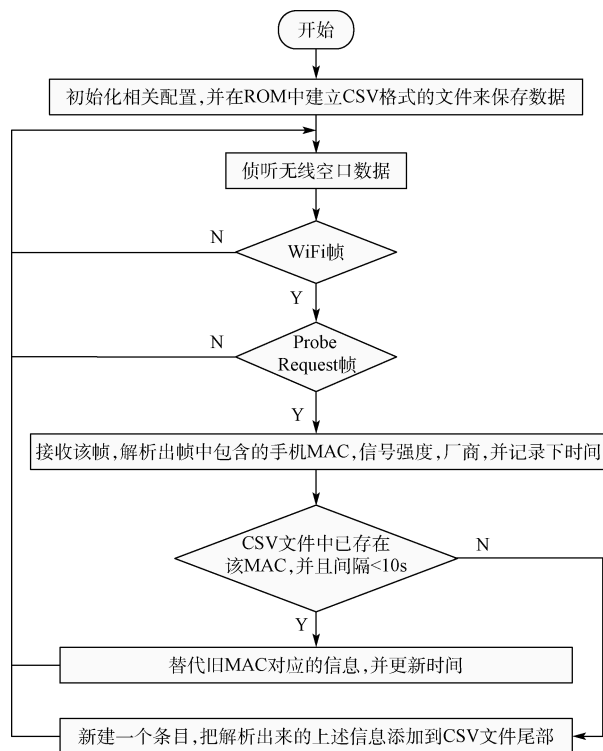


图7 线程执行流程

4 实验测试

为测试系统的功能实现情况,将开发的感知系统放置于中国科学院上海微系统所嘉定园区新微大厦 A 座三楼的办公区,对该区域内的移动终端进行实时监测,实验场景如图 8 所示。实验范围为 50 m×40 m,感知系统被放置在该区域中央位置的天花板上,距地面约 2.6 m。程序运行时间从 09:00~18:30,采集到数据 8 717 条,记录在 phonemacs.csv 文件中,部分数据如图 9 所示。

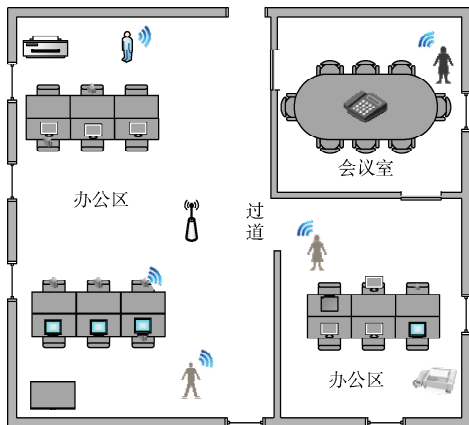


图 8 实验场景示意

终端MAC地址	终端厂商	信号强度	解析时间
f4:37:b7:c6:08:b7	Apple	-84	17:01:56 2016/1/7
dc:37:14:5a:67:f7	Apple	-78	17:02:01 2016/1/7
a0:18:28:06:40:ce	Apple	-82	17:02:16 2016/1/7
1c:7b:21:98:08:18	Sony	-84	17:02:17 2016/1/7
7c:7a:91:70:9a:04	Intel	-86	17:02:28 2016/1/7
8c:34:fd:3f:94:e5	Huawei	-66	17:02:33 2016/1/7
bc:79:ad:d4:5a:64	Samsung	-70	17:02:45 2016/1/7
00:26:c7:5a:52:fa	Intel	-86	17:02:48 2016/1/7
bc:79:ad:d4:5a:64	Samsung	-64	17:03:02 2016/1/7
68:df:dd:dc:71:62	Xiaomi	-84	17:03:11 2016/1/7
00:1e:64:2a:2d:e4	Intel	-82	17:03:21 2016/1/7
2c:cc:15:0b:23:78	Nokia	-64	17:03:34 2016/1/7
54:72:4f:5b:53:54	Apple	-84	17:03:40 2016/1/7
68:3e:34:26:5a:65	Meizu	-66	17:03:49 2016/1/7
7c:7d:3d:b9:96:70	Unkown	-88	17:03:51 2016/1/7
8c:34:fd:37:88:38	Huawei	-86	17:03:54 2016/1/7

图 9 系统采集的部分数据

上述数据显示,感知系统在该区域检测到的无线终端除了移动终端还有笔记本电脑(Intel)等。终端所属厂商是由该终端的 MAC 地址前 3 个字节确定的,IEEE 分配给厂商 MAC 地址的数据文件可以从 IEEE 官方网站下载。此外,从收集的数据可以看出,不同品牌的移动终端,发送 Probe Request 帧的时间间隔有一定的差别,而且同一个移动终端,Probe Request 帧的发送间隔也不是固定不变的。

5 应用前景

本文设计的 WiFi 移动终端感知系统,除了用于估计某片区域的人员密度以外,通过改进应用层的软件,还有其它

的一些潜在应用,列举如下:

商场分析顾客行为 在当今移动互联网时代,WiFi 成了购物中心延长顾客逗留时间的重要选择。如若在商场布置基于 WiFi 的移动终端感知系统,可以统计客流量和顾客停留时间以及他们的日常、每周或者每月的消费趋势。这样,商家就可以掌握商场的实时在线人数、新老顾客比、消费水平、客户集中到店时间等信息,这些数据对管理者制定商业战略大有裨益。

智能家居 智能家居利用计算机技术、网络通信技术、无线传感器等先进技术的发展理念,通过智能化的控制和管理,以实现“以人为本”的家居生活^[14]。它融合了人的个性需求,节能和温馨服务是智能家居的重要内容。将移动终端感知系统与家庭自动化系统相连,可以使电视、日光灯以及空调等根据人的出现和离开信息实现智能开关,提供人性化服务的同时,也可以达到节能的目的,并且这些应用会随着形形色色的智能可穿戴设备不断涌现将会得到进一步发展。

智能安防 将移动终端感知系统和和家里或公司的安全系统相结合,可以提升安防系统的效率。当上述区域被陌生人“入侵”而陌生人移动终端又开着 WiFi 时,感知系统将会检测到陌生的移动终端 MAC 地址,然后产生报警信号,并反馈给安全系统,进而让安保人员知悉,从而实现更有效的监控。同时,安全系统也可以启动应急程序,以保证公司或个人财产免受损失,保存下来的 MAC 地址以后也可以作为破案的线索。

智能考勤 随着科技的发展进步,公司的管理制度也应与时俱进,智能考勤也将是移动终端感知系统的一个潜在应用领域。长期下来,打卡将会占用员工很多的时间,考勤机安放位置不当也会给员工打卡带来不便。而且公司有时也存在员工代打卡、漏打卡的情况,给公司管理带来不必要的麻烦。利用感知系统可以在上班时间内对员工移动终端进行追踪,这样显然可以提高公司的管理效率以及员工的生产率。

6 结 论

在移动互联网蓬勃发展的今天,每人都有智能移动终端,随着智慧城市的发展,公共场所的 WiFi 热点越来越多,人们在公共场所移动终端开启 WiFi 的现象也越来越普遍,这就为对某片区域的移动终端进行 WiFi 嗅探提供了有利条件。本文提出并实现了一种基于 WiFi 的移动终端感知系统,并从软硬件实现方面给出了详细说明。该系统具有成本低廉、方便易用、易于扩展等优点,在 WiFi 越来越普及的今天,将其进行扩展完善,可实现更广泛的应用,具有良好的应用前景。

参考文献

- [1] 崔文婷,陈恒,高远,等. 基于 Android 的 LED 控制设计[J]. 电子测量技术, 2014,37(12): 102-104.

- [2] 彭宇, 庞景月, 刘大同, 等. 大数据: 内涵, 技术体系与展望[J]. 电子测量与仪器学报, 2015, 29(4): 469-482.
- [3] 叶斌, 李玉榕, 陈建国, 等. 基于 Android 系统的 KOA 运动监测系统研究[J]. 仪器仪表学报, 2014, 35(12): 7-11.
- [4] 吴松, 雒江涛, 周云峰, 等. 基于移动网络信令数据的实时人流量统计方法[J]. 计算机应用研究, 2014, 31(3): 776-779.
- [5] 胡斌杰, 詹益旺. 基于移动终端信令的道路交通流量状态识别及预测[J]. 移动通信, 2015, 39(10): 17-21.
- [6] FREUDIGER J. How talkative is your mobile device?: an experimental study of WiFi probe requests[C]. Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks. New York, USA: ACM, 2015: 8-14.
- [7] BARBERA M V, EPASTO A, MEI A, et al. Signals from the crowd: uncovering social relationships through smartphone probes [C]. Proceedings of the 2013 Conference on Internet Measurement Conference. Barcelona, Spain: ACM, 2013: 265-276.
- [8] CUNCHE M, KAAFAR M A, BORELI R. Linking wireless devices using information contained in WiFi probe requests[J]. Pervasive and Mobile Computing, 2014, 11(10): 56-69.
- [9] KONINGS B, BACHMAIER C, SCHAUB F, et al. Device names in the wild: Investigating privacy risks of zero configuration networking[C]. 2013 IEEE 14th International Conference on Mobile Data Management (MDM), Milan, Italy: IEEE, 2013: 51-56.
- [10] IEEE Standard Association. 802.11TM-2012. IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY)[S].
- [11] BHASCAR A, QU M, NANTES A, et al. Is bus overrepresented in Bluetooth MAC scanner data? Is MAC-ID really unique [J]. International Journal of Intelligent Transportation Systems Research, 2014, 13(2): 119-130.
- [12] 路青起, 席丹丹. 嵌入式 Linux 系统移植[J]. 国外电子测量技术, 2014, 33(12): 78-81.
- [13] KAO K F, LIAO I E, LYU J S. An indoor location-based service using access points as signal strength data collectors[C]. 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN), IEEE, 2010: 1-6.
- [14] 汪浩, 田丰, 杨溢, 等. 多模智能家居系统设计与实现[J]. 电子测量技术, 2014, 37(10): 20-24

作者简介

赵飞飞, 硕士研究生, 主要研究方向为无线传感器网络、嵌入式无线通信。

E-mail: hfutzff@163.com

金彦亮, 副教授, 博士, 主要研究方向为宽带无线通信、无线传感器网络。

E-mail: jinyanliang@staff.shu.edu.cn

熊勇(通讯作者), 博士, 研究员, 主要研究方向为移动通信和物联网技术。

E-mail: yong.xiong@mail.sim.ac.cn