

骚扰电话监控系统的研究与设计^①

殷春祥¹ 朱晓民¹ 彭刚²

(1.北京邮电大学网络与交换技术国家重点实验室 北京 100876; 2.东信北邮信息技术有限公司 北京 100191)

摘要: 摘要骚扰电话监控系统是一个对骚扰电话进行监控和拦截的系统,它对减少骚扰电话,提高电话用户的使用体验有着很重要的意义。首先给出了骚扰电话监控系统的组网方案和总体架构,然后重点设计了骚扰电话监控系统的监控和拦截的流程。

关键词: 骚扰电话; 系统结构; 数据采集

Study and Design of Spam Call Control System

YIN Chun-Xiang¹, ZHU Xiao-Min¹, PENG Gang²

(1. State Key Lab of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China; 2. EB Information Technology Co. Ltd., Beijing 100191, China)

Abstract: The Spam Call Control System is a system to find out those spam calls and intercepts them. It is effective in reducing spam calls and making customers get better mobile services. This paper firstly proposes a network topology and an overall architecture of Spam Call Control System. Secondly, it focuses on designing the procedure of finding out and intercepting spam calls.

Keywords: spam call; system hierarchy; data collection

1 引言

随着通信技术的不断发展,移动通信网络建设成本和手机终端成本的不断下降,移动通信业务的不断丰富,人们对移动通信的依赖越来越大,使用的频度越来越高。移动通信的飞速发展在给人们带来方便的同时,也使得有些人出于商业目的利用移动通信宣传和传播一些骚扰信息,导致了骚扰电话的泛滥,人们的生活带来很大的骚扰。骚扰电话主要表现为:不法用户对移动客户大范围拨打,振铃一次后挂机,等客户回拨时呼转到音信台或录音电话,形成骚扰和欺诈,在主观上违背手机用户意志并且在客观上对用户通信自由、生活安宁造成侵害或者蒙蔽用户的呼叫。尤其是近年来,骚扰电话越来越多,部分不法用户又开始使用高科技手段进行骚扰,这主要体现在使用专用的外呼软件来代替以往的手动外呼。这类软件的功能强大,在普通的PC机上运行即可达到每秒钟几百

次的外呼量,并且还具有自动生成呼叫号码、呼叫频率可调等特点。而且,随着号码资源成本的不断降低,骚扰号码也体现出经常变换,难以捕捉的特点;部分用户以手机、电话作为媒介,进行电话营销,影响到人们正常的工作、生活和正常通话。

2 骚扰电话监控系统研究与设计

2.1 骚扰电话监控系统特点

针对上述恶意骚扰电话,我们设计了骚扰电话监控系统,能对骚扰电话进行有效的监控和拦截。该系统主要有如下的几个特点:

(1) 使用“呼叫可疑度”评估算法,对用户呼叫进行“实时监控”与“非实时监控”,两种监控方式相辅相成,可以有效、快速、准确地识别出各种特点的骚扰电话(包括恶意骚扰和电话营销等),进而对骚扰电话进行拦截与封锁,有效阻止了其广大用户的骚

^① 基金项目:国家杰出青年科学基金(60525110);国家973计划(2007CB307100;2007CB307103);国家自然科学基金(60902051);中央高校基本科研业务资助(BUPT2009RC0505);电子信息产业发展基金项目
收稿时间:2009-12-06;收到修改稿时间:2010-12-28

扰，杜绝了可能发现的欺诈及电话营销行为。

(2) 同时，监控系统还支持黑白名单的过滤功能，并且可以针对每个用户号码设置拦截封锁有效期限。

(3) 系统与 BOSS(Business and Operation Support System)系统、汇接局、关口局具有交互接口，能够实现对省内号码、省外号码以及互联号码的有效拦截封锁。

2.2 系统总体结构

骚扰电话监控系统硬件上分为采集分析节点和管理节点。采集分析节点负责从数据源平台采集原始话单数据，对数据进行分析处理，形成疑似号码清单，定期提供给 BOSS 并且从 BOSS 获取黑白名单。管理节点负责系统管理，提供给操作员进行业务查询、统计，监控和配置系统参数等。

2.2.1 网络结构设计

骚扰电话监控系统采用双网络集群应用环境。硬件上采用备份冗余机制，保证服务不间断。

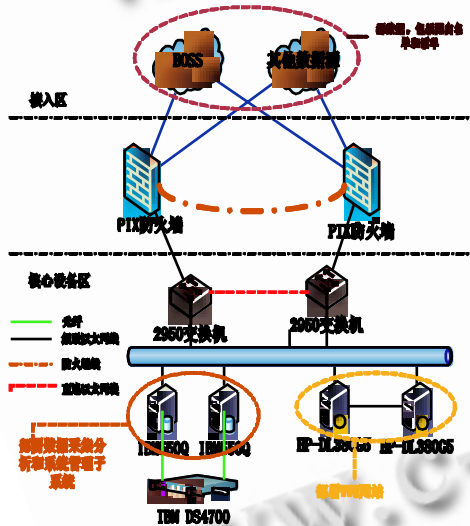


图1 系统网络结构图

网络拓扑说明：

① 接入区为各个外部网元实体。包括：BOSS、数据源平台等。(本图按照某省实际情况绘制，实际的外部网元可能很多，包括：短信中心、彩信中心等)

② 外部接口和骚扰电话监控系统通过防火墙进行连接，保证系统安全性。

③ 核心设备由两台 IBM550Q 和 HP DL380G5 组成。IBM550Q 部署数据采集、行为分析、系统管理子系统；HPDL380G5 部署 WWW(World Wide Web)

接入子系统，提供基于 Web 的统一操作界面。

④ 数据库数据、业务文件数据和系统运行关键配置数据保存到磁盘阵列上(IBM DS4700)。数据库安装在 IBM550Q 上。

采用双网络结构。所有网络设备(防火墙、交换机、路由器)、主机系统俩俩备份。

2.2.2 逻辑结构设计

系统逻辑结构主要用来说明系统逻辑层次划分。骚扰电话监控系统系统从上至下分为三层，包括：界面层，业务层，数据层。上下层之间通过制定的接口，使得下层为上层提供服务。逻辑结构图如图 2 所示。

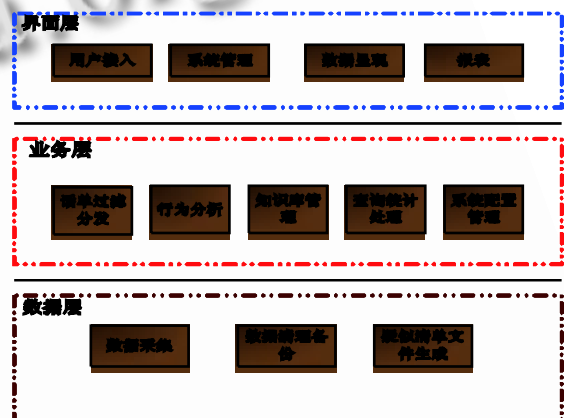


图2 系统逻辑结构示意图

骚扰电话监控系统系统逻辑结构图说明：

① 界面层：负责实现与用户交互的界面，提供 UI(User Interface)让用户方便直观地接入和管理系统，并采用丰富的展现形式显示用户操作结果和(定期或者实时)报表数据。

② 业务层：负责实现具体的业务处理逻辑，完成用户指定的操作，并返回相应的处理结果，从功能实现上包括业务投诉处理，业务统计处理，业务查询处理，业务分析处理，知识库管理，系统配置管理等。

③ 数据层：负责采集、数据备份，文件提供，包括 CDR(Call Detail Record)文件采集、备份；黑白名单文件的采集和备份；疑似清单文件的定期生成。

2.2.3 软件总体结构设计

系统软件总体结构包括骚扰电话监控系统的内部各个子系统划分以及子系统间的通信，骚扰电话监控系统的外部业务系统的接口，包括各种通信接口(TCP(Transmission Control Protocol)/IP(Internet

Protocol)、FTP(File Transfer Protocol)等和数据库接口。总体结构图如图 3 所示:

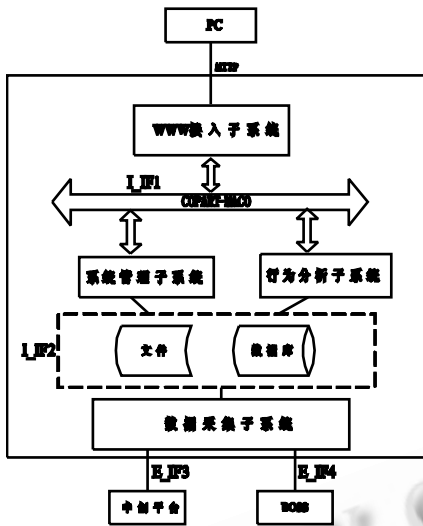


图 3 系统软件总体结构图

系统总体结构图说明:

① 外部系统与接口:

外部相连的系统目前有中创信令检测平台和 BOSS 系统。接口包括 E_IF3 和 E_IF4,都为 FTP 接口。E_IF3: 骚扰电话监控系统的数据采集子系统作为客户端,从中创平台上通过 FTP 下载话单文件,保存到本地,提供给行为分析子系统使用; E_IF4: 骚扰电话监控系统的数据采集子系统作为服务器,向 BOSS 提供疑似清单文件,并接受 BOSS 上传的黑白名单文件。

用户终端用于用户通过浏览器接入骚扰电话监控系统系统,经过身份验证后,可以触发相应的业务操作。

② 内部系统与接口:

WWW 接入子系统: 用户界面采用 WWW 访问方式。WWW 接入子系统是面向操作员的系统操作和访问接口,提供一个图形化、友好的操作界面,实现各种业务逻辑的结果展示。另外 WWW 还需要实现自身的管理,包括:用户管理、权限管理等系统管理功能。

行为分析子系统: 完成对骚扰电话、垃圾短信话单的分析,过滤出其中疑似的异常号码,然后将这些疑似号码输入到疑似号码的数据库。

系统管理子系统: 主要实现系统管理、查询统计、客户知识库管理的功能。

数据采集子系统: 完成对原始话单数据采集,向

BOSS 提供疑似清单,定期清除过期记录和文件。

DB(Data Base)(I_IF2 接口): 骚扰电话监控系统系统数据库,存放各种数据,包括:业务数据、统计数据、管理数据和系统配置数据等。骚扰电话监控系统系统的行为分析子系统,系统管理子系统,数据采集子系统和数据库有接口,使用 EB 公司开发的数据库访问接口 ecfuncs。

文件接口(I_IF2 接口): 实现话单文件、疑似清单文件、黑白名单文件接口。

COPART-MACO(I_IF1 接口): EB 通用消息平台,是一种通信中间件。提供系统内部模块之间的通信,并为建立在中间件基础上的模块提供了操作维护,进程守护等管理机制。

2.3 系统业务流程

通过对大量骚扰电话的拨打电话记录分析得知,骚扰电话有如下的几个显著特征:

- ① 短时间振铃
- ② 用户通常无法正常接听
- ③ 超频呼叫
- ④ 不产生话单记录

根据上述的骚扰电话特征,骚扰电话监控系统的总体思路如下:



图 4 骚扰电话监控拦截流程图

2.3.1 源数据的采集

数据采集功能的采集粒度可以灵活设置,采集的数据主要包括如下的内容:

- ① 时间(起始时间,终止时间,提交时间)
- ② 号码)主叫号码,被叫号码,IMSI,MSRN)
- ③ 主备叫位置信息

2.3.1.1 分析数据的采集

- ① 从计费系统采集端试呼话单获取数据源
- ② 通过开发接口从信令监测系统采集 CDR 获取数据源

2.3.1.2 黑白名单的采集

从 BOSS 平台采集骚扰电话的黑白名单。黑名单是那些早已确定为骚扰电话的号码，白名单是那些确认为不是骚扰电话的号码，比如移动的 10086 等。黑白名单是用来在骚扰电话分析模型对源数据进行分析之前对源数据进行过滤用的，去除确定是骚扰电话和确定不是骚扰电话的号码。对于那些已经确认的号码不用再分析了，节省分析时间，提高分析的效率^[1]。

2.3.2 骚扰电话的分析

通话时长和通话频率是各种电话呼叫的主要特征^[2]。通过对上述采集的数据进行通话时长和通话频率的挖掘分析，根据统计分析的结果对骚扰电话的号码的各个属性设置相应的阈值，满足这个阈值时则该号码被视为疑似号码，送至局方进行确认^[3]。对号码的统计分析主要包括如下几个方面：

- ① 主叫特征，包括主叫号码、呼叫频次（如分钟最大呼叫次数、小时最大呼叫次数、天最大呼叫次数）、主叫归属地、主叫位置信息。
- ② 被叫特征，包括连续号码、同一号段、随机号码、号码归属地。
- ③ 其他特征，包括呼叫时间间隔、呼叫时间段、呼叫时长、振铃时长等。

2.3.3 骚扰电话的确认

现阶段由于骚扰电话还没有官方的权威的定义，只能暂时通过人工拨测电话号码来确定。虽然骚扰电话监控系统暂时无法准确的确定某个号码是否是骚扰电话，但是通过骚扰分析模型的分析之后可以大大减少拨测号码的数量，同时有针对性地提高了号码拨测的准确性^[4]。

2.3.4 拦截策略

- ① 省内骚扰电话：在 BOSS 系统进行拦截，关闭该拨打电话的功能，同时发送短信提醒，规避法律险。
- ② 外省骚扰电话：设置拦截期限，从软交换汇接局拦截，同时派发投诉单至外省移动，履行告知义务，在拦截期限之后自动解除拦截。
- ③ 外省漫游至本省的骚扰电话：根据骚扰电话监控平台提供的黑名单和用户位置信息，设置拦截期限，对用户所在 VMSC/VLR 进行拦截，同时派发投诉单至外省移动，履行告知义务，在拦截期限之后自动解除拦截。
- ④ 互联互通骚扰电话：设置拦截期限，对互联互通的骚扰电话在关口局进行拦截，并对互联互通投诉，在拦截期限之后自动解除拦截。

3 小结

骚扰电话监控系统通过对骚扰电话的统计特征进行分析，建立骚扰电话模型，对骚扰电话进行了有效的分析和拦截，大大减少了骚扰电话对人们正常生活的负面影响，提高了用户对移动电话的使用体验。同时由于骚扰电话到现在还没有一个明确的定义，对骚扰电话的监控和拦截涉及到一些法律方面的问题，所以骚扰电话监控平台还有待进一步的发展和改进。

参考文献

- 1 管乐,王纯.多维关联规则挖掘在彩铃推荐中的应用.计算机系统应用,2009,18(4):155-157.
- 2 张奇支,廖建新.移动智能网用户的通话时长模型分析.高技术通讯,2006,16(4):353-357.
- 3 Han Jw, Kamber Mi.范明,孟小峰译.数据挖掘概念与技术.2nd ed.,北京:机械工业出版社,2007.67-69.
- 4 David Freedman 等著.魏宗舒,施锡铨等译.统计学.2nd ed.北京:中国统计出版社,1999.50-55.