



违规语音电话的治理方法与实践

朱宏毅¹, 林建洪², 王彦青², 张锦然², 王敬宇²

(1. 中国移动通信集团浙江有限公司 杭州 310051;

2. 东信北邮信息技术有限公司 北京 100191)

摘要: 电信运营商一直致力于违规语音电话治理方法的研究, 包括分析违规语音电话的风险来源、确定违规语音电话的治理重点、制定违规语音电话的防治策略。通过建立基于各类呼叫信令采集的大数据分析平台, 研究了各类违规语音电话的呼叫特点, 建设了违规语音电话(特别是诈骗电话)的防治系统, 实现了对违规语音电话全方位的管控, 并取得了明显的效果。

关键词: 通信诈骗; 骚扰电话; 违规语音电话

doi: 10.11959/j.issn.1000-0801.2015187

Approach and Practice of Illegal Voice Call Governance

Zhu Hongyi¹, Lin Jianhong², Wang Yanqing², Zhang Jinran², Wang Jingyu²

(1.China Mobile Group Zhejiang Co., Ltd., Hangzhou 310051, China;

2.Eastcom-Bupt Information Technology Co., Ltd., Beijing 100191, China)

Abstract: Telecom operators have been committed to the research of illegal voice call governance approach, which includes analyzing the risk origin, determining the management focus and formulating the prevention strategy. Through the establishment of big data analysis platform based on all kinds of call signaling collection, the call features of all kinds of illegal voice call were researched, the control system of illegal voice call (especially fraud call) was constructed, an all-around control was implemented, and the effect was obvious when the system was applied actually.

Key words: communication fraud, telephone harassment, illegal voice call

1 引言

近年来, 各类以诈骗电话、骚扰电话、虚假主叫电话为主的违规语音电话在网内急速蔓延, 尤其是涉嫌诈骗的语音电话, 给广大移动用户带来了严重损失。“搜狗号码通”数据报告显示, 2014 年全国骚扰电话呼叫总数高达 270 亿个, 主要为“响一声”, 涉嫌违法和诈骗、广告推销和房产中

介等^[1]。另据公安部统计, 2014 年我国通信诈骗发案 40 万余起, 群众损失超过 107 亿元。2015 年度的央视 3·15 晚会也报道了虚假主叫严重侵犯消费者合法权益的案例。

电信诈骗案件一般会涉及用户、运营商、银行和公安部门。银行及公安部门已经针对电信诈骗结合自身专长进行了积极治理, 也取得了一定的效果。运营商具备对通信网络的统一管控能力, 可借助最新技术手段从通信网络中

收稿日期: 2015-06-03; 修回日期: 2015-07-15

论文引用格式: 朱宏毅, 林建洪, 王彦青等. 违规语音电话的治理方法与实践. 电信科学, 2015187

Zhu H Y, Lin J H, Wang Y Q, et al. Approach and practice of illegal voice call governance. Telecommunications Science, 2015187

提前发现电信诈骗并制止诈骗行为,对违规语音电话的源头进行反制。中国移动通信集团浙江有限公司(以下简称浙江移动)近年来对此进行了积极探索和创新实践,承担起维护通信安全、肃清通信环境的社会责任,为广大移动用户创造了一个安全和谐的通信环境。

2 违规语音电话的特点和治理思路

电信运营商一直致力于骚扰电话和虚假主叫类的违规语音电话治理,随着 VoIP(voice over internet protocol,网络电话协议)和移动互联网技术的兴起,不法分子借助改号设备和大数据技术使得违规语音电话的花样不断翻新,更具隐蔽性和迷惑性,违规语音电话的种类和结果也在不断变化中。经过全面调查和近一年的试点验证,总结出现有网络环境中违规语音电话存在以下特点。

(1) 违规语音电话来源复杂,重点是国际诈骗

违规语音电话的风险来源如图 1 所示。违规语音电话主要有 4 个风险来源:一是来自网间国际电话的虚假改号,从用户的投诉内容来看,目前的虚假主叫号码绝大多数来自跨网的呼叫^[2];二是来自国际长途局国际电话的虚假改号,目前通信网上出现的虚假主叫主要是通过改号方式对主叫号码进行伪装来实现,其改号功能是指一种通过网络改号平台进行号码转换的服务功能,能使用户呼出电话时在对方手机或座机上显示任意号码的一种号码变换方式^[3];三是来自网间/网内的语音专线,部分用户开展违

规外呼活动,这些业务平台在经济利益的驱使下,把设置主叫号码的权限作为业务的一个特色来吸引用户使用,从而导致虚假主叫的产生^[4];四是来自移动网内的骚扰电话,目前针对骚扰电话的治理手段已经比较成熟^[5,6]。

根据治理试点验证来自国际的虚拟改号来电占诈骗电话总量的 80%以上,对人民群众的危害也最大,所以对国际来电诈骗电话的防范是当前亟须加强的重中之重,也是违规语音电话治理的首要目标。

(2) 仿冒公检法、特服号码居多

诈骗电话呼叫多为高频呼叫,且仿冒特服号码、公检法权威部门电话等公众号码居多。因此,在治理前期,从仿冒公众号码的诈骗号码切入,通过分析此类诈骗电话的特征,积累主叫用户行为,探索违规语音电话治理的长效机制。

(3) 在一定周期内使用某固定主叫号码进行诈骗

诈骗电话呼叫,通常在一定时间周期内使用某固定主叫号码对不同被叫用户进行诈骗。如果系统能及时发现诈骗的主叫号码,并对其进行封堵拦截,可以大大减少该号码后续造成的危害。针对这类情况,目前运营商的做法是设置黑名单,在关口局和汇接局对这些号码进行封堵拦截。但是这些号码一旦被拦截,诈骗分子很快就会变换号码。考虑到诈骗分子掌握号码变换能力和路由能力,所以需要管控平台引入动态监测和多重封堵拦截来治理诈骗电话^[7]。

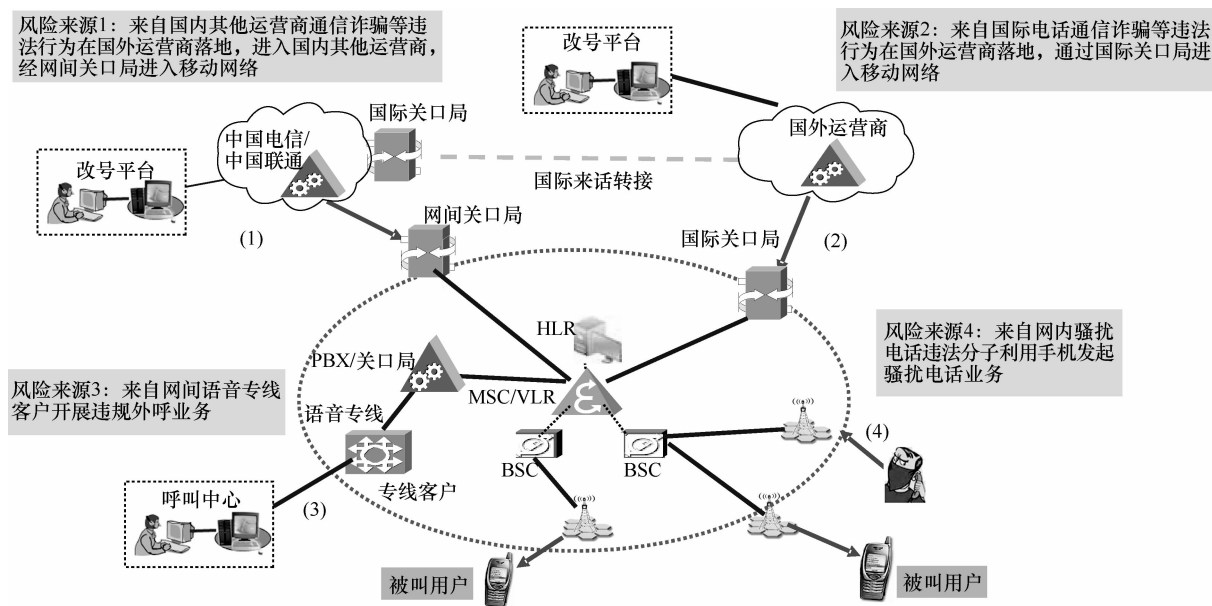


图 1 移动违规语音电话的风险来源示意



3 业务治理系统建设

按照“突出重点、速度优先”的治理思路,浙江移动构建了由呼叫信息采集、大数据分析、取证审核、拦截管控、信息提醒及反馈 5 个节点组成的违规语音电话管控平台,综合治理骚扰电话和诈骗电话相关的违规语音电话。违规语音电话管控平台结构如图 2 所示。

3.1 呼叫信息采集

呼叫信息采集是对违规语音电话疑似号码进行分析的基础和前提。目前运营商通用的做法是采集 CDR(call detail record, 呼叫详细记录)数据,网内的任意一条呼叫记录都包含大量全面的信息。CDR 中包含呼叫摘机时间、应答时间、挂机时间、主叫号码、被叫号码、原被叫号码、网元信息等和用户行为相关的有关信息^[8]。采集的信令数据越全,治理的范围越广,达到的治理效果才能越好。目前,平台已采集的 CDR 信令包括网间关口局的智能网触发信令数据、省内签约智能信息(T_CSI)用户信令数据、长途汇接局信令数据、互联互通关口局的 ISUP(ISDN user part)信令数据以及省内所有的 A 接口信令数据。这些数据通过实时/准实时接口将数据送到管控平台分析模块,以进行后

续的数据挖掘工作。

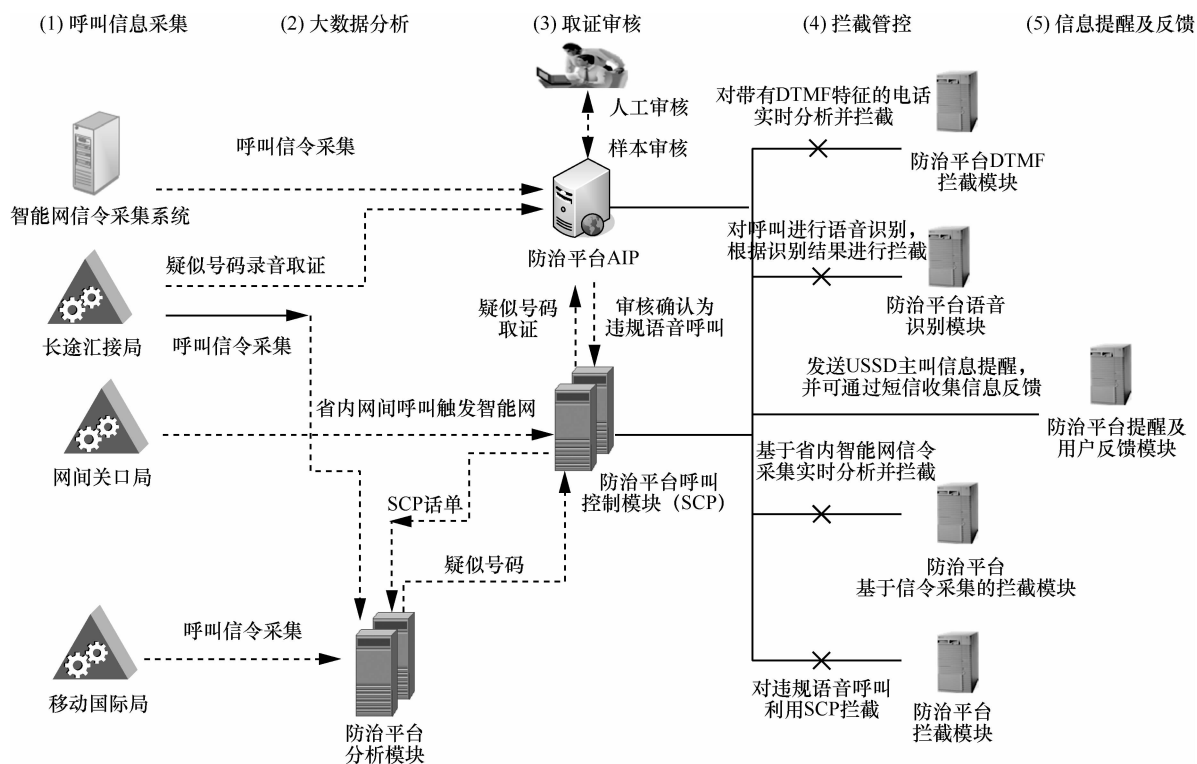
3.2 大数据分析

通信网内实时的用户行为数据及用户通话的信令数据都具有大数据 4V 特点(规模大、变化速度快、价值高、类型复杂),充分利用大数据技术,通过数据挖掘与分析工作,可实现对违规语音电话的识别和管控^[9]。中国电信股份有限公司广东分公司(以下简称广东电信)已经采用基于大数据处理技术的虚假主叫甄别的技术,取得了一定的效果^[3]。用大数据技术进行实时分析,根据对应的规则可发现高疑似的诈骗号码,违规语音电话识别模型如图 3 所示。根据已确认的诈骗电话行为特征,通过挖掘算法匹配识别出具有同样特征的疑似诈骗电话号码,并与其语音通话记录信息进行对比识别分析,最终识别出疑似的违规语音电话。

在违规语音电话识别模型构建环节,目前采用 3 种算法模型:号码池对比规则模型、决策树和逻辑回归模型、聚类算法模型。大数据分析节点的核心是在海量的信令数据中能够精准过滤出疑似违规语音号码。

3.2.1 号码池比对规则模型

针对网间关口局数据库采用号码池比对规则,主要是



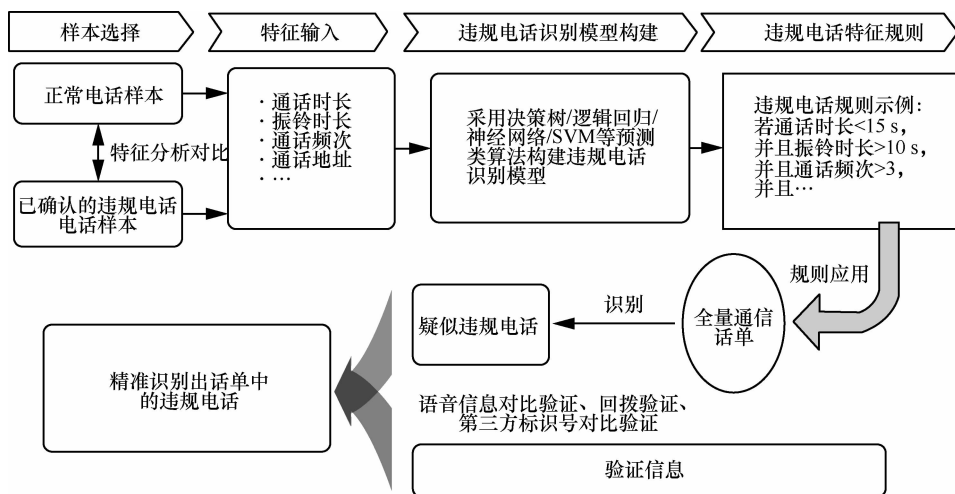


图3 违规语音电话识别模型

将呼叫信令中的主叫号码与号码池进行比对,并结合呼叫频次、被叫号码离散度等规则筛选出高疑似号码,该规则主要用于疑似诈骗号码的识别。

由于诈骗电话多是仿冒公检法、特服号码,因此,已知的公检法、特服号码可以作为号码池中的号码。号码池号码分为:110 尾号、公检法号码、特服号码、外部资源库号码。号码池对比规则模型实现如图 4 所示。

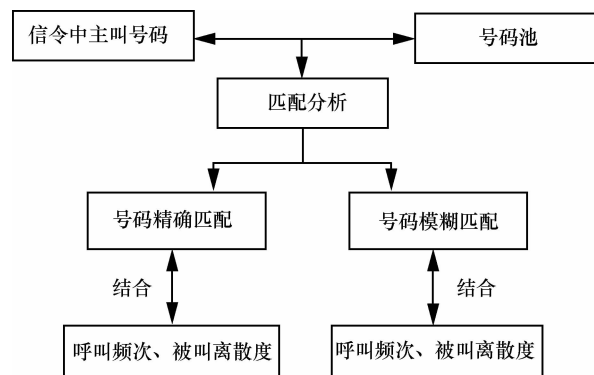


图4 号码池对比规则模型实现

号码精确匹配识别方法:规则主要判断信令中的主叫号码是否与号码池中的特殊号码完全匹配,若完全匹配,且呼叫频次及被叫离散度分别大于某一既定的阈值,则判定该主叫号码为疑似号码。

号码模糊匹配识别方法:规则主要判断信令中主叫号码尾号或中间连续若干位与号码池中特殊号码是否完全匹配,若完全匹配,且呼叫频次及被叫离散度分别大于某一既定的阈值,则判定该主叫号码为疑似号码。

针对省内 A 口数据源,采用预测类算法构建决策树和逻辑回归组合模型对违规电话进行识别,通过用户的通话行为信息(包括通话时长、振铃时长、通话开始时间、通话结束时间等)、用户地址等信息对数据进行模型构建。决策树和逻辑回归模型实现流程如图 5 所示。

模型具体构建方法介绍如下。

(1)数据源选择。抽取连续 3 日 A 口忙时的所有话单记录。

(2)建立模型构建所需数据表。统计话单记录中每个主叫号码的若干个特征指标的值,所述特征指标包括但不限于:总通话次数、总通话时长、无应答次数、总振铃时长、主动释放总次数、被动释放总次数等,并判断主叫号码是否在拦截号码表中。如果是,将该主叫号码的违规标识值设为 1;如果否,将该主叫号码的违规标识值设为 0。

(3)算法评估选择。随机抽取数据表中 100 000 个号码信息数据(主叫号码类型包含 0 和 1)作为训练样本,利用统计分析软件 SPSS Clementine 进行模型试运行,选择可用来预测识别分析的若干算法作为候选。在此,选择决策树-C5.0、逻辑回归、神经网络算法作为候选,评估每种算法生成结果的准确性、覆盖率、运行时长、顽健性、可扩展性及易理解性,将其作为算法选择准则。各算法的比较见表 1。

从表 1 的结果可以看出,决策树算法模型与逻辑回归算法模型的识别准确率均较佳,为增强模型稳健性,平台采用决策树和逻辑回归算法组合的方式构建模型。

(4)特征指标选择。为有效避免变量之间因相互关联

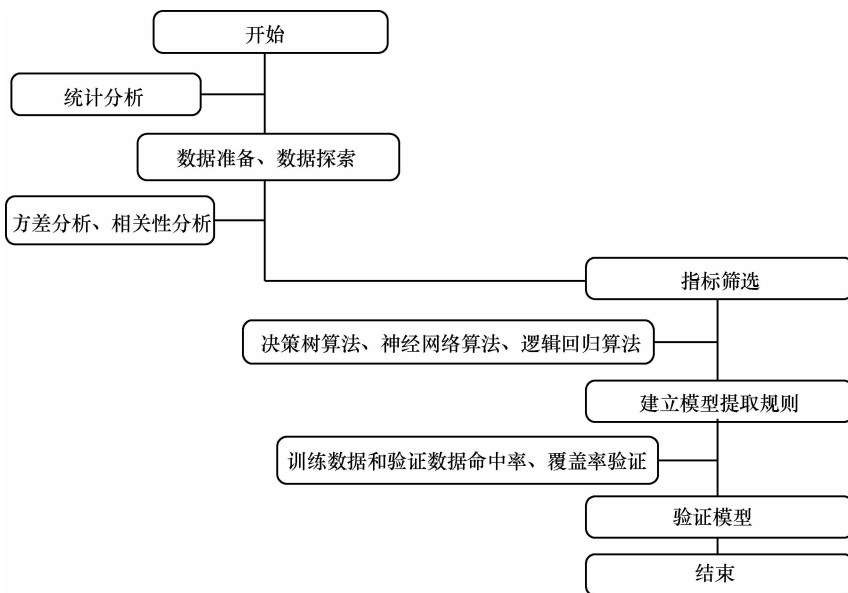


图5 决策树和逻辑回归模型实现流程

表1 算法比较

算法	准确率	覆盖率	时长/s	顽健性	扩展性	理解性
决策树算法	89.7%	43.4%	38	好	一般	很好
逻辑回归算法	75.6%	39.7%	39	一般	很好	很好
神经网络算法	51.2%	36.5%	74	好	一般	差

性较大及各变量与目标之间无相关导致的模型误差,建模前通过相关性分析对变量进行筛选,过滤掉部分冗余变量,相关系数的计算方法为:

$$r = \frac{\sum (R_{xi} - \bar{R}_x)(R_{yi} - \bar{R}_y)}{\sqrt{\sum (R_{xi} - \bar{R}_x)^2 \cdot \sum (R_{yi} - \bar{R}_y)^2}} \quad (1)$$

其中, R_{xi} 和 R_{yi} 分别表示第 i 个 X 变量和 Y 变量经过排序后的次序, $\bar{R}_x$ 和 $\bar{R}_y$ 分别表示 R_{xi} 和 R_{yi} 的均值。

不同的经验 $|r|$ 值表示的不同程度的线性相关关系为:

- $|r| < 0.3$, 表示低度线性相关;
- $0.3 \leq |r| < 0.5$, 表示中低度线性相关;
- $0.5 \leq |r| < 0.8$, 表示中度线性相关;
- $0.8 \leq |r| < 1.0$, 表示高度线性相关。

(5) 决策树线算法模型构建, 识别每个主叫号码是否为疑似违规电话。决策树算法是一种类似于流程图的树结构, 核心是在决策树的各级节点上选择属性, 用信息增益作为属性选择标准式。

信息量计算式为:

$$I(p, n) = -\frac{p}{p+n} \cdot \lg \frac{p}{p+n} - \frac{n}{p+n} \cdot \lg \frac{n}{p+n} \quad (2)$$

信息期望计算式为:

$$E(A) = \sum_{i=1}^p \frac{p_i + n_i}{p+n} I(p_i, n_i) \quad (3)$$

信息增益计算式为:

$$\text{gain}(A) = I(p, n) - E(A) \quad (4)$$

其中, p 表示数据集中正例的个数, n 表示数据集中反例的个数, p_i 表示属性 A 下第 i 个子集中正例的个数, n_i 表示属性 A 下第 i 个子集中反例的个数, $I(p_i, n_i)$ 表示属性 A 下第 i 个子集的信息熵。

(6) 逻辑回归算法模型构建。通过模型输出的逻辑回归方程式可计算出每个主叫号码的疑似违规程度, 即违规概率值, 该概率值为 $[0, 1]$ 内的数值。模型经多次评估验证, 当方程式大于或等于 0.9 时, 所识别出的违规电话在准确率和查全率方面均为最优, 因此, 在实际应用中, 将疑似违规程度大于或等于 0.9 的主叫号码认定为疑似违规号码。模型输出计算式为:

$$P(Y=1 | X=x) = \exp(a^x + b^{x_1} + c^{x_2} + \dots + \beta) / (1 + \exp(a^x + b^{x_1} + c^{x_2} + \dots + \beta)) \quad (5)$$

其中, x_1, x_2, x_3 为模型输入指标, β 为极大似然估计。也就是说, 当违规概率 $P(Y=1 | X=x) \geq 0.9$ 时, 认定为疑似违规号码。

(7) 模型组合, 更精准地判断所有主叫号码是否为违规电话。当全量主叫号码信息同时满足决策树-C5.0 算法模型生成的规则与逻辑回归算法模型生成的方程时, 则判定为疑似违规电话。

总之, 该模型主要依据已确认的违规语音电话在拦截前的通信行为特征, 采用模型输出的规则及方程, 以组合的形式进行匹配并识别出具有相似特征的疑似违规语音电话, 再与其语音通话记录进行比对分析, 最终识别出违规语音电话。

3.2.2 聚类算法模型

根据研究发现, 违规电话、疑似违规电话都具有较明显的特征差异, 例如: 违规电话具有忙时高频呼叫、被叫用户相对集中、呼叫时间间隔离散度较高的特征; 疑似违规电话具有高频呼叫、被叫用户相对分散、呼叫圈重合度较高、呼叫时间离散度较高的特征; 非违规电话具有低频呼叫且时间集中、呼叫圈重合度较低、主叫呼叫行为较少、忙时基本无呼叫行为的特征。因此, 本模型采用聚类算法, 根据 SCP(service control point, 业务控制点) 话单记录中所有主叫号码的多个特征指标值来对主叫号码进行特征分类, 将具有相同或相似特征的主叫号码分到一个簇中, 也就是说, 将全部用户分为具有明显特征差异的多个簇, 再通过和已确认违规电话的特征进行对比, 从而找到和已确认违规电话特征最接近的违规电话簇和较接近的疑似违规电

话簇。对于违规电话簇和疑似违规电话簇, 还可以进一步采用逻辑回归算法来甄选出其中的违规电话和疑似违规电话, 从而实现最大范围内违规电话的精准识别和拦截。模型整体的构建流程如图 6 所示。

模型的具体构建方法介绍如下。

(1) 计算所有主叫号码在一定时间周期内的特征指标值, 并为所有主叫号码分别构建对应的特征指标集: $X_i = (x_{i1}, x_{i2}, \dots, x_{iN})$, 其中, X_i 是主叫号码的特征指标集, $x_{i1}, x_{i2}, \dots, x_{iN}$ 分别是主叫号码 z_i 的若干个特征指标值, N 是特征指标数。

(2) 构建 3 个簇 (例如簇 1、簇 2、簇 3), 并将所有主叫号码随机划分到这 3 个簇中, 每个主叫号码都唯一地归属于一个簇。

(3) 计算每个簇的特征指标中心值集 $C_j: (c_1^j, c_2^j, \dots, c_N^j)$, 其中, C_j 是簇 j 的特征指标中心值集, j 是 1 到 3 之间的一个自然数, c_t^j 是 C_j 中的第 t 个特征指标中心值, t 是 0

到 N 之间的一个自然数, 且 $c_t^j = \frac{\sum_{i=1}^M x_{it}^j}{M}$, i 是 1 到 M 之间的一个自然数, M 是簇 j 中的主叫号码数, x_{it}^j 是簇 j 中的主叫号码 z_i^j 的第 t 个特征指标值。

(4) 计算所有主叫号码的平方误差和: $E = \sum_{j=1}^3 \sum_{i=1}^M \sum_{t=1}^N |x_{it}^j - c_t^j|^2$,

并判断 E 是否小于或等于 E 的阈值。如果是, 本流程结束; 如果否, 计算每个主叫号码和所有簇的特征指标中心值集之间的距离, 并从中挑选距离的最小值, 然后将对应的主叫号码重新划分到距离的最小值所对应的簇中。主叫号

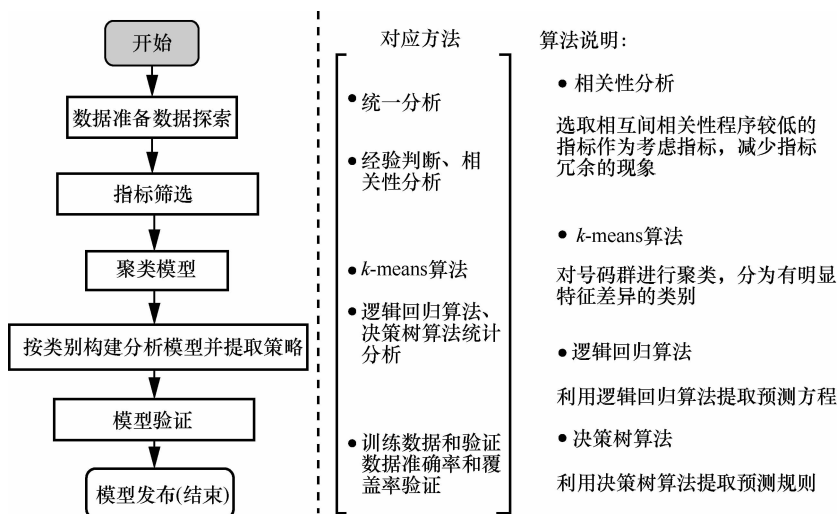


图 6 聚类算法模型流程



码 z_i 和簇 j 的特征指标中心值集之间的距离的计算式为:

$$D_{ij} = \sqrt{\sum_{t=1}^N (x_{it} - c_{jt})^2}, \text{ 其中, } x_{it} \text{ 是主叫号码 } z_i \text{ 的第 } t \text{ 个特征指标值。}$$

(5) 聚类完成后,再通过和已确认违规电话的特征进行对比,从而找到和已确认违规电话特征最接近的违规电话簇、较接近的疑似违规电话簇。采用逻辑回归算法,逐一计算违规号码簇或疑似违规号码簇中每个主叫号码的违规可疑度指数 P : $P = \frac{e^Y}{1+e^Y}$, 其中, Y 是主叫号码 z_i 的违规特征值,

$$Y = \sum_{t=1}^N \alpha_t x_{it} + \beta, N \text{ 是特征指标数, } \alpha_t \text{ 是特征指标 } t \text{ 的权重系数,}$$

x_{it} 是主叫号码 z_i 的特征指标 t 的值, β 是极大似然估计值,然后判断主叫号码的违规可疑度指数是否大于违规可疑度指数的阈值。如果是,说明该主叫号码是违规电话或疑似违规电话;如果否,说明该主叫号码不是违规号码或疑似违规号码,从违规号码簇或疑似违规号码簇中删除所述主叫号码。

3.3 取证审核

取证审核是指将通过大数据分析得出的疑似号码进行取证和审核确认,审核分为人工审核和自动审核两种。通过引入语音识别技术对取证文件进行判别,可以提高系统的审核效率,目前市场上也已经出现基于语音识别的骚扰电话呼叫检测技术^[10]。

人工审核是业务开展前期的一种审核手段,违规语音电话管控平台通过大数据分析得出高疑似号码,再通过 SCP 将呼叫实时路由到取证平台,对这些疑似号码进行取证,由客服中心通过回拨、第三方互联网公司对比等手段进行审核确认,确认为诈骗号码的进入拦截管控环节。人工审核准确率高,但是效率相对较低。

自动审核功能采用了频谱分析比对技术,速度快效率高,是人工审核的必要补充。频谱分析比对技术是使用基于能量和时间双维度语音对比的核心技术,将疑似号码的频谱与样本库进行比对,当某号码的全部频谱特征都与样本库成功匹配时,将该号码视为确定的诈骗号码,后续该号码的所有来电由拦截模块进行拦截管控。利用频谱分析比对技术进行自动审核,大大提高了审核的周期和效率,放大了人工审核的价值。

3.4 拦截管控

拦截管控节点是违规语音电话治理的关键环节,在试点期间采用直接治理和间接治理两种治理方式,都取得了一定的效果。

3.4.1 直接治理

直接治理的思路是将违规语音电话直接拦截,目前违规语音电话管控平台能实现按网元拦截、号首+号长拦截和 DTMF(double tone multifrequency, 双音多频,即被叫的按键行为)拦截等多种拦截手段。

按网元拦截对已经确认的诈骗号码来电进行拦截,诈骗电话监测系统可以根据来电触发的 GT 码区分网元,分别进行拦截。

号首+号长拦截针对不规范国际主叫,开发号首+号长匹配拦截,即对类似 0057 开头的号段、位长不超过 10 的主叫号码进行拦截。

DTMF 拦截针对具有群呼行为的诈骗电话进行拦截,先播放一段录音,诱导被叫按键转人工服务,通过检测被叫按键行为进行拦截。系统将可能带有 DTMF 信号的疑似异常呼叫路由至 DTMF 检测平台,根据 DTMF 消息产生时间和次数等灵活策略来准确定位诈骗电话,并进行实时拦截。

3.4.2 间接治理

间接治理是一种放行转接功能,放行转接是将部分诈骗号码来电路由到 IIP(independent intelligent peripheral, 独立智能外设)进行接通,对主叫放音。一方面,麻痹主叫侧的犯罪分子,以减少主叫频繁换号给系统增加的追踪难度;另一方面,通过接通诈骗来电对诈骗通话进一步分析,发现更多可疑特征。

3.5 信息提醒及反馈

信息提醒及反馈包括:对省内移动用户在接听电话的过程中进行事中提醒;挂机后被被叫用户下发短信,并鼓励用户进行众包举报。

(1) 事中提醒

移动用户在接听违规语音来电过程中,平台获取呼叫事件后下发 USSD (unstructured supplementary service data, 非结构化补充数据)或闪信提醒,手机在振铃过程中即刻弹出提示信息,提醒用户当前来电是疑似诈骗或者被多少人举报为诈骗、骚扰电话,帮助用户警惕通话中的诈骗行为。当被叫用户处于关机、正忙或不可接听等状态时,系统不会下发提醒。

(2) 众包举报

被叫用户接听违规语音来电后,平台在获取通话结束

表 2 浙江移动违规语音业务统计(一周数据)

统计量	骚扰电话			诈骗电话	网间虚假移动主叫
	一次振铃电话	违规语音电话	携带 DTMF		
甄别及拦截呼叫数量/次	12 137 970	3 406 777	9 276	960 868	831 024

表 3 区分来源的违规语音电话统计(一周数据)

统计量	他网				长途电话		省内网内		
	骚扰电话	诈骗电话	虚假移动主叫	DTMF 诈骗	诈骗电话	DTMF 诈骗	骚扰电话	复制卡	骚扰电话
甄别及拦截呼叫数量/次	14 644 747	940 736	831 024	7 900	20 132	1 376	909 842	5 500	1 664

事件后下发短信,提醒之前的来电是疑似诈骗或者被多少人举报为诈骗、骚扰电话,并邀请用户参与举报互动,通过系统的自学习能力来发现更多的诈骗号码,从而提高系统的诈骗号码识别率。

事中提醒和众包举报功能利用了运营商特有的用户资源来探测未知特征的可疑号码,拓宽采集过程覆盖的号码范围,以追求 100%的覆盖率。同时与审核系统 100%的准确率相配合,使系统在覆盖率和准确率这两个相互矛盾的问题上实现双一百的效益。

4 业务治理实施效果

浙江移动自 2014 年 5 月启动国际诈骗电话治理以来,取得了明显的效果。针对国际诈骗电话治理的良好效果,2014 年 9 月将诈骗电话的治理扩大试点到全国 10 个重点省份,截至 2015 年 3 月底,平台共拦截各类违规呼叫逾 4 亿次,发现国际诈骗电话 9 289 个,拦截诈骗电话呼叫达 3.4 亿次,平均每日拦截诈骗电话呼叫近 200 万次、400 呼叫 11 万次、虚假主叫 13 万次。

同期针对浙江移动的违规语音综合治理也取得了明显效果,针对骚扰电话、诈骗电话、网间虚假主叫的治理效果见表 2。每周甄别大量的违规业务号码,同时对这些违规号码进行实时拦截和管控,最大程度保护移动用户的通信安全。区分来源的违规语音电话统计见表 3。由表 3 可知,来自他网的诈骗电话和骚扰电话相对较多,所以网间来电是违规语音电话治理的重点。

5 结束语

浙江移动利用大数据技术进行违规语音电话治理,建立集中分析、全面布控的违规语音管控平台,取得了一定的效果。随着违规语音电话治理的不断深入,违规语音电话的剧本和技术手段也在不断更新,也更具有迷惑性,治

理的难度也在逐渐增加,所以,违规语音电话的治理是一个长期动态的过程,需要长期与不法分子斗智斗勇。另外,目前运营商网内的数据源基本上还是“部门所有制”^[11],今后大数据平台将采集更全的数据进行主被叫行为分析,更有助于违规语音电话的治理。

由于通信网络是由各运营商互联互通组成的,一点接入,全球可达。因此,要实现违规语音电话的根本治理,还有待于各运营商分别落实网内治理责任,从呼叫源头拒绝违规业务接入,银行和公安各相关部门各司其职,多部门协同作战,全方位保护人民群众利益。

参考文献

1 杜磊. 骚扰电话何时休. 计算机与网络, 2015, 41(3)
Du L. Harassing phone calls. Computer and Network, 2015, 41(3)

2 王文超, 房一童. 网间虚假主叫号码处理方案研究. 电信网技术, 2010(9): 65~68
Wang W C, Fang Y T. Study on the treatment of inter network false calling number. Telecommunications Network Technology, 2010(9): 65~68

3 李程, 李立卡. 基于大数据处理技术的虚假主叫号码甄别探讨. 电信技术, 2014(10): 15~18
Li C, Li L K. Study on data processing technology based on the false calling number screening. Telecommunications Technology, 2014(10): 15~18

4 安茂波. 诈骗电话特点分析及防范. 中国信息安全, 2014(5): 86~89
An M B. Characteristics of telephone fraud and its prevention. China Information Security, 2014(5): 86~89

5 计国华, 邵四清. 上海移动骚扰电话分析拦截系统应用案例. 电信网技术, 2011(5): 82~87
Ji G H, Shao S Q. Case of annoyance call analysis and interception system of shanghai mobile. Telecommunications Network Technology, 2011(5): 82~87



- 6 彭华东. 某地移动骚扰电话整治总结. 电脑知识与技术, 2014(15): 3647~3661
Peng H D. Summary of mobile phone and harassing phone calls in a certain place. Computer Knowledge and Technology, 2014(15): 3647~3661
- 7 李立卡, 马泽雄, 陈庆年等. 电话诈骗防治技术解决方案与运维对策研究. 电信科学, 2014, 30(11): 166~172
Li L K, Ma Z X, Chen Q N, *et al.* Research of technology solutions and operation countermeasures to telephone fraud prevention and control. Telecommunications Science, 2014, 30(11): 166~172
- 8 刘碧波. 电信网内恶意呼叫的识别方法. 电信技术, 2014(10): 67~68
Liu B B. Identification method of malicious call in telecommunication network. Telecommunications Technology, 2014(10): 67~68
- 9 黄勇军, 冯明, 丁圣勇等. 电信运营商大数据发展策略探讨. 电信科学, 2013, 29(3): 7~11
Huang Y J, Feng M, Ding S Y, *et al.* Discussion on the development strategy of telecom operators. Telecommunications Science, 2013, 29(3): 7~11
- 10 杜海涛, 张峰, 高曼颖等. 基于语音识别的骚扰电话呼叫检测技术分析. 电信工程技术与标准化, 2014(12): 5~9
Du H T, Zhang F, Gao M Y, *et al.* Analysis and experiment of disturbing call detection technology based on speech recognition. Telecommunications Engineering Technology and Standardization, 2014(12): 5~9

- 11 胡舜耕, 魏进武. 大数据及其在电信运营中的应用研究. 电信技术, 2015(1): 14~17
Hu S G, Wei J W. Research on large data and its application in telecom operation. Telecommunications Technology, 2015(1): 14~17

[作者简介]



朱宏毅,男,中国移动通信集团浙江有限公司网络管理中心高级工程师,主要从事交换网络维护管理工作。

林建洪,男,东信北邮信息技术有限公司工程中心技术专家,主要从事项目管理和新产品研究工作。

王彦青,男,东信北邮信息技术有限公司数据业务产品中心数据分析师,主要从事算法研究工作。

张锦然,女,东信北邮信息技术有限公司电信业务产品中心高级工程师,主要从事新产品研究工作。

王敬宇,男,东信北邮信息技术有限公司战略研究部技术专家,主要从事网络和大数据技术研究工作。

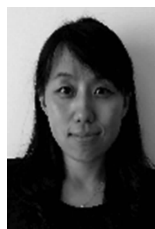
2015187-9

(上接 2015173-6)

[作者简介]



蔡永顺,男,中国电信股份有限公司北京研究院工程师,主要研究方向为服务器虚拟化、云资源管理与监控、定制化服务器等。



司伟,女,中国电信股份有限公司北京研究院高级工程师,主要从事虚拟化安全、云计算系统架构、大数据等工作。



杨明川,男,中国电信股份有限公司北京研究院云计算与大数据产品线总监,主要研究方向为大数据平台与业务关键技术。



饶少阳,男,中国电信股份有限公司北京研究院高级工程师,主要从事云计算技术与产品研发工作。

2015173-7