

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2017 年 5 月 11 日 (11.05.2017)



(10) 国际公布号
WO 2017/076314 A1

- (51) 国际专利分类号:
H04M 1/663 (2006.01) H04L 29/08 (2006.01)
- (21) 国际申请号: PCT/CN2016/104504
- (22) 国际申请日: 2016 年 11 月 4 日 (04.11.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510752441.1 2015 年 11 月 6 日 (06.11.2015) CN
- (71) 申请人: 北京奇虎科技有限公司 (BEIJING QIHOO TECHNOLOGY COMPANY LIMITED) [CN/CN]; 中国北京市西城区新街口外大街 28 号 D 座 112 室 (德胜园区), Beijing 100088 (CN)。奇智软件 (北京) 有限公司 (QIZHI SOFTWARE (BEIJING) COMPANY LIMITED) [CN/CN]; 中国北京市朝阳区酒仙桥路 6 号院 2 号楼 B 座 2 层、3 层 301-306 室, Beijing 100015 (CN)。
- (72) 发明人: 王务志 (WANG, Wuzhi); 中国北京市朝阳区酒仙桥路 6 号院 2 号楼 B 座 2 层、3 层 301-306 室, Beijing 100015 (CN)。王军 (WANG, Jun); 中国

北京市朝阳区酒仙桥路 6 号院 2 号楼 B 座 2 层、3 层 301-306 室, Beijing 100015 (CN)。

(74) 代理人: 北京市浩天知识产权代理事务所 (普通合伙) (HYLANDS LAW FIRM); 中国北京市朝阳区朝阳门外大街 18 号丰联广场 A 座 15 层 1511, Beijing 100020 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO,

[见续页]

(54) Title: PROCESSING METHOD AND SYSTEM FOR ADAPTIVE UNWANTED CALL IDENTIFICATION

(54) 发明名称: 自适应识别骚扰电话的处理方法及系统

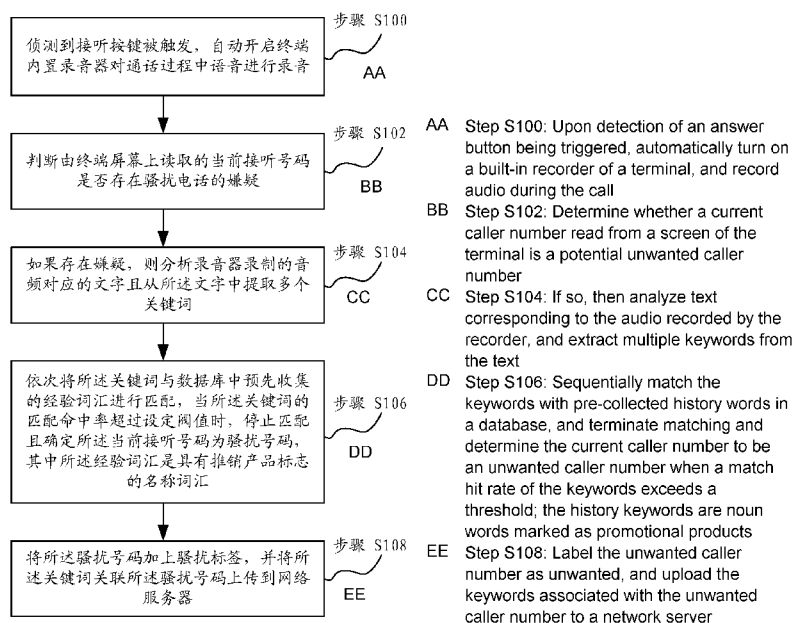


图 1

(57) Abstract: The present application discloses a processing method and system for adaptive unwanted call identification. The method comprises: upon detection of an answer button being triggered, automatically turning on a built-in recorder of a terminal, and recording audio during the call; determining whether a current caller number read from a screen of the terminal is a potential unwanted caller number; if so, then analyzing text corresponding to the audio recorded by the recorder, and extracting multiple keywords from the text; sequentially matching the keywords with pre-collected history words in a database; and terminating matching and determining the current caller number to be an unwanted caller number when a match hit rate of the keywords exceeds a threshold. The present application identifies and configures whether an unknown number is an unwanted caller number by recording a call and extracting keywords from the recording to be matched.

(57) 摘要: 本申请公开了一种自适应识别骚扰电话的处理方法及系统, 包括: 检测到接听按键被触发, 自动开启终端内置录音器对通话过程中语音进行录音;

判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑, 如果存在嫌疑, 则分析录音器录制的音频对应的文字且从所述文字中提取多个关键词, 依次将所述关键词与数据库中预先收集的经验词汇进行匹配, 当所述关键词的匹配命中率超过设定阈值时, 停止匹配且确定所述当前接听号码为骚扰号码。本申请通过对通话过程录音并提取录音中关键词进行匹配, 判别和设置未知号码是否为骚扰电话。



RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, **本国际公布:**

CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, — 包括国际检索报告(条约第 21 条(3))。
TG)。

自适应识别骚扰电话的处理方法及系统

相关申请的交叉参考

本申请主张在 2015 年 11 月 6 日在中国提交的中国专利申请号 No. 201510752441.1 的优先权，其全部内容通过引用包含于此。

技术领域

本申请属于移动通信领域，具体地说，涉及一种自适应识别骚扰电话的处理方法及系统。

10

背景技术

随着通信技术的发展，越来越多的商家、不法分子为了达到盈利的目的，借助于手机、固定电话、网络等通信技术进行广告推销、非法宣传和金融诈骗，造成很多电话用户经常被骚扰，甚至可能给个人带来经济损失。而现有技术主要是通过记录拦截电话号码实现拦截，在业务控制点 SCP 上保存记录表，该记录表中包括一个或多个待拦截电话号码，在接收业务交换点发送的包括主叫号码和被叫号码的初始检测点消息时，业务控制点根据记录表判断该主叫号码是否为配置的待拦截电话号码，获取判断结果，再根据判断结果对该呼叫进行处理，以拦截该被叫号码对应用户所需要拦截的呼叫，但是上述拦截方法只能拦截被叫用户设定的黑名单号码，不能判别和自动设置未设定的未知电话号码是否为骚扰电话。

20

因此，如何解决无法判别和自动设置未设定的未知电话号码是否为骚扰电话，便成为亟待解决的技术问题。

发明内容

25

有鉴于此，本申请所要解决的技术问题是提供了一种自适应识别骚扰电话的处理方法及系统，以解决现有技术中无法判别和自动设置未设定的未知电话号码是否为骚扰电话的问题。

为了解决上述技术问题，本申请公开了一种自适应识别骚扰电话的处理方法，应用于移动终端，包括：侦测到接听按键被触发，自动开启终端内置录音器对通话过程中语音进行录音；判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，如果存在嫌疑，则分析录音器录制的音频对应的文字且从所述文字中提取多个关键词，依次将所述关键词与数据库中预先收集的经验词汇进行匹配，当所述关键词的匹配命中率超过设定阈值时，停止匹配且确定所述当前接听号码为骚扰号码，其中所述经验词汇是骚扰电话的通话内容中的标志性词汇；将所述骚扰号码加上骚扰标签，并将所述关键词关联所述骚扰号码上传到网络服务器。

10 本申请公开了一种自适应识别骚扰电话的处理系统，应用于移动终端，包括：触发录音模块，用于侦测到接听按键被触发，自动开启终端内置录音器对通话过程中语音进行录音；语音匹配模块，用于判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，如果存在嫌疑，则分析录音器录制的音频对应的文字且从所述文字中提取多个关键词，依次将所述关键词与数据库中预先收集的经验词汇进行匹配，当所述关键词的匹配命中率超过设定阈值时，停止匹配且确定所述当前接听号码为骚扰号码，其中所述经验词汇是骚扰电话的通话内容中的标志性词汇；关联存储模块，用于将所述骚扰号码加上骚扰标签，并将所述关键词关联所述骚扰号码上传到网络服务器。

20 本申请公开了一种计算机程序，其包括计算机可读代码，当计算机可读代码在计算设备上运行时，导致计算设备执行上述的自适应识别骚扰电话的处理方法。

本申请公开了一种计算机可读介质，其中存储了上述计算机程序。

与现有技术相比，本申请可以获得包括以下技术效果：

25 1) 通过对通话过程录音，并从录音中提取多个关键词，如果关键词与预先收集的经验词汇匹配命中率高于设定阈值，确定该号码为骚扰电话，实现了判别和自动设置未设定的未知电话号码是否为骚扰电话。

2) 从录音中提取多个关键词，如果关键词与预先收集的经验词汇匹配命中率高于设定阈值，确定该号码为骚扰电话，多个关键词的匹配降低了误判骚扰电话的概率。

3) 将所述骚扰号码加上骚扰标签, 并将所述关键词关联的所述骚扰号码上传到网络服务器, 使得其它终端再次被该号码呼叫时可以提示, 通过网络实现将各终端的骚扰号码汇总分享且提示。

当然, 实施本申请的任一产品必不一定需要同时达到以上所述的所有技术效果。

附图概述

此处所说明的附图用来提供对本申请的进一步理解, 构成本申请的一部分, 本申请的示意性实施例及其说明用于解释本申请, 并不构成对本申请的不当限定。在附图中:

图 1 是本申请实施例的识别骚扰电话的处理方法;

图 2 是本申请实施例的又一识别骚扰电话的处理方法;

图 3 是本申请实施例的又一识别骚扰电话的处理方法;

图 4 是本申请实施例的识别骚扰电话的处理系统的架构图;

图 5 是本申请实施例的用于执行根据本发明实施例的自适应识别骚扰电话的处理方法的计算设备的框图;

图 6 是本申请实施例的用于保持或者携带实现根据本发明实施例的自适应识别骚扰电话的处理方法的程序代码的存储单元。

20 本发明的较佳实施方式

以下将配合附图及实施例来详细说明本申请的实施方式, 藉此对本申请如何应用技术手段来解决技术问题并达成技术功效的实现过程能充分理解并据以实施。

当前对骚扰电话的拦截, 主要是通过拦截被叫用户设定的黑名单号码, 不能判别和自动设置未设定的未知电话号码是否为骚扰电话, 而本申请通过对通话过程录音, 并从录音中提取多个关键词, 如果关键词与预先收集的经验词汇匹配命中率高于设定阈值, 确定该号码为骚扰电话, 并将骚扰号码加上骚扰标签关联关键词上传到网络服务器, 不仅可以判别和自动设置未设定

的未知电话号码是否为骚扰电话，还可以通过网络实现将各终端的骚扰号码汇总分享且提示。

图 1 所示，为本申请一个实施例所述的识别骚扰电话的处理方法，包括：

- 5 步骤 S100，侦测到接听按键被触发，自动开启终端内置录音器对通话过程中语音进行录音；

来电时，侦测到用户触发接听按键后，自动开启终端内置录音器或者录音软件对通话过程中语音进行录音，用户自行设置是由终端内置录音器或者录音软件进行录音，可以选择录音音频质量较好的设备进行录音；

- 10 步骤 S102，判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，如果存在嫌疑，则执行步骤 S104；

- 15 通话时，录音器会对通话过程中语音进行录音，无论来电是正常电话还是骚扰电话，都需要通过对录制的对话音频分析才能得知，而且当接听电话明显不是骚扰电话时，依然需要对录制的对话音频进行分析操作，相当于重复操作，增加了系统运行负担和存储负担，所以在对录制的对话音频进行分析操作之前可以进行嫌疑判断，可以应用以下几种方式：

- 20 （1）来电话时用户看到陌生号码，尤其是另类号码，第一反应就是认为可能是推销或者诈骗等骚扰电话，因此，本申请采取方式 1：将所述当前接听号码与终端中存储号码逐一进行匹配，若匹配不成功则所述接听号码为陌生号码，且确定所述当前接听号码存在骚扰电话的嫌疑，其中，终端中存储号码包括通讯录、通话记录和短信箱中存储的所有号码。遍历匹配通讯录、通话记录和短信箱中的所有号码，可以及时通过陌生号码的条件确定存在骚扰电话的嫌疑，而且可以防止将用户的最新通讯未存储的号码认为是陌生号码。

- 25 用户接听电话录音过程中或录音结束后（比如：极短时间挂断电话录音也就随之结束），将当前接听号码先与通讯录中的所有电话号码依次进行匹配，当匹配成功时则不存在骚扰嫌疑且停止匹配；若当前接听号码与通讯录中所有号码都没有匹配成功，则继续与通话记录中的所有号码进行匹配，当匹配成功时则不存在骚扰嫌疑且停止匹配；若当前接听号码与通话记录中所

有号码都没有匹配成功，则继续与短信箱中的所有号码进行匹配，当匹配成功时则不存在骚扰嫌疑且停止匹配，若当前接听号码与短信箱中所有号码都没有匹配成功，则确定所述当前接听号码存在骚扰电话的嫌疑；如果电话号码不是陌生号码时，与通讯录中号码匹配成功度最高，按照通讯录、通话记录5 和短信箱的依次匹配顺序，可以降低匹配工作强度。

(2) 用户接到骚扰电话时，由于对其讲述内容并不感兴趣或者存有防范心理，通常用户会选择很快挂断电话以示拒绝，因此，本申请采取方式 2：判断接听通话时间是否小于预设时间阈值，如果是，则确定存在骚扰电话的嫌疑，其中预设时间阈值是统计骚扰电话挂断时间的平均值。

10 用户接听电话后录音器对通话过程中语音进行录音同时计时，当计时时间超过预设时间阈值，则确定不存在骚扰电话的嫌疑；当录音结束（用户挂断电话后录音自行结束，或者用户手动操作停止录音）且计时时间小于预设时间阈值，如果是则确定存在骚扰电话的嫌疑。通过计时通话时间长短判断是否存在骚扰电话的嫌疑，一方面是操作简单易于获取时间值，另一方面是15 不存在任何分析阶段误差，得到的结果更加直接。

例如：用户接听电话后，手机屏幕上计时器会自动对通话时间进行计时并显示，如果对方是用户熟悉或者知道的人，用户对其讲述内容自然不会毫无兴趣而很快挂断电话，如果对方是推销或者诈骗的潜在人员，用户通常会对其讲述内容并不感兴趣从而想尽快结束通话，继续做用户正在做的事情，20 所以，利用判断通话时间长短的方式来确定骚扰电话的嫌疑，不仅操作直接而且准确度较高。

(3) 用户接到骚扰电话时，由于对其讲述内容并不感兴趣或者存有防范心理，出于礼貌用户通常会委婉拒绝后挂断电话，但是有时基于对方不断讲述不同的新内容，用户可能会对讲述内容进行了解后再次进行委婉拒绝，25 总之，无论用户对讲述内容是否了解，用户的讲话过程中一定会出现委婉拒绝词汇，因此，本申请采取方式 3：将麦克风采集的用户声音片段与数据库中预存音频片段进行匹配，若音频匹配相似度大于预设相似阈值，则确定存在骚扰电话的嫌疑，其中，所述预存音频片段是含有拒绝标志词汇的音频。将麦克风识别到的用户声音直接做语义分析，由于麦克风直接识别到的用户

声音更加立体且直接，所以语义分析后对应的文字准确度更高，而且使得语义分析和音频匹配的过程更加便捷。

预存音频片段是含有拒绝标志词汇的音频，例如：用户说“谢谢、不需要、不用了、没必要”等委婉拒绝词汇。

- 5 用户可以预先设置对麦克风采集的用户声音是定时多段录音还是一片全程录音，若用户设定的是定时多段录音（用户每次讲话停顿时间较长时为一录音片段），则将每段录音依次与数据库中预存音频片段进行匹配，当出现录音匹配相似度大于预设相似阈值时，则确定存在骚扰电话的嫌疑且停止剩余音频片段的匹配；若用户设定的是一片全程录音，将该全程录音与数据库
- 10 中预存音频片段进行匹配，如音频匹配相似度大于预设相似阈值，则确定存在骚扰电话的嫌疑。录制的音频片段与预存音频片段的匹配实际上就是音频片段中拒绝词汇的匹配，用户根据自己的实际需求设定是定时多段录音还是一片全程录音，定时多段录音匹配过程较快但匹配准确度较低，一片全程录音匹配过程较慢但匹配准确度较高。

- 15 例如：用户接听电话后，如果发现对方想要推销自己的产品，用户可能委婉说出拒绝词汇，在用户拒绝的过程中对方会尽可能讲到吸引用户的内容，从而用户可能讲述对其产品进行了解的话语，而在用户讲话的同时，麦克风实时采集用户声音，系统在后台将采集到的声音片段语义分析为对应文字，并将该采集到的声音片段与含有拒绝标志性词汇的音频进行匹配，当相似
- 20 度较高时确定存在骚扰电话的嫌疑，以便进行后续处理。

当然，上述三种方式都存在一定的误差概率，因此可以考虑通过以上方式协同组合的方式进行判断，来降低误差，具体组合方式如下：

方式 1+方式 2

- 25 将所述当前接听号码与终端中存储号码逐一进行匹配，若匹配不成功则所述接听号码为陌生号码；再判断接听通话时间是否小于预设时间阈值，如果是则确定存在骚扰电话的嫌疑。

首先判断当前号码是否为陌生号码，如果当前号码不是陌生号码那么必然不是骚扰电话，所以也无需再进行任何后续操作来确定该号码是否为骚扰电话；如果当前号码是陌生号码，再继续判断其通话时间是否小于预设时间

阈值，若当前号码是陌生号码且通话时间小于预设时间阈值，则确定存在骚扰电话的嫌疑。

- 5 例如：当前号码不是陌生号码而是手机中存储号码时，也可能因为时间仓促、讲述内容简单等出现挂断电话时间短的现象，通过将判断陌生号码和判断通话时间相结合的方式来确定是否存在骚扰电话的嫌疑，可以避免在这种情况下误将当前号码确定存在骚扰电话的嫌疑，提高判断准确度。

方式 1+方式 3

- 10 将所述当前接听号码与终端中存储号码逐一进行匹配，若匹配不成功则所述接听号码为陌生号码；再将麦克风采集的用户声音片段与数据库中预存音频片段进行匹配，若音频匹配相似度大于预设相似阈值，则确定存在骚扰电话的嫌疑。

- 15 首先判断当前号码是否为陌生号码，如果当前号码不是陌生号码那么必然不是骚扰电话，所以也无需再进行任何后续操作确定该号码是否为骚扰电话；如果当前号码是陌生号码，再继续将麦克风识别的用户声音片段与预存音频片段进行匹配，若当前号码是陌生号码且音频匹配相似度大于预设相似阈值，则确定存在骚扰电话的嫌疑。

- 20 例如：当前号码不是是陌生号码而是手机中存储号码时，也可能因为对方出游想给用户带些礼物等出现用户委婉拒绝的现象，通过将判断陌生号码和判断音频匹配相结合的方式来确定是否存在骚扰电话的嫌疑，可以避免在这种情况下误将当前号码确定存在骚扰电话的嫌疑，提高判断准确度。

步骤 S104, 分析录音器录制的音频对应的文字且从所述文字中提取多个关键词；

- 25 无论是通过上述哪种方式来判断骚扰电话的嫌疑，为了避免重复判断的操作，只有在确定当前号码存在骚扰电话嫌疑的情况下，才继续分析录音器录制的通话过程双方说话的音频。而且如果通话时间极短（录音随着通话结束而结束），那么分析录制音频的过程发生在用户挂断电话后，如果用户通话时间不是极短的情况下；那么分析录制音频的过程发生在录音过程中，实时对录音进行分析，但无论是发生在挂断电话后还录音过程中，分析录制音频的过程如下：

利用音频转换器将录制的音频转换成对应的文本，然后分析文本中文字并对所述文字进行分词，根据数据库中的词性对照表判断所述每个分词的词性，选取一定数量的词性为动词和名词的分词作为关键词，对关键词的数量和词性分别进行限制，提高了关键词质量，进而使得匹配准确度较高。

- 5 例如:音频转换后的文本为“您好，我们是某某理财公司，我们公司新推出针对不同工薪族的多种理财方案”，我们可以一句一句对该文本进行分词划分为：您、好，我们、是、某某理财公司，我们、公司、新、推出、针对、不同工薪族、的、多种、理财方案；然后对划分之后得到的分词进行去重处理，再根据数据库中的词性对照表识别每个分词的词性，由于动词和名词作
10 为一个文本的关键词的可能性较大，所以从动词和名词中选取一定数量的分词作为关键词。

- 步骤 S106，依次将所述关键词与数据库中预先收集的经验词汇进行匹配，当所述关键词的匹配命中率超过设定阈值时，停止匹配且确定所述当前接听号码为骚扰号码，其中所述经验词汇是骚扰电话的通话内容中的标志性
15 词汇；

当通话时间过于简短时，则录制音频转化为的文字势必会较少，进而会出现从转化后的文字中提取的关键词更少，而利用一个或两个等较少关键词匹配的结果不够客观性，准确度势必不高，所以，在将关键词与经验词汇进行匹配时还包括如下步骤：

- 20 首先将提取的关键词个数与预设最小值进行比较，如果关键词个数小于等于预设最小值时，则判断结果的准确度较低，即放弃判断当前号码是否为骚扰电话，如果关键词个数大于预设最小值时，则将提取的多个关键词依次与数据库中预先收集的经验词汇进行匹配，每个关键词都分别与所有经验词汇匹配，如匹配成功则停止匹配，否则继续匹配直到所有经验词汇匹配完毕
25 后认为匹配不成功；如此继续进行下一个关键词与所有经验词汇的匹配，直到关键词的匹配命中率超过预先设定阈值时，停止匹配且确定所述当前接听号码为骚扰号码；提取多个关键词与预先收集的经验词汇匹配命中率高于设定阈值，确定该号码为骚扰电话，多个关键词的匹配降低了误判骚扰电话的概率。

其中，数据库中预先收集的经验词汇可以是保险、房屋、理财、金融产品、中奖等推销或诈骗的标志性词汇。

步骤 S108，将所述骚扰号码加上骚扰标签，并将所述关键词关联所述骚扰号码上传到网络服务器。

- 5 将所述骚扰号码加上骚扰标签以用来提示该号码为骚扰电话，并将所有匹配成功的关键词关联的所述骚扰号码上传到网络服务器，网络服务器会统计接收到该骚扰号码信息的次数，同时还需识别发送该骚扰号码信息的终端是否为同一终端，如果所有次数的信息都是由同一终端发送的，则有可能存在恶意破坏，这种情况下网络服务器不会对该骚扰号码信息进行任何操作，
- 10 如果网络服务器接收到该骚扰号码信息的次数较多并且分别来自不同的发送终端，则网络服务器对该骚扰号码及其对应的关键词进行汇总后统一生成黑名单，以便网络服务器将生成的最新黑名单及时下发到其它终端，提示其它终端该号码为骚扰号码，且通过网络实现将各终端的骚扰号码汇总分享且提示。

- 15 对于图 1 中步骤 S108 中将所述骚扰号码加上骚扰标签，并将所述关键词关联的所述骚扰号码上传到网络服务器之后，如图 2 所示，还包括如下步骤：

- 步骤 S200，接收所述网络服务器侧发来的骚扰电话黑名单，其中，所述骚扰电话黑名单为所述网络服务器根据骚扰标签对骚扰号码及对应关键词进行汇总后统一生成。
- 20

- 终端接收到网络服务器下发的最新黑名单后，及时将最新黑名单存储至终端中并删除原有的黑名单，且将该最新黑名单设置为应用拦截提示；网络服务器对各终端发送来的相同骚扰号码的发送次数进行统计，如果所有次数的信息都是由同一终端发送的，则有可能存在恶意破坏，这种情况下网络服务器不会对该骚扰号码信息进行任何操作；如果该号码的发送次数较多且分别来自不同的发送终端，则网络服务器会根据终端发送来的骚扰标签对骚扰号码及对应关键词进行汇总后统一生成骚扰电话黑名单，网络服务器不断对骚扰电话黑名单进行更新处理，将最新骚扰电话黑名单下发到各个移动终端，以便提示其它终端该号码为骚扰号码。步骤 S202，当侦测到电话呼入
- 25

时, 识别呼入的电话号码, 并在所述骚扰电话黑名单中进行遍历比对, 如命中, 则针对所述电话号码提示骚扰标签或依据用户预先设定直接挂断呼入电话并提示所述电话号码为骚扰电话。

5 当侦测到终端有电话呼入时, 识别呼入的电话号码, 并在终端接收的最新骚扰电话黑名单中进行遍历比对, 如果该呼入号码与黑名单中骚扰号码一致, 则可以利用该呼入号码对应的骚扰标签对其进行提示, 还可以依据用户预先设定直接挂断呼入电话并提示所述电话号码为骚扰电话, 通过网络实现将各终端的骚扰号码汇总分享且提示。

10 图 3 所示, 为本申请又一个实施例所述的识别骚扰电话的处理方法, 包括:

步骤 S300, 侦测到接听按键被触发, 自动开启终端内置录音器对通话过程中语音进行录音;

来电时, 侦测到用户触发接听按键后, 自动开启终端内置录音器或者录音软件对通话过程中语音进行录音;

15 步骤 S302, 判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑, 如果不存在嫌疑, 执行步骤 S304, 否则执行步骤 S306;

步骤 S304, 如果不存在嫌疑, 则删除录音器录制的相应音频;

不存在嫌疑时删除录制的相应音频, 既可以减轻移动终端的存储压力, 又可以防止终端音频存储混乱。

20 步骤 S306, 分析录音器录制的音频对应的文字且从所述文字中提取多个关键词;

25 步骤 S308, 依次将所述关键词与数据库中预先收集的经验词汇进行匹配, 当所述关键词的匹配命中率超过设定阈值时, 停止匹配且确定所述当前接听号码为骚扰号码, 其中所述经验词汇是骚扰电话的通话内容中的标志性词汇;

步骤 S3010, 将所述骚扰号码加上骚扰标签, 并将所述关键词关联所述骚扰号码上传到网络服务器。

网络服务器会统计接收到该骚扰号码信息的次数, 同时还需识别发送该

骚扰号码信息的终端是否为同一终端，如果所有次数的信息都是由同一终端发送的，则有可能存在恶意破坏，这种情况下网络服务器不会对该骚扰号码信息进行任何操作，如果网络服务器接收到该骚扰号码信息的次数较多并且分别来自不同的发送终端，则网络服务器对该骚扰号码及其对应的关键词进行汇总后统一生成黑名单，然后网络服务器将生成的最新黑名单及时下发到其它终端；

终端接收到网络服务器下发的最新黑名单后，及时将最新黑名单存储至终端中并删除原有的黑名单，且将该最新黑名单设置为应用拦截提示，比如：当识别到呼入陌生电话号码时，并在终端接收的最新骚扰电话黑名单中进行遍历比对，如果该呼入号码与黑名单中骚扰号码一致，则可以利用该呼入号码对应的骚扰标签对其进行提示，还可以依据用户预先设定直接挂断呼入电话并提示所述电话号码为骚扰电话，通过网络实现将各终端的骚扰号码汇总分享且提示。

图 3 所示的实施例之后也可以结合图 2 所示实施例内容，形成一新的实施例。

如图 4 所示，为本申请一个实施例所述的识别骚扰电话的处理系统，包括：

触发录音模块 400，用于侦测到接听按键被触发，自动开启终端内置录音器对通话过程中语音进行录音；

语音匹配模块 402，用于判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，如果存在嫌疑，则分析录音器录制的音频对应的文字且从所述文字中提取多个关键词，依次将所述关键词与数据库中预先收集的经验词汇进行匹配，当所述关键词的匹配命中率超过设定阈值时，停止匹配且确定所述当前接听号码为骚扰号码，其中所述经验词汇是骚扰电话的通话内容中的标志性词汇；

关联存储模块 404，用于将所述骚扰号码加上骚扰标签，并将所述关键词关联所述骚扰号码上传到网络服务器。

语音匹配模块 402，进一步用于将所述当前接听号码与终端中存储号码逐一进行匹配，若匹配不成功则所述接听号码为陌生号码，且确定所述当前

接听号码存在骚扰电话的嫌疑，其中所述终端中存储号码包括通讯录、通话记录和短信箱中存储的所有号码。

语音匹配模块 402，进一步用于判断接听通话时间是否小于预设时间阈值，如果是则确定存在骚扰电话的嫌疑。

- 5 语音匹配模块 402，进一步用于将麦克风采集的用户声音片段与数据库中预存音频片段进行匹配，若音频匹配相似度大于预设相似阈值，则确定存在骚扰电话的嫌疑，其中，所述预存音频片段是含有拒绝标志词汇的音频。

- 10 语音匹配模块 402，进一步用于对所述文字进行分词，根据数据库中的词性对照表判断所述每个分词的词性，选取一定数量的词性为动词和名词的分词作为关键词。

语音匹配模块 402，进一步用于如果不存在嫌疑，则删除录音器录制的相应音频。

- 15 关联存储模块 404，还用于接收所述网络服务器侧发来的骚扰电话黑名单，其中，所述骚扰电话黑名单为所述网络服务器根据骚扰标签对骚扰号码及对应关键词进行汇总后统一生成。

关联存储模块 404，进一步用于当侦测到电话呼入时，识别呼入的电话号码，并在所述骚扰电话黑名单中进行遍历比对，如命中，则针对所述电话号码提示骚扰标签或依据用户预先设定直接挂断呼入电话并提示所述电话号码为骚扰电话。

- 20 图 4 中所示出的识别骚扰电话的处理系统，与图 1-3 中所示的方法实施例特征和方案相互对应，上述对于快捷操作系统中描述的不足之处，请参考图 1-3 中所示的方法实施例的描述。

- 25 本申请通过接听电话后对通话过程进行录音，并从录音中提取多个关键词，如果关键词与预先收集的经验词汇匹配命中率高于设定阈值，确定该号码为骚扰电话，实现了判别和自动设置未设定的未知电话号码是否为骚扰电话，即可以自行判断黑名单和终端中存储号码以外的陌生号码是否为骚扰电话。

图 5 示出了可以实现根据本发明的自适应识别骚扰电话的处理方法的计

算设备。该计算设备传统上包括处理器 510 和以存储设备 520 形式的计算机程序产品或者计算机可读介质。存储设备 520 可以是诸如闪存、EEPROM(电可擦除可编程只读存储器)、EPROM、硬盘或者 ROM 之类的电子存储器。存储设备 520 具有存储用于执行上述方法中的任何方法步骤的程序代码 531 的存储空间 530。例如, 存储程序代码的存储空间 530 可以包括分别用于实现上面的方法中的各种步骤的各个程序代码 531。这些程序代码可以从一个或者多个计算机程序产品中读出或者写入到这一个或者多个计算机程序产品中。这些计算机程序产品包括诸如硬盘、紧致盘(CD)、存储卡或者软盘之类的程序代码载体。这样的计算机程序产品通常为例如图 6 所示的便携式或者固定存储单元。该存储单元可以具有与图 5 的计算设备中的存储设备 520 类似布置的存储段、存储空间等。程序代码可以例如以适当形式进行压缩。通常, 存储单元包括用于执行根据本发明的方法步骤的计算机可读代码 531', 即可以由诸如 510 之类的处理器读取的代码, 当这些代码由计算设备运行时, 导致该计算设备执行上面所描述的方法中的各个步骤。

如在说明书及权利要求当中使用了某些词汇来指称特定组件。本领域技术人员应可理解, 硬件制造商可能会用不同名词来称呼同一个组件。本说明书及权利要求并不以名称的差异来作为区分组件的方式, 而是以组件在功能上的差异来作为区分的准则。如在通篇说明书及权利要求当中所提及的“包含”为一开放式用语, 故应解释成“包含但不限于”。“大致”是指在可接收的误差范围内, 本领域技术人员能够在一定误差范围内解决所述技术问题, 基本达到所述技术效果。此外, “耦接”一词在此包含任何直接及间接的电性耦接手段。因此, 若文中描述一第一装置耦接于一第二装置, 则代表所述第一装置可直接电性耦接于所述第二装置, 或通过其他装置或耦接手段间接地电性耦接至所述第二装置。说明书后续描述为实施本申请的较佳实施方式, 然所述描述乃以说明本申请的一般原则为目的, 并非用以限定本申请的范围。本申请的保护范围当视所附权利要求所界定者为准。

还需要说明的是, 术语“包括”、“包含”或者其任何其他变体意在涵盖非排除性的包含, 从而使得包括一系列要素的商品或者系统不仅包括那些要素, 而且还包括没有明确列出的其他要素, 或者是还包括为这种商品或者系

统所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的商品或者系统中还存在另外的相同要素。

- 上述说明示出并描述了本申请的若干优选实施例，但如前所述，应当理解本实用新型并非局限于本文所披露的形式，不应看作是对其他实施例的排除，而可用于各种其他组合、修改和环境，并能够在本文所述发明构想范围内，通过上述教导或相关领域的技术或知识进行改动。而本领域人员所进行的改动和变化不脱离本申请的精神和范围，则都应在本申请所附权利要求的保护范围内。
- 5

权 利 要 求 书

1、一种自适应识别骚扰电话的处理方法，应用于移动终端，其特征在于，包括：

5 侦测到接听按键被触发，自动开启终端内置录音器对通话过程中语音进行录音；

判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，如果存在嫌疑，则分析录音器录制的音频对应的文字且从所述文字中提取多个关键词，依次将所述关键词与数据库中预先收集的经验词汇进行匹配，当所述关键词的匹配命中率超过设定阈值时，停止匹配且确定所述当前接听号码为
10 骚扰号码，其中所述经验词汇是骚扰电话的通话内容中的标志性词汇；

将所述骚扰号码加上骚扰标签，并将所述关键词关联所述骚扰号码上传到网络服务器。

2、如权利要求 1 所述的方法，其特征在于，从所述文字中提取多个关键词，进一步包括：对所述文字进行分词，根据数据库中的词性对照表判断
15 所述每个分词的词性，选取一定数量的词性为动词和名词的分词作为关键词。

3、如权利要求 1 所述的方法，其特征在于，判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，进一步包括：将所述当前接听号码与终端中存储号码逐一进行匹配，若匹配不成功则所述接听号码为陌生号
20 码，且确定所述当前接听号码存在骚扰电话的嫌疑。

4、如权利要求 3 所述的方法，其特征在于，所述终端中存储号码包括通讯录、通话记录和短信箱中存储的所有号码。

5、如权利要求 1 所述的方法，其特征在于，判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，进一步包括：判断接听通话时间是否小于预设时间阈值，如果是则确定存在骚扰电话的嫌疑。
25

6、如权利要求 1 所述的方法，判断当前接听号码是否存在骚扰电话的嫌疑，进一步包括：将麦克风采集的用户声音片段与数据库中预存音频片段进行匹配，若音频匹配相似度大于预设相似阈值，则确定存在骚扰电话的嫌

疑，其中，所述预存音频片段是含有拒绝标志词汇的音频。

7、如权利要求 1 所述的方法，接收所述网络服务器侧发来的骚扰电话黑名单，其中，所述骚扰电话黑名单为所述网络服务器根据骚扰标签对骚扰号码及对应关键词进行汇总后统一生成。

5 8、如权利要求 7 所述的方法，当侦测到电话呼入时，识别呼入的电话号码，并在所述骚扰电话黑名单中进行遍历比对，如命中，则针对所述电话号码提示骚扰标签或依据用户预先设定直接挂断呼入电话并提示所述电话号码为骚扰电话。

10 9、如权利要求 1 所述的方法，判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，进一步包括：如果不存在嫌疑，则删除录音器录制的相应音频。

10、一种自适应识别骚扰电话的处理系统，应用于移动终端，包括：

触发录音模块，用于侦测到接听按键被触发，自动开启终端内置录音器对通话过程中语音进行录音；

15 语音匹配模块，用于判断由终端屏幕上读取的当前接听号码是否存在骚扰电话的嫌疑，如果存在嫌疑，则分析录音器录制的音频对应的文字且从所述文字中提取多个关键词，依次将所述关键词与数据库中预先收集的经验词汇进行匹配，当所述关键词的匹配命中率超过设定阈值时，停止匹配且确定所述当前接听号码为骚扰号码，其中所述经验词汇是骚扰电话的通话内容中的
20 的标志性词汇；

关联存储模块，用于将所述骚扰号码加上骚扰标签，并将所述关键词关联所述骚扰号码上传到网络服务器。

11、如权利要求 10 所述的系统，所述语音匹配模块，进一步用于对所述文字进行分词，根据数据库中的词性对照表判断所述每个分词的词性，选
25 取一定数量的词性为动词和名词的分词作为关键词。

12、如权利要求 10 所述的系统，所述语音匹配模块，进一步用于将所述当前接听号码与终端中存储号码逐一进行匹配，若匹配不成功则所述接听号码为陌生号码，且确定所述当前接听号码存在骚扰电话的嫌疑。

13、如权利要求 12 所述的系统，所述语音匹配模块，进一步用于所述终端中存储号码包括通讯录、通话记录和短信箱中存储的所有号码。

14、如权利要求 10 所述的系统，所述语音匹配模块，进一步用于判断接听通话时间是否小于预设时间阈值，如果是则确定存在骚扰电话的嫌疑。

5 15、如权利要求 10 所述的系统，所述语音匹配模块，进一步用于将麦克风采集的用户声音片段与数据库中预存音频片段进行匹配，若音频匹配相似度大于预设相似阈值，则确定存在骚扰电话的嫌疑，其中，所述预存音频片段是含有拒绝标志词汇的音频。

10 16、如权利要求 10 所述的系统，所述关联存储模块，还用于接收所述网络服务器侧发来的骚扰电话黑名单，其中，所述骚扰电话黑名单为所述网络服务器根据骚扰标签对骚扰号码及对应关键词进行汇总后统一生成。

15 17、如权利要求 16 所述的系统，所述关联存储模块，进一步用于当侦测到电话呼入时，识别呼入的电话号码，并在所述骚扰电话黑名单中进行遍历比对，如命中，则针对所述电话号码提示骚扰标签或依据用户预先设定直接挂断呼入电话并提示所述电话号码为骚扰电话。

18、如权利要求 10 所述的系统，所述语音匹配模块，进一步用于如果不存在嫌疑，则删除录音器录制的相应音频。

20 19、一种计算机程序，包括计算机可读代码，当所述计算机可读代码在计算设备上运行时，导致所述计算设备执行根据权利要求 1-9 中的任一项所述的自适应识别骚扰电话的处理的方法。

20、一种计算机可读介质，其中存储了如权利要求 19 所述的计算机程序。

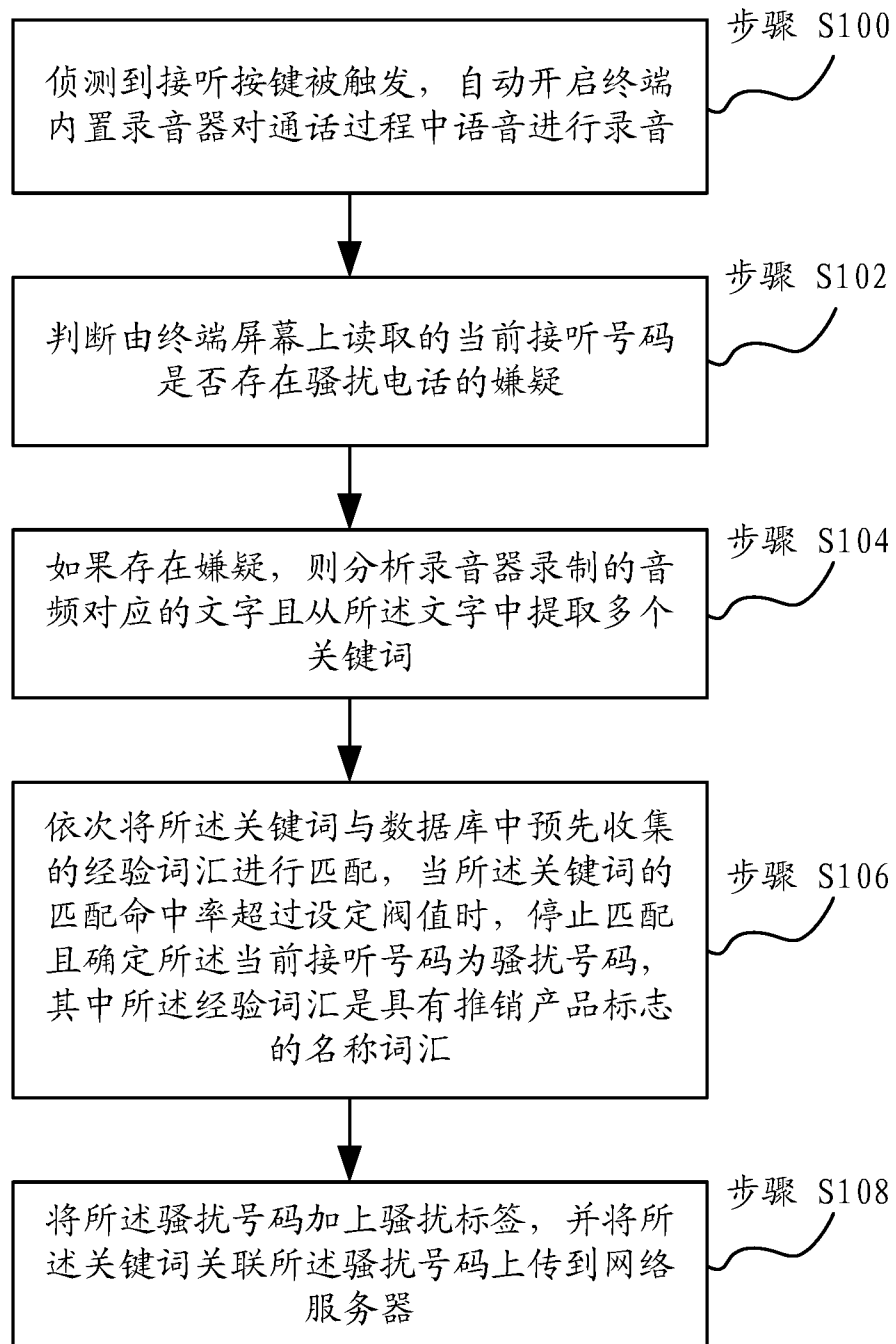


图 1

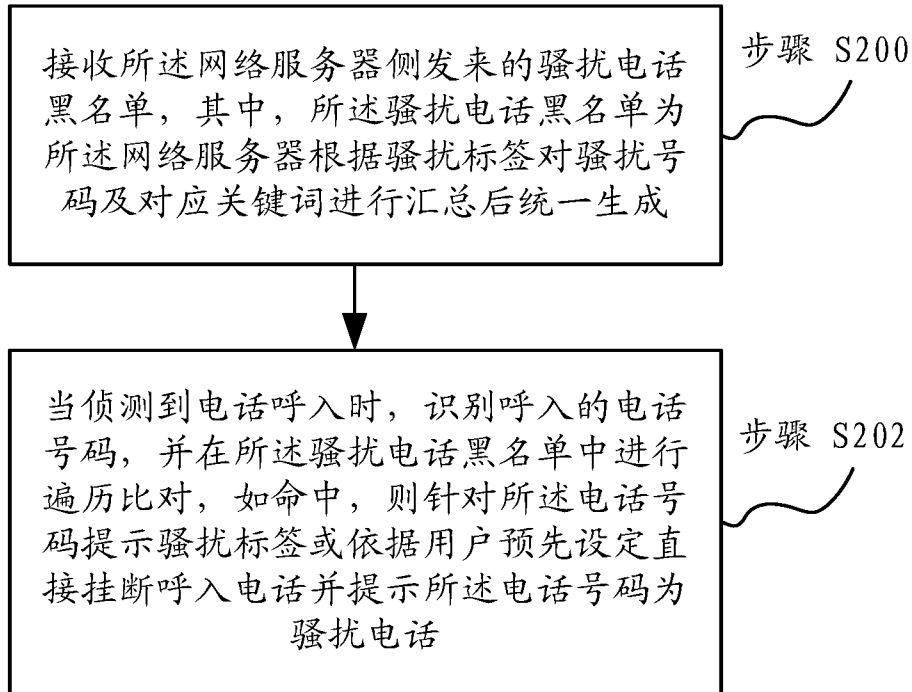


图 2

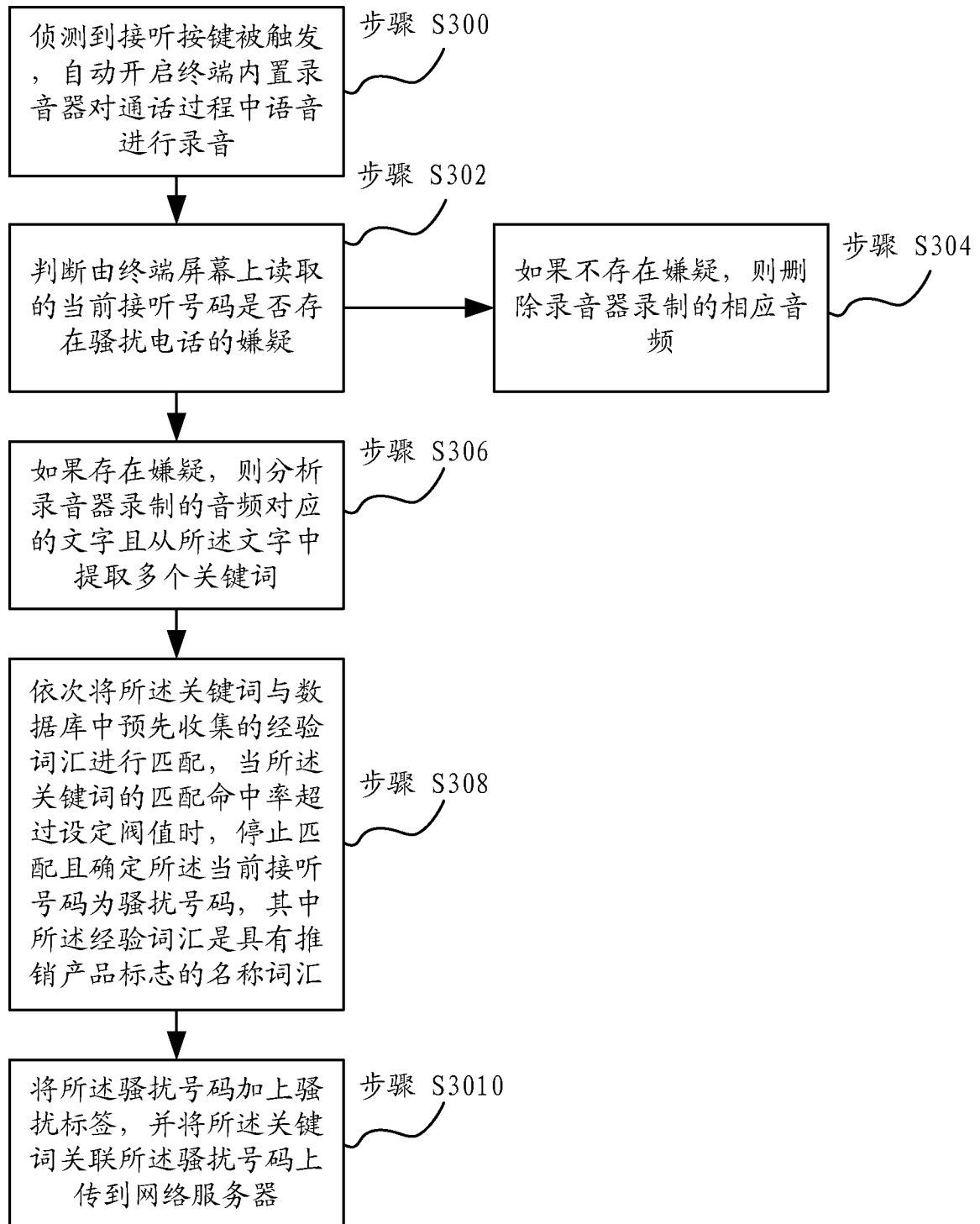


图 3

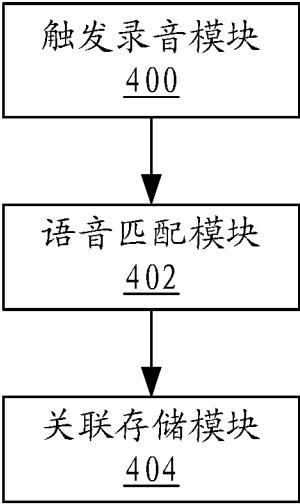


图 4

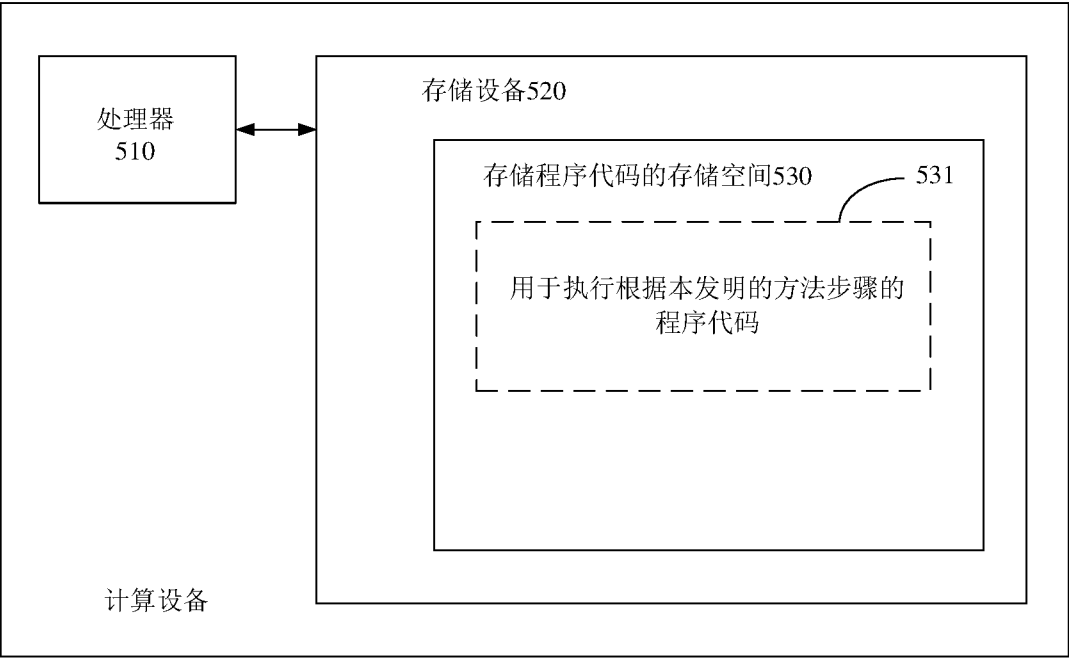


图 5

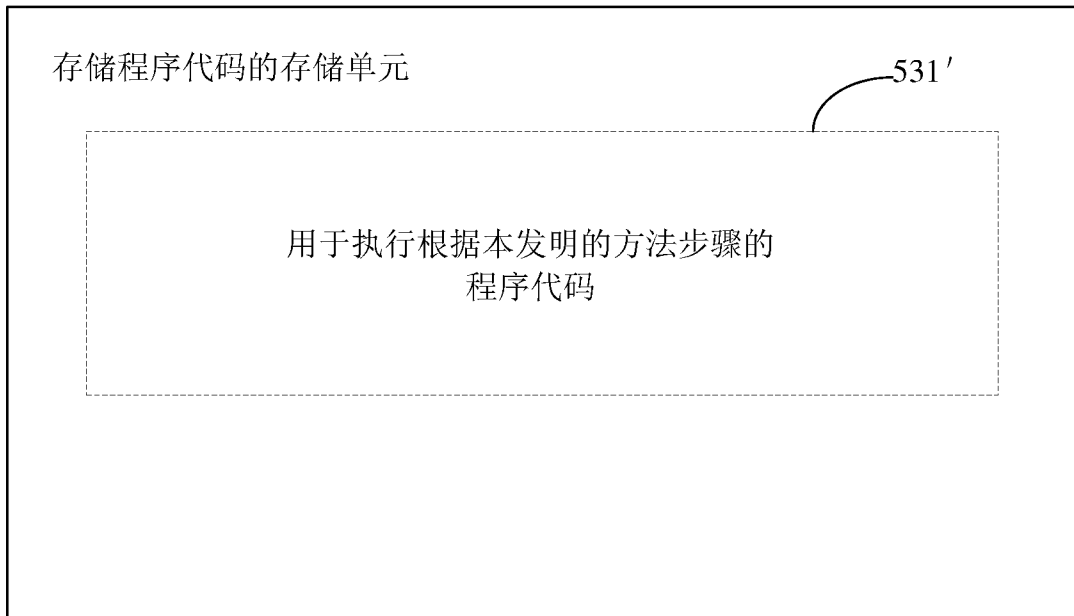


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/104504

A. CLASSIFICATION OF SUBJECT MATTER

H04M 1/663 (2006.01) i; H04L 29/08 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04M; H04Q; H04W; H04L; G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: malice, disturb, incoming call, discrimination, analysis, sound recording, phone, prank, call, identification, annoy, illegal, incoming, match, record

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105338157 A (XIAOMI TECHNOLOGY CO., LTD.), 17 February 2016 (17.02.2016), description, paragraphs [0064]-[0148], and figures 1-11	1-20
Y	CN 104253869 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.), 31 December 2014 (31.12.2014), description, paragraphs [0089]-[0141], and figure 1	1-20
Y	CN 101112078 A (NIPPON TELEGRAPH & TELEPHONE CORPORATION), 23 January 2008 (23.01.2008), description, page 54, line 4 to page 63, line 8, and figure 30	1-20
A	CN 104113466 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.), 22 October 2014 (22.10.2014), the whole document	1-20
A	CN 101827367 A (ZTE CORP.), 08 September 2010 (08.09.2010), the whole document	1-20

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 25 November 2016 (25.11.2016)	Date of mailing of the international search report 04 February 2017 (04.02.2017)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer RAN, Jian'guo Telephone No.: (86-10) 61648270

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2016/104504

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105338157 A	17 February 2016	None	
CN 104253869 A	31 December 2014	None	
CN 101112078 A	23 January 2008	WO 2006085565 A1	17 August 2006
		EP 1848192 A1	24 October 2007
		US 2009018832 A1	15 January 2009
		JP 4423327 B2	03 March 2010
CN 104113466 A	22 October 2014	WO 2014169688 A1	23 October 2014
		US 2014328478 A1	06 November 2014
CN 101827367 A	08 September 2010	None	

国际检索报告

国际申请号

PCT/CN2016/104504

A. 主题的分类

H04M 1/663 (2006.01) i; H04L 29/08 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

H04M; H04Q; H04W; H04L; G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNPAT, CNKI, WPI, EPODOC, 骚扰, 恶意, 非法, 打扰, 电话, 来电, 识别, 鉴别, 辨别, 辨识, 分析, 匹配, 录音, phone, prank, call, identification, annoy, illegal, incoming, match, record

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 105338157 A (小米科技有限责任公司) 2016年 2月 17日 (2016 - 02 - 17) 说明书第[0064]-[0148]段, 附图1-11	1-20
Y	CN 104253869 A (北京搜狗科技发展有限公司) 2014年 12月 31日 (2014 - 12 - 31) 说明书第[0089]-[0141]段, 附图1	1-20
Y	CN 101112078 A (日本电信电话株式会社) 2008年 1月 23日 (2008 - 01 - 23) 说明书第54页第4行至第63页第8行, 附图30	1-20
A	CN 104113466 A (腾讯科技深圳有限公司) 2014年 10月 22日 (2014 - 10 - 22) 全文	1-20
A	CN 101827367 A (中兴通讯股份有限公司) 2010年 9月 8日 (2010 - 09 - 08) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2016年 11月 25日

国际检索报告邮寄日期

2017年 2月 4日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

冉建国

传真号 (86-10) 62019451

电话号码 (86-10) 61648270

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/104504

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	105338157	A	2016年 2月 17日	无			
CN	104253869	A	2014年 12月 31日	无			
CN	101112078	A	2008年 1月 23日	WO	2006085565	A1	2006年 8月 17日
				EP	1848192	A1	2007年 10月 24日
				US	2009018832	A1	2009年 1月 15日
				JP	4423327	B2	2010年 3月 3日
CN	104113466	A	2014年 10月 22日	WO	2014169688	A1	2014年 10月 23日
				US	2014328478	A1	2014年 11月 6日
CN	101827367	A	2010年 9月 8日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)