

(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
2019 年 1 月 31 日 (31.01.2019)



(10) 国际公布号
WO 2019/019778 A1

(51) 国际专利分类号:
H04M 3/523 (2006.01)

(21) 国际申请号: PCT/CN2018/087327

(22) 国际申请日: 2018 年 5 月 17 日 (17.05.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710612991.2 2017 年 7 月 25 日 (25.07.2017) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY(SHENZHEN)CO.,LTD.) [CN/CN];
中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。

(72) 发明人: 陆诗(LU, Shi); 中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。 项同德(XIANG, Tongde); 中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。 杨辛未(YANG, Xinwei); 中国广东省深圳市福田区八卦岭八卦三

路平安大厦六楼, Guangdong 518052 (CN)。 张旭(ZHANG, Xu); 中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518052 (CN)。

(74) 代理人: 广州华进联合专利商标代理有限公司(ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区花城大道 85 号 3901 房, Guangdong 510623 (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, JO, JP, KE, KG, KH, 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。

(54) Title: METHOD AND APPARATUS FOR PROCESSING CALL DATA, STORAGE MEDIUM AND COMPUTER DEVICE

(54) 发明名称: 通话数据处理方法、装置、存储介质和计算机设备

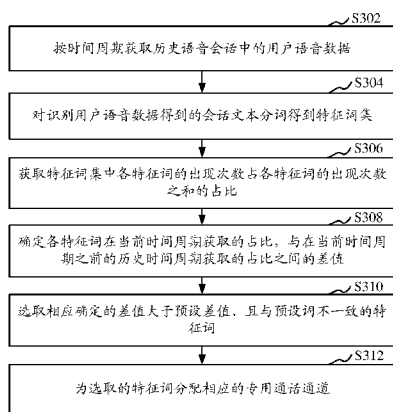


图 3

(57) Abstract: A method for processing call data, comprising: acquiring, according to a time cycle, user voice data in a historical voice session; carrying out word segmentation on a session text obtained by recognizing the user voice data, so as to obtain a set of feature words; acquiring a ratio of the number of occurrences of each feature word in the set of feature words to the sum of the number of occurrences of each feature word; determining a difference value between a ratio of each feature word acquired during the current time cycle and a ratio acquired during a historical time cycle prior to the current time cycle; selecting a feature word having a correspondingly determined difference value greater than a pre-set difference value and inconsistent with a pre-set word; and allocating a corresponding dedicated call channel for the selected feature word.

(57) 摘要: 一种通话数据处理方法, 包括: 按时间周期获取历史语音会话中的用户语音数据; 对识别所述用户语音数据得到的会话文本分词得到特征词集; 获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比; 确定各所述特征词在当前时间周期获取的占比, 与在当前时间周期之前的历史时间周期获取的占比之间的差值; 选取相应确定的差值大于预设差值、且与预设词不一致的特征词; 为选取的所述特征词分配相应的专用通话通道。

- S302 Acquire, according to a time cycle, user voice data in a historical voice session
S304 Carry out word segmentation on a session text obtained by recognizing the user voice data, so as to obtain a set of feature words
S306 Acquire a ratio of the number of occurrences of each feature word in the set of feature words to the sum of the number of occurrences of each feature word
S308 Determine a difference value between a ratio of each feature word acquired during the current time cycle and a ratio acquired during a historical time period prior to the current time cycle
S310 Select a feature word having a correspondingly determined difference value greater than a pre-set difference value and inconsistent with a pre-set word
S312 Allocate a corresponding dedicated call channel for the selected feature word

(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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

发明名称：通话数据处理方法、装置、存储介质和计算机设备**相关申请的交叉引用**

本申请要求于 2017 年 07 月 25 日提交中国专利局，申请号为 2017106129912，发明名称为“通话数据处理方法、装置、存储介质和计算机设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及计算机技术领域，特别是涉及一种通话数据处理方法、装置、存储介质和计算机设备。

背景技术

随着计算机技术的发展，使用电话与电话客服人员进行交互，通过远程通话来实现各种类型的业务办理和咨询越来越常用。

然而，发明人意识到，在传统的人工交互中，用户需要在与电话客服人员进行通话后，主动告知电话客服人员自己的需求，客服人员再被动根据用户的语音内容为用户办理相关的业务，使得电话交互的效率较低。

发明内容

根据本申请公开的各种实施例，提供一种通话数据处理方法、装置、存储介质和计算机设备。

一种通话数据处理方法，包括：

按时间周期获取历史语音会话中的用户语音数据；

对识别所述用户语音数据得到的会话文本分词得到特征词集；

获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；

确定各所述特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；

选取相应确定的差值大于预设差值、且与预设词不一致的特征词；及

为选取的所述特征词分配相应的专用通话通道。

一种通话数据处理装置，包括：

第一获取模块，用于按时间周期获取历史语音会话中的用户语音数据；

分词模块，用于对识别所述用户语音数据得到的会话文本分词得到特征词集；

第二获取模块，用于获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；

确定模块，用于确定各所述特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；

选取模块，用于选取相应确定的差值大于预设差值、且与预设词不一致的特征词；及

分配模块，用于为选取的所述特征词分配相应的专用通话通道。

一个或多个存储有计算机可读指令的非易失性存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

按时间周期获取历史语音会话中的用户语音数据；

对识别所述用户语音数据得到的会话文本分词得到特征词集；

获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；

确定各所述特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；

选取相应确定的差值大于预设差值、且与预设词不一致的特征词；及

为选取的所述特征词分配相应的专用通话通道。

一种计算机设备，包括存储器和一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述一个或多个处理器执行以下步骤：

按时间周期获取历史语音会话中的用户语音数据；

对识别所述用户语音数据得到的会话文本分词得到特征词集；

获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；

确定各所述特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；

选取相应确定的差值大于预设差值、且与预设词不一致的特征词；及

为选取的所述特征词分配相应的专用通话通道。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其它特征、

目的和优点将从说明书、附图以及权利要求书变得明显。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其它的附图。

图1为根据一个或多个实施例中通话数据处理方法的应用环境图。

图2为根据一个或多个实施例中计算机设备的框图。

图3为根据一个或多个实施例中通话数据处理方法的流程示意图。

图4为根据另一个实施例中通话数据处理方法的流程示意图。

图5为根据一个或多个实施例中通话数据处理装置的框图。

图6为根据另一个实施例中通话数据处理装置的框图。

具体实施方式

为了使本申请的技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

图1为一个实施例中通话数据处理方法的应用环境图。参照图1，该通话数据处理方法应用于通话数据处理系统。通话数据处理系统包括用户终端110、服务器120和坐席终端130，用户终端110和坐席终端130均可通过网络与服务器120通信。用户终端110可以是用户使用的移动终端，也可以是固定通话终端。服务器120可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群。坐席终端130为固定通话终端，每个坐席终端对应着一个固定的人工坐席。多个坐席终端对应一条专用通话通道。服务器120按时间周期，从用户在历史语音会话中的用户语音数据提取代表用户需求的特征词，获取这些特征词各自的出现次数占这些特征词的出现次数之和的占比，并为在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值大于预设差值、且与预设词不一致的特征词分配专用通话通道，以使用户享受坐席人员提供的咨询、业务办理等服务。

图2为一个实施例中计算机设备的框图。该计算机设备可以是图1中服务器120。如

图 2 所示, 该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。存储器包括非易失性存储介质和内存。该计算机设备的非易失性存储介质可存储操作系统和计算机可读指令, 该计算机可读指令被执行时, 可使得处理器执行一种通话数据处理方法。该计算机设备的处理器用于提供计算和控制能力, 支撑整个计算机设备的运行。该内存中可储存有计算机可读指令, 该计算机可读指令被处理器执行时, 该计算机可读指令用于实现以下各实施例所提供的一种通话数据处理方法。该计算机设备还可通过网络与用户终端和/或坐席终端连接, 接收用户终端和/或坐席终端发送的语音数据。本领域技术人员可以理解, 图 2 中示出的结构, 仅仅是与本申请方案相关的部分结构的框图, 并不构成对本申请方案所应用于其上的终端的限定, 具体的终端可以包括比图中所示更多或更少的部件, 或者组合某些部件, 或者具有不同的部件布置。

如图 3 所示, 在其中一个实施例中, 提供了一种通话数据处理方法。本实施例主要以该方法应用于上述图 1 中的服务器 120 来举例说明。参照图 3, 该通话数据处理方法具体包括如下步骤:

S302, 按时间周期获取历史语音会话中的用户语音数据。

语音会话是坐席人员与用户之间进行语音对答的会话。用户语音数据是在语音会话过程中产生的用户的语音数据。历史语音会话是已经发生的语音会话。时间周期是指进行数据获取的周期, 如一个或多个日、周或者月。

具体地, 服务器在制定时间周期后, 判断当下时间是否达到符合时间周期的周期性时间点。周期性时间点如每日预设时刻或者每月预设日。服务器在判定当下时间达到周期性时间点时, 获取历史语音会话中的用户语音数据。用户语音数据存储在服务器的数据库、文件或者缓存等。

在其中一个实施例中, 用户终端可通过拨打服务号码的方式, 与服务器通过电话网络建立语音会话, 采集用户在语音会话中产生的语音数据, 将采集的用户语音数据通过电话网络发送至服务器。服务器可将接收到的用户语音数据存储于数据库、缓存或者文件中, 以在需要的时候读取。

在其中一个实施例中, 用户终端也可以通过发起语音会话的网络请求, 与服务器建立互联网连接, 从而建立基于互联网的语音会话, 采集用户在语音会话中产生的语音数据, 将采集的用户语音数据通过互联网发送至服务器。服务器可将接收到的用户语音数据存储于数据库、缓存或者文件中, 以在需要的时候读取。

S304, 对识别用户语音数据得到的会话文本分词得到特征词集。

分词是指将一个连续的字符序列切分成多个单独的字符或者字符序列。特征词是指具有语义表达功能的字符或者字符序列。

具体地, 服务器可对用户语音数据进行特征提取, 获得待识别的用户语音特征数据, 然后基于声学模型对待识别的用户语音特征数据进行语音分帧处理得到多个音素, 根据候选字库中候选字与音素的对应关系, 将处理得到的多个音素转化为字符序列, 再利用语言模型调整转化得到的字符序列, 从而得到符合自然语言模式的会话文本。

进一步地, 服务器可采用预设的分词方式对会话文本进行分词处理, 得到多个字符或者字符序列, 从得到的字符序列中筛选出具有实际语义的字符或者字符序列作为特征词, 形成特征词集。特征词集可以包括一个或多个特征词。预设的分词方式可以是基于字符匹配、基于语义理解或者基于统计的分词方式。

更进一步地, 服务器从得到的字符或字符序列中筛选出具有实际语义的字符或者字符序列作为特征词时, 具体可从得到的字符或字符序列中过滤掉停用词。停用词是指自然语言中包括的一种功能字符或者字符序列, 这类功能字符或者字符序列并无实际语义, 包括代表语气的语气字符或字符序列和表示某种逻辑关系连接字符或字符序列等。具体地, 语气字符比如“吗”或者“呢”等, 连接字符比如“的”或“在”等, 语气字符序列比如“而已”或者“就是了”等, 连接字符序列比如“至于”或“然后”等。

S306, 获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比。

特征词的出现次数是特征词在由用户语音数据识别得到的会话文本中出现的次数。具体地, 服务器在对识别用户语音数据得到的会话文本分词得到特征词集后, 可统计特征词集中各特征词在由用户语音数据识别得到的会话文本中出现的次数, 得到各特征词的出现次数, 再计算得到各特征词的出现次数占各特征词的出现次数之和的占比。

举例说明, 特征词集中包括“健康险”、“一卡通”和“医疗保险”三个特征词。特征词“健康险”在当前时间周期获取由用户语音数据识别得到的会话文本中出现的次数为185次, 特征词“一卡通”在当前时间周期获取由用户语音数据识别得到的会话文本中出现的次数为260次, 特征词“医疗保险”在当前时间周期获取由用户语音数据识别得到的会话文本中出现的次数为97次。那么, 特征词“健康险”的出现次数占各特征词的出现次数之和的占比为: $185/(185+160+97)=41.86\%$ 。

S308, 确定各特征词在当前时间周期获取的占比, 与在当前时间周期之前的历史时

间周期获取的占比之间的差值。

具体地，服务器可对于在当前时间周期得到的特征词集中的每个特征词，获取在当前时间周期时，出现次数占各特征词的出现次数之和的占比，以及在当前时间周期之前的历史时间周期时，出现次数占各特征词的出现次数之和的占比，再计算两个占比之间的差值。在当前时间周期之前的历史时间周期可以是与当前时间周期相邻的上一时间周期，也可以是时间在当前时间周期之前的任一时间周期。

S310，选取相应确定的差值大于预设差值、且与预设词不一致的特征词。

预设差值是预先设置的特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值。预设差值被配置为选取分配专用通话通道的特征词的条件。当特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值大于预设差值，则判定该特征词满足被分配专用通话通道的条件。预设词是预先设置的特征词，且服务器已为预设词分配相应的专用通话通道。

具体地，服务器可在对于在当前时间周期得到的特征词集中的每个特征词确定该特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值后，将确定的各差值逐一与预设差值进行比较，筛选出相应确定的差值大于预设差值的特征词。服务器可再将筛选出的特征词与预设词比较，选取与预设词不一致的特征词。

S312，为选取的特征词分配相应的专用通话通道。

专用通话通道是专用于某一特定业务的通话通道。比如，专用于健康险的通话通道，或者专用于医疗保险的通话通道。具体地，服务器选取的特征词代表用户近期关注的重点，服务器可在选取特征词后为该特征词分配相应的通道，已提供相应的专属服务。

上述通话数据处理方法，按时间周期从用户在历史语音会话中的用户语音数据提取代表用户需求的特征词，获取这些特征词各自的出现次数占这些特征词的出现次数之和的占比，并确定各特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值。当特征词相应的差值大于预设差值、且与预设词不一致时，判定近期该特征词相关的业务量明显增多，自动为该特征词分配专用通话通道，这样即可通过分配的专用通话通道主动为用户提供特征词相关业务的信息，从而提高电话交互的效率。

在其中一个实施例中，步骤 S304 包括：对识别用户语音数据得到的会话文本分词得到单独词；按单独词所属语义对单独词分类得到单独词子集；挑选各单独词子集中出现次

数最高的单独词作为特征词得到特征词集。步骤 S306 包括：对于特征词集中每个特征词，获取当前特征词挑选自的单独词子集中各单独词的出现次数之和，作为当前特征词的出现次数；确定各特征词的出现次数占各特征词的出现次数之和的占比。

单独词是表示单一语义的字符或者字符序列。

在其中一个实施例中，服务器可采用基于字符匹配的分词方式进行分词处理，将会话文本按照从前到后或者从后到前的顺序逐一切分出单个字符，再将该单个字符与标准词库进行匹配。如果匹配成功，则获取该字符作为一个单独词；若匹配失败，则通过增加一个字符继续进行匹配，直至会话文本中包括的字符全部匹配完成。

在其中一个实施例中，服务器也可同时对会话文本进行正向匹配分词和逆向匹配分词。在两种分词方式的分词结果相同时，将分词得到的多个单独的字符或者字符序列作为单独词。在两种分词方式的分词结果不相同，分别计算两种分词方式得到的单独的字符或者字符序列的数量，选取计算的数量少的分词方式得到的单独的字符或者字符序列作为单独词。

进一步地，服务器可确定得到的各单独词所属的语义，将所属语义相同的单独词划分为同一类，得到多个单独词子集。对于得到的每一个单独词子集，服务器可再从中挑选出出现次数最高的单独词作为特征词，来代表相应的单独词子集，从而得到特征词集。对于特征词集中每个特征词，由于每个单独词子集中的单独词属于相同的语义，服务器可将各特征词挑选自的单独词子集中各单独词的出现次数之和，作为各特征词的出现次数；再计算各特征词的出现次数占各特征词的出现次数之和的占比。

上述实施例中，将属于相同语义的单独词划分为同一类，并从中选出出现次数最高的单独词作为代表，然后将作为代表的特征词挑选自的单独词子集中，各单独词的出现次数之和作为该特征词的出现次数，使得对特征词出现次数的统计更合理，从而使得挑选出的特征词更能反映用户近期关注的重点。

在其中一个实施例中，该通话数据处理方法还包括：收集每个坐席标识对应的历史服务数据；确定各坐席标识的历史服务数据与选取的特征词之间的相关度；按照相关度降序对相应的坐席标识进行排序；从排序的坐席标识首位起选取预设比例的坐席标识；将选取的坐席标识对应的坐席终端关联至专用通话通道。

坐席标识用于唯一标识一个坐席人员。坐席标识可以是包括数字、字母和符号中的至少一种字符的字符串。历史服务数据是指坐席人员在历史时间内为用户提供在线业务咨

询和业务办理的在线通话记录。历史服务数据包括所接通的在线通话数量、每通在线通话相应的通话记录 and 用户反馈等数据。

具体地，服务器可获取每个坐席标识对应的历史服务数据，计算各坐席标识的历史服务数据与选取的特征词之间的相关度，从而根据该相关度来对相应的坐席标识进行排序。排序时，按照相关度降序排序，相关度高的靠前而相关度低的靠后。服务器可再从排序的坐席标识中，从相关度最高的坐席标识开始，按照排序的坐席标识总量的预设百分比来选取坐席标识。比如共有 X 个坐席标识排序，则取排名靠前的 $X*10\%$ 的坐席标识，此时预设比例可为 10%。服务器可再将选取的坐席标识对应的坐席终端，与选取的特征词相应的专用通话通道关联。

在本实施例中，以坐席标识的历史服务数据与选取的特征词之间的相关度为依据，为选取的特征词分配的专用通话通道关联坐席终端，通过熟悉的坐席人员为用户提供服务，从而提高电话交互的效率。

在其中一个实施例中，该通话数据处理方法还包括：与用户终端建立通信连接后向用户终端发送语音引导信息；实时接收用户终端根据引导信息选择的专用通话通道；建立用户终端与选择的专用通话通道对应的坐席终端之间的通信连接。

语音引导信息是引导用户选择业务服务的信息。比如，健康险咨询请执行 XX 操作，万能险账户查询请执行 XX 操作等。

具体地，服务器可调用 IVR 系统向用户终端发送语音引导信息，该语音引导信息包括对业务类型的选择的语音提示信息。用户终端可播放该语音提示信息，将所获取的用户的操作信息发送至服务器。服务器在接收到用户终端发送的操作信息后，根据预设的操作信息和特征词之间的对应关系确定对应的特征词，并建立用户终端与确定的特征词相应的专用通话通道对应的坐席终端之间的通信连接。

在本实施例中，服务器通过语音引导信息获取用户选择的专用通话通道，实时建立用户终端与该专用通话通道的对应的坐席终端之间的通信连接，从而实时连接到为该专用通话通道分配的专业的坐席人员与用户进行电话交互，提高电话交互效率。

在其中一个实施例中，该通话数据处理方法还包括：确定选取的特征词所属的业务类型；从与业务类型对应的业务数据库中获取业务知识信息；根据业务知识信息生成与业务类型对应的应答模板；将应答模板发送至选取的坐席标识对应的坐席终端。

业务类型包括信用卡业务类型、银行业务类型、证券业务类型以及保险业务类型等

其中的多种。而保险业务类型又可包括寿险业务类型、健康险业务类型、投连险业务类型等子业务类型。

具体地，坐席业务数据库中存储了多个业务类型的特征词。并设置了每种业务类型与一个或多个特征词的对应关系，与业务类型具有对应关系的特征词，即为从属于该业务类型的特征词。服务器在确定了用户终端选取的特征词后，可根据该特征词相应业务类型的对应关系，确定其所属的业务类型。

进一步地，服务器还针对不同业务类型设置了与业务数据库对应的调用接口，并为不同的业务类型分配了对应相同或不同数量的接口调用资源。在确定了业务类型后，服务器可根据调用与该业务类型对应的接口，与该业务类型对应的业务数据库建立连接。

更进一步地，业务数据库中存储了该业务数据库相应业务类型的业务知识信息。业务知识信息包括业务相关知识和问题与回答的知识信息。知识信息包括某个专业名词解释、计算公式、业务流程等信息。服务器在与该业务类型对应的业务数据库建立连接后，从该业务数据库中获取业务知识信息。服务器可再根据业务知识信息生成与业务类型对应的应答模板，并将应答模板发送至选取的坐席标识对应的坐席终端。

在本实施例中，通过进一步向坐席终端自动提供查询出的与特征词相关的业务知识信息与应答模板，使坐席人员在需要检索相关的业务知识信息来解答用户咨询的问题时，提高对问题答案的获取效率。

如图4所示，在其中一个具体的实施例中，通话数据处理方法包括以下步骤：

S402，按时间周期获取历史语音会话中的用户语音数据。

S404，对识别用户语音数据得到的会话文本分词得到单独词；按单独词所属语义对单独词分类得到单独词子集；挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集。

S406，对于特征词集中每个特征词，获取当前特征词挑选自的单独词子集中各单独词的出现次数之和，作为当前特征词的出现次数；确定各特征词的出现次数占各特征词的出现次数之和的占比。

S408，确定各特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值。

S410，选取相应确定的差值大于预设差值、且与预设词不一致的特征词；为选取的特征词分配相应的专用通话通道。

S412, 收集每个坐席标识对应的历史服务数据; 确定各坐席标识的历史服务数据与选取的特征词之间的相关度; 按照相关度降序对相应的坐席标识进行排序; 从排序的坐席标识首位起选取预设比例的坐席标识; 将选取的坐席标识对应的坐席终端关联至专用通话通道。

S414, 确定选取的特征词所属的业务类型; 从与业务类型对应的业务数据库中获取业务知识信息; 根据业务知识信息生成与业务类型对应的应答模板; 将应答模板发送至选取的坐席标识对应的坐席终端。

S416, 与用户终端建立通信连接后向用户终端发送语音引导信息; 实时接收用户终端根据引导信息选择的专用通话通道; 建立用户终端与选择的专用通话通道对应的坐席终端之间的通信连接。

在本实施例中, 按时间周期从用户在历史语音会话中的用户语音数据提取代表用户需求的特征词, 获取这些特征词各自的出现次数占这些特征词的出现次数之和的占比, 并确定各特征词在当前时间周期获取的占比, 与在当前时间周期之前的历史时间周期获取的占比之间的差值。当特征词相应的差值大于预设差值、且与预设词不一致时, 判定近期该特征词相关的业务量明显增多, 自动为该特征词分配专用通话通道, 这样即可通过分配的专用通话通道主动为用户提供特征词相关业务的信息, 从而提高电话交互的效率。

其次, 为该特征词分配的专用通话通道关联了熟悉高的坐席人员对应的坐席终端, 并提供查询出的与特征词相关的业务知识信息与应答模板, 使坐席人员在需要检索相关的业务知识信息来解答用户咨询的问题时, 提高对问题答案的获取效率, 进一步提高了电话交互的效率。

应该理解的是, 上述流程图中的各个步骤按照箭头的指示依次显示, 但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明, 这些步骤的执行并没有严格的顺序限制, 这些步骤可以以其它的顺序执行。而且, 上述流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段, 这些子步骤或者阶段并不必然是在同一时刻执行完成, 而是可以在不同的时刻执行, 这些子步骤或者阶段的执行顺序也不必然是依次进行, 而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

如图 5 所示, 在其中一个实施例中, 提供了一种通话数据处理装置 500, 该通话数据处理装置 500 包括: 第一获取模块 501、分词模块 502、第二获取模块 503、确定模块 504、

选取模块 505 以及分配模块 506。

第一获取模块 501, 用于按时间周期获取历史语音会话中的用户语音数据。

分词模块 502, 用于对识别用户语音数据得到的会话文本分词得到特征词集。

第二获取模块 503, 用于获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比。

确定模块 504, 用于确定各特征词在当前时间周期获取的占比, 与在当前时间周期之前的历史时间周期获取的占比之间的差值。

选取模块 505, 用于选取相应确定的差值大于预设差值、且与预设词不一致的特征词。

分配模块 506, 用于为选取的特征词分配相应的专用通话通道。

上述通话数据处理装置 500, 按时间周期从用户在历史语音会话中的用户语音数据提取代表用户需求的特征词, 获取这些特征词各自的出现次数占这些特征词的出现次数之和的占比, 并确定各特征词在当前时间周期获取的占比, 与在当前时间周期之前的历史时间周期获取的占比之间的差值。当特征词相应的差值大于预设差值、且与预设词不一致时, 判定近期该特征词相关的业务量明显增多, 自动为该特征词分配专用通话通道, 这样即可通过分配的专用通话通道主动为用户提供特征词相关业务的信息, 从而提高电话交互的效率。

在其中一个实施例中, 分词模块 502 还用于对识别用户语音数据得到的会话文本分词得到单独词; 按单独词所属语义对单独词分类得到单独词子集; 挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集。第二获取模块 503 还用于对于特征词集中每个特征词, 获取当前特征词挑选自的单独词子集中各单独词的出现次数之和, 作为当前特征词的出现次数; 确定各特征词的出现次数占各特征词的出现次数之和的占比。

如图 6 所示, 在其中一个实施例中, 通话数据处理装置 500 还包括: 关联模块 507。

关联模块 507, 用于收集每个坐席标识对应的历史服务数据; 确定各坐席标识的历史服务数据与选取的特征词之间的相关度; 按照相关度降序对相应的坐席标识进行排序; 从排序的坐席标识首位起选取预设比例的坐席标识; 将选取的坐席标识对应的坐席终端关联至专用通话通道。

在其中一个实施例中, 关联模块 507 还用于与用户终端建立通信连接后向用户终端发送语音引导信息; 实时接收用户终端根据引导信息选择的专用通话通道; 建立用户终端与选择的专用通话通道对应的坐席终端之间的通信连接。

在其中一个实施例中，关联模块 507 还用于确定选取的特征词所属的业务类型；从与业务类型对应的业务数据库中获取业务知识信息；根据业务知识信息生成与业务类型对应的应答模板；将应答模板发送至选取的坐席标识对应的坐席终端。

关于数据处理装置的具体限定可以参见上文中对于数据处理方法的限定，在此不再赘述。上述数据处理装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

一个或多个存储有计算机可读指令的非易失性存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：按时间周期获取历史语音会话中的用户语音数据；对识别用户语音数据得到的会话文本分词得到特征词集；获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；确定各特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；选取相应确定的差值大于预设差值、且与预设词不一致的特征词；为选取的特征词分配相应的专用通话通道。

在一个实施例中，对识别用户语音数据得到的会话文本分词得到特征词集，包括：对识别用户语音数据得到的会话文本分词得到单独词；按单独词所属语义对单独词分类得到单独词子集；挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集。获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比，包括：对于特征词集中每个特征词，获取当前特征词挑选自的单独词子集中各单独词的出现次数之和，作为当前特征词的出现次数；确定各特征词的出现次数占各特征词的出现次数之和的占比。

在一个实施例中，计算机可执行指令被处理器执行时，还使得处理器执行以下步骤：收集每个坐席标识对应的历史服务数据；确定各坐席标识的历史服务数据与选取的特征词之间的相关度；按照相关度降序对相应的坐席标识进行排序；从排序的坐席标识首位起选取预设比例的坐席标识；将选取的坐席标识对应的坐席终端关联至专用通话通道。

在一个实施例中，计算机可执行指令被处理器执行时，还使得处理器执行以下步骤：与用户终端建立通信连接后向用户终端发送语音引导信息；实时接收用户终端根据引导信息选择的专用通话通道；建立用户终端与选择的专用通话通道对应的坐席终端之间的通信连接。

在一个实施例中，计算机可执行指令被处理器执行时，还使得处理器执行以下步骤：

确定选取的特征词所属的业务类型；从与业务类型对应的业务数据库中获取业务知识信息；根据业务知识信息生成与业务类型对应的应答模板；将应答模板发送至选取的坐席标识对应的坐席终端。

一种计算机设备，包括存储器和一个或多个处理器，存储器中储存有计算机可读指令，计算机可读指令被处理器执行时，使得一个或多个处理器执行以下步骤：按时间周期获取历史语音会话中的用户语音数据；对识别用户语音数据得到的会话文本分词得到特征词集；获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；确定各特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；选取相应确定的差值大于预设差值、且与预设词不一致的特征词；为选取的特征词分配相应的专用通话通道。

在一个实施例中，对识别用户语音数据得到的会话文本分词得到特征词集，包括：对识别用户语音数据得到的会话文本分词得到单独词；按单独词所属语义对单独词分类得到单独词子集；挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集。获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比，包括：对于特征词集中每个特征词，获取当前特征词挑选自的单独词子集中各单独词的出现次数之和，作为当前特征词的出现次数；确定各特征词的出现次数占各特征词的出现次数之和的占比。

在一个实施例中，计算机可执行指令被处理器执行时，还使得处理器执行以下步骤：收集每个坐席标识对应的历史服务数据；确定各坐席标识的历史服务数据与选取的特征词之间的相关度；按照相关度降序对相应的坐席标识进行排序；从排序的坐席标识首位起选取预设比例的坐席标识；将选取的坐席标识对应的坐席终端关联至专用通话通道。

在一个实施例中，计算机可执行指令被处理器执行时，还使得处理器执行以下步骤：与用户终端建立通信连接后向用户终端发送语音引导信息；实时接收用户终端根据引导信息选择的专用通话通道；建立用户终端与选择的专用通话通道对应的坐席终端之间的通信连接。

在一个实施例中，计算机可执行指令被处理器执行时，还使得处理器执行以下步骤：确定选取的特征词所属的业务类型；从与业务类型对应的业务数据库中获取业务知识信息；根据业务知识信息生成与业务类型对应的应答模板；将应答模板发送至选取的坐席标识对应的坐席终端。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通

过计算机可读指令来指令相关的硬件来完成,所述的计算机可读指令可存储于一非易失性计算机可读取存储介质中,该计算机可读指令在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程 ROM (PROM)、电可编程 ROM (EPROM)、电可擦除可编程 ROM (EEPROM) 或闪存。易失性存储器可包括随机存取存储器(RAM) 或者外部高速缓冲存储器。作为说明而非局限,RAM 以多种形式可得,诸如静态 RAM (SRAM)、动态 RAM (DRAM)、同步 DRAM (SDRAM)、双数据率 SDRAM (DDRSDRAM)、增强型 SDRAM (ESDRAM)、同步链路 (Synchlink) DRAM (SLDRAM)、存储器总线 (Rambus) 直接 RAM (RDRAM)、直接存储器总线动态 RAM (DRDRAM)、以及存储器总线动态 RAM (RDRAM) 等。

以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但并不能因此而理解为对本申请专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

权利要求书

1、一种通话数据处理方法，包括：

按时间周期获取历史语音会话中的用户语音数据；

对识别所述用户语音数据得到的会话文本分词得到特征词集；

获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；

5 确定各所述特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；

选取相应确定的差值大于预设差值、且与预设词不一致的特征词；及

为选取的所述特征词分配相应的专用通话通道。

2、根据权利要求1所述的方法，其特征在于，所述对识别所述用户语音数据得到的
10 会话文本分词得到特征词集，包括：

对识别所述用户语音数据得到的会话文本分词得到单独词；

按所述单独词所属语义对所述单独词分类得到单独词子集；及

挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集；

所述获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比，包括：

15 对于特征词集中每个特征词，获取当前特征词挑选自的单独词子集中各单独词的出现次数之和，作为当前特征词的出现次数；及

确定各特征词的出现次数占各特征词的出现次数之和的占比。

3、根据权利要求1所述的方法，其特征在于，还包括：

收集每个坐席标识对应的历史服务数据；

20 确定各坐席标识的历史服务数据与选取的所述特征词之间的相关度；

按照相关度降序对相应的坐席标识进行排序；

从排序的坐席标识首位起选取预设比例的坐席标识；及

将选取的坐席标识对应的坐席终端关联至所述专用通话通道。

4、根据权利要求3所述的方法，其特征在于，还包括：

25 与用户终端建立通信连接后向所述用户终端发送语音引导信息；

实时接收所述用户终端根据所述引导信息选择的专用通话通道；及

建立所述用户终端与选择的所述专用通话通道对应的坐席终端之间的通信连接。

5、根据权利要求3所述的方法，其特征在于，还包括：

确定选取的所述特征词所属的业务类型;

从与所述业务类型对应的业务数据库中获取业务知识信息;

根据所述业务知识信息生成与所述业务类型对应的应答模板; 及

将所述应答模板发送至选取的所述坐席标识对应的坐席终端。

5 6、一种通话数据处理装置, 包括:

第一获取模块, 用于按时间周期获取历史语音会话中的用户语音数据;

分词模块, 用于对识别所述用户语音数据得到的会话文本分词得到特征词集;

第二获取模块, 用于获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比;

10 确定模块, 用于确定各所述特征词在当前时间周期获取的占比, 与在当前时间周期之前的历史时间周期获取的占比之间的差值;

选取模块, 用于选取相应确定的差值大于预设差值、且与预设词不一致的特征词;
及

分配模块, 用于为选取的所述特征词分配相应的专用通话通道。

15 7、根据权利要求 6 所述的装置, 其特征在于, 所述分词模块还用于对识别所述用户语音数据得到的会话文本分词得到单独词; 按所述单独词所属语义对所述单独词分类得到单独词子集; 挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集;

所述第二获取模块还用于对于特征词集中每个特征词, 获取当前特征词挑选自的单独词子集中各单独词的出现次数之和, 作为当前特征词的出现次数; 确定各特征词的出现
20 次数占各特征词的出现次数之和的占比。

8、根据权利要求 6 所述的装置, 其特征在于, 所述装置还包括:

关联模块, 用于收集每个坐席标识对应的历史服务数据; 确定各坐席标识的历史服务数据与选取的所述特征词之间的相关度; 按照相关度降序对相应的坐席标识进行排序; 从排序的坐席标识首位起选取预设比例的坐席标识; 将选取的坐席标识对应的坐席终端关
25 联至所述专用通话通道。

9、根据权利要求 8 所述的装置, 其特征在于, 所述关联模块还用于与用户终端建立通信连接后向所述用户终端发送语音引导信息; 实时接收所述用户终端根据所述引导信息选择的专用通话通道; 及建立所述用户终端与选择的所述专用通话通道对应的坐席终端之间的通信连接。

10、根据权利要求 8 所述的装置，其特征在于，所述关联模块还用于确定选取的所述特征词所属的业务类型；从与所述业务类型对应的业务数据库中获取业务知识信息；根据所述业务知识信息生成与所述业务类型对应的应答模板；及将所述应答模板发送至选取的所述坐席标识对应的坐席终端。

- 5 11、一个或多个存储有计算机可读指令的非易失性计算机可读存储介质，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行以下步骤：

按时间周期获取历史语音会话中的用户语音数据；

对识别所述用户语音数据得到的会话文本分词得到特征词集；

获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；

- 10 确定各所述特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；

选取相应确定的差值大于预设差值、且与预设词不一致的特征词；及

为选取的所述特征词分配相应的专用通话通道。

- 15 12、根据权利要求 11 所述的存储介质，其特征在于，所述对识别所述用户语音数据得到的会话文本分词得到特征词集，包括：

对识别所述用户语音数据得到的会话文本分词得到单独词；

按所述单独词所属语义对所述单独词分类得到单独词子集；及

挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集；

所述获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比，包括：

- 20 对于特征词集中每个特征词，获取当前特征词挑选自的单独词子集中各单独词的出现次数之和，作为当前特征词的出现次数；及

确定各特征词的出现次数占各特征词的出现次数之和的占比。

13、根据权利要求 11 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

- 25 收集每个坐席标识对应的历史服务数据；

确定各坐席标识的历史服务数据与选取的所述特征词之间的相关度；

按照相关度降序对相应的坐席标识进行排序；

从排序的坐席标识首位起选取预设比例的坐席标识；及

将选取的坐席标识对应的坐席终端关联至所述专用通话通道。

14、根据权利要求 13 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

与用户终端建立通信连接后向所述用户终端发送语音引导信息；

实时接收所述用户终端根据所述引导信息选择的专用通话通道；及

5 建立所述用户终端与选择的所述专用通话通道对应的坐席终端之间的通信连接。

15、根据权利要求 13 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

确定选取的所述特征词所属的业务类型；

从与所述业务类型对应的业务数据库中获取业务知识信息；

10 根据所述业务知识信息生成与所述业务类型对应的应答模板；及

将所述应答模板发送至选取的所述坐席标识对应的坐席终端。

16、一种计算机设备，包括存储器及一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述一个或多个处理器执行时，使得所述一个或多个处理器执行以下步骤：

15 按时间周期获取历史语音会话中的用户语音数据；

对识别所述用户语音数据得到的会话文本分词得到特征词集；

获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比；

确定各所述特征词在当前时间周期获取的占比，与在当前时间周期之前的历史时间周期获取的占比之间的差值；

20 选取相应确定的差值大于预设差值、且与预设词不一致的特征词；及

为选取的所述特征词分配相应的专用通话通道。

17、根据权利要求 16 所述的计算机设备，其特征在于，所述对识别所述用户语音数据得到的会话文本分词得到特征词集，包括：

对识别所述用户语音数据得到的会话文本分词得到单独词；

25 按所述单独词所属语义对所述单独词分类得到单独词子集；及

挑选各单独词子集中出现次数最高的单独词作为特征词得到特征词集；

所述获取特征词集中各特征词的出现次数占各特征词的出现次数之和的占比，包括：

对于特征词集中每个特征词，获取当前特征词挑选自的单独词子集中各单独词的出现次数之和，作为当前特征词的出现次数；及

确定各特征词的出现次数占各特征词的出现次数之和的占比。

18、根据权利要求 16 所述的计算机设备，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

收集每个坐席标识对应的历史服务数据；

5 确定各坐席标识的历史服务数据与选取的所述特征词之间的相关度；

按照相关度降序对相应的坐席标识进行排序；

从排序的坐席标识首位起选取预设比例的坐席标识；及

将选取的坐席标识对应的坐席终端关联至所述专用通话通道。

10 19、根据权利要求 18 所述的计算机设备，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

与用户终端建立通信连接后向所述用户终端发送语音引导信息；

实时接收所述用户终端根据所述引导信息选择的专用通话通道；及

建立所述用户终端与选择的所述专用通话通道对应的坐席终端之间的通信连接。

15 20、根据权利要求 18 所述的计算机设备，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

确定选取的所述特征词所属的业务类型；

从与所述业务类型对应的业务数据库中获取业务知识信息；

根据所述业务知识信息生成与所述业务类型对应的应答模板；及

将所述应答模板发送至选取的所述坐席标识对应的坐席终端。

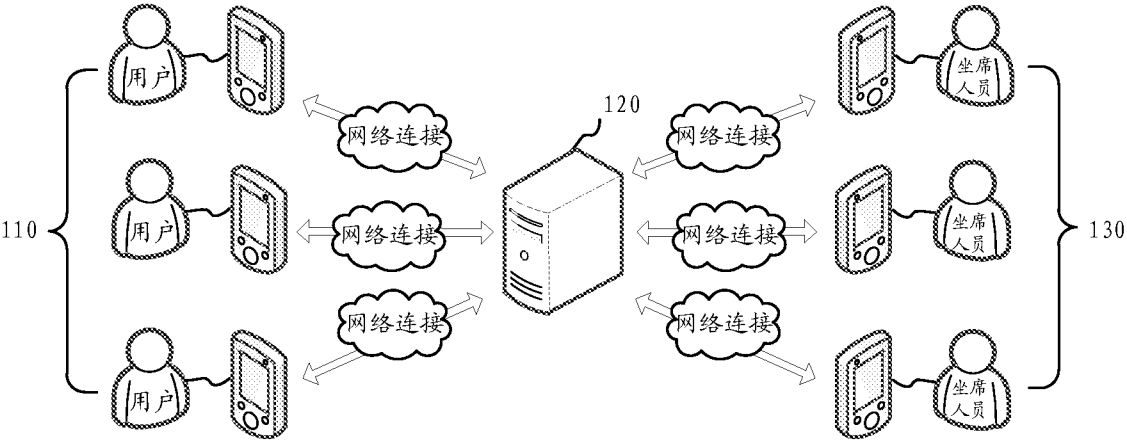


图 1

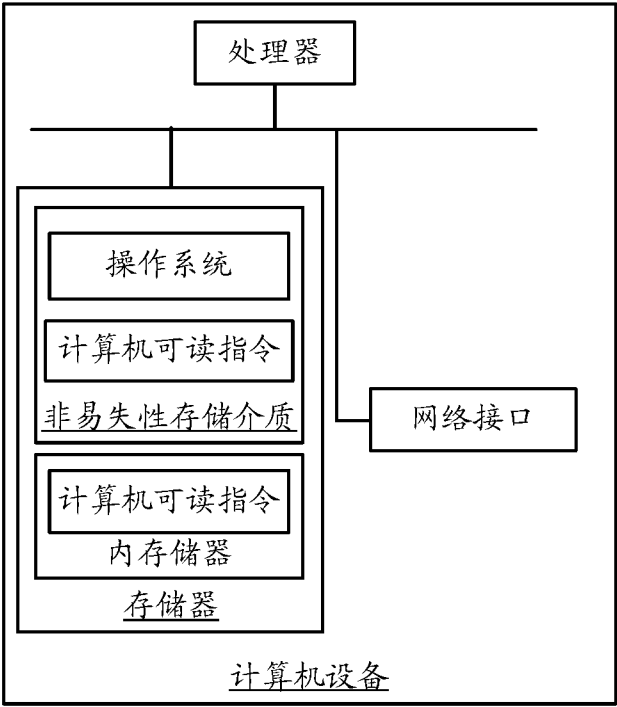


图 2

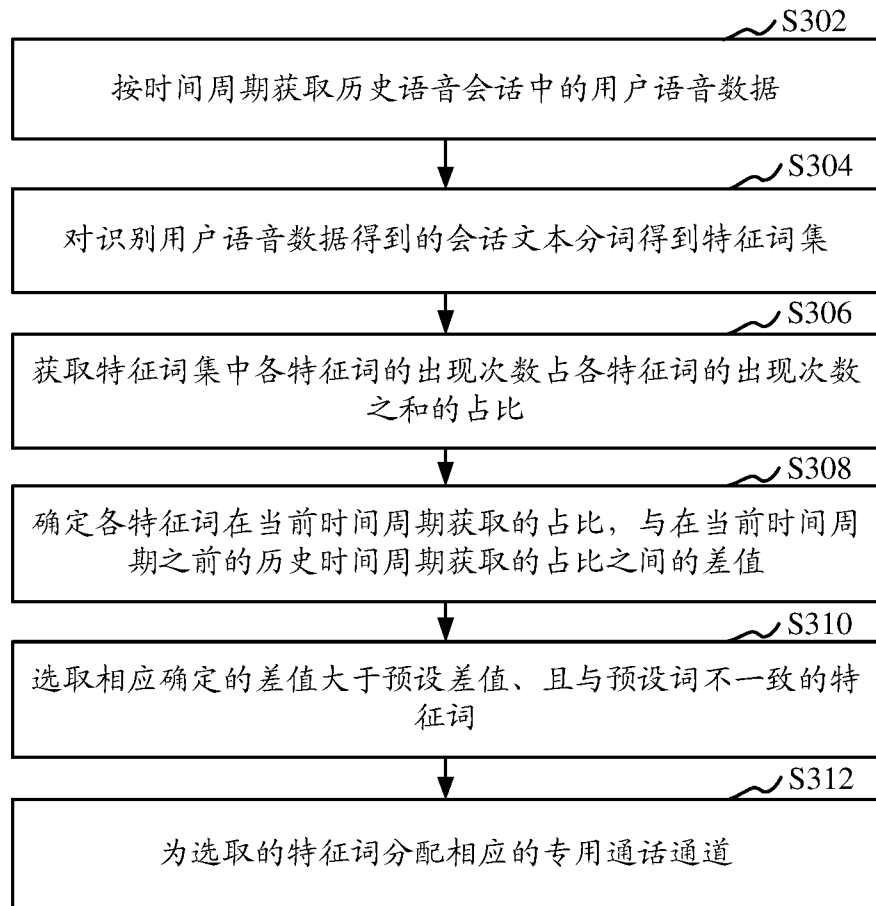


图 3

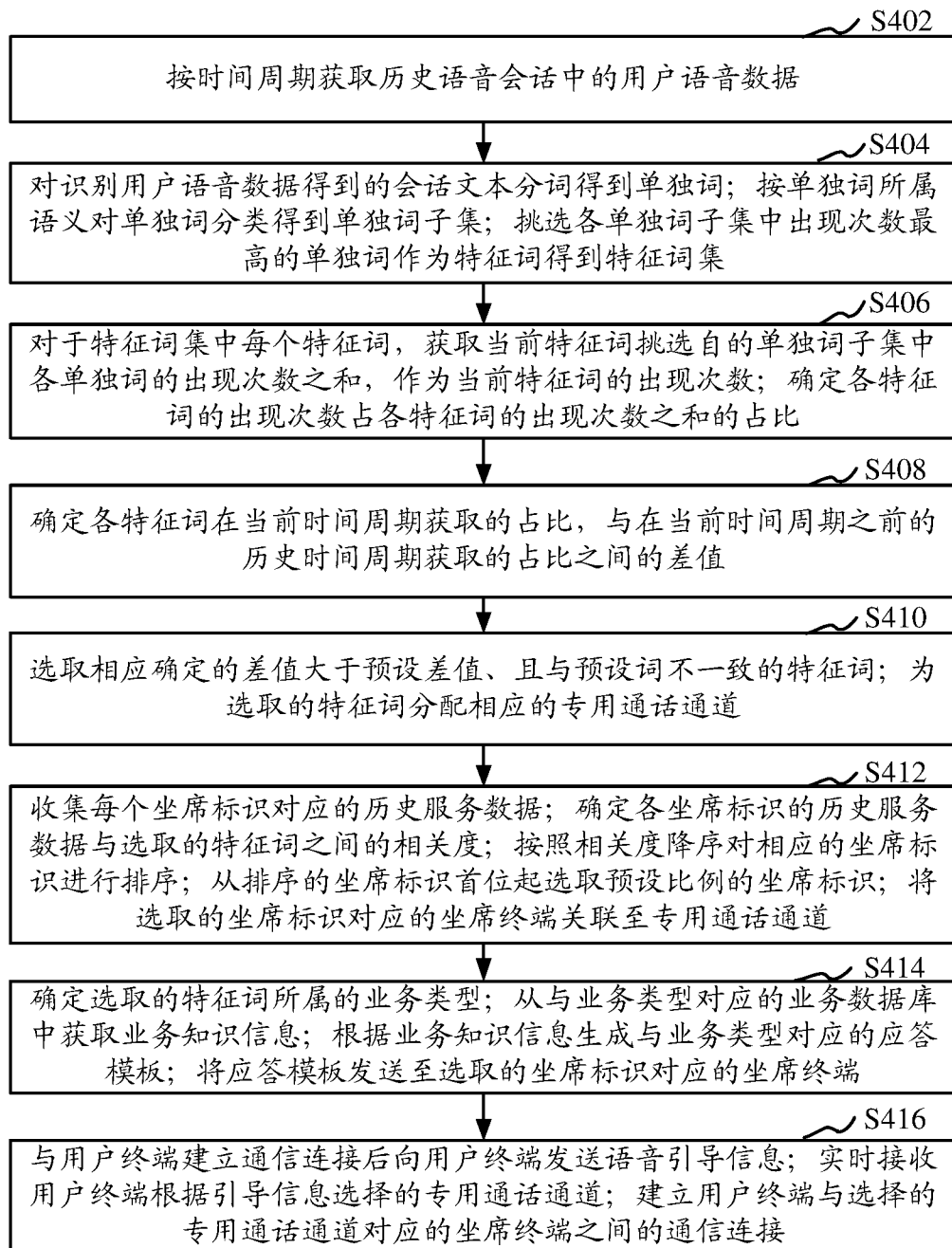


图 4

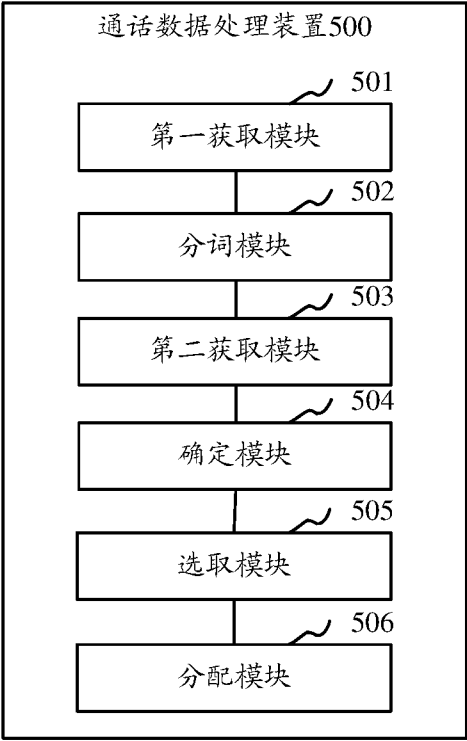


图 5

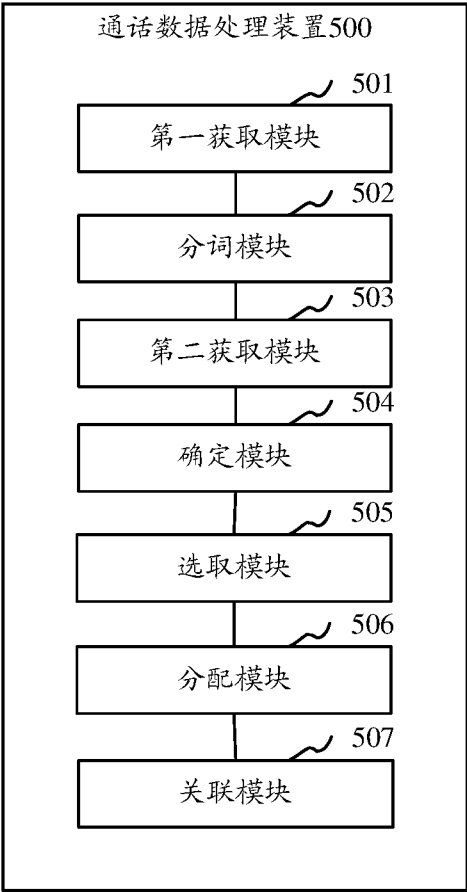


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/087327

A. CLASSIFICATION OF SUBJECT MATTER

H04M 3/523 (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, G10L, G06F, H04L, H04W

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: 分词, 特征, 需求, 意图, 坐席, 客服, 话务员, 呼叫中心, 互动式语音应答, 交互式语音应答, 增加, 专用, 占比, 比值, 百分比, 概率, 次数, 频率, 比例, 比率, 差, 减, 历史, 当前, 通话, 语音, tokenized, characteristic, demand, requirement, seat, call center, IVR, ratio, number, amount, difference, historical, current, call, voice

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107800900 A (PING AN TECH SHENZHEN CO., LTD.), 13 March 2018 (13.03.2018), description, paragraphs [0066]-[0150], and claims 1-10	1-20
Y	CN 101094271 A (HUAWEI TECHNOLOGIES CO., LTD.), 26 December 2007 (26.12.2007), description, page 1, line 5 to page 3, line 6 and page 3, line 27 to page 11, line 5	1-20
Y	CN 104598632 A (BEIHANG UNIVERSITY), 06 May 2015 (06.05.2015), description, paragraphs [0052]-[0085]	1-20
A	CN 104375847 A (HUAWEI TECHNOLOGIES CO., LTD.), 25 February 2015 (25.02.2015), entire document	1-20
A	WO 2015047272 A1 (INTERACTIVE INTELLIGENCE, INC.), 02 April 2015 (02.04.2015), entire document	1-20
A	US 9036806 B1 (XEROX CORPORATION), 19 May 2015 (19.05.2015), entire 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 16 July 2018	Date of mailing of the international search report 01 August 2018
Name and mailing address of the ISA 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 ZHANG, Fan Telephone No. 86-(10)-53961651

International application No.
PCT/CN2018/087327

Form PCT/ISA/210 (patent family annex) (January 2015)

国际检索报告

国际申请号

PCT/CN2018/087327

A. 主题的分类 H04M 3/523 (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 G10L G06F H04L H04W 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: 分词, 特征, 需求, 意图, 坐席, 客服, 话务员, 呼叫中心, 交互式语音应答, 交互式语音应答, 增加, 专用, 占比, 比值, 百分比, 概率, 次数, 频率, 比例, 比率, 差, 减, 历史, 当前, 通话, 语音, tokenized, characteristic, demand, requirement, seat, call center, IVR, ratio, number, amount, difference, historical, current, call, voice																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 107800900 A (平安科技深圳有限公司) 2018年 3月 13日 (2018 - 03 - 13) 说明书第[0066]-[0150]段, 权利要求1-10</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 101094271 A (华为技术有限公司) 2007年 12月 26日 (2007 - 12 - 26) 说明书第1页第5行到第3页第6行, 第3页第27行到第11页第5行</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 104598632 A (北京航空航天大学) 2015年 5月 6日 (2015 - 05 - 06) 说明书第[0052]-[0085]段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 104375847 A (华为技术有限公司) 2015年 2月 25日 (2015 - 02 - 25) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>WO 2015047272 A1 (INTERACTIVE INTELLIGENCE, INC.) 2015年 4月 2日 (2015 - 04 - 02) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 9036806 B1 (XEROX CORPORATION) 2015年 5月 19日 (2015 - 05 - 19) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 107800900 A (平安科技深圳有限公司) 2018年 3月 13日 (2018 - 03 - 13) 说明书第[0066]-[0150]段, 权利要求1-10	1-20	Y	CN 101094271 A (华为技术有限公司) 2007年 12月 26日 (2007 - 12 - 26) 说明书第1页第5行到第3页第6行, 第3页第27行到第11页第5行	1-20	Y	CN 104598632 A (北京航空航天大学) 2015年 5月 6日 (2015 - 05 - 06) 说明书第[0052]-[0085]段	1-20	A	CN 104375847 A (华为技术有限公司) 2015年 2月 25日 (2015 - 02 - 25) 全文	1-20	A	WO 2015047272 A1 (INTERACTIVE INTELLIGENCE, INC.) 2015年 4月 2日 (2015 - 04 - 02) 全文	1-20	A	US 9036806 B1 (XEROX CORPORATION) 2015年 5月 19日 (2015 - 05 - 19) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 107800900 A (平安科技深圳有限公司) 2018年 3月 13日 (2018 - 03 - 13) 说明书第[0066]-[0150]段, 权利要求1-10	1-20																					
Y	CN 101094271 A (华为技术有限公司) 2007年 12月 26日 (2007 - 12 - 26) 说明书第1页第5行到第3页第6行, 第3页第27行到第11页第5行	1-20																					
Y	CN 104598632 A (北京航空航天大学) 2015年 5月 6日 (2015 - 05 - 06) 说明书第[0052]-[0085]段	1-20																					
A	CN 104375847 A (华为技术有限公司) 2015年 2月 25日 (2015 - 02 - 25) 全文	1-20																					
A	WO 2015047272 A1 (INTERACTIVE INTELLIGENCE, INC.) 2015年 4月 2日 (2015 - 04 - 02) 全文	1-20																					
A	US 9036806 B1 (XEROX CORPORATION) 2015年 5月 19日 (2015 - 05 - 19) 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 2018年 7月 16日		国际检索报告邮寄日期 2018年 8月 1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 传真号 (86-10) 62019451		受权官员 张凡 电话号码 86-(10)-53961651																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/087327

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	107800900	A	2018年 3月 13日	无	
CN	101094271	A	2007年 12月 26日	WO 2009012689 A1	2009年 1月 29日
CN	104598632	A	2015年 5月 6日	CN 104598632 B	2017年 12月 1日
CN	104375847	A	2015年 2月 25日	CN 104375847 B	2017年 12月 5日
WO	2015047272	A1	2015年 4月 2日	无	
US	9036806	B1	2015年 5月 19日	无	