

(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
2020 年 1 月 9 日 (09.01.2020)



(10) 国际公布号
WO 2020/006878 A1

(51) 国际专利分类号:
G10L 15/01 (2013.01)

(21) 国际申请号: PCT/CN2018/106256

(22) 国际申请日: 2018 年 9 月 18 日 (18.09.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810735475.3 2018 年 7 月 6 日 (06.07.2018) 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福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

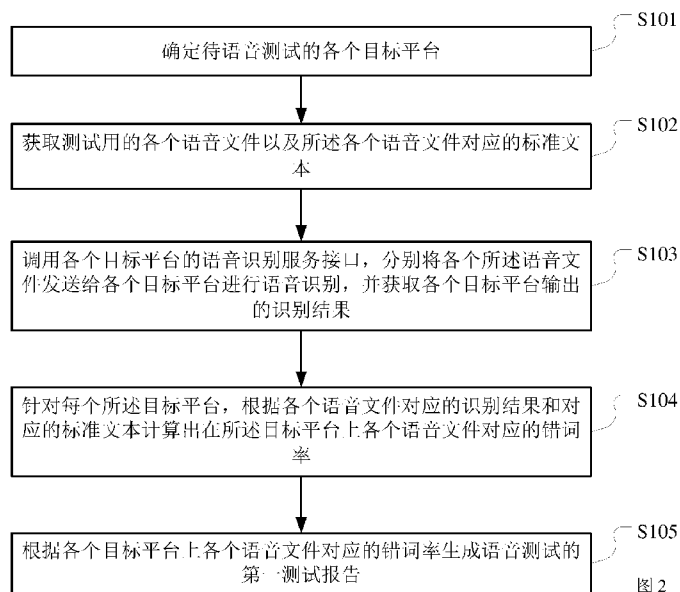
(72) 发明人: 岳鹏昱(YUE, Pengyu); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 李秀丰(LI, Xiufeng); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳众鼎专利商标代理事务所(普通合伙)(SHENZHEN ZHONGDING INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市龙岗区龙城街道中心城清林路 546 号城投商务中心 4 层/B, Guangdong 518172 (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,

(54) Title: VOICE RECOGNITION TEST METHOD AND APPARATUS, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 语音识别测试方法、装置、计算机设备及存储介质



S101 Determine each target platform for voice testing.
S102 Acquire each voice file for testing and a standard text corresponding to each voice file.
S103 Call a voice recognition service interface of each target platform, send each voice file to each target platform for voice recognition respectively, and acquire a recognition result output by each target platform.
S104 For each target platform, calculate an incorrect word rate corresponding to each voice file on the target platform according to the recognition result corresponding to each voice file and the corresponding standard text.
S105 Generate a first test report of the voice test according to the incorrect word rate corresponding to each voice file on each target platform.

(57) Abstract: Provided are a voice recognition test method and apparatus, a computer device and a storage medium, which are used for solving the problem that the existing voice recognition test process is easily interfered with by human factors and is difficult to standardize and unify. The method comprises: determining each target platform for voice testing (S101); acquiring each voice file for testing and a standard text corresponding to each voice file (S102); calling a voice recognition service interface of each target platform, sending each voice file to each target platform for voice recognition respectively, and acquiring a recognition result output by each target

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。

(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))。

platform (S103); for each target platform, calculating an incorrect word rate corresponding to each voice file on the target platform according to the recognition result corresponding to each voice file and the corresponding standard text (S104); and generating a first test report of the voice test according to the incorrect word rate corresponding to each voice file on each target platform (S105).

(57) 摘要: 一种语音识别测试方法、语音识别测试装置、计算机设备及存储介质, 用于解决现有语音识别测试流程容易受到人为因素干扰, 难以做到规范统一的问题。方法包括: 确定待语音测试的各个目标平台 (S101); 获取测试用的各个语音文件以及各个语音文件对应的标准文本 (S102); 调用各个目标平台的语音识别服务接口, 分别将各个语音文件发送给各个目标平台进行语音识别, 并获取各个目标平台输出的识别结果 (S103); 针对每个目标平台, 根据各个语音文件对应的识别结果和对应的标准文本计算出在目标平台上各个语音文件对应的错词率 (S104); 根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告 (S105)。

语音识别测试方法、装置、计算机设备及存储介质

本申请以 2018 年 07 月 06 日提交的申请号为 201810735475.3，名称为“语音识别测试方法、装置、计算机设备及存储介质”的中国发明专利申请为基础，并要求其优先权。

技术领域

本申请涉及语音识别技术领域，尤其涉及语音识别测试方法、装置、计算机设备及存储介质。

背景技术

目前，语音识别技术已经应用极为广泛，很多平台可以对外提供语音识别的服务接口，用户可以通过平台提供的服务接口使用各个平台的语音识别功能。但是，不同平台使用的语音识别模型一般都会存在差异，这就导致同一个声音文件在不同的平台上识别出来的结果往往并不相同，有时候甚至差别很大。为此，平台在开发语音识别模型时，需要测试不同平台的语音识别效果与自身的进行对比，分析其中的优劣，以便更好地改进平台的语音识别功能。

然而，现有语音识别测试流程上不同环节需要由不同的处理人员进行处理，由于不同处理人员往往有各自的测试习惯，难以做到规范统一，给语音识别的测试工作带来诸多不便，不利于测试流程的数据记录和测试数据的横向比对。

发明内容

基于此，有必要针对上述技术问题，提供一种可以提高语音识别测试的效率、排除人为干扰因素的语音识别测试方法、装置、计算机设备及存储介质。

一种语音识别测试方法，包括：

确定待语音测试的各个目标平台；

获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

一种语音识别测试装置，包括：

平台确定模块，用于确定待语音测试的各个目标平台；

文件文本获取模块，用于获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

语音识别模块，用于调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

错词率计算模块，用于针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

第一报告生成模块，用于根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现上述语音识别测试方法的步骤。

一个或多个存储有计算机可读指令的非易失性可读存储介质，所述计算机可读存储介质存储有计算机可读指令，使得所述一个或多个处理器执行上述语音识别测试方法的步骤。

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，本申请的其他特征和优点将从说明书、附图以及权利要求变得明显。

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的描述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性的前提下，还可以根据这些附图获得其他的附图。

图 1 是本申请一实施例中语音识别测试方法的一应用环境示意图；

图 2 是本申请一实施例中语音识别测试方法的一流程图；

图 3 是本申请一实施例中语音识别测试方法步骤 S105 在一个应用场景下的流程示意图；

图 4 是本申请一实施例中语音识别测试方法在一个应用场景下测试目标平台上语音识别模型的流程示意图；

图 5 是本申请一实施例中语音识别测试方法在一个应用场景下确定语音识别模型的

弱支持领域的流程示意图；

图 6 是本申请一实施例中语音识别测试装置的结构示意图；

图 7 是本申请一实施例中计算机设备的一示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请提供的语音识别测试方法可以通过语音测试平台实施，可应用在如图 1 的应用环境中，其中，该语音测试平台的服务器通过网络与各个目标平台的服务器进行通信。其中，上述服务器可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

在一实施例中，如图 2 所示，提供一种语音识别测试方法，以该方法应用在图 1 中语音测试平台的服务器为例进行说明，包括如下步骤：

S101、确定待语音测试的各个目标平台；

本方案中，在开始语音测试时，需要先确定出本次待语音测试的各个目标平台。一般来说，所述各个目标平台为多个各不相同的语音识别平台，这样有利于横向对比这些目标平台的语音识别效果。具体地，本方案的语音测试平台可以提供一界面窗口给测试人员选择，在测试窗口中，测试人员可以具体选择本次测试所针对的各个目标平台有哪些，在一次语音测试中可以选择一个、两个或多个目标平台。

S102、获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

可以理解的是，测试用的各个语音文件以及与这些语音文件对应的标准文本可以是预先收集好并存储在语音测试平台上的，当然，也可以由测试人员在本次语音测试开始之前手动上传或者通过网络从其它数据库中提取得到，对此本实施例不作具体限定。

需要说明的是，标准文本是指所述语音文件的音频内容对应的文本内容，该标准文本在本次语音测试之前均经过修正、确认，确保了文本内容与对应语音文件的音频所表达的内容一致。

S103、调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

在确定本次语音测试的各个目标平台以及获取到测试用的各个语音文件、标准文本之后，可以调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别。可以理解的是，各个目标平台对外提供语音识别服务接口，本方案的语音测试平台通过对接这些接口，可以使用各个目标平台提供的语音识别功能。在将各个语音文件发送给各个目标平台之后，各个语音文件会分别在各个目标平台中完成音转字，然后各个目标平台分别输出各个语音文件对应的识别结果，也即识别得到的文本。

进一步地，在调用语音识别服务接口对语音文件进行语音识别之前，还可以先对各个语音文件进行预处理，比如降噪处理。在预处理之前，测试人员可以手动选择是否降噪以及降噪的具体处理方式；或者，语音测试平台也可以设置默认的降噪处理方式，并默认对各个语音文件进行降噪处理。

S104、针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

具体地，针对每个语音文件，可以将该语音文件对应的识别结果与该语音文件对应的标准文本进行比对，计算识别结果相对标准文本的错词率，也即文本内容的误差比例。而对于每个目标平台来说，则需要分别计算各个语音文件对应的错词率。因此，假设一共 N 个参与本次语音测试的语音文件， M 个目标平台，则针对每个目标平台计算得到的错词率个数为 N 个，对 M 个目标平台则总共需要计算 $M*N$ 个错词率。

S105、根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

可以理解的是，在得到各个目标平台上各个语音文件对应的错词率之后，本次语音测试基本上完成，最后需要生成本次语音测试的第一测试报告提供给测试人员查阅。需要说明的是，该第一测试报告的具体内容可以根据实际情况来确定。比如，若本次语音测试关注的重点是各个目标平台平均错词率的排行，则可以根据各个目标平台上各个语音文件对应的错词率分别计算出各个目标平台的平台错词率，然后依据平台错词率从大到小或从小到大对各个目标平台进行排序，并在生成的第一测试报告中体现。这样，测试人员查看该第一测试报告即可清楚地得知各个目标平台在错词率方面的优劣。

进一步地，如图 3 所示，所述步骤 S105 具体可以包括：

S201、针对每个目标平台，根据所述目标平台上各个语音文件对应的错词率计算所述目标平台的平台错词率；

S202、在计算得到各个目标平台对应的平台错词率后，按照预设的排序规则对所述各

个目标平台对应的平台错词率进行排序；

S203、生成空白的第一测试报告，并将排序后的所述各个目标平台对应的平台错词率导入所述第一测试报告。

对于上述步骤 S201，可以理解的是，对于一个目标平台来说，该目标平台的平台错词率具体可以是各个语音文件对应的错词率的平均值，假设有 N 个语音文件，则每个语音文件对应的错词率表示为 W_n ，则该目标平台的平台错词率等于 $(W_1+W_2+\cdots+W_n)/N$ 。

对于步骤 S202，该预设的排序规则具体可以是按照目标平台的预设平台顺序进行排序，也可以按照平台错词率进行排序。可以理解的是，本次语音测试中，测试人员有可能更加关注某些著名平台的平台错词率，因此，测试人员可以预设设置各个平台之间的排序，将重点关注的平台排序在前，其它平台排序在后，从而在生成的第一测试报告可以更加突出反映测试人员重点关注的几个平台的平台错词率。或者，测试人员若希望对比查看这些目标平台之间哪些目标平台的平台错词率更高或者更低，通过按照平台错词率进行排序，最后生成的第一测试报告可以更突出反映各个目标平台之间在平台错词率方面的优劣。

对于步骤 S203，可知，在对各个目标平台对应的平台错词率进行排序之后，即可依据排序后的顺序将各个目标平台对应的平台错词率以及相应的目标平台信息导入至该第一测试报告中。

通过上述步骤 S201-S203，本实施例的语音识别测试方法可以分析各个目标平台的平台错词率，并按照预设的排序规则对各个目标平台进行排序，这样，测试人员可以依据平台错词率对比这些目标平台上语音识别服务的优劣，使得各个目标平台之间的横向对比更加突出和高效。

进一步地，本方法还可以在测试时，指定平台上的特定语音识别模型进行测试。具体地，如图 4 所示，所述语音识别测试方法还包括：

S301、分别确定所述各个目标平台的指定语音识别模型；

S302、调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别，并获取各个目标平台输出的识别文本；

S303、针对每个所述目标平台，根据各个语音文件对应的识别文本和对应的标准文本计算出在各个目标平台上所述指定语音识别模型对应的错词率；

S304、根据各个目标平台上所述指定语音识别模型对应的错词率生成语音测试的第二测试报告。

对于步骤 S301，可以理解的是，在本次语音测试时，可以指定各个目标平台上用来

测试的语音识别模型具体是哪个，被指定的语音识别模型即为上述的指定语音识别模型。具体地，这些指定语音识别模型可以由语音测试平台指定，也可以由测试人员手动选择各个目标平台上的语音识别模型作为指定语音识别模型，还可以由各个目标平台自行指定语音识别模型，对此不做具体限定。

步骤 S302 与上述步骤 S103 类似，在调用各个目标平台的语音识别服务接口对各个所述语音文件进行语音识别时，发送给各个目标平台上的指定语音识别模型来识别即可，识别后可以获取到各个目标平台反馈的识别文本。

对于步骤 S303，在得到各个目标平台输出的识别文本后，与上述步骤 S104 类似，可以根据各个语音文件对应的识别文本和对应的标准文本计算出在各个目标平台上所述指定语音识别模型对应的错词率。

对于步骤 S304，在得到各个目标平台上所述指定语音识别模型对应的错词率之后，与上述步骤 S105 同理，可以由此生成语音测试的第二测试报告。其中，与步骤 S105 存在区别的地方是，步骤 S304 生成的第二测试报告主要针对各个目标平台上所述指定语音识别模型的，该第二测试报告重点横向对比的是不同平台上指定语音识别模型之间的差异和优劣，因此在生成的第二测试报告中需要体现出各个指定语音识别模型对应的错词率。

通过上述步骤 S301-S304，本方法在本次语音测试时，不仅可以从第一维度（即平台的维度）确定各个目标平台，还可以从第二维度（即平台上语音识别模型的维度）确定每个目标平台下的语音识别模型，通过两个维度的确定，使得本次语音测试的结果可以反映出具体语音识别模型的效果和优劣。

进一步地，在针对第二维度，即目标平台下具体的语音识别模型进行测试时，还可以具体评估这些语音识别模型在词语领域的支持力度如何。目前，在语音识别领域，由于语音识别模型往往需要极其大量的训练样本完成训练，而语言的领域过于宽泛，为了提高训练效率同时也为了提升语音识别模型的准确性，目前训练的语音识别模型均会区分不同的应用领域，比如财经、新闻、体育、电影对话等。因此，测试语音识别模型对某个应用领域的支持力度存在重大意义。为此，如图 5 所示，在步骤 S302 之后，该语音识别测试方法还可以包括：

S401、针对每个指定语音识别模型，获取各个所述语音文件输入所述指定语音识别模型进行语音识别后得到的各个所述语音文件所对应识别文本；

S402、针对每个指定语音识别模型，将各个所述语音文件所对应识别文本与标准文本进行对比，确定各个所述标准文本的易错词，所述易错词是指识别文本与标准文本对比后、

所述标准文本中对比不一致的文本；

S403、针对每个指定语音识别模型，确定各个所述标准文本的易错词所属的词语领域；

S404、针对每个指定语音识别模型，将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域，所述弱支持领域是指语音识别模型用于语音识别的效果未达到期望值的词语领域。

上述步骤 S401~S404 均是分别针对每个指定语音识别模型进行的处理，为了便于说明，下述对步骤 S401~S404 的说明中不再强调。

对于步骤 S401，类似步骤 S103，可以理解的是，将各个所述语音文件输入所述指定语音识别模型进行语音识别，识别后，这些指定语音识别模型会输出对应的识别文本，获取这些识别文本。

对于步骤 S402，在获取到各个语音文件对应识别文本后，将这些识别文本与相应的标准文本进行对比，可以得知标准文本中哪些文本与识别文本不一致，该标准文本中与识别文本不一致的词语可以认为就是该标准文本的易错词。这里说的易错词是指容易识别错误的文本内容。

对于步骤 S403，可以理解的是，不同的文本均有其所述的词语领域，例如，比亚迪、宝马、奔驰、四轮驱动等文本属于汽车的词语领域；七言、静夜思、观沧海等文本属于诗词的词语领域。因此，在确定出各个所述标准文本的易错词之后，可以进一步确定这些易错词所属的词语领域。具体地，可以通过网络检索的方式确定易错词的词语领域，当网络检索无法得到结果时，可以提供给指定工作人员人工确认，本实施例对此不作限定。

对于步骤 S404，对于一个语音识别模型来说，标准文本中的易错词，代表了该语音识别模型对这些易错词的识别效果较差。可以理解的是，由于语音识别模型是通过样本训练得到的，而且往往样本的选择均会区分领域，因此对于特定的语音识别模型来说，存在“偏科”现象是较为常见的。例如，某个语音识别模型训练用的样本均为诗词领域的文本，该领域文本中很少包含数值，从而该语音识别模型识别标准文本的易错词很可能是数值类文本，比如将“一心一意”错误识别为“1 心 1 意”，通过标准文本“一心一意”确定出易错词为“一”，当该语音识别模型识别其它语音文件得出的识别文本也有类似情况时，即其它语音文件对应的标准文本的易错词也包含数值时，则可以确定该语音识别模型对数值类词语领域为弱支持。需要说明的是，弱支持领域是指语音识别模型用于语音识别的效果未达到期望值的词语领域，具体地，可以是该语音识别模型对弱支持领域的识别准确率在预设的阈值以下。

为便于理解，在步骤 S404 将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域之前，可以通过如下步骤确定某个指定语言识别模型的弱支持领域：

S501、针对每个指定语音识别模型，根据各个所述语音文件所对应的识别文本和对应的标准文本计算出所述指定语音识别模型识别各个所述语音文件时的模型错词率；

S502、针对每个指定语音识别模型，筛选模型错词率超过预设错词率阈值的所述语音文件；

S503、针对每个指定语音识别模型，确定出与筛选出的所述语音文件对应的标准文本作为目标标准文本；

在上述步骤 S501-S503 的基础上，所述将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域具体为：将所述目标标准文本的易错词所属的词语领域确定为所述指定语音识别模型的弱支持领域。

对于步骤 S501，针对单个指定语音识别模型，可以根据各个所述语音文件所对应的识别文本和对应的标准文本计算出所述指定语音识别模型识别各个所述语音文件时的模型错词率，这里所说的模型错词率是指这个语音识别模型在识别某个语音文件时的错词率。需要说明的是，步骤 S501 需要分别计算出一个语音识别模型识别各个语音文件时的模型错词率，假设存在 N 个语音文件，则对于一个语音识别模型来说，需要计算出 N 个模型错词率。

对于步骤 S502，可以理解的是，当某个语音文件的模型错词率超过预设错词率阈值时，可以认为该语音识别模型识别这个语音文件时的错词率过高，也即识别准确率过低，这反映了该语音识别模型在识别这个语音文件所述的词语领域时识别准确率过低，因此，筛选出这些模型错词率超过预设错词率阈值的所述语音文件后，即可依据这些筛选出的语音文件来确认该语音识别模型的弱支持领域。

对于步骤 S503，容易理解的是，在筛选出语音文件后，容易确定出与之对应的标准文本。

在步骤 S501-S503 的基础上，确定出目标标准文本之后，由于上述步骤 S403 已经确定出了各个标准文本的易错词所述的词语领域，因此直接可以得知目标标准文本所对应的词语领域，这些词语领域就是该指定语音识别模型的弱支持领域。

进一步地，上述步骤 S103 中得出的识别结果还可以包括各个语音文件对应的识别耗时，识别耗时是指平台对某个语音文件进行语音识别所需的时间。可以理解的是，识别耗时与错词率之间往往是相对的，在优化语音识别模型时，需要考量模型的错词率优先还有

识别速度优先。一般来说，识别耗时越短，即识别速度越快，则往往错词率越高；反之，错词率越低，即识别准确率越高，则往往识别速度越慢、识别耗时越长。本实施例中，可以将各个目标平台的识别耗时加入至该第一测试报告中，使得第一测试报告上可以反映各个目标平台在提供语音识别服务时的耗时，测试人员可以将耗时作为评估平台语音识别能力的参考因素。

进一步地，在步骤 S105 之后，本语音识别测试方法还可以包括：将所述测试报告发送至预先指定的目标终端。可以理解的是，所述目标终端可以预先设定，比如设定邮件地址，在生成得到测试报告后，将该测试报告自动通过邮件发送给该邮件地址，从而测试人员可以方便地接收到测试报告，并方便测试人员查看测试报告。

由上述内容可知，本实施例提供的语音识别测试方法在整个语音测试流程中无需人员处理，不仅提高了语音识别测试的效率，而且排除了人为干扰因素，可以实现测试报告的规范统一；另外，通过在多个目标平台上进行语音文件的语音识别测试并计算各个目标平台识别这些语音文件时的错词率，可以横向对比各个目标平台的测试性能，更直观地反映各个目标平台语音识别功能的优劣。

应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

在一实施例中，提供一种语音识别测试装置，该语音识别测试装置与上述实施例中语音识别测试方法一一对应。如图 6 所示，该语音识别测试装置包括平台确定模块 601、文件文本获取模块 602、语音识别模块 603、错词率计算模块 604 和第一报告生成模块 605。各功能模块详细说明如下：

平台确定模块 601，用于确定待语音测试的各个目标平台；

文件文本获取模块 602，用于获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

语音识别模块 603，用于调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

错词率计算模块 604，用于针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

第一报告生成模块 605，用于根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

进一步地，所述第一报告生成模块可以包括：

平台错词率计算单元，用于针对每个目标平台，根据所述目标平台上各个语音文件对应的错词率计算所述目标平台的平台错词率；

排序单元，用于在计算得到各个目标平台对应的平台错词率后，按照预设的排序规则对所述各个目标平台对应的平台错词率进行排序；

导入报告单元，用于生成空白的第一测试报告，并将排序后的所述各个目标平台对应的平台错词率导入所述第一测试报告。

进一步地，所述语音识别测试装置还可以包括：

指定模块确定模块，用于分别确定所述各个目标平台的指定语音识别模型；

模型识别模块，用于调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别，并获取各个目标平台输出的识别文本；

模型错词率计算模块，用于针对每个所述目标平台，根据各个语音文件对应的识别文本和对应的标准文本计算出在各个目标平台上所述指定语音识别模型对应的错词率；

第二报告生成模块，用于根据各个目标平台上所述指定语音识别模型对应的错词率生成语音测试的第二测试报告。

进一步地，所述语音识别测试装置还可以包括：

识别文本获取模块，用于针对每个指定语音识别模型，获取各个所述语音文件输入所述指定语音识别模型进行语音识别后得到的各个所述语音文件所对应识别文本；

易错词确定模块，用于针对每个指定语音识别模型，将各个所述语音文件所对应识别文本与标准文本进行对比，确定各个所述标准文本的易错词，所述易错词是指识别文本与标准文本对比后、所述标准文本中对比不一致的文本；

词语领域确定模块，用于针对每个指定语音识别模型，确定各个所述标准文本的易错词所属的词语领域；

弱支持领域确定模块，用于针对每个指定语音识别模型，将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域，所述弱支持领域是指语音识别模型用于语音识别的效果未达到期望值的词语领域。

进一步地，所述语音识别测试装置还可以包括：

各模型错词计算模块，用于针对每个指定语音识别模型，根据各个所述语音文件所对应的识别文本和对应的标准文本计算出所述指定语音识别模型识别各个所述语音文件时的模型错词率；

语音文件筛选模块，用于针对每个指定语音识别模型，筛选模型错词率超过预设错词率阈值的所述语音文件；

目标标准文本确定模块，用于针对每个指定语音识别模型，确定出与筛选出的所述语

音文件对应的标准文本作为目标标准文本；

所述弱支持领域确定模块具体可以用于将所述目标标准文本的易错词所属的词语领域确定为所述指定语音识别模型的弱支持领域。

关于语音识别测试装置的具体限定可以参见上文中对于语音识别测试方法的限定，在此不再赘述。上述语音识别测试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。

在一个实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 7 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、网络接口和数据库。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储语音识别测试方法中涉及到的数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种语音识别测试方法。

在一个实施例中，提供了一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机可读指令，处理器执行计算机可读指令时实现上述实施例中语音识别测试方法的步骤，例如图 2 所示的步骤 S101 至步骤 S105。或者，处理器执行计算机可读指令时实现上述实施例中语音识别测试装置的各模块/单元的功能，例如图 6 所示模块 601 至模块 605 的功能。为避免重复，这里不再赘述。

在一个实施例中，提供了一种计算机可读存储介质，该一个或多个存储有计算机可读指令的非易失性可读存储介质，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计算机可读指令时实现上述方法实施例中语音识别测试方法的步骤，或者，该一个或多个存储有计算机可读指令的非易失性可读存储介质，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计算机可读指令时实现上述装置实施例中语音识别测试装置中各模块/单元的功能。为避免重复，这里不再赘述。

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

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。

所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元、模块完成，即将所述装置的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。

以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围，均应包含在本申请的保护范围之内。

权 利 要 求 书

1. 一种语音识别测试方法，其特征在于，包括：

确定待语音测试的各个目标平台；

获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

2. 根据权利要求 1 所述的语音识别测试方法，其特征在于，所述根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告包括：

针对每个目标平台，根据所述目标平台上各个语音文件对应的错词率计算所述目标平台的平台错词率；

在计算得到各个目标平台对应的平台错词率后，按照预设的排序规则对所述各个目标平台对应的平台错词率进行排序；

生成空白的第一测试报告，并将排序后的所述各个目标平台对应的平台错词率导入所述第一测试报告。

3. 根据权利要求 1 或 2 所述的语音识别测试方法，其特征在于，所述语音识别测试方法还包括：

分别确定所述各个目标平台的指定语音识别模型；

调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别，并获取各个目标平台输出的识别文本；

针对每个所述目标平台，根据各个语音文件对应的识别文本和对应的标准文本计算出在各个目标平台上所述指定语音识别模型对应的错词率；

根据各个目标平台上所述指定语音识别模型对应的错词率生成语音测试的第二测试报告。

4. 根据权利要求 3 所述的语音识别测试方法，其特征在于，在分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别之后，还包括：

针对每个指定语音识别模型，获取各个所述语音文件输入所述指定语音识别模型进行

语音识别后得到的各个所述语音文件所对应识别文本；

针对每个指定语音识别模型，将各个所述语音文件所对应识别文本与标准文本进行对比，确定各个所述标准文本的易错词，所述易错词是指识别文本与标准文本对比后、所述标准文本中对比不一致的文本；

针对每个指定语音识别模型，确定各个所述标准文本的易错词所属的词语领域；

针对每个指定语音识别模型，将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域，所述弱支持领域是指语音识别模型用于语音识别的效果未达到期望值的词语领域。

5. 根据权利要求 4 所述的语音识别测试方法，其特征在于，在将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域之前，还包括：

针对每个指定语音识别模型，根据各个所述语音文件所对应的识别文本和对应的标准文本计算出所述指定语音识别模型识别各个所述语音文件时的模型错词率；

针对每个指定语音识别模型，筛选模型错词率超过预设错词率阈值的所述语音文件；

针对每个指定语音识别模型，确定出与筛选出的所述语音文件对应的标准文本作为目标标准文本；

所述将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域具体为：将所述目标标准文本的易错词所属的词语领域确定为所述指定语音识别模型的弱支持领域。

6. 一种语音识别测试装置，其特征在于，包括：

平台确定模块，用于确定待语音测试的各个目标平台；

文件文本获取模块，用于获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

语音识别模块，用于调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

错词率计算模块，用于针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

第一报告生成模块，用于根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

7. 根据权利要求 6 所述的语音识别测试装置，其特征在于，所述第一报告生成模块包括：

平台错词率计算单元，用于针对每个目标平台，根据所述目标平台上各个语音文件对

应的错词率计算所述目标平台的平台错词率；

排序单元，用于在计算得到各个目标平台对应的平台错词率后，按照预设的排序规则对所述各个目标平台对应的平台错词率进行排序；

导入报告单元，用于生成空白的第一测试报告，并将排序后的所述各个目标平台对应的平台错词率导入所述第一测试报告。

8. 根据权利要求 6 或 7 所述的语音识别测试装置，其特征在于，所述语音识别测试装置还包括：

指定模块确定模块，用于分别确定所述各个目标平台的指定语音识别模型；

模型识别模块，用于调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别，并获取各个目标平台输出的识别文本；

模型错词率计算模块，用于针对每个所述目标平台，根据各个语音文件对应的识别文本和对应的标准文本计算出在各个目标平台上所述指定语音识别模型对应的错词率；

第二报告生成模块，用于根据各个目标平台上所述指定语音识别模型对应的错词率生成语音测试的第二测试报告。

9. 根据权利要求 8 所述的语音识别测试装置，其特征在于，所述语音识别测试装置还包括：

识别文本获取模块，用于针对每个指定语音识别模型，获取各个所述语音文件输入所述指定语音识别模型进行语音识别后得到的各个所述语音文件所对应识别文本；

易错词确定模块，用于针对每个指定语音识别模型，将各个所述语音文件所对应识别文本与标准文本进行对比，确定各个所述标准文本的易错词，所述易错词是指识别文本与标准文本对比后、所述标准文本中对比不一致的文本；

词语领域确定模块，用于针对每个指定语音识别模型，确定各个所述标准文本的易错词所属的词语领域；

弱支持领域确定模块，用于针对每个指定语音识别模型，将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域，所述弱支持领域是指语音识别模型用于语音识别的效果未达到期望值的词语领域。

10. 根据权利要求 9 所述的语音识别测试装置，其特征在于，所述语音识别测试装置还包括：

各模型错词计算模块，用于针对每个指定语音识别模型，根据各个所述语音文件所对

应的识别文本和对应的标准文本计算出所述指定语音识别模型识别各个所述语音文件时的模型错词率；

语音文件筛选模块，用于针对每个指定语音识别模型，筛选模型错词率超过预设错词率阈值的所述语音文件；

目标标准文本确定模块，用于针对每个指定语音识别模型，确定出与筛选出的所述语音文件对应的标准文本作为目标标准文本；

所述弱支持领域确定模块具体可以用于将所述目标标准文本的易错词所属的词语领域确定为所述指定语音识别模型的弱支持领域。

11. 一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，其特征在于，所述处理器执行所述计算机可读指令时实现如下步骤：

确定待语音测试的各个目标平台；

获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

12. 根据权利要求 11 所述的计算机设备，其特征在于，所述根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告包括：

针对每个目标平台，根据所述目标平台上各个语音文件对应的错词率计算所述目标平台的平台错词率；

在计算得到各个目标平台对应的平台错词率后，按照预设的排序规则对所述各个目标平台对应的平台错词率进行排序；

生成空白的第一测试报告，并将排序后的所述各个目标平台对应的平台错词率导入所述第一测试报告。

13. 根据权利要求 11 或 12 所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还实现如下步骤：

分别确定所述各个目标平台的指定语音识别模型；

调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送至各个目标平

台上的指定语音识别模型进行语音识别，并获取各个目标平台输出的识别文本；

针对每个所述目标平台，根据各个语音文件对应的识别文本和对应的标准文本计算出在各个目标平台上所述指定语音识别模型对应的错词率；

根据各个目标平台上所述指定语音识别模型对应的错词率生成语音测试的第二测试报告。

14. 根据权利要求 13 所述的计算机设备，其特征在于，在分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别之后，所述处理器执行所述计算机可读指令时还实现如下步骤：

针对每个指定语音识别模型，获取各个所述语音文件输入所述指定语音识别模型进行语音识别后得到的各个所述语音文件所对应识别文本；

针对每个指定语音识别模型，将各个所述语音文件所对应识别文本与标准文本进行对比，确定各个所述标准文本的易错词，所述易错词是指识别文本与标准文本对比后、所述标准文本中对比不一致的文本；

针对每个指定语音识别模型，确定各个所述标准文本的易错词所属的词语领域；

针对每个指定语音识别模型，将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域，所述弱支持领域是指语音识别模型用于语音识别的效果未达到期望值的词语领域。

15. 根据权利要求 14 所述的计算机设备，其特征在于，在将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域之前，所述处理器执行所述计算机可读指令时还实现如下步骤：

针对每个指定语音识别模型，根据各个所述语音文件所对应的识别文本和对应的标准文本计算出所述指定语音识别模型识别各个所述语音文件时的模型错词率；

针对每个指定语音识别模型，筛选模型错词率超过预设错词率阈值的所述语音文件；

针对每个指定语音识别模型，确定出与筛选出的所述语音文件对应的标准文本作为目标标准文本；

所述将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域具体为：将所述目标标准文本的易错词所属的词语领域确定为所述指定语音识别模型的弱支持领域。

16. 一个或多个存储有计算机可读指令的非易失性可读存储介质，其特征在于，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如下步骤：

确定待语音测试的各个目标平台；

获取测试用的各个语音文件以及所述各个语音文件对应的标准文本；

调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送给各个目标平台进行语音识别，并获取各个目标平台输出的识别结果；

针对每个所述目标平台，根据各个语音文件对应的识别结果和对应的标准文本计算出在所述目标平台上各个语音文件对应的错词率；

根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告。

17. 根据权利要求 16 所述的非易失性可读存储介质，其特征在于，所述根据各个目标平台上各个语音文件对应的错词率生成语音测试的第一测试报告包括：

针对每个目标平台，根据所述目标平台上各个语音文件对应的错词率计算所述目标平台的平台错词率；

在计算得到各个目标平台对应的平台错词率后，按照预设的排序规则对所述各个目标平台对应的平台错词率进行排序；

生成空白的第一测试报告，并将排序后的所述各个目标平台对应的平台错词率导入所述第一测试报告。

18. 根据权利要求 16 或 17 所述的非易失性可读存储介质，其特征在于，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还执行如下步骤：

分别确定所述各个目标平台的指定语音识别模型；

调用各个目标平台的语音识别服务接口，分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别，并获取各个目标平台输出的识别文本；

针对每个所述目标平台，根据各个语音文件对应的识别文本和对应的标准文本计算出在各个目标平台上所述指定语音识别模型对应的错词率；

根据各个目标平台上所述指定语音识别模型对应的错词率生成语音测试的第二测试报告。

19. 根据权利要求 18 所述的非易失性可读存储介质，其特征在于，在分别将各个所述语音文件发送至各个目标平台上的指定语音识别模型进行语音识别之后，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还执行如下步骤：

针对每个指定语音识别模型，获取各个所述语音文件输入所述指定语音识别模型进行语音识别后得到的各个所述语音文件所对应识别文本；

针对每个指定语音识别模型，将各个所述语音文件所对应识别文本与标准文本进行对比，确定各个所述标准文本的易错词，所述易错词是指识别文本与标准文本对比后、所述

标准文本中对比不一致的文本；

针对每个指定语音识别模型，确定各个所述标准文本的易错词所属的词语领域；

针对每个指定语音识别模型，将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域，所述弱支持领域是指语音识别模型用于语音识别的效果未达到期望值的词语领域。

20. 根据权利要求 19 所述的非易失性可读存储介质，其特征在于，在将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域之前，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还执行如下步骤：

针对每个指定语音识别模型，根据各个所述语音文件所对应的识别文本和对应的标准文本计算出所述指定语音识别模型识别各个所述语音文件时的模型错词率；

针对每个指定语音识别模型，筛选模型错词率超过预设错词率阈值的所述语音文件；

针对每个指定语音识别模型，确定出与筛选出的所述语音文件对应的标准文本作为目标标准文本；

所述将确定出的所述词语领域确定为所述指定语音识别模型的弱支持领域具体为：将所述目标标准文本的易错词所属的词语领域确定为所述指定语音识别模型的弱支持领域。

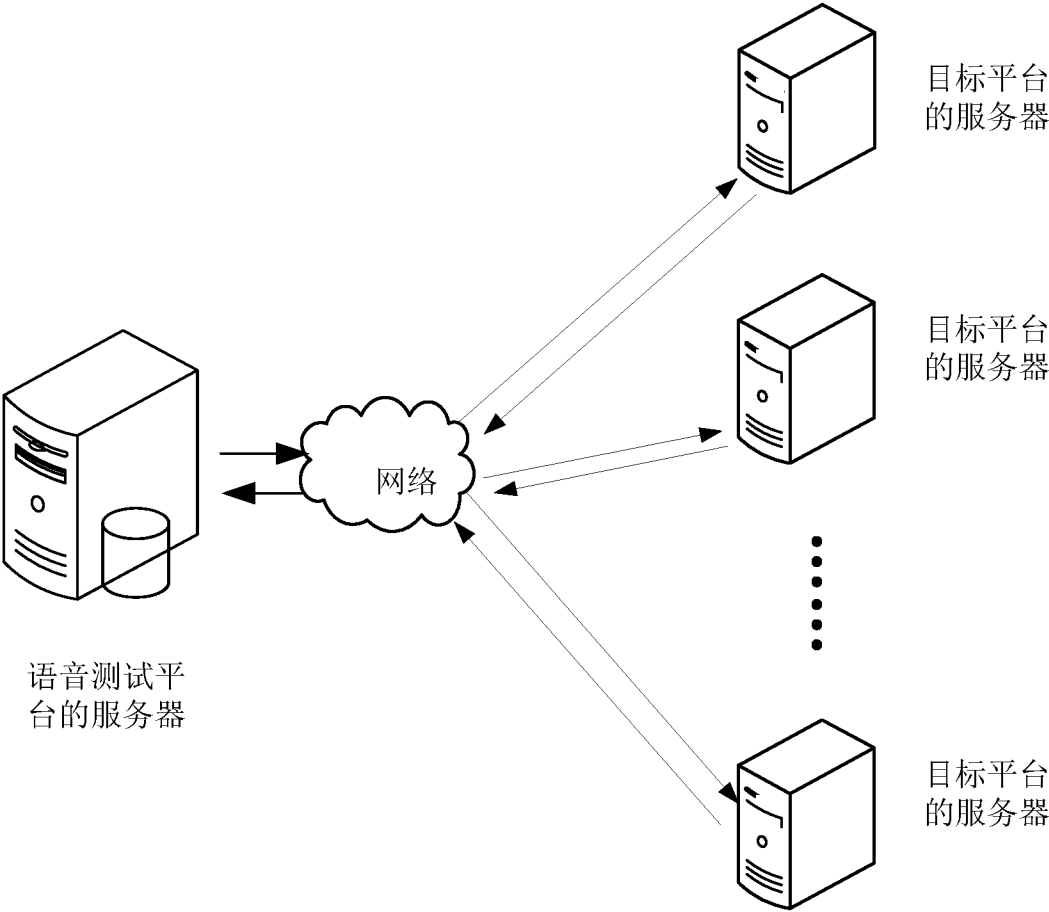


图 1

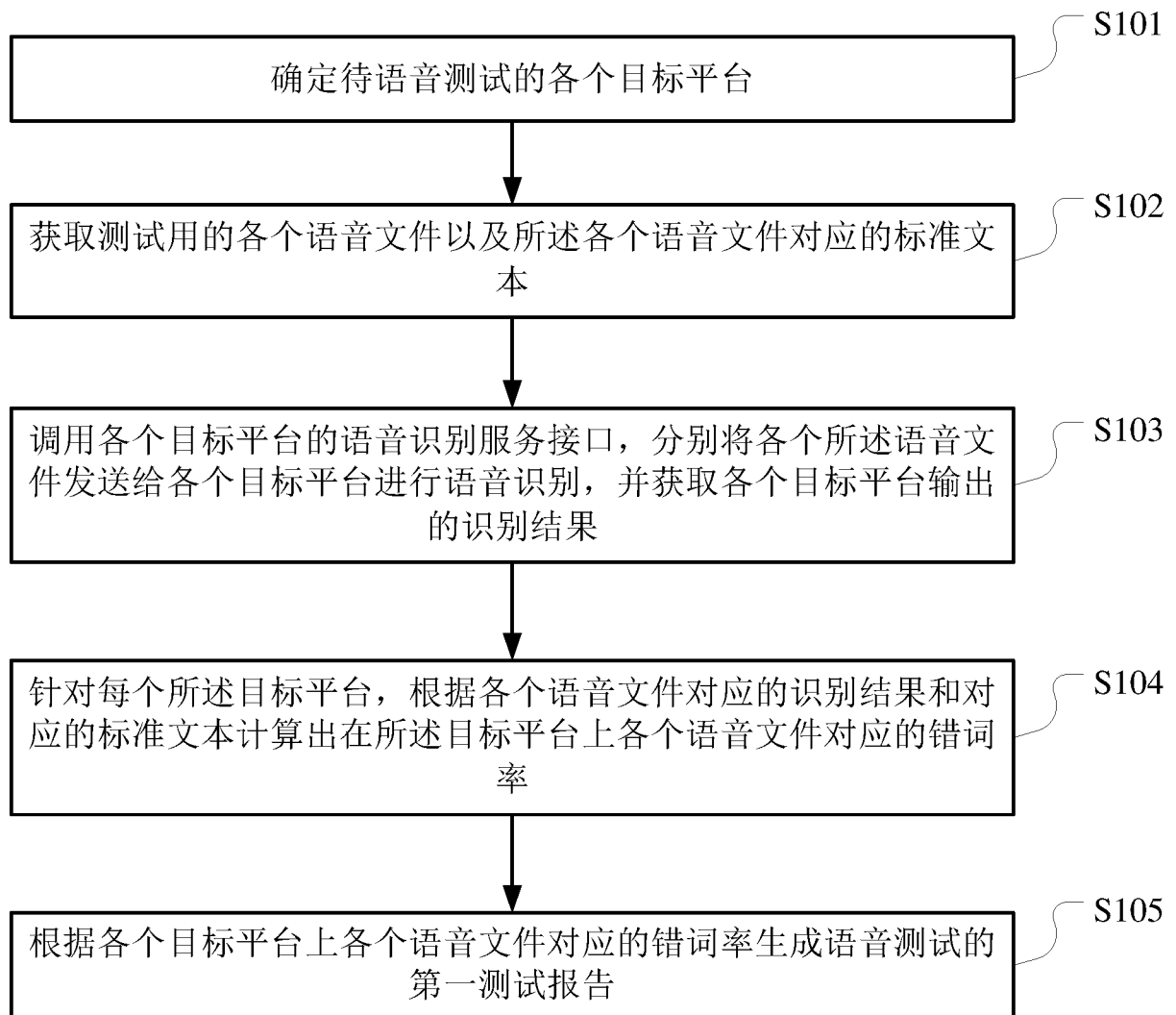


图 2

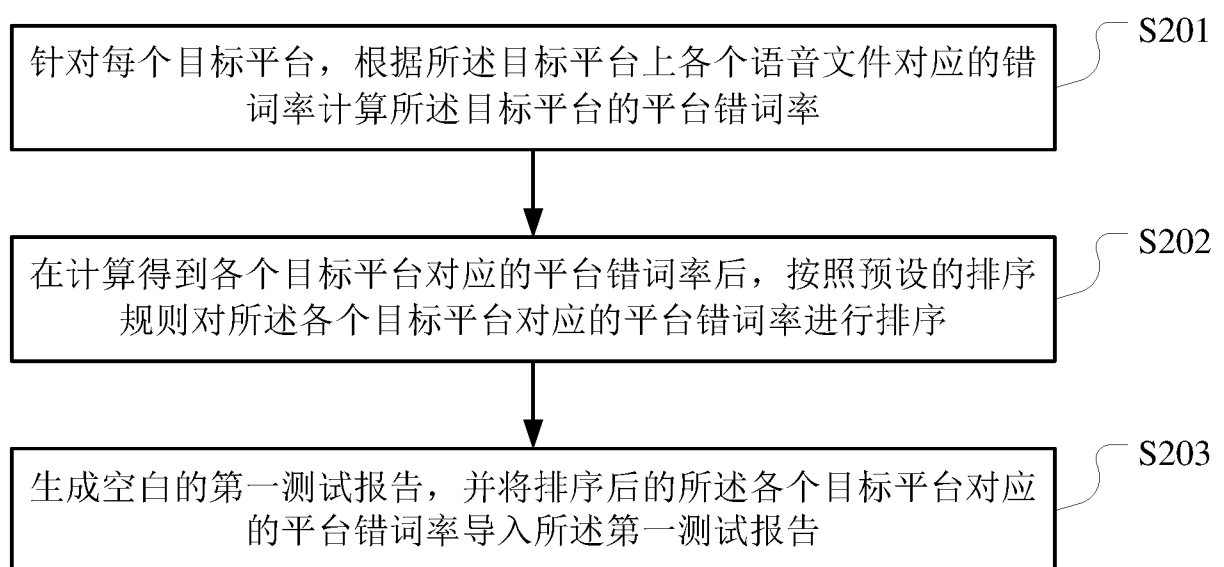


图 3

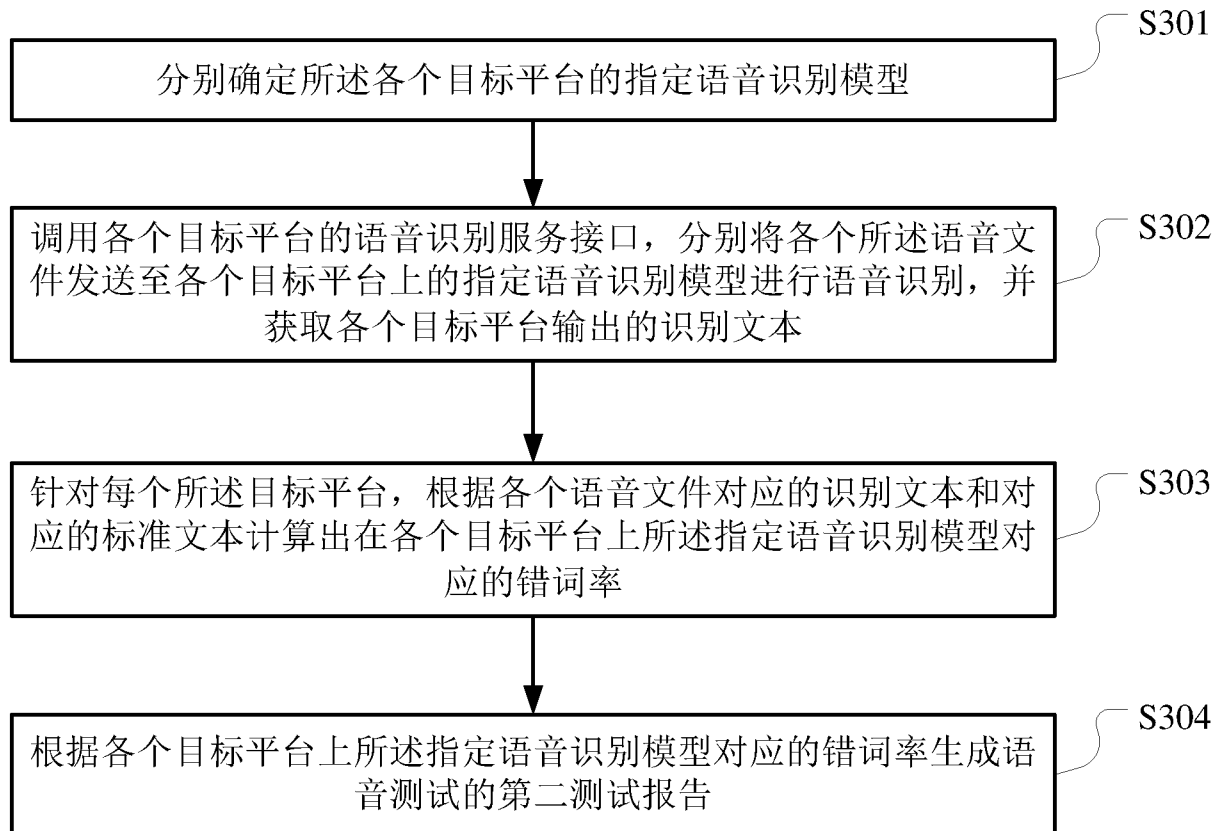


图 4

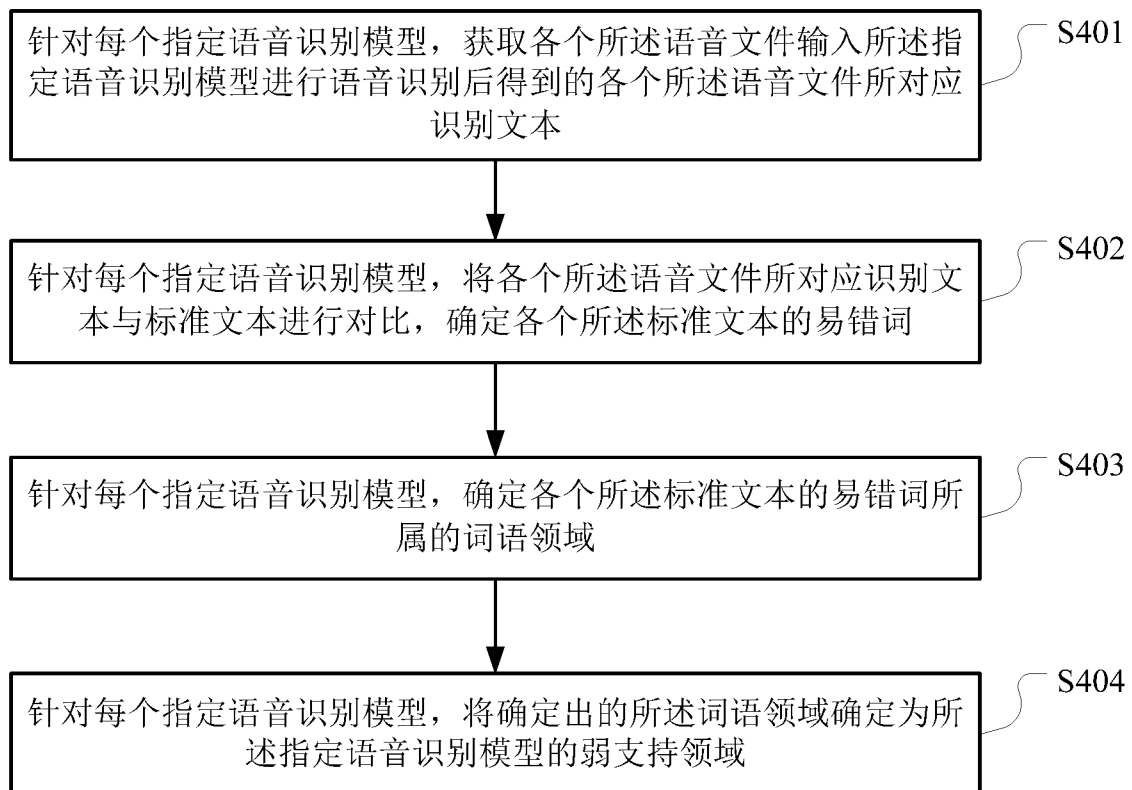


图 5

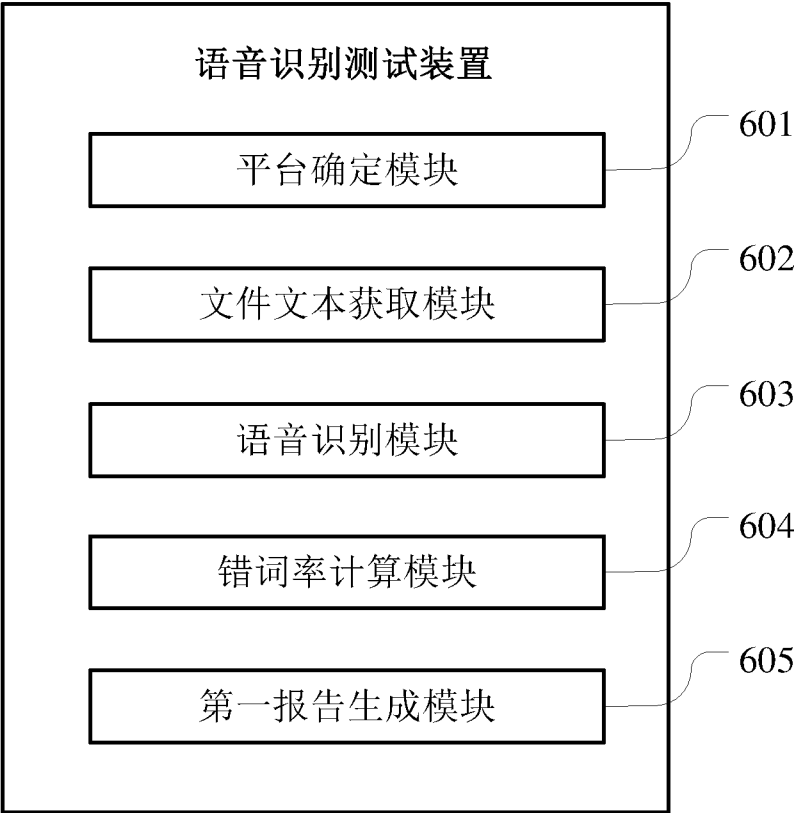


图 6

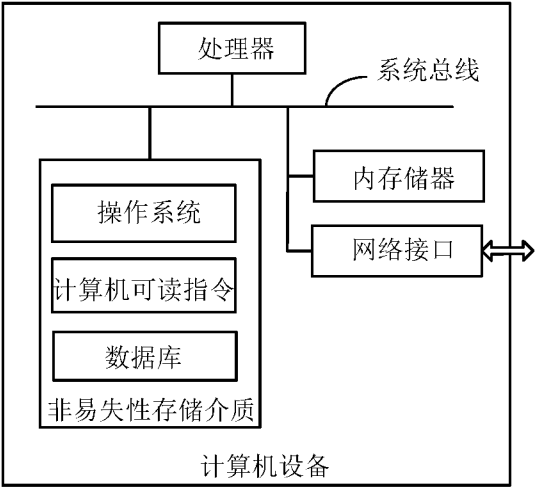


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/106256

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/01(2013.01)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)

G10L 15/-

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: 平安科技, 语音, 识别, 测试, 评估, 文本, 模型, speech, voice, recogni+, test+, evaluat+, estimat+, text, model+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104538042 A (ACOSTEK CO., LTD) 22 April 2015 (2015-04-22) description, paragraphs [0019]-[0022], and figures 1 and 2	1-20
X	CN 103578463 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 12 February 2014 (2014-02-12) description, paragraphs [0028]-[0041], and figures 1-3	1, 2, 6, 7, 11, 12, 16, 17
X	CN 103745731 A (ANHUI USTC IFLYTEK CO., LTD.) 23 April 2014 (2014-04-23) claims 1-3	1, 2, 6, 7, 11, 12, 16, 17
A	CN 107039050 A (ALIBABA GROUP HOLDING LIMITED) 11 August 2017 (2017-08-11) entire document	1-20
A	CN 106228986 A (BEIJING QIHOO TECHNOLOGY CO., LTD. ET AL.) 14 December 2016 (2016-12-14) entire document	1-20
A	JP 2014202781 A (NIPPON TELEGRAPH & TELEPHONE CORP.) 27 October 2014 (2014-10-27) entire document	1-20
A	KR 20110061987 A (KOREA IND. TECH. INST.) 10 June 2011 (2011-06-10) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“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)

“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

“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

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search

12 March 2019

Date of mailing of the international search report

02 April 2019

Name and mailing address of the ISA/CN

**National Intellectual Property Administration, PRC (ISA/
CN)**
**No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China**

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/106256

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	104538042	A	22 April 2015	None			
CN	103578463	A	12 February 2014	CN	103578463	B	01 December 2017
CN	103745731	A	23 April 2014	CN	103745731	B	19 October 2016
CN	107039050	A	11 August 2017	None			
CN	106228986	A	14 December 2016	None			
JP	2014202781	A	27 October 2014	JP	6078402	B2	08 February 2017
KR	20110061987	A	10 June 2011	KR	101145401	B1	16 May 2012

国际检索报告

国际申请号

PCT/CN2018/106256

A. 主题的分类 G10L 15/01 (2013.01) 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) G10L 15/- 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, EP0D0C: 平安科技, 语音, 识别, 测试, 评估, 文本, 模型, speech, voice, recogni+, test+, evaluat+, estimat+, text, model+																										
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>X</td> <td>CN 104538042 A (南京声准科技有限公司) 2015年 4月 22日 (2015 - 04 - 22) 说明书第[0019]-[0022]段、附图1, 2</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 103578463 A (腾讯科技深圳有限公司) 2014年 2月 12日 (2014 - 02 - 12) 说明书第[0028]-[0041]段、附图1-3</td> <td>1, 2, 6, 7, 11, 12, 16, 17</td> </tr> <tr> <td>X</td> <td>CN 103745731 A (安徽科大讯飞信息科技股份有限公司) 2014年 4月 23日 (2014 - 04 - 23) 权利要求1-3</td> <td>1, 2, 6, 7, 11, 12, 16, 17</td> </tr> <tr> <td>A</td> <td>CN 107039050 A (阿里巴巴集团控股有限公司) 2017年 8月 11日 (2017 - 08 - 11) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 106228986 A (北京奇虎科技有限公司 等) 2016年 12月 14日 (2016 - 12 - 14) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>JP 2014202781 A (NIPPON TELEGRAPH & TELEPHONE CORP.) 2014年 10月 27日 (2014 - 10 - 27) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>KR 20110061987 A (KOREA IND. TECH. INST.) 2011年 6月 10日 (2011 - 06 - 10) 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	X	CN 104538042 A (南京声准科技有限公司) 2015年 4月 22日 (2015 - 04 - 22) 说明书第[0019]-[0022]段、附图1, 2	1-20	X	CN 103578463 A (腾讯科技深圳有限公司) 2014年 2月 12日 (2014 - 02 - 12) 说明书第[0028]-[0041]段、附图1-3	1, 2, 6, 7, 11, 12, 16, 17	X	CN 103745731 A (安徽科大讯飞信息科技股份有限公司) 2014年 4月 23日 (2014 - 04 - 23) 权利要求1-3	1, 2, 6, 7, 11, 12, 16, 17	A	CN 107039050 A (阿里巴巴集团控股有限公司) 2017年 8月 11日 (2017 - 08 - 11) 全文	1-20	A	CN 106228986 A (北京奇虎科技有限公司 等) 2016年 12月 14日 (2016 - 12 - 14) 全文	1-20	A	JP 2014202781 A (NIPPON TELEGRAPH & TELEPHONE CORP.) 2014年 10月 27日 (2014 - 10 - 27) 全文	1-20	A	KR 20110061987 A (KOREA IND. TECH. INST.) 2011年 6月 10日 (2011 - 06 - 10) 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																								
X	CN 104538042 A (南京声准科技有限公司) 2015年 4月 22日 (2015 - 04 - 22) 说明书第[0019]-[0022]段、附图1, 2	1-20																								
X	CN 103578463 A (腾讯科技深圳有限公司) 2014年 2月 12日 (2014 - 02 - 12) 说明书第[0028]-[0041]段、附图1-3	1, 2, 6, 7, 11, 12, 16, 17																								
X	CN 103745731 A (安徽科大讯飞信息科技股份有限公司) 2014年 4月 23日 (2014 - 04 - 23) 权利要求1-3	1, 2, 6, 7, 11, 12, 16, 17																								
A	CN 107039050 A (阿里巴巴集团控股有限公司) 2017年 8月 11日 (2017 - 08 - 11) 全文	1-20																								
A	CN 106228986 A (北京奇虎科技有限公司 等) 2016年 12月 14日 (2016 - 12 - 14) 全文	1-20																								
A	JP 2014202781 A (NIPPON TELEGRAPH & TELEPHONE CORP.) 2014年 10月 27日 (2014 - 10 - 27) 全文	1-20																								
A	KR 20110061987 A (KOREA IND. TECH. INST.) 2011年 6月 10日 (2011 - 06 - 10) 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 2019年 3月 12日		国际检索报告邮寄日期 2019年 4月 2日																								
ISA/CN的名称和邮寄地址 中国国家知识产权局 (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 王馨宁 电话号码 86- (10) -53962649																								

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/106256

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	104538042	A	2015年 4月 22日	无	
CN	103578463	A	2014年 2月 12日	CN 103578463	B 2017年 12月 1日
CN	103745731	A	2014年 4月 23日	CN 103745731	B 2016年 10月 19日
CN	107039050	A	2017年 8月 11日	无	
CN	106228986	A	2016年 12月 14日	无	
JP	2014202781	A	2014年 10月 27日	JP 6078402	B2 2017年 2月 8日
KR	20110061987	A	2011年 6月 10日	KR 101145401	B1 2012年 5月 16日