(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 7 月 2 日 (02.07.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/134896 A1

(51) 国际专利分类号:
G10L 13/02 (2013.01) *H04L 29/08* (2006.01)

(21) 国际申请号: PCT/CN2019/122545

(22) 国际申请日: 2019 年 12 月 3 日 (03.12.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811596879.5 2018年12月26日 (26.12.2018) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (**ALIBABA GROUP HOLDING LIMITED**) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: 韩喆 (**HAN, Zhe**); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 王磊 (**WANG, Lei**); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 傅春霖 (**FU, Chunlin**); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京博思佳知识产权代理有限公司 (**BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION**); 中国北京市海淀区上地三街9号嘉华大厦B座409, Beijing 100085 (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,

(54) **Title:** METHOD AND DEVICE FOR INVOKING SPEECH SYNTHESIS FILE

(54) 发明名称: 一种语音合成文件的调用方法及装置

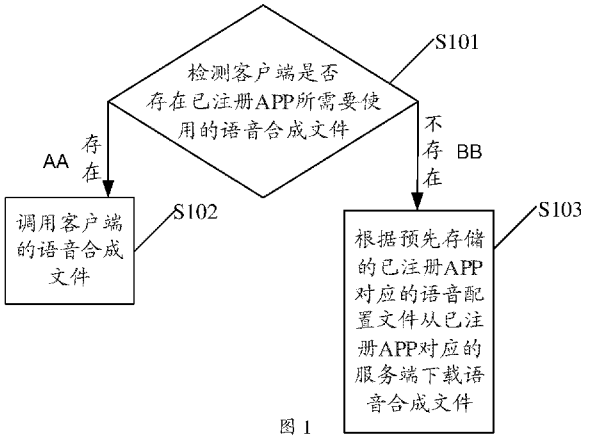


图 1

S101 Detect whether a client has a speech synthesis file needed to be used by a registered app
S102 Invoke the speech synthesis file of the client
S103 Download, according to a pre-stored speech configuration file corresponding to the registered app, a speech synthesis file from a server side corresponding to the registered app
AA Having
BB Not having

(57) **Abstract:** A method and a device for invoking a speech synthesis file. Said method comprises: detecting whether a client has a speech synthesis file needed to be used by a registered app (S101), the registered app being a pre-registered app requiring the use of a speech synthesis file; if it is detected that the client does not have a speech synthesis file, downloading, according to a pre-stored speech configuration file corresponding to the registered app, a speech synthesis file from a server side corresponding to the registered app (S103), the speech configuration file including the download address of the speech synthesis file; and if it is detected that the client has the speech synthesis file, invoking the speech synthesis file of the client (S102), so that the registered app performs speech playback according to the speech synthesis file. When a registered app needs to use a speech synthesis file, whether a client has a speech synthesis file is detected, and if the client has a speech synthesis file, the speech synthesis file cached in the client is preferentially invoked, reducing the response time of the entire speech system.

WO 2020/134896 A1

[见续页]

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))。

(57) 摘要：一种语音合成文件的调用方法及装置，方法包括：检测客户端是否存在已注册APP所需要使用的语音合成文件（S101），已注册APP为预先注册需要使用语音合成文件的APP；若检测出客户端不存在语音合成文件，根据预先存储的已注册APP对应的语音配置文件从已注册APP对应的服务端下载语音合成文件（S103），语音配置文件内置有语音合成文件的下载地址；若检测出客户端存在语音合成文件，调用客户端的语音合成文件（S102），以供已注册APP根据语音合成文件进行语音播放。已注册APP需要使用语音合成文件时，检测客户端是否存在语音合成文件，在客户端存在语音合成文件时优先调用客户端缓存的语音合成文件，减少整个语音系统的响应时间。

一种语音合成文件的调用方法及装置

技术领域

[01] 本说明书涉及计算机领域，尤其是涉及一种语音合成文件的调用方法及装置。

背景技术

[02] 随着互联网的发展，多方合作已经体现在越来越多的方面。建设一个大型语音系统时，终端的框架和服务端由运行商进行搭建，但终端的应用需要多个 ISV（独立软件开发商）来共同完成。

[03] 现有的大型语音系统中，ISV 开发的 APP 调用语音合成文件进行语音播放时，每次都需要由服务端合成该语音合成文件，再将该语音合成文件下载至终端进行调用，整个过程使得系统的响应时间增加，严重的还会造成整个系统的瘫痪，从而影响系统的正常运行。

发明内容

[04] 本说明书实施例提供一种语音合成文件的调用方法及装置，解决了上述背景技术提出的问题。

[05] 为解决上述技术问题，本说明书实施例是这样实现的：

[06] 本说明书实施例提供的一种语音合成文件的调用方法，该方法包括：

[07] 检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP；

[08] 若检测出客户端不存在所述语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地址；

[09] 若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

[10] 可选的，所述检测客户端是否存在已注册 APP 所需要使用的语音合成文件之前，所述方法还包括：

[11] 向所述已注册 APP 对应的服务端拉取所述语音配置文件;

[12] 接收所述已注册 APP 对应的服务端下发的语音配置文件，下发的所述语音配置文件包括所述已注册 APP 对应的服务端对下发的所述语音配置文件进行加密后，分配给所述已注册 APP 对应的第一验证信息;

[13] 判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配;

[14] 在判断出所述第一验证信息与客户端预先保存的第二验证信息匹配时，则验证下发的所述语音配置文件正确。

[15] 可选的，判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配，具体包括:

[16] 根据所述已注册 APP 的标识从内置于客户端安全运行环境中预先保存的与所述已注册 APP 对应的第二验证信息;

[17] 判断所述第一验证信息与第二验证信息是否匹配。

[18] 可选的，所述向所述已注册 APP 对应的服务端拉取所述语音配置文件之前，所述方法还包括:

[19] 向所述已注册 APP 对应的服务端发送所述 APP 开发者提供的反映所述 APP 开发者特征的语音数据，以便所述已注册 APP 对应的服务端通过内置的语音基础训练模型训练出所述 APP 开发者定制化的语音模型，并将预先储存的文本输入所述 APP 开发者定制化的语音模型以生成已注册 APP 需要的语音合成文件，所述语音基础训练模型为根据所述已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型。

[20] 可选的，所述已注册 APP 根据所述语音合成文件进行语音播放之前，所述方法还包括:

[21] 计算所述语音合成文件对应的第一摘要值;

[22] 判断根据所述语音配置文件内预先储存的所述语音合成文件对应的第二摘要值与所述第一摘要值是否相同;

[23] 若判断出所述第二摘要值与所述第一摘要值相同时，则所述已注册 APP 根据所述语音合成文件进行语音播放。

[24] 可选的，所述已注册 APP 根据所述语音合成文件进行语音播放，具体包括：所述

已注册 APP 对应的服务端根据预设规则对所述语音合成文件进行加密；所述加密的语音合成文件根据内置解密模块解密后，由所述已注册 APP 进行语音播放。

[25] 本说明书实施例提供的一种语音合成文件的调用装置，所述装置包括：

[26] 检测单元，用于检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP；

[27] 下载单元，用于若检测出客户端不存在所述语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地址；

[28] 调用单元，用于若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

[29] 可选的，所述装置还包括：

[30] 拉取单元，用于向所述已注册 APP 对应的服务端拉取所述语音配置文件；

[31] 接收单元，用于接收所述已注册 APP 对应的服务端下发的语音配置文件，下发的所述语音配置文件包括所述已注册 APP 对应的服务端对下发的所述语音配置文件进行加密后，分配给所述已注册 APP 对应的第一验证信息；

[32] 判断单元，用于判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配；

[33] 验证单元，用于在判断出所述第一验证信息与客户端预先保存的第二验证信息匹配时，验证下发的所述语音配置文件正确。

[34] 可选的，所述判断单元具体用于：

[35] 根据所述已注册 APP 的标识从内置于客户端安全运行环境中预先保存的与所述已注册 APP 对应的第二验证信息；

[36] 判断所述第一验证信息与第二验证信息是否匹配。

[37] 可选的，所述装置还包括：

[38] 训练单元，用于向所述已注册 APP 对应的服务端发送所述 APP 开发者提供的反映所述 APP 开发者特征的语音数据，以便所述已注册 APP 对应的服务端通过内置的语音基础训练模型训练出所述 APP 开发者定制化的语音模型，并根据预先储存的文本由所

述 APP 开发者定制化的语音模型生成已注册 APP 对应的语音合成文件，所述语音基础训练模型为根据所述已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型。

[39] 可选的，所述装置还包括:

[40] 计算单元，用于计算所述语音合成文件对应的第一摘要值;

[41] 所述判断单元还用于判断根据所述语音配置文件内预先储存的所述语音合成文件对应的第二摘要值与所述第一摘要值是否相同;

[42] 所述判断单元若判断出所述第二摘要值与所述第一摘要值相同时，则所述已注册 APP 根据所述语音合成文件进行语音播放。

[43] 可选的，所述已注册 APP 根据所述语音合成文件进行语音播放，具体包括：所述已注册 APP 对应的服务端根据预设规则对所述语音合成文件进行加密；所述加密的语音合成文件根据内置解密模块解密后，由所述已注册 APP 进行语音播放。

[44] 本说明书实施例提供的一种语音系统，包括终端、服务器，终端包括运行在终端中的语音 SDK、已注册 APP 以及 APP 开发者端;

[45] 所述 APP 开发者端用于向所述已注册 APP 对应的服务端发送所述 APP 开发者提供的反映所述 APP 开发者特征的语音数据;

[46] 所述服务端用于通过内置的语音基础训练模型训练出所述 APP 开发者定制化的语音模型，并将预先储存的文本输入所述 APP 开发者定制化的语音模型以生成已注册 APP 需要的语音合成文件，所述语音基础训练模型为根据所述已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型;

[47] 所述语音 SDK 用于向所述已注册 APP 对应的服务端拉取所述语音配置文件; 接收所述已注册 APP 对应的服务端下发的语音配置文件，所述下发语音配置文件包括所述已注册 APP 对应的服务端对所述下发的语音配置文件进行加密后，分配给所述已注册 APP 对应的第一验证信息; 判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配; 在判断出所述第一验证信息与客户端预先保存的第二验证信息匹配时，则验证所述下发的语音配置文件正确;检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP; 若检测出客户端不存在所述语音合成文件，根据已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地

址;若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

[48] 本说明书实施例提供的一种计算机可读介质，其上存储有计算机可读指令，所述计算机可读指令可被处理器执行以下步骤:

[49] 检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP;

[50] 若检测出客户端不存在所述语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地址;

[51] 若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

[52] 本说明书实施例提供的一种语音合成文件的调用设备，该设备包括用于存储计算机程序指令的存储器和用于执行程序指令的处理器，其中，当该计算机程序指令被该处理器执行时，触发该设备执行以下步骤:

[53] 检测单元，用于检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP;

[54] 下载单元，用于若检测出客户端不存在所述语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地址;

[55] 调用单元，用于若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

[56] 本说明书实施例采用的上述至少一个技术方案能够达到以下有益效果:

[57] 1、已注册 APP 需要使用语音合成文件时，检测客户端是否缓存该语音合成文件，在客户端存在该语音合成文件时优先调用客户端缓存的语音合成文件，减少整个语音系统的响应时间;

[58] 2、APP 开发者通过已注册 APP 对应的服务器可以训练出 APP 开发者定制化的语音模型，再将预先储存的文本输入至 APP 开发者定制化的语音模型以生成 APP 开发者需要使用语音合成文件，已注册 APP 需要使用其中的语音合成文件时再将对应的语音

合成文件下载以供已注册 APP 进行语音播放;

[59] 3、该语音系统可以支持多个已注册 APP，使得该语音系统使用率得到充分利用。

附图说明

[60] 为了更清楚地说明本说明书实施例或现有技术中的技术方案，下面将对实施例或现有技术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说明书中记载的一些实施例，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。

[61] 图 1 为本说明书实施例一提供的语音合成文件的调用方法的流程示意图;

[62] 图 2 为本说明书实施例二提供的语音合成文件的调用方法的流程示意图;

[63] 图 3 为本说明书实施例三提供的语音合成文件的调用装置的结构示意图;

[64] 图 4 为本说明书实施例四提供的语音系统的结构示意图。

具体实施方式

[65] 为了使本技术领域的人员更好地理解本说明书中的技术方案，下面将结合本说明书实施例中的附图，对本说明书实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本说明书实施例，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，都应当属于本申请保护的范围。

[66] 图 1 为本说明书实施例提供的一种语音合成文件的调用方法的流程示意图，该流程示意图包括:

[67] 步骤 S101，检测客户端是否存在已注册 APP 所需要使用的语音合成文件，若存在，则执行步骤 S102，若不存在，则执行步骤 S103。

[68] 在本说明书实施例的步骤 S101 中，可由语音 SDK 执行检测客户端是否存在已注册 APP 所需要使用的语音合成文件的步骤，语音 SDK 设置有同时连接多个 APP 的接口，即 APP 向语音 SDK 进行注册，就是将 APP 数据连接至语音 SDK，已注册 APP 为预先向语音 SDK 注册且需要使用语音合成文件的 APP。其中，在本实施例中语音 SDK 为 APP 开发者在开发软件时的框架。

[69] 在本说明书实施例的步骤S101中，语音合成文件是由已注册APP对应的服务端根据APP开发者的需求训练出的。首先APP开发者向已注册APP对应的服务端发送APP开发者提供的反映APP开发者特征的语音数据，以便已注册APP对应的服务端通过内置的语音基础训练模型训练出APP开发者定制化的语音模型，并将预先储存的文本输入APP开发者定制化的语音模型以生成已注册APP需要的语音合成文件。语音基础训练模型为根据已注册APP播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册APP共用的模型。其中，若干语音样本为已注册APP对应的服务端储存的高质量语音数据。

[70] 进一步的，在本说明书实施例的步骤S101中，语音基础训练模型根据整个语音系统的精确度来确定高质量语音数据的采样时间，在整个语音系统要求的精确度高时时，高质量语音数据的采样时间可以为300小时，但对于整个语音系统要求的精确度不高时，高质量语音数据的采样时间选取100小时。

[71] 在本说明书实施例的步骤S101中，已注册APP对应的服务端在训练出语音基础训练模型后，APP开发者上传反映APP开发者特征的语音数据至已注册APP对应的服务端，通过语音基础训练模型训练出APP开发者定制化的语音模型。反映APP开发者特征的语音数据是根据APP开发者需求的语言环境而录制的语音数据。此时APP开发者只需将少量上传的语音数据上传至已注册APP对应的服务端。其中，语音基础训练模型可以理解为已注册APP对应的服务端为APP开发者提供的数据集很大的中间模型，然后将此中间模型为APP开发者上传的语音数据进行调优训练，从而得出反映APP开发者特征的定制化的语音模型。

[72] 在本说明书实施例的步骤S101中，APP开发者上传的语音数据需要进行审核，在生成反映APP开发者特征的定制化的语音模型后，由该语音系统的管理人员进行审核，此时审核的机制可以为审核通过后该反映APP开发者特征的定制化的语音模型才可以正常使用，也就是说即使生成了反映APP开发者特征的定制化的语音模型，但未经过审核人员审核通过，该反映APP开发者特征的定制化的语音模型也是无法正常使用的；同时审核的机制也可以为不管该反映APP开发者特征的定制化的语音模型的审核结果是否通过，已注册的APP皆可正常使用，但是审核人员一旦检测出该反映APP开发者特征的定制化的语音模型不合格时，该该反映APP开发者特征的定制化的语音模型即失效。

[73] 在本说明书实施例的步骤S101中，若是APP开发者不采用这种方案，而是使用传

统的方法实现定制化的要求，其一是 APP 开发者直接上传反映 APP 开发者特征的语音数据，不经过任何处理，这样做使得鲁棒性低；其二是 APP 开发者单独制作反映 APP 开发者特征的定制化的语音模型，该过程在执行的过程时间长，而且无法保证定制化的语音模型的质量。

[74] 在本说明书实施例的步骤 S101 中，该语音系统也可以应用于视频系统中，即已注册 APP 对应的服务端内储存视频基础训练模型。

[75] 步骤 S102，调用客户端的语音合成文件。

[76] 在本说明书实施例 S102 中，已向语音 SDK 已注册 APP 有需要使用的语音合成文件时，语音 SDK 优先检测客户端是否存在，在客户端存在所需要调用的配置文件时，调用存放在客户端的语音合成文件，已注册 APP 可以根据语音合成文件进行语音播放。

[77] 步骤 S103，根据预先存储的已注册 APP 对应的语音配置文件从已注册 APP 对应的服务端下载语音合成文件。

[78] 在本说明书实施例的步骤 S103 中，语音合成文件是根据预先储存的文本并由 APP 开发者定制化的语音模型所生成。若是在步骤 S102 的判断时语音合成文件不存在，说明该语音合成文件以前从未被已注册 APP 下载。

[79] 在本说明书实施例的步骤 S103 中，语音配置文件内置有语音合成文件的下载地址，已注册 APP 根据该语音合成文件的下载地址下载所需要使用的语音合成文件，以供已注册 APP 根据语音合成文件进行语音播放。

[80] 在本说明书实施例的步骤 S103 中，已注册 APP 根据语音合成文件进行语音播放之前，还需要对语音合成文件进行验证，具体步骤可以为：

[81] 步骤 1、计算语音合成文件对应的第一摘要值。

[82] 在本说明书实施例的步骤 1 中，语音合成文件对应的第一摘要值检查下载的语音合成文件是否出错，或者说下载的语音合成文件是否被篡改的参数值。在本实施例中可以采用 MD5 摘要实现，其中，MD5 是一种被广泛使用的密码散列函数，可以产生出一个 128 位（16 字节）的散列值（hash value），用于确保下载的语音配置文件是否出错，或者下载的语音配置文件是否被篡改。比如，在 Unix 下有很多软件在下载的时候都有一个文件名相同，文件扩展名为.md5 的文件，在这个文件中通常只有一行文本，大致结构如：

[83] MD5 (tanajiya.tar.gz) = 38b8c2c1093dd0fec383a9d9ac940515

[84] 这就是 tanajiya.tar.gz 文件的数字签名。MD5 将整个文件当作一个大文本信息，通过其不可逆的字符串变换算法，产生了这个唯一的 MD5 信息摘要。通俗来说，地球上任何人都有自己独一无二的指纹，这常常成为司法机关鉴别罪犯身份最值得信赖的方法；与之类似，MD5 就可以为任何文件（不管其大小、格式、数量）产生一个同样独一无二的“数字指纹”，如果任何人对文件做了任何改动，其 MD5 值也就是对应的“数字指纹”都会发生变化。下载站点中的 MD5，它的作用就在于我们可以在下载文件后，对下载的文件用专门的软件（如 Windows MD5 Check 等）做一次 MD5 校验，以确保我们获得的文件与该站点提供的文件为同一文件。具体来说文件的 MD5 值就像是这个文件的“数字指纹”。每个文件的 MD5 值是不同的，如果任何人对文件做了任何改动，其 MD5 值也就是对应的“数字指纹”就会发生变化。比如下载服务器针对一个文件预先提供一个 MD5 值，用户下载完该文件后，重新计算下载文件的 MD5 值，通过比较这两个值是否相同，就能判断下载的文件是否出错，或者说下载的文件是否被篡改了。

[85] 在本说明书实施例的步骤 1 中，计算第一摘要值是检查下载的语音合成文件是否出错，或者说下载的语音合成文件是否被篡改，实现对语音合成文件错误的实时检测，一旦语音合成文件内容发生错误，将直观地报出错误信息，防止错误在应用程序中蔓延。此外，检测语音合成文件的检查也可采用 SHA256 摘要实现。

[86] 步骤 2，判断根据语音配置文件内预先储存的语音合成文件对应的第二摘要值与第一摘要值是否相同，若相同，则执行步骤 3，若不相同，则返回步骤 S103。

[87] 步骤 3，已注册 APP 根据语音合成文件进行语音播放。

[88] 在本说明书实施例的步骤 3 中，已注册 APP 对应的服务端可以根据内置的私钥进行加密，加密的语音合成文件播放时需要根据解密模块存储的公钥解密后进行语音播放。

[89] 在本说明书实施例的步骤 S103 中，语音基础训练模型内配置有通用的语音数据库，该通用的语音数据库内包括交易金额、时间的语音播报，即 APP 开发者在文本中输入数字时通过定制化的语音模型可直接转化为交易的金额的语音或时间的语音合成文件，而非单纯的数字朗读，例如，文本中写入 5:00 时，语音合成文件播放出的语音为时间为 5 点。

[90] 上述步骤中，已注册 APP 需要使用语音合成文件时，检测客户端是否缓存该语音合成文件，在客户端存在该语音合成文件时优先调用客户端缓存的语音合成文件，减少整个语音系统的响应时间。

[91] 进一步的，为了该语音系统可以应用在安全的环境下，对于上述实施例做出了改变，图 2 为本说明书实施例提供的一种语音合成文件的调用方法的流程示意图，该流程示意图包括:

[92] 步骤 S201，向已注册 APP 对应的服务端拉取语音配置文件。

[93] 在本说明书实施例的步骤 S201 中，已注册 APP 对应的定制化的语音模型将预先储存的文本转化为语音合成文件，已注册 APP 对应的语音配置文件包括该语音合成文件的语音列表。

[94] 步骤 S202，接收已注册 APP 对应的服务端下发的语音配置文件，下发语音配置文件包括已注册 APP 对应的服务端对下发的语音配置文件进行加密后，分配给已注册 APP 对应的第一验证信息。

[95] 在本说明书实施例的步骤 S202 中，开发者 APP 在语音 SDK 进行注册，语音 SDK 连接有解密模块，该解密模块内可以通过 TSM 下发解密的公钥，该公钥为对应已注册 APP 对应的唯一公钥，服务端配置有对应的私钥，已注册 APP 对应的服务端对下发的语音配置文件由该私钥进行加密。公钥与私钥是一个密钥对，公钥是密钥对中公开的部分，私钥则是非公开的部分。公钥和私钥组成的密钥对能保证在是唯一的。使用这个密钥对的时候，如果用其中一个密钥加密一段数据，必须用另一个密钥解密。比如用公钥加密数据就必须用私钥解密，如果用私钥加密也必须用公钥解密，否则解密将不会成功。

[96] 进一步的，在本说明书实施例的步骤 S202 中，解密模块可以为 SE 模块，SE 模块为确保系统安全的模块，通过安全芯片和芯片操作系统(COS)实现数据安全存储、加解密运算等功能。SE 模块在安全体系里主要功能包括：密钥的安全存储、数据加密运算和信息的安全存放。密钥的安全存储可建立相对完善的密钥管理体系，保证密钥不可被读取。数据加密运算包括对于可靠的安全算法的支持、敏感数据密文传输和数据传输防篡改等。信息安全存放指的是严格的文件访问权限机制和可靠的认证算法和流程。本实施例中是将公钥放置 SE 模块中。SE 模块可封装成各种形式，常见的有智能卡和嵌入式安全模块（eSE）等。本实施例中可以针对语音系统的语音 SDK 植入嵌入式安全模块（eSE），并采用满足 CCEAL5+安全等级要求的智能安全芯片，内置安全操作系统，满足终端的安全密钥存储、数据加密服务的需求。使得该语音系统可广泛应用于金融、地图导航、城市交通、医疗、零售等领域，能保护系统在使用时安全性。

[97] 步骤 S203，判断第一验证信息与客户端预先保存的第二验证信息是否匹配，若是，

则执行步骤 S204，若否，则结束流程。

[98] 在本说明书实施例的步骤 S203 中，根据已注册 APP 的标识从内置于客户端安全运行环境中预先保存的与已注册 APP 对应的第二验证信息；判断第一验证信息与第二验证信息是否匹配。其中，已注册 APP 的标识是该已注册 APP 的身份信息。

[99] 步骤 S204，验证下发的语音配置文件正确。

[100] 步骤 S205，检测客户端是否存在已注册 APP 所需要使用的语音合成文件，若存在，则执行步骤 S206，若不存在，则执行步骤 S207。

[101] 在本说明书实施例的步骤 S205 中，同上述步骤 S101，不再赘述。

[102] 步骤 S206，调用客户端的语音合成文件。

[103] 在本说明书实施例的步骤 S206 中，同上述步骤 S102，不再赘述。

[104] 步骤 S207，根据已注册 APP 对应的语音配置文件从已注册 APP 对应的服务端下载语音合成文件。

[105] 在本说明书实施例的步骤 S207 中，同上述步骤 S103，不再赘述。

[106] 进一步的，本实施例中的语音系统还存在着服务端和已注册 APP 的同步问题，为了解决该问题，可以支持服务端主动推送的方式，即客户端的语音合成文件发生变化时服务端主动向客户端进行推送。

[107] 图 3 为本说明书实施例提供的一种语音合成文件的调用装置的结构示意图，该结构示意图包括：检测单元 1、调用单元 2、下载单元 3、拉取单元 4、接收单元 5、判断单元 6、验证单元 7、训练单元 8 与计算单元 9。

[108] 检测单元 1 用于检测客户端是否存在已注册 APP 所需要使用的语音合成文件，已注册 APP 为预先注册需要使用语音合成文件的 APP。

[109] 调用单元 2 用于若检测出客户端存在语音合成文件，调用客户端的语音合成文件，以供已注册 APP 根据语音合成文件进行语音播放。

[110] 下载单元 3 用于若检测出客户端不存在语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从已注册 APP 对应的服务端下载语音合成文件，语音配置文件内置有语音合成文件的下载地址。

[111] 拉取单元 4 用于向已注册 APP 对应的服务端拉取语音配置文件。

[112] 接收单元 5 用于接收已注册 APP 对应的服务端下发的语音配置文件，下发语音配置文件包括已注册 APP 对应的服务端对下发的语音配置文件进行加密后，分配给已注册 APP 对应的第一验证信息。

[113] 判断单元 6 用于判断第一验证信息与客户端预先保存的第二验证信息是否匹配;

[114] 验证单元 7 用于在判断出第一验证信息与客户端预先保存的第二验证信息匹配时，验证下发的语音配置文件正确。

[115] 判断单元 6 具体用于:

[116] 根据已注册 APP 的标识从内置于客户端安全运行环境中预先保存的与已注册 APP 对应的第二验证信息;

[117] 判断第一验证信息与第二验证信息是否匹配。

[118] 训练单元 8 用于向已注册 APP 对应的服务端发送 APP 开发者提供的反映 APP 开发者特征的语音数据，以便已注册 APP 对应的服务端通过内置的语音基础训练模型训练出 APP 开发者定制化的语音模型，并根据预先储存的文本由 APP 开发者定制化的语音模型生成已注册 APP 对应的语音合成文件，语音基础训练模型为根据已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型。

[119] 计算单元 9 用于计算语音合成文件对应的第一摘要值;

[120] 判断单元 6 还用于判断根据语音配置文件内预先储存的语音合成文件对应的第二摘要值与第一摘要值是否相同;

[121] 判断单元 6 若判断出第二摘要值与第一摘要值相同时，则已注册 APP 根据语音合成文件进行语音播放。

[122] 已注册 APP 根据语音合成文件进行语音播放，具体包括：已注册 APP 对应的服务端根据预设规则对语音合成文件进行加密；加密的语音合成文件根据内置解密模块解密后，由已注册 APP 进行语音播放。

[123] 本说明书实施例还提供了一种计算机可读介质，其上存储有计算机可读指令，计算机可读指令可被处理器执行以下步骤:

[124] 检测客户端是否存在已注册 APP 所需要使用的语音合成文件，已注册 APP 为预先注册需要使用语音合成文件的 APP;

[125] 若检测出客户端不存在语音合成文件，根据预先存储的已注册 APP 对应的语音配

置文件从已注册 APP 对应的服务端下载语音合成文件，语音配置文件内置有语音合成文件的下载地址；

[126] 若检测出客户端存在语音合成文件，调用客户端的语音合成文件，以供已注册 APP 根据语音合成文件进行语音播放。

[127] 本说明书实施例还提供一种语音合成文件的调用设备，该设备包括用于存储计算机程序指令的存储器和用于执行程序指令的处理器，其中，当该计算机程序指令被该处理器执行时，触发该设备执行以下步骤：

[128] 检测单元，用于检测客户端是否存在已注册 APP 所需要使用的语音合成文件，已注册 APP 为预先注册需要使用语音合成文件的 APP；

[129] 下载单元，用于若检测出客户端不存在语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从已注册 APP 对应的服务端下载语音合成文件，语音配置文件内置有语音合成文件的下载地址；

[130] 调用单元，用于若检测出客户端存在语音合成文件，调用客户端的语音合成文件，以供已注册 APP 根据语音合成文件进行语音播放。

[131] 本说明书实施例提供的一种语音系统，包括终端、服务器，终端包括运行在终端中的语音 SDK、已注册 APP 以及 APP 开发者端；

[132] APP 开发者端用于向已注册 APP 对应的服务端发送 APP 开发者提供的反映 APP 开发者特征的语音数据；

[133] 服务端用于通过内置的语音基础训练模型训练出 APP 开发者定制化的语音模型，并将预先储存的文本输入 APP 开发者定制化的语音模型以生成已注册 APP 需要的语音合成文件，语音基础训练模型为根据已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型；

[134] 语音 SDK 用于向已注册 APP 对应的服务端拉取语音配置文件；接收已注册 APP 对应的服务端下发的语音配置文件，下发语音配置文件包括已注册 APP 对应的服务端对下发的语音配置文件进行加密后，分配给已注册 APP 对应的第一验证信息；判断第一验证信息与客户端预先保存的第二验证信息是否匹配；在判断出第一验证信息与客户端预先保存的第二验证信息匹配时，则验证下发的语音配置文件正确；检测客户端是否存在已注册 APP 所需要使用的语音合成文件，已注册 APP 为预先注册需要使用语音合成文件的 APP；若检测出客户端不存在语音合成文件，根据已注册 APP 对应的语音配置

文件从已注册 APP 对应的服务端下载语音合成文件，语音配置文件内置有语音合成文件的下载地址;若检测出客户端存在语音合成文件，调用客户端的语音合成文件，以供已注册 APP 根据语音合成文件进行语音播放。

[135] 本领域内的技术人员应明白，本发明的实施例可提供为方法、系统、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

[136] 本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和／或方框图来描述的。应理解可由计算机程序指令实现流程图和／或方框图中的每一流程和／或方框、以及流程图和／或方框图中的流程和／或方框的结合。可提供这些计算机程序指令到计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的装置。

[137] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能。

[138] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的步骤。

[139] 在一个典型的配置中，计算设备包括一个或多个处理器（CPU）、输入/输出接口、网络接口和内存。

[140] 内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器（RAM）和/或非易失性内存等形式，如只读存储器（ROM）或闪存（flashRAM）。内存是计算机可读介质的示例。

[141] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法

或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存（PRAM）、静态随机存取存储器（SRAM）、动态随机存取存储器（DRAM）、其他类型的随机存取存储器（RAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、快闪记忆体或其他内存技术、只读光盘只读存储器（CD-ROM）、数字多功能光盘（DVD）或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体（transitory media），如调制的数据信号和载波。

[142] 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他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，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过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括要素的过程、方法、商品或者设备中还存在另外的相同要素。

[143] 以上仅为本说明书的实施例而已，并不用于限制本说明书。对于本领域技术人员来说，本说明书可以有各种更改和变化。凡在本说明书的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本说明书的权利要求范围之内。

权利要求书

1、一种语音合成文件的调用方法，其特征在于，所述方法包括:

检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP;

若检测出客户端不存在所述语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地址;

若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

2、根据权利要求 1 所述的语音合成文件的调用方法，其特征在于，所述检测客户端是否存在已注册 APP 所需要使用的语音合成文件之前，所述方法还包括:

向所述已注册 APP 对应的服务端拉取所述语音配置文件;

接收所述已注册 APP 对应的服务端下发的语音配置文件，下发的所述语音配置文件包括所述已注册 APP 对应的服务端对下发的所述语音配置文件进行加密后，分配给所述已注册 APP 对应的第一验证信息;

判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配;

在判断出所述第一验证信息与客户端预先保存的第二验证信息匹配时，则验证下发的所述语音配置文件正确。

3、根据权利要求 2 所述的语音合成文件的调用方法，其特征在于，判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配，具体包括:

根据所述已注册 APP 的标识从内置于客户端安全运行环境中预先保存的与所述已注册 APP 对应的第二验证信息;

判断所述第一验证信息与第二验证信息是否匹配。

4、根据权利要求 2 所述的语音合成文件的调用方法，其特征在于，向所述已注册 APP 对应的服务端拉取所述语音配置文件之前，所述方法还包括:

向所述已注册 APP 对应的服务端发送所述 APP 开发者提供的反映所述 APP 开发者特征的语音数据，以便所述已注册 APP 对应的服务端通过内置的语音基础训练模型训练出所述 APP 开发者定制化的语音模型，并将预先储存的文本输入所述 APP 开发者定制化的语音模型以生成已注册 APP 需要的语音合成文件，所述语音基础训练模型为根据所述已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型。

5、根据权利要求 1 所述的语音合成文件的调用方法，其特征在于，所述已注册 APP 根据所述语音合成文件进行语音播放之前，所述方法还包括:

计算所述语音合成文件对应的第一摘要值;

判断根据所述语音配置文件内预先储存的所述语音合成文件对应的第二摘要值与所述第一摘要值是否相同;

若判断出所述第二摘要值与所述第一摘要值相同时，则所述已注册 APP 根据所述语音合成文件进行语音播放。

6、根据权利要求 1 所述的语音合成文件的调用方法，其特征在于，所述已注册 APP 根据所述语音合成文件进行语音播放，具体包括：所述已注册 APP 对应的服务端根据预设规则对所述语音合成文件进行加密；所述加密的语音合成文件根据内置解密模块解密后，由所述已注册 APP 进行语音播放。

7、一种语音合成文件的调用装置，其特征在于，所述装置包括:

检测单元，用于检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP;

下载单元，用于若检测出客户端不存在所述语音合成文件，根据预先存储的已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地址;

调用单元，用于若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

8、根据权利要求 7 所述的语音合成文件的调用装置，其特征在于，所述装置还包括:

拉取单元，用于向所述已注册 APP 对应的服务端拉取所述语音配置文件;

接收单元，用于接收所述已注册 APP 对应的服务端下发的语音配置文件，下发的所述语音配置文件包括所述已注册 APP 对应的服务端对下发的所述语音配置文件进行加密后，分配给所述已注册 APP 对应的第一验证信息;

判断单元，用于判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配;

验证单元，用于在判断出所述第一验证信息与客户端预先保存的第二验证信息匹配时，验证下发的所述语音配置文件正确。

9、根据权利要求 8 所述的语音合成文件的调用装置，其特征在于，所述判断单元具体用于:

根据所述已注册 APP 的标识从内置于客户端安全运行环境中预先保存的与所述已

注册 APP 对应的第二验证信息;

判断所述第一验证信息与第二验证信息是否匹配。

10、根据权利要求 8 所述的语音合成文件的调用装置，其特征在于，所述装置还包括:

训练单元，用于向所述已注册 APP 对应的服务端发送所述 APP 开发者提供的反映所述 APP 开发者特征的语音数据，以便所述已注册 APP 对应的服务端通过内置的语音基础训练模型训练出所述 APP 开发者定制化的语音模型，并根据预先储存的文本由所述 APP 开发者定制化的语音模型生成已注册 APP 对应的语音合成文件，所述语音基础训练模型为根据所述已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型。

11、根据权利要求 7 所述的语音合成文件的调用装置，其特征在于，所述装置还包括:

计算单元，用于计算所述语音合成文件对应的第一摘要值;

判断单元还用于判断根据所述语音配置文件内预先储存的所述语音合成文件对应的第二摘要值与所述第一摘要值是否相同;

所述判断单元若判断出所述第二摘要值与所述第一摘要值相同时，则所述已注册 APP 根据所述语音合成文件进行语音播放。

12、根据权利要求 7 所述的语音合成文件的调用装置，其特征在于，所述已注册 APP 根据所述语音合成文件进行语音播放，具体包括:

所述已注册 APP 对应的服务端根据预设规则对所述语音合成文件进行加密；所述加密的语音合成文件根据内置解密模块解密后，由所述已注册 APP 进行语音播放。

13、一种语音系统，其特征在于，包括终端、服务器，终端包括运行在终端中的语音 SDK、已注册 APP 以及 APP 开发者端;

所述 APP 开发者端用于向所述已注册 APP 对应的服务端发送所述 APP 开发者提供的反映所述 APP 开发者特征的语音数据;

所述服务端用于通过内置的语音基础训练模型训练出所述 APP 开发者定制化的语音模型，并将预先储存的文本输入所述 APP 开发者定制化的语音模型以生成已注册 APP 需要的语音合成文件，所述语音基础训练模型为根据所述已注册 APP 播放语音的需要利用预先提供的若干语音样本训练得到的、可供已注册 APP 共用的模型;

所述语音 SDK 用于向所述已注册 APP 对应的服务端拉取所述语音配置文件；接收所述已注册 APP 对应的服务端下发的语音配置文件，所述下发语音配置文件包括所述

已注册 APP 对应的服务端对所述下发的语音配置文件进行加密后，分配给所述已注册 APP 对应的第一验证信息；判断所述第一验证信息与客户端预先保存的第二验证信息是否匹配；在判断出所述第一验证信息与客户端预先保存的第二验证信息匹配时，则验证所述下发的语音配置文件正确;检测客户端是否存在已注册 APP 所需要使用的语音合成文件，所述已注册 APP 为预先注册需要使用语音合成文件的 APP；若检测出客户端不存在所述语音合成文件，根据已注册 APP 对应的语音配置文件从所述已注册 APP 对应的服务端下载所述语音合成文件，所述语音配置文件内置有所述语音合成文件的下载地址;若检测出客户端存在所述语音合成文件，调用客户端的所述语音合成文件，以供所述已注册 APP 根据所述语音合成文件进行语音播放。

14、一种计算机可读介质，其上存储有计算机可读指令，所述计算机可读指令可被处理器执行以实现权利要求 1 至 6 中任一项所述的方法。

15、一种语音合成文件的调用设备，该设备包括用于存储计算机程序指令的存储器和用于执行程序指令的处理器，其中，当该计算机程序指令被该处理器执行时，触发该设备执行权利要求 1 至 6 中任一项所述的方法。

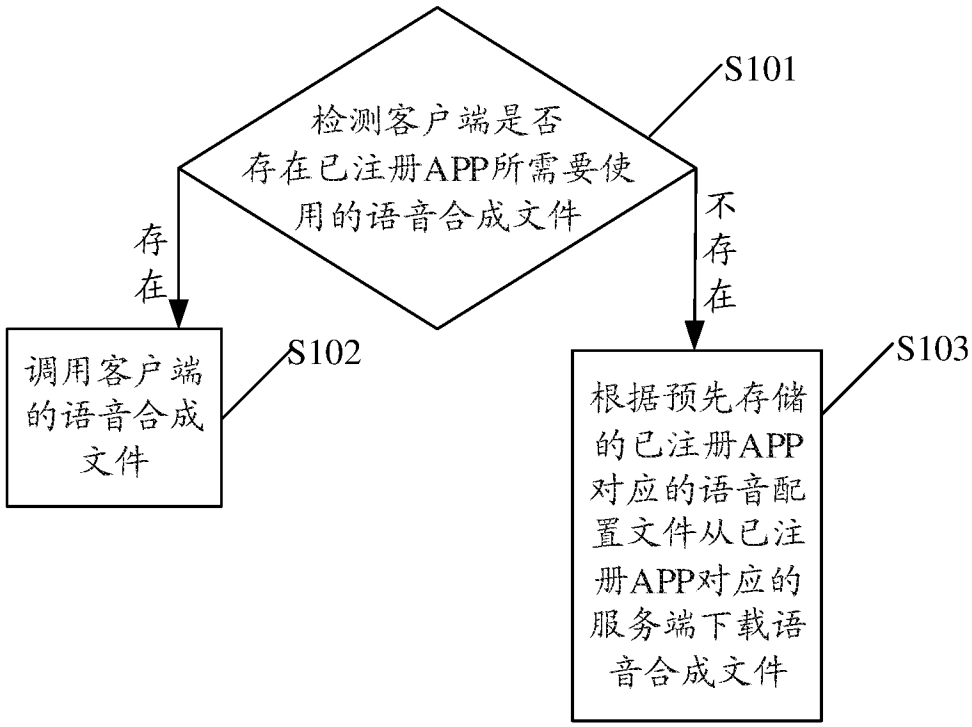

检测客户端是否存在已注册APP所需要使用的语音合成文件
S101
存在
不存在
调用客户端的语音合成文件
S102
根据预先存储的已注册APP对应的语音配置文件从已注册APP对应的服务端下载语音合成文件
S103

图 1

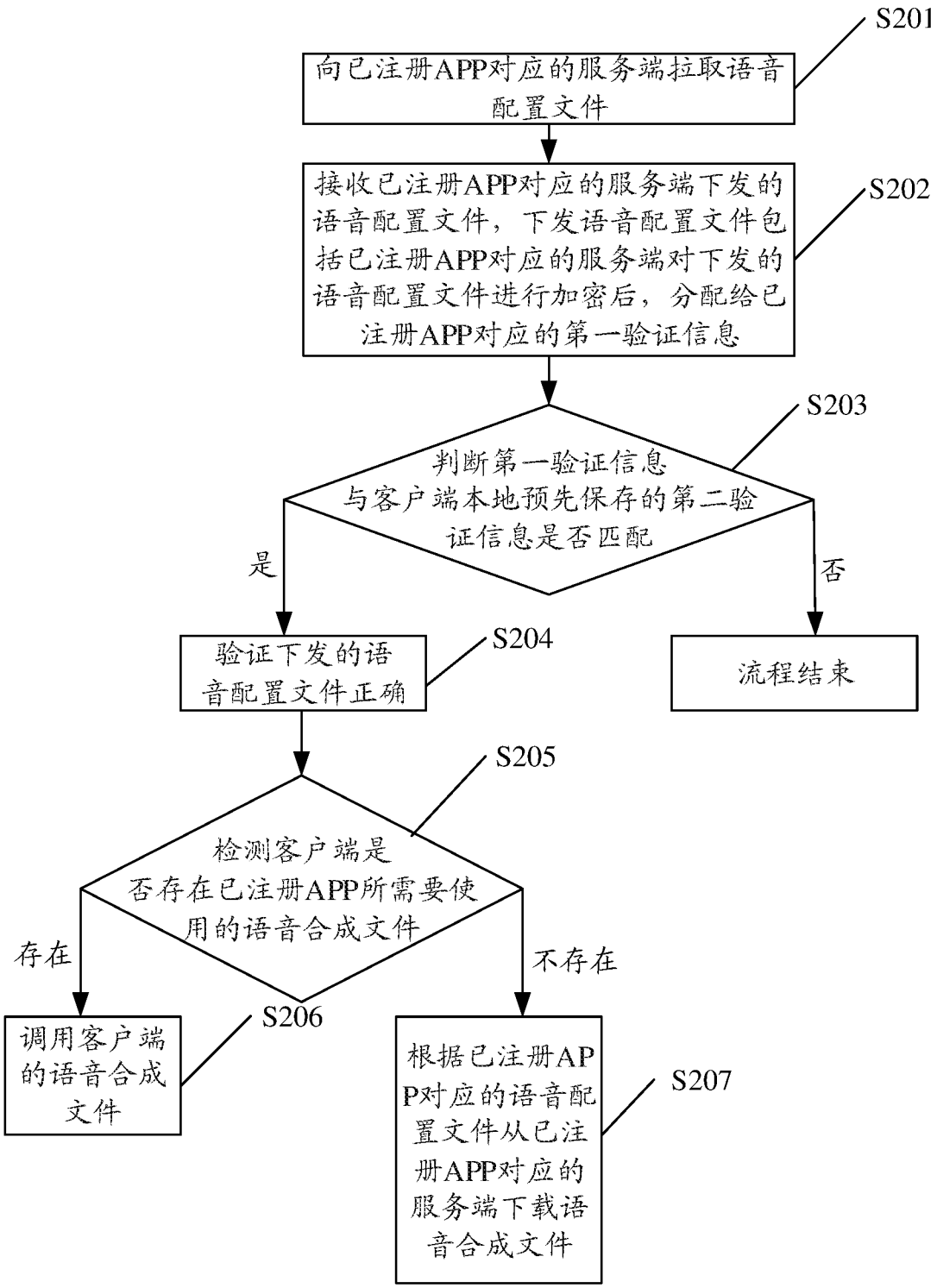

S201
向已注册APP对应的服务端拉取语音配置文件
接收已注册APP对应的服务端下发的语音配置文件，下发语音配置文件包括已注册APP对应的服务端对下发的语音配置文件进行加密后，分配给已注册APP对应的第一验证信息
S202
S203
判断第一验证信息与客户端本地预先保存的第二验证信息是否匹配
是
否
验证下发的语音配置文件正确
S204
流程结束
S205
检测客户端是否存在已注册APP所需要使用的语音合成文件
存在
不存在
调用客户端的语音合成文件
S206
根据已注册APP对应的语音配置文件从已注册APP对应的服务端下载语音合成文件
S207

图 2

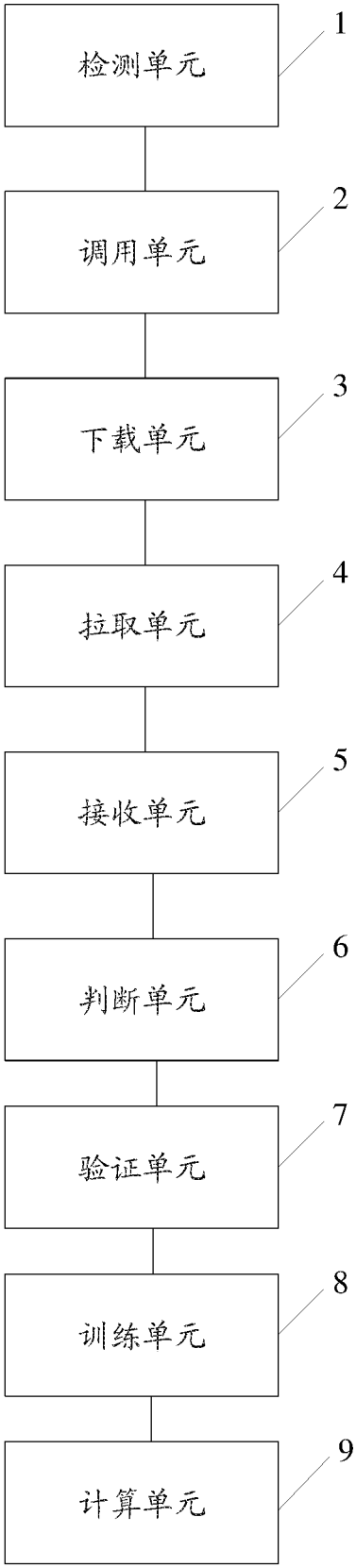


图 3

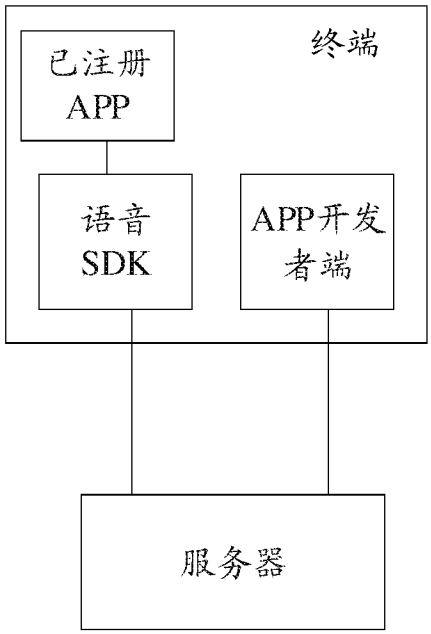


图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/122545

A. CLASSIFICATION OF SUBJECT MATTER

G10L 13/02(2013.01)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)

G10L13/-; H04L29/-

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)

WPI, EPODOC, CNPAT, CNKI: 阿里巴巴, 韩喆, 王磊, 傅春霖, 语音, 合成, 文本, 应用, 服务器, 服务端, 下载, 播放; speech, voice, syntheti+, TTS, text, APP, application, server, download+, play+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110021291 A (ALIBABA GROUP HOLDING LIMITED) 16 July 2019 (2019-07-16) claims 1-15, description, paragraphs [0001]-[0143], and figures 1-4	1-15
A	CN 105354096 A (THE 28TH RESEARCH INSTITUTE OF CHINA ELECTRONICS TECHNOLOGY GROUP CORP.) 24 February 2016 (2016-02-24) description, paragraphs [0029]-[0095], and figures 1-3	1-15
A	CN 105161091 A (BEIJING WISELINK SOFTWARE CO., LTD.) 16 December 2015 (2015-12-16) entire document	1-15
A	CN 104992703 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 21 October 2015 (2015-10-21) entire document	1-15
A	CN 101098507 A (ZTE CORPORATION) 02 January 2008 (2008-01-02) entire document	1-15
A	US 2003187658 A1 (SELIN, J et al.) 02 October 2003 (2003-10-02) entire document	1-15
A	US 2018075838 A1 (MASON, P. W.) 15 March 2018 (2018-03-15) entire document	1-15

☑ 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	Date of mailing of the international search report
23 February 2020	**03 March 2020**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/122545

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2005002160 A1 (NORTEL NETWORKS LTD.) 06 January 2005 (2005-01-06) entire document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/122545

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110021291	A	16 July 2019	None			
CN	105354096	A	24 February 2016	CN	105354096	B	10 July 2018
CN	105161091	A	16 December 2015	CN	105161091	B	26 March 2019
CN	104992703	A	21 October 2015	CN	104992703	B	03 October 2017
				WO	2017016135	A1	02 February 2017
CN	101098507	A	02 January 2008	CN	100544463	C	23 September 2009
US	2003187658	A1	02 October 2003	None			
US	2018075838	A1	15 March 2018	US	9830903	B2	28 November 2017
				US	2017133005	A1	11 May 2017
WO	2005002160	A1	06 January 2005	EP	1642431	A1	05 April 2006
				US	2004267531	A1	30 December 2004

国际检索报告

国际申请号
PCT/CN2019/122545

A. 主题的分类

G10L 13/02(2013.01)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)

G10L13/-; H04L29/-

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））

WPI, EPODOC, CNPAT, CNKI: 阿里巴巴,韩喆,王磊,傅春霖,语音,合成,文本,应用,服务器,服务端,下载,播放; speech, voice, syntheti+, TTS,text, APP, application, server, download+, play+

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 110021291 A (阿里巴巴集团控股有限公司) 2019年 7月 16日 (2019-07-16) 权利要求1-15、说明书第[0001]-[0143]段、图1-4	1-15
A	CN 105354096 A (中国电子科技集团公司第二十八研究所) 2016年 2月 24日 (2016-02-24) 说明书第[0029]-[0095]段、图1-3	1-15
A	CN 105161091 A (北京开元智信通软件有限公司) 2015年 12月 16日 (2015-12-16) 全文	1-15
A	CN 104992703 A (百度在线网络技术北京有限公司) 2015年 10月 21日 (2015-10-21) 全文	1-15
A	CN 101098507 A (中兴通讯股份有限公司) 2008年 1月 2日 (2008-01-02) 全文	1-15
A	US 2003187658 A1 (SELIN, Jari 等) 2003年 10月 2日 (2003-10-02) 全文	1-15
A	US 2018075838 A1 (MASON, Paul Wendell) 2018年 3月 15日 (2018-03-15) 全文	1-15

☑ 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	国际检索报告邮寄日期
2020年 2月 23日	2020年 3月 3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 陈红红 电话号码 86-(10)-53962642

国际检索报告

国际申请号
PCT/CN2019/122545

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
A	WO 2005002160 A1 (NORTEL NETWORKS LTD.) 2005年 1月 6日 (2005-01-06) 全文	1-15

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/122545

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	110021291	A	2019年 7月 16日		无		
CN	105354096	A	2016年 2月 24日	CN	105354096	B	2018年 7月 10日
CN	105161091	A	2015年 12月 16日	CN	105161091	B	2019年 3月 26日
CN	104992703	A	2015年 10月 21日	CN	104992703	B	2017年 10月 3日
				WO	2017016135	A1	2017年 2月 2日
CN	101098507	A	2008年 1月 2日	CN	100544463	C	2009年 9月 23日
US	2003187658	A1	2003年 10月 2日		无		
US	2018075838	A1	2018年 3月 15日	US	9830903	B2	2017年 11月 28日
				US	2017133005	A1	2017年 5月 11日
WO	2005002160	A1	2005年 1月 6日	EP	1642431	A1	2006年 4月 5日
				US	2004267531	A1	2004年 12月 30日