

文件编号: ISTMaker-SWC2020-20200175

受控状态: 受控 非受控

保密级别: 公司级 部门级 项目级 普通级

采纳标准: CMMI DEV V1.2



AI—智慧方言

AI—Intelligent Dialect

项目开发文档

Version 4.0

2020.5.25

Written by ISTMaker



All Rights Reserved

目录

1	项目概述.....	1
1.1	项目背景.....	1
1.2	项目定位.....	1
1.2.1	应用场景.....	1
1.2.2	目标人群.....	3
1.3	项目方案.....	3
1.4	项目目标.....	7
1.5	项目价值.....	8
2	开发计划.....	8
2.1	最终呈现形式.....	9
2.2	主要功能描述.....	10
2.3	运行环境.....	10
2.4	验收标准.....	11
2.5	关键问题.....	11
2.6	进度安排.....	11
2.7	开发预算.....	12
3	可行性分析.....	12
3.1	技术可行性分析.....	12
3.2	资源可行性分析.....	13
3.3	市场可行性分析.....	13
4	需求分析.....	14
4.1	数据需求.....	14
4.1.1	静态数据.....	14
4.1.2	动态数据.....	14
4.1.3	数据采集.....	14
4.2	功能需求.....	14
4.2.1	功能模块.....	14
4.3	性能需求.....	19
4.3.1	时间特性.....	19
4.3.2	适应性.....	19
4.4	界面需求.....	19
4.5	接口需求.....	20
4.5.1	硬件接口.....	20
4.5.2	软件接口.....	20
4.6	其他需求.....	20
4.6.1	优秀的扩展性和可维护性.....	20
4.6.2	高可靠性.....	21
4.6.3	灵活的配置和高效的性能.....	21
4.6.4	安全性.....	21
4.6.5	可重用性.....	21

5	概要设计	21
5.1	处理流程.....	21
5.2	总体结构设计.....	22
5.3	功能设计.....	25
5.4	用户界面设计.....	27
5.5	数据结构设计.....	30
5.6	接口设计.....	30
5.6.1	外部接口.....	30
5.6.2	内部接口.....	31
5.7	错误/异常处理设计.....	31
5.7.1	错误/异常输出信息.....	31
5.7.2	错误/异常处理对策.....	31
5.8	系统配置策略.....	32
5.9	系统工作流程:.....	32
5.10	系统部署方案.....	33
5.10.1	系统表示层.....	33
5.10.2	系统应用层.....	33
5.11	其他相关技术与方案.....	34
5.11.1	5. 11. 1 均衡负载技术.....	34
5.11.2	云计算.....	34
5.11.3	三重数据审核算法.....	34
6	数据库设计	35
6.1.1	系统数据层.....	35
6.1.2	系统数据库 Mysql.....	35
6.1.3	开源的数据库连接池.....	35
6.1.4	数据库内容.....	35
7	详细设计	36
7.1	终端功能模块.....	36
7.1.1	功能描述.....	36
7.1.2	性能描述.....	37
7.1.3	输入.....	38
7.1.4	输出.....	38
7.1.5	程序逻辑.....	38
7.1.6	限制条件.....	39

记录更改历史

序号	更改原因	版本	作者	更改日期	备注
1	撰写文档	1.0	A	2019.11.1	无
2	修改功能模块	2.0	A	2019.11.5	无
3	修改时间延迟	3.0	A	2019.11.11	无
4	增加文档完整性	4.0	B	2020.4.15	无
5	完善开发过程	5.0	A	2020.5.20	无

1 项目概述

本项目《AI—智慧方言》，采用模块化设计，即快应用模块、云服务器深度学习模块、智能家居应用模块搭建方言采集与应用平台。三大模块相互独立又紧密联系，在每个模块各自完成其任务之时，又通过 WIFI，云服务器等手段进行数据交互，构成一个紧密的整体。不仅如此，我们考察多种家具属性，抽象出具有更高代表性的机器人，极大提高了产品的普适性，紧紧贴合“万物互融”的指导思想。

云服务器，机器人模块，快应用模块三者关系如下：

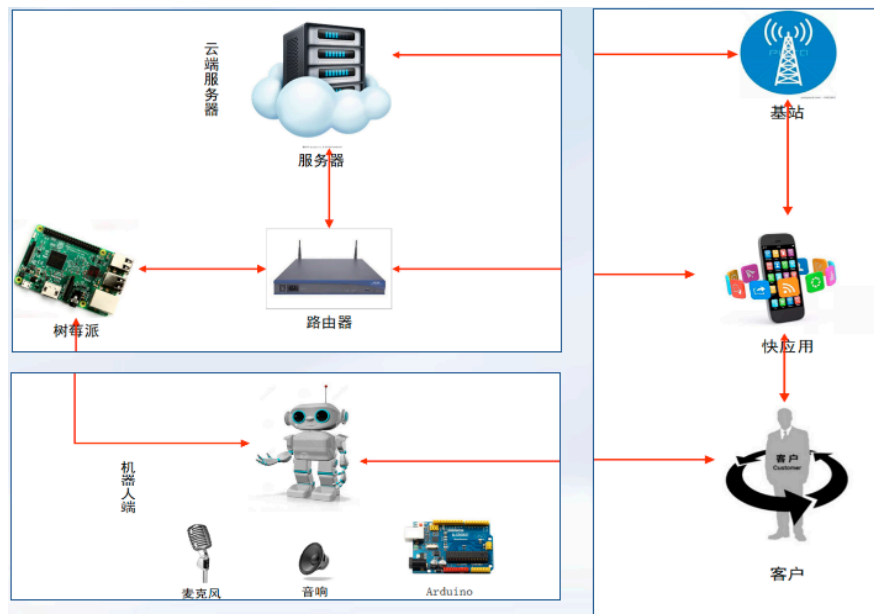


图 1 系统总体架构

1.1 项目背景

1. 方言作为一种文化正逐渐消逝，其保护已迫在眉睫
2. 智能语音服务难以惠及方言用户群体
3. 缺乏高效方式进行方言数据的收集方式
4. 方言产品的针对性较低

1.2 项目定位

1.2.1 应用场景

1. 方言保护领域

中国是一个多民族多语言国家，但近几年语言资源和文化资源流失严重。在我国大力推动普通话普及的时代背景下，方言的处境愈发危险。但政府也逐渐意识到，方言其本身就是一种非物质的文化遗产，是我们正亟需保护的传统文化的一种重要形式。正如中央民族大学教授滕星所说，“语言是文化的载体，一种语言消失了，它所承载的文化也就消失了。”可见，方言的保护已迫在眉睫。我组产品可以帮助人们重新看见方言魅力。在某些景点部署服务，不仅可以使游客感受更为浓烈的乡土气息，还可以使游客了解方言，为方言的保护奉献自己的一份力。

2. 方言服务领域

中国拥有七大方言区，大小门类方言共数百种。方言用户群体的数目十分可观，占全国人口的百分之二十以上。但现阶段，在这个语音功能服务日益普及的当下，方言群体却难以享受到语音服务为人们生活所带来的诸多便利。使用我组产品，可以帮助方言使用者找到家的感觉，帮助方言学习者快速了解多种方言，提升中华文化认同感。

3. 方言交流

由于方言使用者相对较少，故目前的语音助手基本都是面向普通话使用者，这样一来，某些方言使用者因为对普通话不了解，故无法享受科技带来的乐趣。我组产品不仅可以使现代科技丰富方言使用者的科技生活，还可以帮助非方言使用者“听懂”没有听说过的方言。

4. 方言数据收集

方言数据来源是现阶段相关工作的巨大阻碍，受制于各方言使用人群所处的地域，使用人数，个体差异等客观条件，很难采用现阶段普适的各类数据收集方式。因此也就难以大面积的获取相关数据。现阶段主要采用招募符合条件的方言掌握者，进行语音的录制，从而获得数据。但该方法需要消耗大量的人力物力成本，获取的数据量却极为有限。我组产品使用自收集方式，在遵守隐私协议的前提下，使用为用户服务时得到的数据，从中挑出高品质数据，对神经网络进行优化。

1.2.2 目标人群

1. 想要得到科技乐趣的方言使用者；
2. 希望学习方言的方言学习者。
3. 希望学习 NLP 的研究人员。

1.3 项目方案

本项目《AI—智慧方言》，采用模块化设计，即快应用模块、云服务器深度学习模块、智能家居应用模块搭建方言采集与应用平台。三大模块相互独立又紧密联系，在每个模块各自完成其任务之时，又通过 WIFI，云服务等手段进行数据交互，构成一个紧密的整体。不仅如此，我们考察多种家具属性，抽象出具有更高代表性的机器人，极大提高了产品的普适性，紧紧贴合“万物互融”的指导思想。

在产品使用过程中，我们打造方言版的语音助手对方言用户进行服务。而这样的快应用除了是我们服务用户的手段，同时也是我们数据收集的重要来源。原有的依靠人工进行数据收集的方式，数据包形成后更新周期长，成本高，用户发现语音识别不准后，也无法及时反馈，最后导致了用户对于方言产品的失望。而我们团队希望充分尊重并最大化用户的作用。在使用我们产品同时，我们将用户的语音数据进行记录，在云端进行数据的筛选，一旦用户发现识别不准，可以立即校正，我们在后台审核处理后，并可以实现数据的实时更新，从而形成闭环。

快应用界面如下：



图 2 快应用界面

另一方面，我们创新性的提出深度挖掘自媒体所蕴含的巨大价值。我们都知道随着移动互联网局面的形成，抖音，快手这样的自媒体平台火遍大街小巷。其中方言类的娱乐内容占据了其中很大一部分的比重。那么我们为什么不把视频中的方言信息进行利用呢？一方面，短视频中的内容通常是大家喜闻乐见的，通过方言把网络上流行的内容以新颖的方式表达出来，这样可以确保我收集到的信息具有时新性，紧跟时代的发展；另一方面，视频中的方言就是最贴近于百姓生活的语言，因为百姓本身就是创作者，我们通过传统的人工数据收集方式时，总是苦于数据和实际中使用的方言存在差异，如果采用自媒体提取，就使得我们的数据来源自动的下沉到了民间。



图3 自媒体能力巨大



图4 方言数据广泛

为了使方言走进大众的生活，我们希望将终端技术用于各种家居设备，来提高用户的参与度，让科技方便人们的生活。让那些仅仅只能服务普通话用户的产品，也拥有方言版本。但是，每种家居都有其固有属性，在展示时我们无法展示所有家居的使用方式，于是，在考察了多种家居的属性后，我

们自主设计了机器人，该机器人可以认为是各种家居的抽象，同时，机器人技术也可以很方便地迁移到各种家居中，这拓展了产品的使用面。在使用时，智能家具可以根据用户说出的方言，用对应的方言回复用户并做出方言内容指示的动作。当用户不在智能家具附近时，用户也可以通过快应用来方便地通过语音控制智能家具。团队自主设计机器人如图 6 所示：

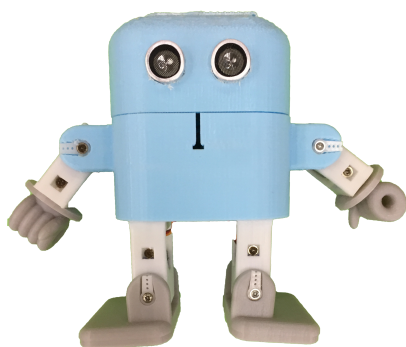


图 5.自主设计通过 3D 打印制作的机器人

除此之外，为了使这些方言真正用到实处，我们通过搭建快应用和数据管理系统，在遵循隐私协议的前提下，收集用于服务客户而产生的海量高品质数据集。通过使用 Keras 深度学习框架，使用 WaveNet 神经网络架构，利用所搜集到的海量方言数据集训练出一个准确率较高的高效网络。利用该模型，即可对包含长沙话、四川话、合肥话在内的八种方言进行分类。我们选择将网络部署到云端和本地，达成功耗和方便性的统一：用户在多种网络环境下均可享受服务——在网络环境较好时，可以通过云服务器进行服务；当网络环境较差时，可以通过同一局域网下的本地服务器进行服务。

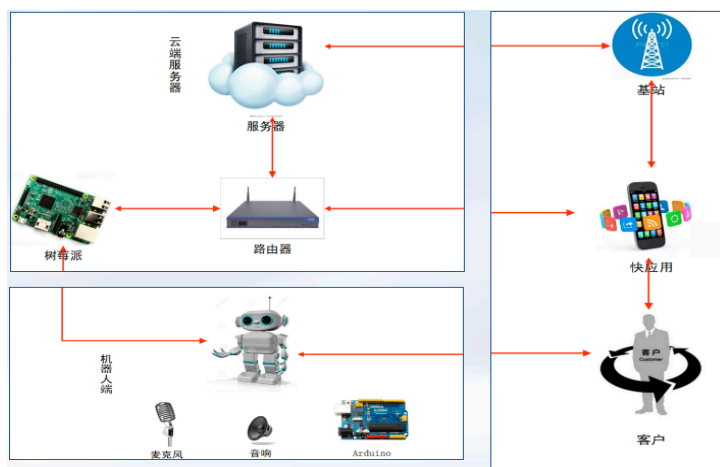


图 6.系统总体架构

1. 项目工作流程展示

- 在用户使用快应用时，有三种功能：

(1) 方言版语音助手

用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，得到方言内容后，调用本地语音助手进行服务。（该功能因为快应用没有开放语音助手接口，将在开放后进行完善）。

(2) 方言翻译助手

用户点击快应用中的方言翻译功能，通过按住按钮让持方言人说出方言，快应用根据网络环境智能选择服务器进行服务，服务器端经过谱减法降噪后对数据进行判断，得到方言内容和种类后，将方言中种类和内容显示在屏幕上告知客户。

(3) 反馈纠正

如果用户在使用的过程中，发现自己所说的方言没有被正确的理解，或者某一段语句无法识别出来，可以点击方言的反馈部分，对相关的语句进行录制，我们将以最快的速度在后台进行处理，这样就使得数据库可以快速的进行更新，同时用户的体验也可以得到很大的提升。

(4) 家居助手

用户通过点击快应用中的家居助手功能，选择希望控制的家电，通过按住按钮说出控制方言，快应用根据网络环境智能选择服务器进行服务，服务器经过谱减法降噪后对数据进行判断，随后对家中电器进行控制。使用该功能时，无论用户是在家中或者户外，都可以在连接互联网的情况下对家居进行控制。（其中家电将被抽象为机器人）

- 在用户想要直接控制家具时：

(1) 面对面家居控制

用户通过触摸开关对家具（机器人）说出方言后，家具（机器人）通过本地服务器获取方言结果，随后进行对应的动作，完成对客户的服务。

1.4 项目目标

我们希望通过该项目可以使方言用户得到现代科技的乐趣，惠及万千方言使用者，具体实现方式如下：

1. 方言版语音助手

用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，在服务器端进行谱减法降噪后进行方言内容的识别和种类的分类，在得到方言内容后，调用本地语音助手进行服务。（该功能因为快应用没有开放语音助手接口，将在开放后进行完善）。

2. 方言翻译助手

用户点击快应用中的方言翻译功能，通过按住按钮让持方言人说出方言，快应用根据网络环境智能选择服务器进行服务，在服务器端进行谱减法降噪后进行方言内容的识别和种类的分类，得到方言内容和种类后，将方言中种类和内容显示在屏幕上告知客户。

3. 方言家居助手

用户通过点击快应用中的家居助手功能，选择希望控制的家电，通过按住按钮说出控制方言，快应用根据网络环境智能选择服务器进行服务，对家中电器进行控制。使用该功能时，无论用户是在家中或者户外，都可以在连接互联网的情况下对家居进行控制。（其中家电将被抽象为机器人）

4. 面对面家居控制

用户通过触摸开关对家居说出方言后，家具通过本地服务器获取方言结果，随后进行对应的动作，完成对客户的服务。

5. 自动网络更新

在每周日晚上 2 点时各本地服务器上传方言数据，在云服务器中进行神经网络的训练，在训练完成后，自动将神经网络参数部署到各本地服务器，完

成神经网络的更新。

1.5 项目价值

1. 助力方言保护

随着时代的发展，越来越多的人背井离乡，走上外出打工的路，但是，在经济发展的同时，以前经常出现我们生活中的方言却不见了踪影。中国教育部语言文字信息管理司司长李宇明日前指出，中国是一个多民族多语言国家，但近几年语言资源和文化资源流失严重。有分析指出，政府在过去大力推广普及普通话，这在一定程度上加速了方言的流失，但令人尴尬的是，那些身处农村地区的民众在逐渐失去自己方言的同时，并没有学会说一口流利的普通话。同时，北京中央民族大学教授滕星认为，语言是文化的载体，一种语言消失了，它所承载的文化也就消失了。可见，方言的保护迫在眉睫。我组项目通过落到实处的技术来进行方言的保护，让人工智能科技走向大众，走到方言保护的第一线。

2. 服务方言使用者

在中国，还存在很多方言使用者。因为当前科技公司开发的语音助手都更加关注于普通话用户的使用方式，而对方言用户的需求较为漠视。我组作品将现代科技带到方言使用者面前，让他们使用方言也可以得到科技的乐趣。其中不仅包括，使用方言与手机对话（待快应用开放本地语音助手接口后实现），使用方言控制家中的电器。不仅如此，那些想要保护方言的志愿者们也可以通过我组产品学习方言，在学习初期也“听懂”方言，让志愿保护方言的路上少了一些障碍，让中国传统文化的的保护走的更快。

2 开发计划

表 2.1 进度安排表

项目重要里程碑	预计完成时间
项目市场调研，需求分析，可行性分析，快应用界面设计	2019-11-10
获取方言音频数据	2020-1-15

搭建训练模型，采集数据构建数据集	2020-1-30
机器人终端设计，测试机器人相关性能	2020-2-15
将各部分连接融合，测试整体性能	2020-3-25
完成项目所有材料	2020-5-25

2.1 最终呈现形式

用户通过快应用，机器人与系统继续交互，具体展现形式如下：

1. 方言语音助手

用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，在服务器端进行谱减法降噪后进行方言内容的识别和种类的分类，得到方言内容后，调用本地语音助手进行服务。（该功能因为快应用没有开放语音助手接口，将在开放后进行完善）。

2. 方言翻译助手

用户点击快应用中的方言翻译功能，通过按住按钮让持方言人说出方言，快应用根据网络环境智能选择服务器进行服务，在服务器端进行谱减法降噪后进行方言内容的识别和种类的分类，得到方言内容和种类后，将方言中种类和内容显示在屏幕上告知客户。

3. 智能家居助手

用户通过点击快应用中的家居助手功能，选择希望控制的家电，通过按住按钮说出控制方言，快应用根据网络环境智能选择服务器进行服务，对家中电器进行控制。使用该功能时，无论用户是在家中或者户外，都可以在连接互联网的情况下对家居进行控制。（其中家电将被抽象为机器人）

4. 面对面方言服务

用户通过触摸开关对家居说出方言后，家具通过本地服务器获取方言结果，随后进行对应的动作，完成对客户的服务。

2.2 主要功能描述

1. 方言语音助手

用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，在服务器端进行谱减法降噪后进行方言内容的识别和种类的分类，在得到方言内容后，调用本地语音助手进行服务。（该功能因为快应用没有开放语音助手接口，将在开放后进行完善）。

2. 方言翻译助手

用户点击快应用中的方言翻译功能，通过按住按钮让持方言人说出方言，快应用根据网络环境智能选择服务器进行服务，在服务器端进行谱减法降噪后进行方言内容的识别和种类的分类，得到方言内容和种类后，将方言中种类和内容显示在屏幕上告知客户。

3. 智慧家居助手

用户通过点击快应用中的家居助手功能，选择希望控制的家电，通过按住按钮说出控制方言，快应用根据网络环境智能选择服务器进行服务，对家中电器进行控制。使用该功能时，无论用户是在家中或者户外，都可以在连接互联网的情况下对家居进行控制。（其中家电将被抽象为机器人）

4. 面对面家居控制

用户通过触摸开关对家居说出方言后，家具通过本地服务器获取方言结果，随后进行对应的动作，完成对客户的服务。

5. 网络参数自动更新

在每周日晚上 2 点时各本地服务器上传方言数据，在云服务器中进行神经网络的训练，在训练完成后，自动将神经网络参数部署到各本地服务器，完成神经网络的更新。

2.3 运行环境

1. 快应用：快应用 1060 及以上版本

2. 机器人端：搭载 Ubuntu16.04 的树莓派 4B 型号，python 版本为 python3.6.8。Arduino 为 Arduino UNO。
3. 云服务器：搭载 Ubuntu16.04 的 1 核 2GB 40GB 系统盘的轻量应用服务器，python 版本为 3.6.8，Tensorflow 版本为 1.3.0

2.4 验收标准

1. 快应用端可以正常进行方言数据的录入和方言内容和分类的取回。
2. 机器人端可以正常被正常激活，同时，完成用户方言数据的输入，并将数据完整传输到云端进行数据的进一步分析。在完成数据的分析得到结果后，可以正确做出相应动作，回复对应方言的回答。
3. 云服务器可以正确完成数据的分析，并将结果返回给机器人端。

2.5 关键问题

1. 云服务器在进行转录任务时，可能会出现转录效果稍差的情况。
解决方案：使用交叉训练策略，同时优化网络架构，提高模型准确度；同时使用高效的降噪措施，将损失降到最低。
2. 在快应用平台应用后期，对于高并发需求不能完全满足。
解决方案：采用模块化构建组件，减少各模块间通信。
3. 机器人外形会影响信号稳定性。
解决方案：自主设计机器人外形，减少因外形而造成的影响。

2.6 进度安排

表 2.2 进度安排表

项目重要里程碑	预计完成时间
项目市场调研，需求分析，可行性分析，快应用界面设计	2019-11-10
获取方言音频数据	2020-1-15
搭建训练模型，采集数据构建数据集	2020-1-30

机器人终端设计，测试机器人相关性能	2020-2-15
将各部分连接融合，测试整体性能	2020-3-25
完成项目所有材料	决赛

2.7 开发预算

1. 服务器租赁：1200 元左右
2. 机器人打印耗材：200 元左右
3. 树莓派 4B：500 元左右
4. Arduino UNO：100 元左右
5. 其他外设(音响，话麦，esp8266 等)：200 元左右

3 可行性分析

3.1 技术可行性分析

1. 利用快应用作为方言服务终端。

当前的语音助手大多讲究轻量级，科技公司希望用户可以以最便捷的方式打开语音助手并得到服务，当然，这在普通话方面应用的很不错。但是，科技公司在方言领域涉猎甚少，我们使用快应用这样一种可以方便使用的功能入口，不仅可以极大方便用户的使用方式，同时，因为快应用的联盟效应，多数用户可以享受到科技带来的乐趣。那些不会说普通话的人，可以因此在一定程度上走进现代生活。

为了提升用户体验，我们应用后端也做出了大量的努力。

我们采用了先进的高并发与高可用技术：

缓存层：缓存热点数据，减小数据库压力。

业务数据层：数据库主从双备，读写分离改善数据库负载压力。通过 CDN 和反向代理层缓存，减轻后端服务器负载压力，提高并发能力。冗余部署，反向代理自动选择可用服务，提高可用性。

在上述努力带来的良好体验下，方言用户黏性将会更高。不仅如此，本平台

始终保持着以平台用户为核心，以服务用户为基础，以扩大平台规模为核心竞争力，通过多样化的操作提升用户体验感，保留用户群，防止用户的流失，提高用户黏性。

2. 万物互联：使用快应用控制多种家电

为了使方言的使用真正走进人们的生活，同时实现真正的万物互联，我们决定拓展使用方言来对家电进行控制的方式。使用云服务器或本地服务器进行信息传递的中介，用户便可以方便地通过快应用或通过家电（机器人）来方便对家电进行控制。

目前，存在大量使用语音助手控制家电的方式，但是，美中不足的是，他们都是使用普通话进行控制的，我组作品采用方言控制方式，在一定程度上弥补了不足。

3.2 资源可行性分析

1. 快应用作为一个广泛被使用的社交类 APP，其独特特性一直深受用户欢迎，我们以快应用为载体，完全可以将产品覆盖面拓宽到全安卓用户，同时，得益于快应用随点随用的特性，产品的可使用性会得到进一步的提高。不仅如此，随着用户量的增大，产品的可靠性也会逐渐增强，以此完成产品的闭环和提升。
2. 实体的两足机器人自主设计，采用 3D 打印技术进行打印，其上的单片机较为廉价，但同时又可以完成所需任务。在智慧家居场景下的本地服务器价格较为低廉，仅为当前主流智能家居整体费用的一半不到。
3. 目前，很多厂商都推出了 PaaS 云服务，这些云服务可以较为低廉的价格得到，同时，因为云服务的可拓展特性，随着应用用户的逐渐增多，我们也可以自行拓展服务器配置，完成功能的升级，为用户提供更为优秀的服务。

3.3 市场可行性分析

1. 目前，很多科技公司都推出了基于安卓平台的手机语音助手，但是，这些语音助手只有普通话的版本，用户仅仅可以通过普通话享受到这些方言服务，这对一些方言的使用者来说，显得极其的不友好。由此可见，方言应用市场

广阔，服务方言使用者，我们在行动。

2. 目前，使用手机点击控制家电成为主流方式，然而，当用户不方便使用手机的点击功能时，控制家中电器成为了一件难事。我组提出的，通过方言语音控制家中电器的方式在一定程度上解决了这个问题，用户在家中或是在公司都可以通过快应用来方便地控制家电。这样一来，方言用户的科技生活又上一层楼。不仅如此，该功能在普通话版本的语音助手中普及度也较低，我们完全也可以将该技术应用到普通话版的家居助手中。由此可见，市场应用前景广阔。

4 需求分析

4.1 数据需求

4.1.1 静态数据

1. 用于进行方言识别模型训练的方言数据。
2. 用于矫正机器人动作的动作参数。

4.1.2 动态数据

1. 用户在使用该产品时所上传的方言数据。（在遵从隐私协议的前提下，我们会采用无差别方式进行模型训练）

4.1.3 数据采集

1. 走进方言使用者的居住区，实地采集第一手方言数据。
2. 利用快应用早期高质量用户上传的方言数据。
3. 对机器人进行多次多方位数据采样。

4.2 功能需求

4.2.1 功能模块

AI-智慧方言功能如下：

1. 用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，在服务器端进行谱减法降噪后进行方言

内容的识别和种类的分类,在得到方言内容后,调用本地语音助手进行服务。
(该功能因为快应用没有开放语音助手接口,将在开放后进行完善)。

2. 用户点击快应用中的方言翻译功能,通过按住按钮让持方言人说出方言,快应用根据网络环境智能选择服务器进行服务,在服务器端进行谱减法降噪后进行方言内容的识别和种类的分类,得到方言内容和种类后,将方言中种类和内容显示在屏幕上告知客户。
3. 用户通过点击快应用中的家居助手功能,选择希望控制的家电,通过按住按钮说出控制方言,快应用根据网络环境智能选择服务器进行服务,对家中电器进行控制。使用该功能时,无论用户是在家中或者户外,都可以在连接互联网的情况下对家居进行控制。(其中家电将被抽象为机器人)
4. 用户通过触摸开关对家居说出方言后,家具通过本地服务器获取方言结果,随后进行对应的动作,完成对客户的服务。
5. 在每周日晚上2点时各本地服务器上传方言数据,在云服务器中进行神经网络的训练,在训练完成后,自动将神经网络参数部署到各本地服务器,完成神经网络的更新。

应用平台的功能模块结构如图所示:

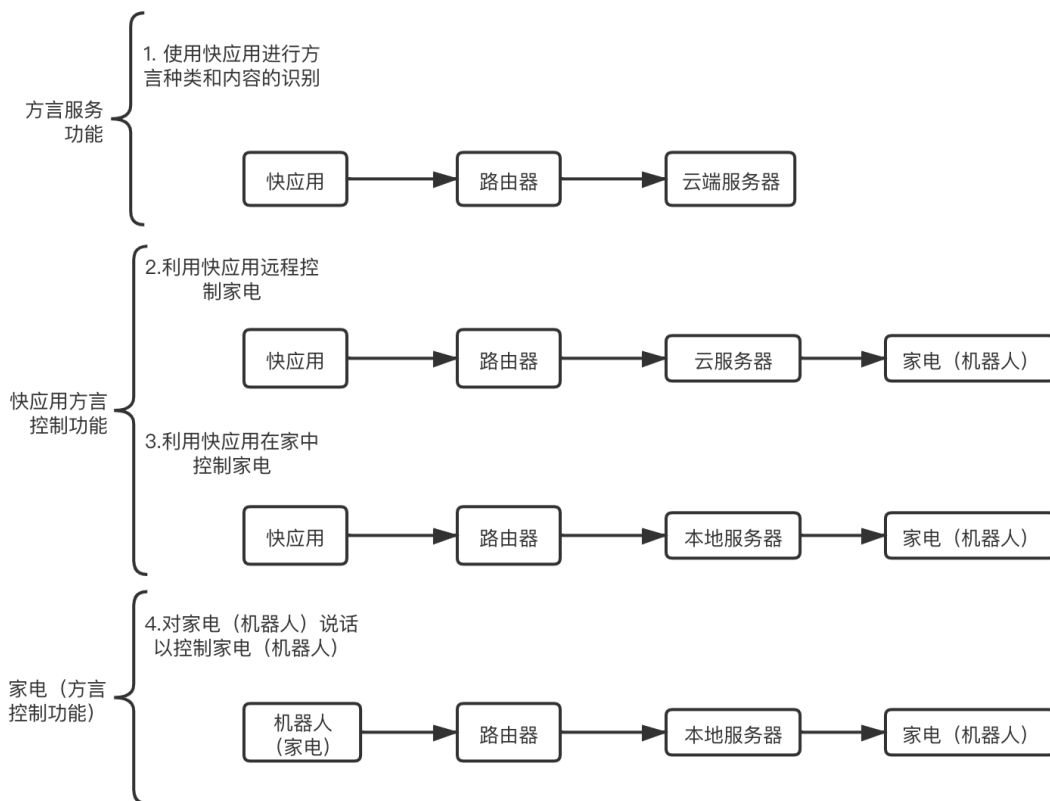


图 7 AI-智慧方言功能实例图

快应用的用户用例图如图所示：

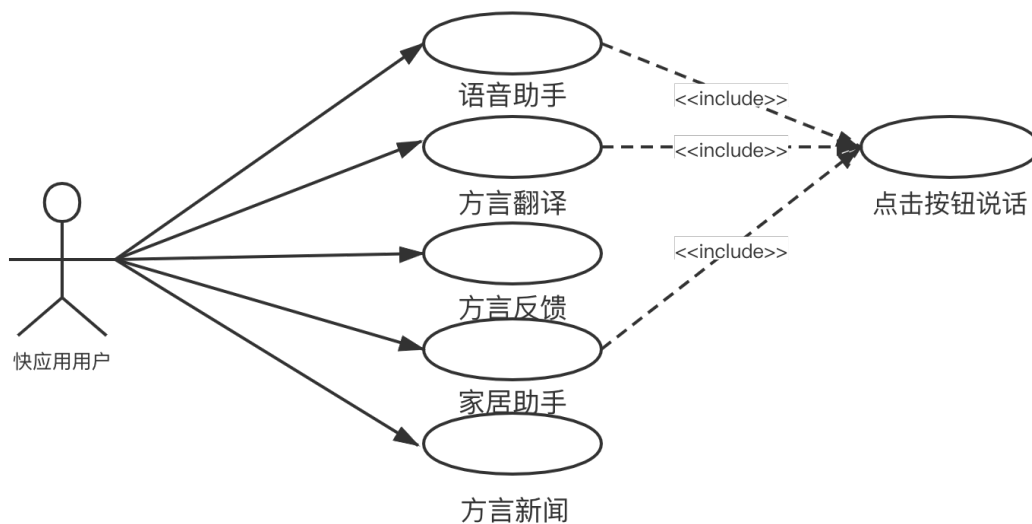


图 8 本系统快应用用例图

业务流程图如图所示：

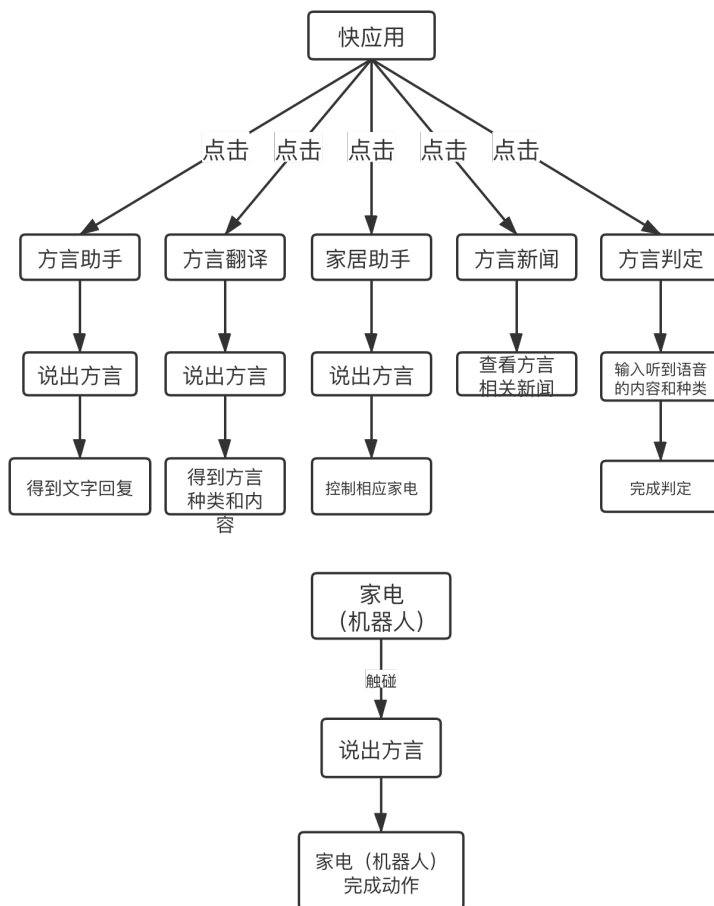


图 9 业务流程图

1. 方言助手模块

方言助手为用户提供了方言版的语音助手服务，所有准备工作已经完成，当快应用开放系统底层语音助手接口后即可上线。

表 1 方言助手表

一级模块	二级模块	用户群	优先级
主界面	方言助手	所有用户	高

2. 方言翻译功能

方言翻译功能是为用户提供对应方言数据的翻译和内容。

表 2 方言翻译表

一级模块	二级模块	用户群	优先级
------	------	-----	-----

主页面	方言判定	所有用户	高
	方言翻译	所有用户	高

3. 家居助手模块

家居助手模块为用户提供了方便的通过快应用来在不同网络环境下控制家中电器（机器人）的高效方式。

表 3 家居助手表

一级模块	二级模块	用户群	优先级
主界面	家居助手	所有用户	高

4. 方言数据审核

任务审核功能由两部分组成，具体可执行操作如下表 4.3 所示：

表 4 任务审核表

一级模块	二级模块	用户群	优先级
任务审核	机器审核	APP 自动执行	高
	人工审核	所有用户	高

- 机器审核：对于用户提交的采集任务进行审核
- 人工审核：对机器审核中未能达到标准的任务发布标识任务，用户可以在平台接取标识任务进行审核。

5. 方言新闻

方言新闻为用户提供时下最新的方言新闻供人们阅读。

表 5 方言新闻表

一级模块	二级模块	用户群	优先级
主页面	方言新闻	所有用户	中
	平台推广	所有用户	中

方言新闻：为用户提供时下最新的方言新闻供人们阅读。

平台推广：用户可以通过分享二维码或者链接的方式对平台任务进行推广，

推广成功平台会有相应奖励。

6. 机器人端

用户通过触摸向机器人说话，机器人得到数据后上传到服务器进行数据判断，随后得到方言内容和种类，并完成相应内容。

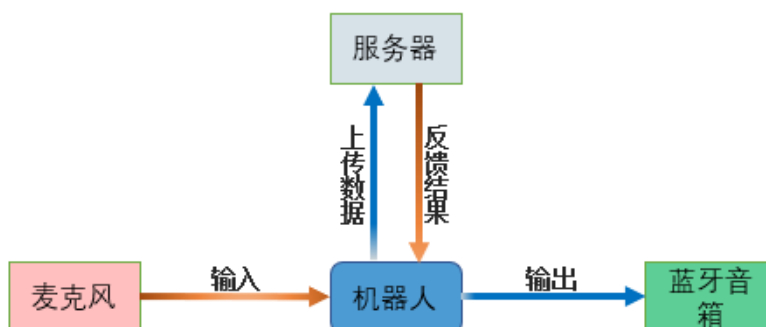


图 10 机器人端示意图

4.3 性能需求

4.3.1 时间特性

1. 服务器响应时延：< 1s
2. 服务器方言预测时延：< 1s
3. 数据传输时延：< 0.5s
4. 机器人回复时延：< 2s

4.3.2 适应性

1. 快应用运行在快应用 1060 以上版本上系统无异常。
2. 服务器方言预测模型运行在 CentOS7 及以上系统中的 python3.6.8 上，系统无异常。
3. 树莓派内程序运行在 Ubuntu16.04 及以上系统中的 python3.6.8 上，系统无异常。

4.4 界面需求

1. 主界面

2. 方言翻译界面
3. 家居助手界面
4. 语音助手界面
5. 方言新闻界面

4.5 接口需求

4.5.1 硬件接口

1. 机器人舵机：Arduino UNO 进行有线控制
2. Esp8266 WiFi: 局域网内 UNO 控制传输
3. Arduino UNO: 使用蓝牙进行调用
4. 树莓派 4B: 云服务器通过网络层协议调用

4.5.2 软件接口

1. 云服务端使用 http 服务方式为快应用提供方言的上传接口。
2. 本地服务器以 http 服务方式为快应用及机器人提供方言数据的上传接口。
3. 云服务端使用 http 服务方式为快应用提供提取方言内容和分类的接口。
4. 本地服务器使用 http 服务方式为快应用提供提取方言内容和分类的接口。
5. 云服务器和本地服务器为机器人提供了 http 的访问数据方式和数据上传方式。

4.6 其他需求

4.6.1 优秀的扩展性和可维护性

本系统是由许多功能独立的组件完成，组件的链接通过固定的接口定义来实现，这使得对某一组件的改动对其他组件的影响最小，使得对平台的维护变得简洁而灵活，并且可以根据实际情况，有选择的开发出零客户端管理的应用平台，减少日常维护的工作量，另外随着业务需求的变动，平台需要局部改变，只需要修改或添加相关组件即可，其他功能部分的组件不需要任何改动，使得平台的扩

展性和灵活性得到更大的提高。

4.6.2 高可靠性

本系统的可靠性通过组建的功能独立性得以提高,某一组件的故障不会影响其他功能组件的正常运行,对某一故障组件的修复也不会中断其他平台功能的正常使用。另外,可以使用组件的多份部署来降低平台对由机器或网络故障所引起的损失提高平台的容错性。

4.6.3 灵活的配置和高效的性能

本系统在项目的第二期会选择分布式的组件,即让组件分布于多台机器上运行,使得以多台廉价的低性能服务器达到单台高性能服务器的应用效果的想法得以实现,并可以使得计算复杂统计的时候更加快速。当并发量继续稳步增长时,我们计划每增加规定的并发量就增加相应的应用服务器来提高平台性能,而需要对应用系统所作的工作只是重新调整组件的分布,不需要对现有的平台进行任何的编码改动。

4.6.4 安全性

系统安全: 采用主机加固手段加固主机;对系统重要文件进行及时备份、加密来保障系统文件的安全;严格控制权限加强对主机的访问控制;日志分析,及时发现异常。

网络安全: 划分安全子网;定制网络安全检测实现持续安全。

数据加密: 通过数据加密技术保护系统安全,防止数据丢失和数据窃听、窥视。

4.6.5 可重用性

采用了多种设计模式,S2SH 的多层架构,因而具有很好的可重用性。

5 概要设计

5.1 处理流程

本系统的总体运行流程图如图 11 所示:

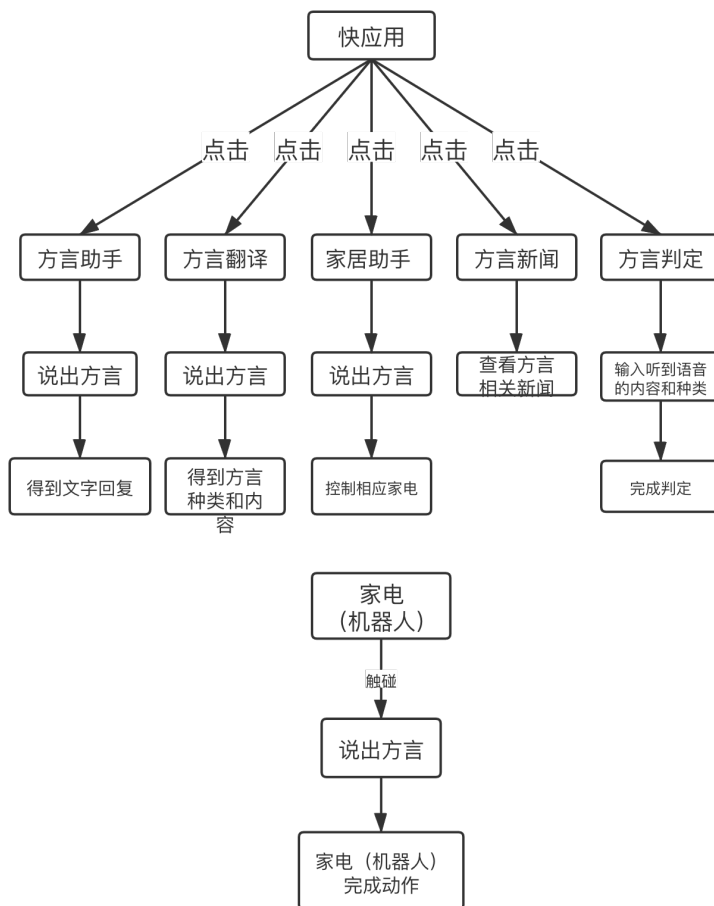


图 11 系统的总体运行流程图

5.2 总体结构设计

本项目《AI—智慧方言》，采用模块化设计，即快应用模块、云服务器深度学习模块、智能家居应用模块搭建方言采集与应用平台。三大模块相互独立又紧密联系，在每个模块各自完成其任务之时，又通过 WIFI，云服务器等手段进行数据交互，构成一个紧密的整体。不仅如此，我们考察多种家具属性，抽象出具有更高代表性的机器人，极大提高了产品的普适性，紧紧贴合“万物互融”的指导思想。

首先，我们选择使用快应用这样一个人人都使用的平台来对用户进行服务，因为快应用的卡片化服务机制，用户使用软件极其方便，这极大简化了使用的流程，不仅如此快应用可以作为子集加入其他的 app 应用，拓宽使用面。

快应用界面如下：



图 12 快应用端

为了使方言走进大众的生活，我们希望将终端技术用于各种家居设备，来提高用户的参与度，让科技方便人们的生活。但是，每种家居都有其固有属性，在展示时我们无法展示所有家居的使用方式，于是，在考察了多种家居的属性后，我们自主设计了机器人，该机器人可以认为是各种家居的抽象，同时，机器人技术也可以很方便地迁移到各种家居中。机器人可以根据用户说出的方言，用对应的方言回复用户并做出方言内容指示的动作。当用户不在机器人附近时，用户也可以通过快应用来方便地通过语音控制机器人。

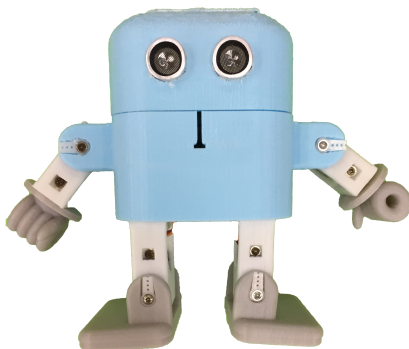


图 13 机器人模型

除此之外，为了使这些方言真正用到实处，我们通过搭建快应用和数据管理系统，在遵循隐私协议的前提下，收集用于服务客户而产生的海量高品质数据集。通过使用 Keras 深度学习框架，使用 WaveNet 神经网络架构，利用所搜集到的海量方言数据集训练出一个准确率较高的高效网络。利用该模型，即可对包含长沙话、四川话、合肥话在内的八种方言进行分类。我们选择将网络部署到云端和本地，达成功耗和方便性的统一：用户在多种网络环境下均可享受服务——在网络

环境较好时，可以通过云服务器进行服务；当网络环境较差时，可以通过同一局域网下的本地服务器进行服务。

三者关系如下：

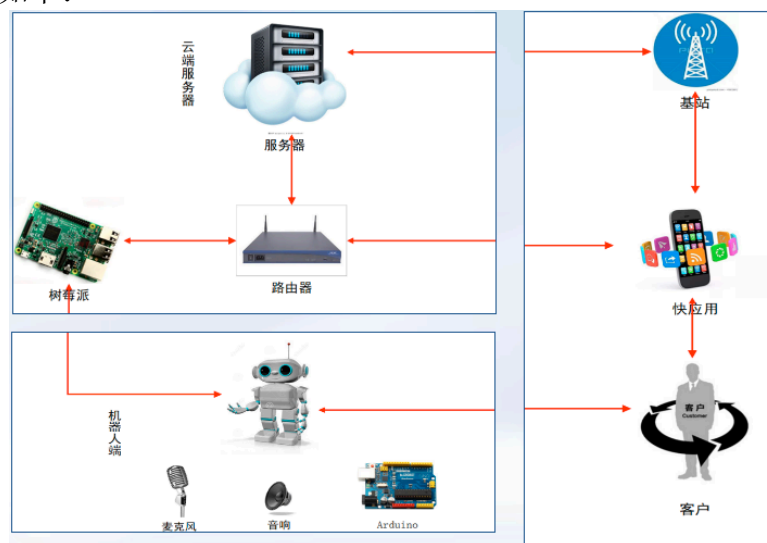


图 14 系统总体架构

系统工作流程：

- 在用户使用快应用时，有三种功能：

(1) 方言版语音助手

用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，得到方言内容后，调用本地语音助手进行服务。（该功能因为快应用没有开放语音助手接口，将在开放后进行完善）。

(2) 方言翻译助手

用户点击快应用中的方言翻译功能，通过按住按钮让持方言人说出方言，快应用根据网络环境智能选择服务器进行服务，得到方言内容和种类后，将方言中种类和内容显示在屏幕上告知客户。

(3) 家居助手

用户通过点击快应用中的家居助手功能，选择希望控制的家电，通过按住按钮说出控制方言，快应用根据网络环境智能选择服务器进行服务，对家中电器进行控制。使用该功能时，无论用户是在家中或者户外，都可以在

连接互联网的情况下对家居进行控制。（其中家电将被抽象为机器人）

- 在用户想要直接控制家电时：

(1) 面对面家居控制

用户通过触摸开关对家电说出方言后，家电通过本地服务器获取方言结果，随后进行对应的动作，完成对客户的服务。

我们会在遵守隐私协议的前提下，对可信度较高的数据，以无差别方式使用，以此为用带来更为优质的服务。

5.3 功能设计

1. 方言助手模块

方言助手为用户提供了方言版的语音助手服务，所有准备工作已经完成，当快应用开放系统底层语音助手接口后即可上线。

表 1 方言助手表

一级模块	二级模块	用户群	优先级
主界面	方言助手	所有用户	高

2. 方言翻译功能

方言翻译功能是为用户提供对应方言数据的翻译和内容。

表 2 方言翻译表

一级模块	二级模块	用户群	优先级
主页面	方言判定	所有用户	高
	方言翻译	所有用户	高

3. 家居助手模块

家居助手模块为用户提供了方便的通过快应用来在不同网络环境下控制家中电器（机器人）的高效方式。

表 3 家居助手表

一级模块	二级模块	用户群	优先级
主界面	家居助手	所有用户	高

4. 方言数据审核

任务审核功能由两部分组成，具体可执行操作如下表 4.3 所示：

表 4 任务审核表

一级模块	二级模块	用户群	优先级
任务审核	机器审核	APP 自动执行	高
	人工审核	所有用户	高

- 机器审核：对于用户提交的采集任务进行审核
- 人工审核：对机器审核中未能达到标准的任务发布标识任务，用户可以在平台接取标识任务进行审核。

5. 方言新闻

方言新闻为用户提供时下最新的方言新闻供人们阅读。

表 5 方言新闻表

一级模块	二级模块	用户群	优先级
主页面	方言新闻	所有用户	中
	平台推广	所有用户	中

方言新闻：为用户提供时下最新的方言新闻供人们阅读。

平台推广：用户可以通过分享二维码或者链接的方式对平台任务进行推广，推广成功平台会有相应奖励。

6. 机器人端

用户通过触摸向机器人说话，机器人得到数据后上传到服务器进行数据判断，随后得到方言内容和种类，并完成相应内容。

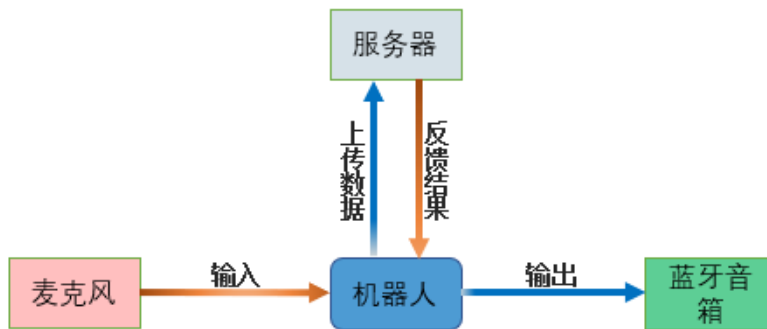


图 15 机器人端示意图

5.4 用户界面设计



图 16 主界面



图 17 主界面

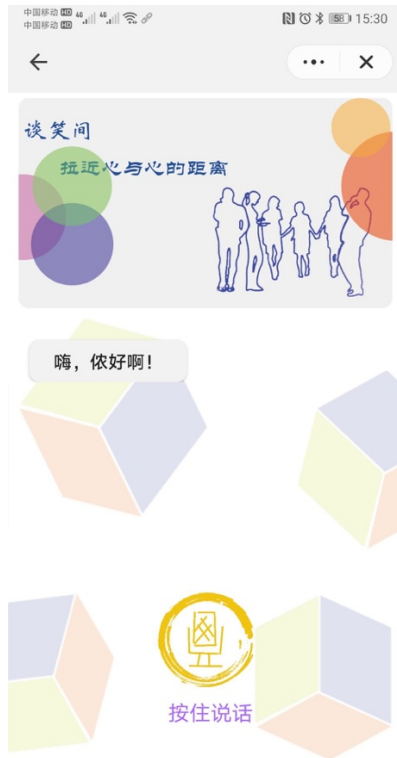


图 18 方言助手



图 19 方言翻译



图 20 方言翻译



图 21 家居助手



图 22 家居助手



图 23 我的空间



图 24 数据反馈

图 25 数据反馈

5.5 数据结构设计

数据层处于系统的底层，一方面负责高效存储数据，另一方面负责接受数据库的请求，调用底层方法获取数据，并对其加工处理。数据层负责完成了所有对基本数据的操作，其读写效率、稳定性、可用性会影响到这个系统的效率和拓展能力。综合来说，对于数据层的选择，要兼顾项目规模、成本预算、性能指标、拓展性要求以及业务场景需要，所以我们采用关系型数据库 Mysql 和非关系型数据库 Redis。

系统采用 Mysql 作为系统数据库，作为主流的关系型数据库，其拥有广阔的社区、详尽的资料、充分的实践、以及规范的 SQL 模型，其支持事务、存储过程、触发器、主从复制等，并发性能好。在较新的版本，无论是数据类型支持的多样性方面，还是在数据库引擎的性能优化方面，都有卓越的进步。考虑到项目规模、性能指标、业务场景需要等因素，我们选择使用关系型数据库 Mysql。

Druid 是目前最好的开源的数据库连接池，在功能、性能、拓展性方面有超过其他数据库连接池，比如 DBCP、C3P0、JBoss 等。其遵从 LRU、支持 PSCache，拥有极佳的性能优势，具有 ExceptionSorter 容错机制，稳定安全；支持 SQL 拦截与解析，提供多种方式监控数据库，拓展性极好。使用 Druid 连接池，在节约系统资源的同时，极大提高了应用访问 Mysql 服务的性能，还提供详尽的监控日志功能，为后期的数据库优化和错误排查提供了有力帮助。

5.6 接口设计

5.6.1 外部接口

表 1 硬件接口：

序号	连接方式	连接设备	
1	Bluetooth5.0	Arduino Nano	Raspberry Pi 4B
2	互联网	Raspberry Pi 4B	aliyun Server
3	针脚连接	MAX9814	Arduino Nano

4	针脚连接	Micro SD Module	Arduino Nano
5	针脚连接	Micro SD Module	Esp8266 Nodemcu

表 2 软件接口

序号	连接方式	连接设备	
1	HC-06 Bluetooth5.0	Arduino Nano	Raspberry Pi 4B
2	Esp8266 nodemcu	Arduino Nano	Raspberry Pi 4B
3	Secure Shell	Raspberry Pi 4B python	aliyun Server python Java
4	互联网	安卓设备	aliyun Server

5.6.2 内部接口

1. 云服务端使用 http 服务方式为快应用提供方言的上传接口。
2. 本地服务器以 http 服务方式为快应用及机器人提供方言数据的上传接口。
3. 云服务端使用 http 服务方式为快应用提供提取方言内容和分类的接口。
4. 本地服务器使用 http 服务方式为快应用提供提取方言内容和分类的接口。
5. 云服务器和本地服务器为机器人提供了 http 的访问数据方式和数据上传方式。

5.7 错误/异常处理设计

5.7.1 错误/异常输出信息

1. 数据上传超时
2. 数据无法获取

5.7.2 错误/异常处理对策

向云端汇报错误信息，并重启应用。

5.8 系统配置策略

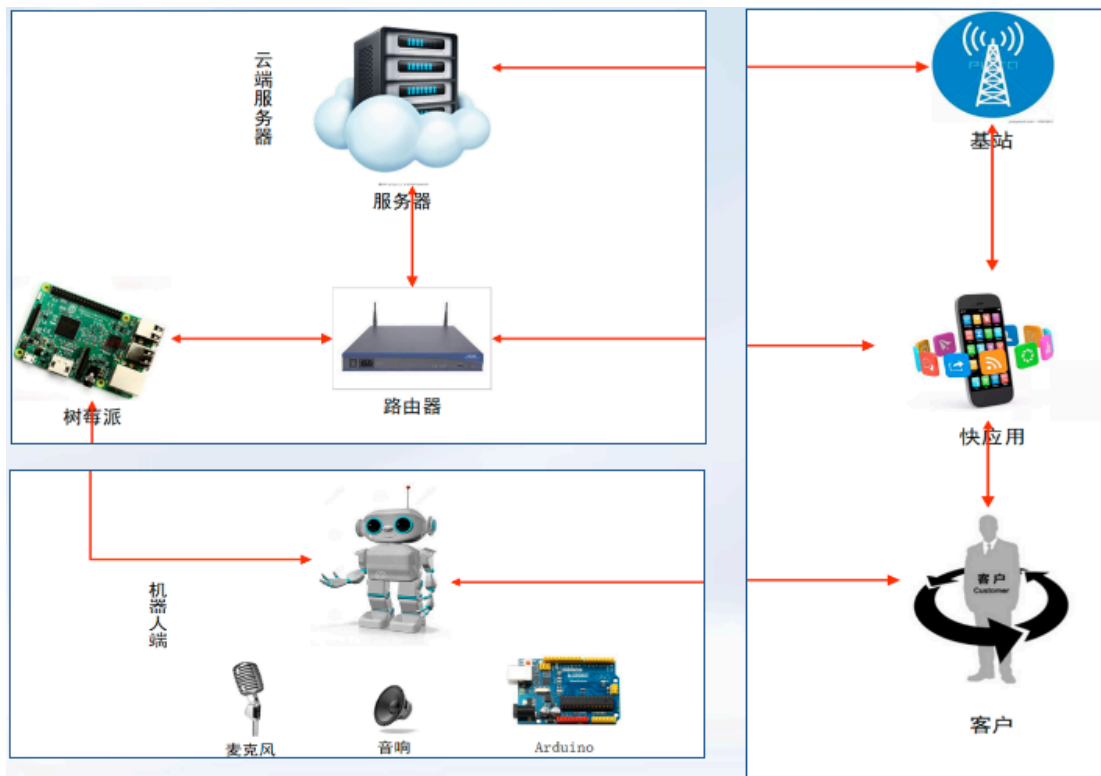


图 26 系统配置策略图

5.9 系统工作流程:

- 在用户使用快应用时，有三种功能：

(1) 方言版语音助手

用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，得到方言内容后，调用本地语音助手进行服务。（该功能因为快应用没有开放语音助手接口，将在开放后进行完善）。

(2) 方言翻译助手

用户点击快应用中的方言翻译功能，通过按住按钮让持方言人说出方言，快应用根据网络环境智能选择服务器进行服务，得到方言内容和种类后，将方言中种类和内容显示在屏幕上告知客户。

(3) 家居助手

用户通过点击快应用中的家居助手功能，选择希望控制的家电，通过按住按钮说出控制方言，快应用根据网络环境智能选择服务器进行服务，对家中电器进行控制。使用该功能时，无论用户是在家中或者户外，都可以在连接互联网的情况下对家居进行控制。（其中家电将被抽象为机器人）

- 在用户想要直接控制家电时：

(1) 面对面家居控制

用户通过触摸开关对家电说出方言后，家电通过本地服务器获取方言结果，随后进行对应的动作，完成对客户的服务。

我们会在遵守隐私协议的前提下，对可信度较高的数据，以无差别方式使用，以此为用带来更为优质的服务。

5.10 系统部署方案

5.10.1 系统表示层

表示层是用户及管理人员直接操作的界面，负责接受用户请求、向应用服务器发送请求、接收并渲染得到可视化结果。系统表示层主要为：移动客户端界面

快应用界面设计遵循 Google 的 Material Design 设计风格，通过平面像素的 Z 轴概念，使 UI 具有二维平面的海拔效果，简单扁平、拟物生动，相比于扁平化，具有更好的用户体验。

机器人方式，我们采用自主设计机器人模式，在一定程度上可以模拟家居，为用户提供高效的智能家居生活。

5.10.2 系统应用层

应用层用于接收用户请求，调用对应的业务层逻辑，将返回结果整合成规范的 json 格式或者渲染成 html，返回给表示层。此外，其需要负责数据校验、身份验证、日志记录、异常处理等工作，其稳定运行直接影响到系统响应的可靠性和及时性。

我们采用 SpringBoot 这个微服务构建框架，以合理的接口为系统表示层提供服务，以此来为用户提供高效和优秀的服务。同时，使用 Keras 构建深度学习

网络，为方言数据的内容和种类的分类判断提供支持。不仅如此，为了保障语音的信息无损，我们使用谱减法降噪技术为语音进行降噪。

5.11 其他相关技术与方案

5.11.1 均衡负载技术

负载均衡其意思就是将负载进行平衡、分摊到多个操作单元上进行执行，如下图所示。

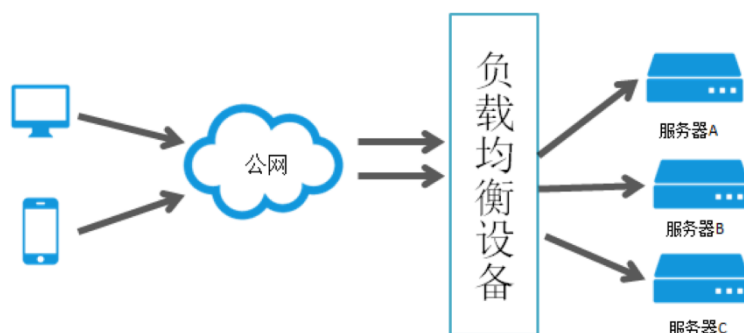


图 27 负载均衡工作原理图

目前主要存在硬件负载均衡、软件负载均衡两种方式，其中硬件负载均衡需要花费较高的费用，软件负载均衡中间件有 Nginx、NodeJs Server 等，使用低成本就能达到企业级负载均衡的效果。遥聆采用软件负载均衡方式。

5.11.2 云计算

云计算是一种商业计算模型。它将计算任务分布在大量计算机构成的资源池上，使各种应用系统能够根据需要获取计算力、存储空间和信息服务。因其高度的水平拓展性、通用型、高可靠性，成本低廉且无需额外维护，非常适合遥聆初期使用。考虑到遥聆未来的市场前景以及规模拓展速率，我们选择阿里云提供的可按需服务的 Linux 服务器以及数据库服务器。

5.11.3 三重数据审核算法

利用全连接 BP 神经网络进行机器学习，设定双阈值区间，合理将数据集分为三类，分别为真数据、疑似数据、假数据。系统自动将疑似数据进行整合并发布标识任务，利用推荐算法邀请平台高分用户对疑似数据进行判别、投票。若最终用户投票的比例(是/否 或 否/是)小于七比三，则该数据将递交给管理员做最终决断。

6 数据库设计

6.1.1 系统数据层

数据层处于系统的底层，一方面负责高效存储数据，另一方面负责接受数据库的请求，调用底层方法获取数据，并对其加工处理。数据层负责完成了所有对基本数据的操作，其读写效率、稳定性、可用性会影响到这个系统的效率和拓展能力。综合来说，对于数据层的选择，要兼顾项目规模、成本预算、性能指标、拓展性要求以及业务场景需要，所以我们采用关系型数据库 Mysql 和非关系型数据库 Redis。

6.1.2 系统数据库 Mysql

系统采用 Mysql 作为系统数据库，作为主流的关系型数据库，其拥有广阔的社区、详尽的资料、充分的实践、以及规范的 SQL 模型，其支持事务、存储过程、触发器、主从复制等，并发性能好。在较新的版本，无论是数据类型支持的多样性方面，还是在数据库引擎的性能优化方面，都有卓越的进步。考虑到项目规模、性能指标、业务场景需要等因素，我们选择使用关系型数据库 Mysql。

6.1.3 开源的数据库连接池

Druid 是目前最好的开源的数据库连接池，在功能、性能、拓展性方面有超过其他数据库连接池，比如 DBCP、C3P0、JBoss 等。其遵从 LRU、支持 PSCache，拥有极佳的性能优势，具有 ExceptionSorter 容错机制，稳定安全；支持 SQL 拦截与解析，提供多种方式监控数据库，拓展性极好。使用 Druid 连接池，在节约系统资源的同时，极大提高了应用访问 Mysql 服务的性能，还提供详尽的监控日志功能，为后期的数据库优化和错误排查提供了有力帮助。

6.1.4 数据库内容

数据库内容 ER 图如下：

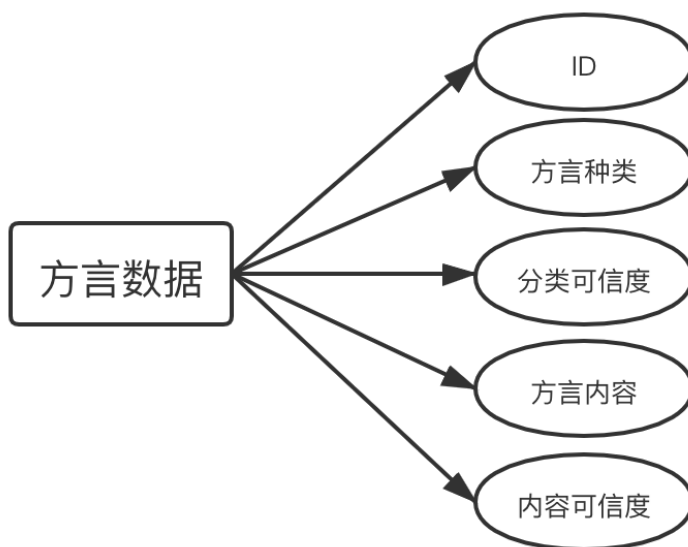


图 28 E-R 图

每个属性意义如下：

1. ID：每条方言唯一标示
2. 方言种类：该方言最终确定的种类。
3. 分类可行度：机器分类后的准确度经校准后的准确度。
4. 方言内容：该方言数据最终确定的方言内容。
5. 内容可信度：机器判断后的准确度经校准后的准确度。

7 详细设计

7.1 终端功能模块

7.1.1 功能描述

- 在用户使用快应用时，有三种功能：

(1) 方言版语音助手

用户点击快应用中方言助手功能后，点击按钮对手机说出方言。快应用根据网络环境智能选择服务器进行服务，得到方言内容后，调用本地语音助手进行服务。（该功能因为快应用没有开放语音助手接口，将在开放后

进行完善)。

(2) 方言翻译助手

用户点击快应用中的方言翻译功能,通过按住按钮让持方言人说出方言,快应用根据网络环境智能选择服务器进行服务,得到方言内容和种类后,将方言中种类和内容显示在屏幕上告知客户。

(3) 家居助手

用户通过点击快应用中的家居助手功能,选择希望控制的家电,通过按住按钮说出控制方言,快应用根据网络环境智能选择服务器进行服务,对家中电器进行控制。使用该功能时,无论用户是在家中或者户外,都可以在连接互联网的情况下对家居进行控制。(其中家电将被抽象为机器人)

- 在用户想要直接控制家电时:

(1) 面对面家居控制

用户通过触摸开关对家电说出方言后,家电通过本地服务器获取方言结果,随后进行对应的动作,完成对客户的服务。

7.1.2 性能描述

通过实验,快应用端可以完成方言内容和分类的输出,不仅如此,我们设计的快应用还可以嵌入到其他的 app 中,从而拓展快应用的使用范围。

在利用流媒体数据采集完成了大量数据的采集之后,我们使用采集到的数据进行深度学习网络的训练,实验证明,高效的网络架构和高效的数据集使得网络的训练较为成功,利用该模型,即可对包含长沙话、四川话、合肥话在内的八种方言进行分类。

在用户使用快应用的方言翻译功能时,最短将在 1s 内得到方言内容和种类的回复。

在用户使用快应用控制家电时,最短将会在 1s 内完成控制动作。

线下的机器人端,可以在接收用户的语音输入之后,在 0.5s 以内对用户的输入的方言进行同种方言的反馈,同时,将会进行一定的动作,来完成与用户的交互。

7.1.3 输入

通过所搭建的快应用和数据管理系统，利用大量移动端用户可以高效、低廉地搜集海量高品质数据集。通过使用 Keras 深度学习框架，不断地堆积一维卷积层和池化层并利用所搜集到的海量方言数据集可训练出一个准确率极高的高效网络。

随后，在使用过程中，用户通过在不同界面下输入方言数据，便可完成服务。同时，我们会在遵守隐私协议的前提下，对可信度较高的数据，以无差别方式使用，以此为用带来更为优质的服务。

7.1.4 输出

快应用在上传数据到服务器后，服务器对数据进行一系列的预处理，如端点检测、降噪、语音增强等，将预处理完成的数据作为输入数据，通过深度学习模型得到方言分类结果，随后快应用会得到关于方言内容和方言种类的返回。

机器人模块中，说话人朝着麦克风说完方言之后，通过机器人端的树莓派将方言数据发送至服务器，并对数据进行一系列的预处理，如端点检测、降噪、语音增强等，将预处理完成的数据作为输入数据，通过深度学习模型得到方言分类结果，并将结果回传树莓派，树莓派控制机器人做出回复，做出相应动作。

7.1.5 程序逻辑

系统总体逻辑如下：

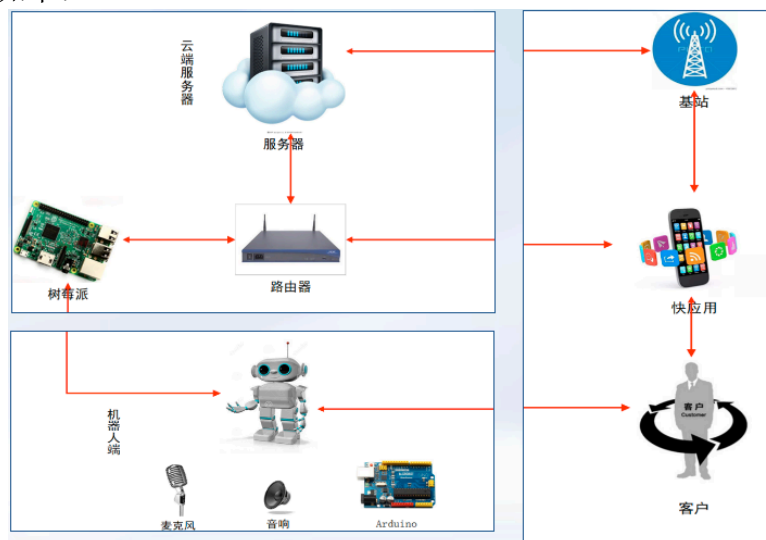


图 29 系统架构图

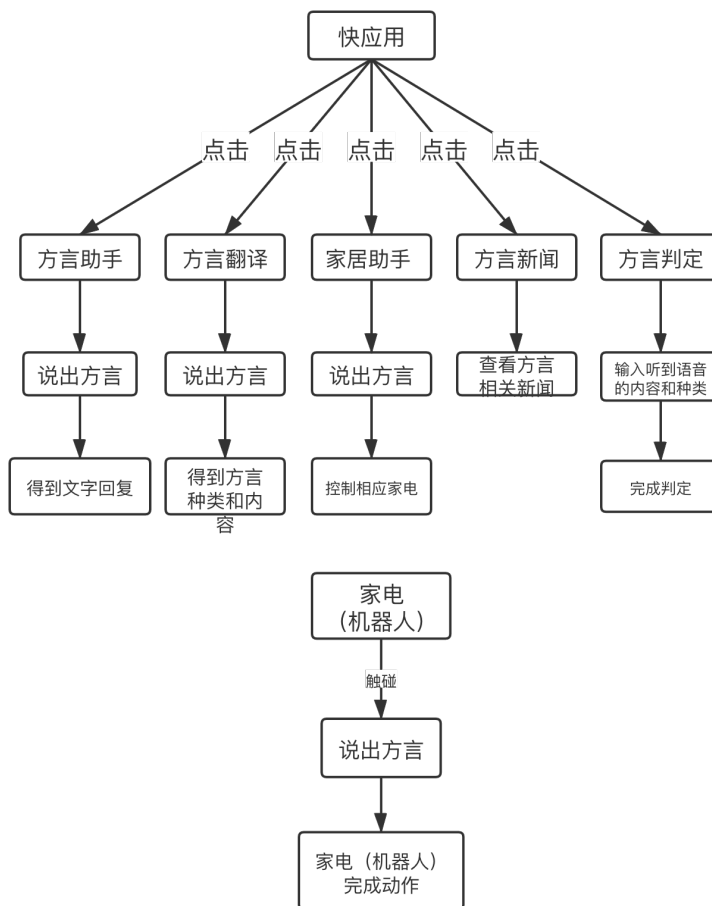


图 30 系统的业务流程图

系统的应用服务架构，主要分为表示层、应用层、业务层、数据层和基础层 5 大方面。

7.1.6 限制条件

在使用使用数据为训练数据时，会使用三重数据审核算法：

利用全连接 BP 神经网络进行机器学习，设定双阈值区间，合理将数据集分为三类，分别为真数据、疑似数据、假数据。系统自动将疑似数据进行整合并发布标识任务，利用推荐算法邀请平台高分用户对疑似数据进行判别、投票。若最终用户投票的比例(是/否 或 否/是)小于七比三，则该数据将递交给管理员做最终决断。