

基于语音转换校核的口令操作安全校核系统设计

对现有的国内外的语音转换技术进行分析，对智能分词技术进行研究，将这两类技术与电网口令操作相结合，设计一种基于语音转换校核的口令操作安全校核系统，主要应用于电网调度控制的日常工作中，提升电网工作的安全性。

叶海峰 陈存林 李玉龙 李有亮 孙仪 黄寅/安徽省电力公司



叶海峰/高级工程师

关键词/Keywords

语音转换 ·
智能分词 ·
口令操作 ·

0 引言

在电网的运作过程中，电网操作直接影响了电网的安全，而调控中心作为电网调度部门对电网进行运营管理的窗口部门，操作指令基本都是从此处发出，因此操作指令的安全性尤为重要。目前，在大多数的情况下，调控员会先写操作票，然后经过一系列的标准化审核过程和相应的校核，保证操作指令的安全后，再进行下达。然而，也存在遇到紧急情况时，调控员直接口令下达，事后再进行补写操作票的情况。目前，在国内没有有关此类口令操作的校核功能的系统。本文中所描述的基于语音转换校核的口令操作安全校核系统主要研究的就是这类口令操作的相关校核工作，补全口令操作票校核的空白，提升调控操

作的安全性。

1 系统设计

本系统主要是用来从电话系统中，采集调控员的下令内容，并实时将语音信号转换为文字内容，再通过对文字内容进行分词操作，解析为电力操作术语，最终将操作解析结果提交到成熟的校核系统中进行校核，在接收到校核的返回结果并确认后，正式进行操作，保证口令操作的安全可靠。图 1 为系统工作的流程图。

1.1 语音采集

语音信号的采集方式比较多样，例如麦克风采集、电话信号收集和录音文件等。但是在本系统中，实时性的要求很高，同时，系统后期展望将实现对对侧的语音信号进行收集转换，以防止调控员的理解发生歧义，所以需要采用直接采集电话信号的方式。

考虑到电话系统的实时性要求以及现有实现功能的独立性，借鉴电话录音系统的实现方式，采用语音卡的方式进行电话语音信号的采集。

语音采集流程如图 2 所示。在本系统中，采用了语音卡进行信号的采集传输，在经过电话系统后，语音信号被传递到语音卡中，语音卡对数字

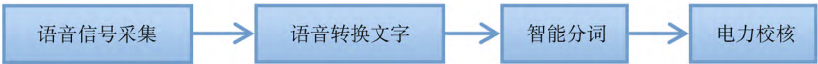


图 1 系统工作流程



图 2 语音采集流程图

信号进行了处理。一方面,语音卡按照调控员的实际业务需求可以选择将信号丢弃或者缓冲传出;另一方面,将所需要转换的音频流保存到文件,以保留备份。

1.2 语言转换

目前,在国内外,存在多种语音转换相关的工具软件。例如 Word 自带语音识别,紫光等各大输入法也都具有语音识别功能,而目前商业应用中比较完善的一个是微信中的语言输入,一个是科大讯飞的 TTS,还有一些其他的语音文字转换工具。

语音转换模块主要是用来将收集到的语音信号转化成文字。该模块的关键点在于,转换的正确性保证与提升,在对各大语音转换工具从使

用性能、准确性、经济性和系统运行环境等方面进行多维的考虑之后,本系统中选择使用科大讯飞的 TTS 进行语音信号的转换。

在 TTS 中,提供了与网络无关的语音包,通过调用本语音包,可以实现麦克风的输入信号转换到文字。同时,在本语音包中,可以兼容方言的差异,大大提升了系统的实用性。但是,电话系统的输出是数字流格式,语音包接收的信号则是音频流,这是两种截然不同的格式。因此需要对语音包进行相应的外部封装和改造。在本系统中,对语音包采用适配器模式进行了改造,接收数字信号,并转换成为语音信号。语音转换包如图 3 所示。

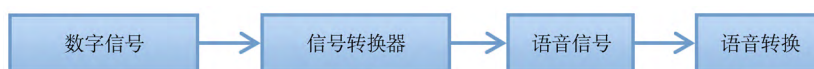


图 3 语音转换包

1.3 智能分词

智能分词作为本系统的核心部分,主要是用来对语音转换出的文字结果进行处理,将操作命令,通过分析其中的关键词确定主语、谓语和宾语后,将主语部分分析成设备,谓语部分分析成动作,宾语部分分析成状态,得到该命令对应的操作设备与动作,以便于后续进行安全校核。

智能分词是智能化的范畴,也是目前在国内外研究相当多的一个主题部分,各类相关的学术、商业产品层出不穷。但是,相关的产品所面向的都是行业无关的背景,所分出的词汇一般都是依据字典、常用语等。而在电力系统中,设备名称、电网操作术语等都存在一些行业特色的、不常用的词汇,这些词汇会导致智能分词的结果的差异性。另一方面,在电网下令的过程中,所使用的术语一般都是固定格式的。例如“将东临 5244 线由运行转为热备用”,在专业分词软件中,东临 5244 线会被分成多个词汇,导致设备名称都不正确。而一个智能分词软件的评价标准主要就是:分词的准确性。鉴于电力术语的格式有限,本系统中自行研制了电力智能分词软件进行相应的分词。

智能分词分为几个关键部分,包括:

1) 对术语结构的自行可扩展定义。即智能

分词所依赖的术语库是可以在后期自行进行定义和扩展的。

2) 术语固定含义描述。对术语中的一些关键词进行了转义,为后续校核工作提供便利。

3) 可维护的电力词库。分词的基础就是词库,对以后出现的一些如设备名称等新的词汇,提供词库可进行维护等。

在模块的实现中,对可能出现的电力词汇进行罗列并典型分析,并对关键词汇进行关联性定义,例如在“将”之后一般是主语,而在主语之后跟有谓语等基本的语法规则。经过验证,本系统所提供的智能分词模块已经能满足电力术语解析的需求。

1.4 安全校核

在智能分词中已经将电力术语解析成了设备、动作和状态变化等关键信息。在本模块中所实现的就是将这些信息所代表的电网变化情况,采用现有的安全校核的功能进行校核,并将校核的结果返回到调用模块或系统,提醒校核结果信息。

其中,本系统中所使用的是可扩展式的安全校核,即除了现有可以调用的安全校核外,随时可以对校核功能进行扩展。目前,在系统中,校

(下转第 26 页)

3.2 优化成效验证

为了验证提出的采集系统优化方案的应用效果,通过在数据采集系统中加入数据筛选与处理模块,实现了上述数据筛查和修正功能。对系统实现优化后 2 个月的共计 155 991 条数据进行分析,可得到以下结论:

1) 可用于后续分析的正常数据条数为 137 116,即数据的准确度为 87.90%,与之前没使用数据处理模块相比,准确度提升了 22.75%。

2) 经过数据处理模块处理后的数据,所有类型的错误数据均有所减少:部分日期异常数据得到修正,修复数据 3.87%;经过运维人员现场消缺排查,充电交易金额、电量及电费异常数据被消除,其他异常数据类型均有所下降。

3) 系统数据处理优化方案的实施,在筛选和修正异常数据、指导现场运维方面都取得了显著成效。

4 结束语

本文针对电动汽车智能充换电运营管理系统的充电交易记录数据质量问题,分析了系统初期存在的异常数据类型和充电站分布情况。基于业务数据具有的数理和逻辑关系,提出了异常数据筛选规则以及修正和标识处理方法,给出了数据采集系统数据质量优化方案。通过在系统中增加数据筛选与处理模块,有效降低了系统异常数据比例,改善后的数据质量优化成效对比结果验证了所提方案的可行性和有效性。

通过运营管理系统数据分析的相关研究和实践发现,提高数据质量一方面需要系统的数据

筛选与处理技术;另一方面还要提高现场设备的数据采集质量和运维管理水平。数据质量分析发现系统和设备问题并给出解决方案,为充电设施运营提供有力支持,并为大数据分析工作打下坚实基础。

参考文献

- [1] SONG Y H, YANG X, LU Z X, et al. Integration of plug-in hybrid and electric vehicles: experience from China [C]. Minneapolis: Proceedings of Power & Engineering Society General Meeting, 2010, July 25-29, 1-5.
- [2] 陈科,王震坡,林程.北京市纯电动汽车的发展与示范运营[J].新材料产业,2006,9:27-30.
- [3] 潘鸣宇,贺家胜,迟忠君,等.电动汽车智能服务平台的设计与实现[J].电力信息与通信技术,2014,12(1):78-81.
- [4] 卢辉.数据挖掘与数据化运营实战:思路、方法、技巧与应用[M].北京:机械工业出版社,2015.
- [5] 王涓,吴旭鸣,王爱凤.应用皮尔逊相关系数算法查找异常电能表用户[J].电力需求侧管理,2014,16(2):52-54.

作者简介

张禄(1984-),男,博士,工程师,研究方向为电动汽车运营服务及大数据分析技术。

潘鸣宇(1985-),男,高级工程师,研究方向为电动汽车运营管理及信息化技术。

(收稿日期:2015-12-10) EA

(编辑:史海疆)

(上接第 21 页)

核的种类已经包含了如设备名称校核、状态校核、潮流校核和五防校核等一系列的校核方式,而由于校核所系统的接口是统一的,随时可以按照接口规范新增其他的校核方式。

2 结束语

在电网运作过程中,口令操作校核是一个空白区域。本文针对口令操作,研究口令操作的流程,并对该流程中的每一个流程点通过对现有的成熟技术进行研究和借鉴提出实现方式,最终提

出并设计了一种基于语音转换校核的口令操作安全校核系统,并初步进行了可行性的探讨。但是本系统在实际中的应用成果还需要进行验证,各个流程节点的技术实现也需要不断进行完善和改进。

作者简介

叶海峰(1978-),男,硕士,高级工程师,主要从事工作领域:电网运行。

(收稿日期:2015-12-10) EA

(编辑:史海疆)