

Rule-based Algorithm for Role Classification and Text Segmentation of Air Traffic Control Voice Text

Bin Long Peng Zhao Weihua Gao

Chongqing Branch of Southwest Regional Air Traffic Management Administration, Chongqing, 401120, China

Abstract

In response to the prevalent issues of mixed voice recognition, role identification, and segmentation difficulties between air traffic controllers and pilots in air traffic control (ATC) speech recognition applications, this paper proposes a rule-based algorithm for ground-air communication speech role classification and communication text segmentation. The algorithm designs a communication template system that includes universal labels such as aircraft call signs and numerical combinations, as well as logical symbols. On the template side, an abstract syntax tree is used to parse the communication templates. On the communication text side, the recognized text is preprocessed first, and then the role determination and automatic segmentation of mixed voice text are achieved through a sequential matching algorithm. This method effectively solves the problems of role identification and mixed dialogue segmentation between air traffic controllers and pilots, providing a technical approach for semantic understanding, normative analysis, and automatic error correction in ATC speech applications.

Keywords

air traffic control speech recognition; role classification; text segmentation; rules; dialog template; matching; Abstract Syntax Tree

基于规则的空管地空通话语音文本角色分类和文本分割算法

龙滨 赵鹏 高卫华

民航西南地区空中交通管理局重庆分局, 中国·重庆 401120

摘要

针对空管语音识别应用中普遍存在的管制员与飞行员语音混合、角色识别和分割困难的问题, 本文提出一种基于规则的地空通话语音角色分类和通话文本分割算法, 该算法设计了一套包含航空器呼号、数字组合等通用标签及逻辑符号的通话模板体系, 在模板侧采用抽象语法树对通话模板进行解析, 在通话文本侧通过对识别文本先进行预处理, 再通过顺序匹配算法实现对混合语音文本的角色判定与自动分割。该方法有效地解决了管制员、飞行员语音的角色识别和混合对话分割问题, 为空管语音的语义理解、规范性分析及自动纠错等应用场景提供了一种技术途径。

关键词

空管语音识别; 角色分类; 文本分割; 规则; 通话模板; 匹配; 抽象语法树

1 引言

在空管中, 我们不少场景需要用到管制语音识别, 如管制员通话用语规范性分析、通话错误提醒、字幕生成等等。在空中交通管制领域, 管制员同飞行员的通话通常是通过话音交换系统, 经 VHF 设备进行地空通信。现有的语音交换系统均提供录音接口, 我们可以使用录音接口通道实现地空对话的语音识别。由于现有的内话系统没有提供管制员和飞行员通话单独的语音输出通道, 经语音识别的话音文本是管制员和飞行员的混合对话, 如果不进行对话角色分离和文本切割, 得到的是一大段文字, 难以区分出说话人的角色, 需要人工识别。这样我们面临一个难题: 能否将管制员和飞行

员的通话角色分离出来?

首先考虑采用 AI 训练的方式对识别的文本进行分离, 我们选取 1800 条真实的管制通话文本进行训练, 每条通话文本对应一种角色: 管制员、飞行员和混合, 分别采用 BERT、BERT+CNN、BERT+RNN、BERT+RCNN、BERT+DPCNN 进行训练, 实测效果非常不理想, 采用 BERT 算法, 总准确率只能达到 37% 左右, 也就是比随机猜测高一点, 其他的算法也没有明显提升。分析原因, 一方面 BERT 算法本质上是单纯的根据分词的关联进行多维距离计算, 也就是根据各种分词的频次和距离进行推算, 没考虑上下文, 不易分出混合情况; 另一方面需要足够的样本, 可能需要几万或十几万以上的样本才能到达较高的百分比。获取大规模的真实语料需要大量的人工标注, 人力成本太高, 因此需要考虑另辟蹊径。

由于管制和飞行的话语在语言表达、用词前后顺序和

【作者简介】龙滨 (1976-), 男, 中国重庆人, 硕士, 高级工程师, 从事空中交通管理通信、导航、监视研究。

句式结构上一些明显的特征和差异,绝大部分对话是可以判断出角色的。我们以行标《空中交通无线电通话用语》(MH/T4014-2003)为蓝本,尝试通过建立一套分类规则算法,以实现管制员、飞行员语音的分类和切割。

2 设计文本角色分类和分割算法需要遵循的原则

(1) 如果某句话不确定是谁说的,就不要纳入规则。避免判断角色错误。(2) 设计规则应该有一定容错性,比如:下高度{某高度}保持,在实际通话时,可能说成:下降到{某高度}保持、下到{某高度}保持、下{某高度}等等。(3) 合理处理对于语气词,比如:啊,呃,嗯,哎之类。由于语气词可能出现句前、句中和句末,它的出现会中断模板匹配。建议删除处理。

3 通话模板设计

为实现分类和文本分割,我们需要设计管制员通话模板和飞行员通话模板。在模板中,我们设计了四种模板符号。

3.1 模板标签

根据作用同,我们设计了一些通用模板标签,作为模板占位符:1) <航空器呼号>,这是由航空公司简称+多个数字或英文字母的组合。我们需要对航空公司名称做一个列表进行枚举,如:东方、国航、南方等。航空器呼号由于无法枚举,宜采用匹配方式,这里的数字和英文字母组合包含幺、两、三、四、五、六、拐、八、九、洞以及26个英文字母组成的数字。2) <数字组合>,用于应答机编码、高度、频率、速度、气压(QNH和QNE)、航向、时间等,比如:应答机编码:五拐两幺,表示二次码为5721;高度:拐两,表示7200米;频率:幺两五点拐,表示频率125.7MHz;速度:两九洞,表示速度290公里/小时。中文数字组合应包含文字:幺、两、三、四、五、六、拐、八、九、洞、一、二、七、点、十、百、千、整点。数字组合应该采用算法对连续出现的数字进行识别,并按连续原则,形成一个数字组合。3) <进离场程序>,用于识别进离场程序,应该配置一个进离场程序列表。如:VYK-1A、NHW-2D等。4) <跑道号>,用于识别跑道,应列表枚举,比如:洞两左、两洞左等。5) <固定点>,比如导航台呼号、航路点、定位点、重要点等固定点名称。如:EKIVI、LKO等。应该配置一个本管制区范围内用到的固定点列表。同一个固定点要有三种表示,比如EKIVI,应包含EKIVI,也应包含ECHO KILO INDIA VICTOR INDIA,和EKIVI ECHO KILO INDIA VICTOR INDIA,这是因为实际通话中三种说法都会出现。6) <航路>,航路名称:如G595。应列表枚举。7) <本管制单位呼号>,比如:成都、区域、进近、重庆进近、塔台、地面、放行、机坪等。应列表枚举。8) <移交管制单位呼号>,比如:西安、万州等。同本管制区分,

是为了更好的匹配飞行员的通话模板。应该配置一个列表。9) <进近种类>,如:目视、VOR、盲降、RNP、LOC等,应该配置一个列表。10) <灯光类别>,如:跑道边灯、跑道入口灯、入口翼排灯、接地地带灯、跑道中线灯和跑道末端灯等。应该配置一个列表。11) <机型>,比如重型、中型、轻型等。应该配置一个列表。12) <滑行位置>,如:ZULU三,用语中如:在<滑行位置>{外}等待。有字母呼号+数字组合组成。由算法识别。13) <飞行情报区>,用语中如:由于<飞行情报区>FIR有干扰。应该配置一个列表。14) <滑行路线>,由滑行位置和滑行道号的多个组合形成。通过算法匹配。15) <字母呼号>,可以用于通报代号、滑行道号等。应该配置一个列表。包含26个字母:ALPHA、BRAVO、CHARLIE等。16) <RNAV-RNP程序>,本管制区用到的RNAV和RNP程序,配置一个列。17) <机型>,如:波音七三七八百、空客三二〇等,通常用于活动通报。18) <xxx>表示识别文本中的不确定文本。用于不确定文字组合的匹配。<xxx>应该是语音识别后,通过空格或符号分割出来的其中一小段文字,而不是包含所有不确定文本。

3.2 模板符号

不同于《空中交通无线电通话用语》中的通话符号,我们自行设计几种符号:1) 标签占位符<...>,如<标签名>,即上述采取的标签占位符名称,用于区分特定用途的文本。2) 中括号[],表示文本中一定会出现的文字,通常配合|使用。比如:[加速|减速]到<数字组合>3) 大括号{},可选片段,表示可能会出现的文字。如:<航班号>预计<进离场航线>{进场}。4) 或|,字面枚举,表示多个文字的或关系。如:<航班号>[上升|下降]率<数字组合>,因为{起飞|落地}间隔,[上升|下降]表示上升或下降的文字必须出现一个;{起飞|落地}表示可能出现三种情况之一:空,起飞、落地。上述的符号可以组合嵌套使用。比如:{[上升|下降]率[小于|大于|不小于|不大于]<数字组合>},<航班号>{现在}[左|右]转飞<固定点>{自主领航},<航班号>离开<数字组合>[上|上升|下|下降]{到}{高度}<数字组合>{保持},注意[收听|守听]相似航班号{预计[进离场航线][进场|进港]}<航班号>。

示例:四川八八三两沿CHARLIE向北滑联系重庆塔台幺两八点两再见|管制员。CHARLIE向北幺幺八两四川八八三两再见|飞行员。对于示例,我们需要各设计一个管制员通话模板和一个飞行员通话模板,可以设计如下:管制员通话模板:<航空器呼号>沿<字母呼号>向[北|南|东|西|东北|西北|西南|东南]滑行联系<本管制单位呼号><数字组合>{再见},飞行员通话模板:{沿}{字母呼号}向[北|南|东|西|东北|西北|西南|东南]<航空器呼号>{再见}

4 地空通话模板文本匹配算法

我们先研究已经进行角色分类文本的匹配算法,在此基础上,再研究分割算法。我们需要对模板侧和文本侧进行分别处理。

4.1 模板侧处理

管制员和飞行员的通话模板采用抽象语法树(AST)进行处理,将每调模板解析为AST节点。也就是把整个模块,映射为一颗树。处理时,将节点分为5类:1)顺序节点SEQ,由多个子节点按顺序组成。2)字面量节点LIT,用于直接匹配具体文本。3)枚举节点ALT:在多个备选文本中二择一或多择一(正则实现为分组)。4)可选节点OPT:子节点有可能整体出现也可能不出现。5)占位符节点PLACEHOLDER:也就是上述的<航空器呼号>。所有的模板处理完毕后,进行正则编译并构建为一个模板集合。方便对文本集中处理。

4.2 文本侧处理

而文本侧的处理应先按照以下顺序,先进行预处理,再进行标签提取。主要包含以下几个过程:1)标点归一化。将中英文中常见的标点替换为空格并压缩空白字符。

2)删除语气词:常用的语气词包括:哦 呃 好 啊 喔 嗯 额 哎 对 欸等,这些词由于可能出现在任何地方,会影响匹配,需要进行删除处理。另外,有些常说的单次,如好的,好嘞,谢谢,收到等词,建议也当做语气词进行处理。实测删除并不会影响匹配。3)重复词压缩。对于相邻的重复字和词,进行去重。重复词一般是通话人员口误。4)识别主要标签:①确定航空器呼号(航班号);②每行对话应该先搜索先根据航空公司名称,如果后面有数字或英文字母的组合(不超过4位),则分割出航空器呼号;③识别本场管制区呼号。由于本管制单位呼号,通常不多且较为明显,容易区分;④识别进离场程序。由于进离场航线包含固定点,应该在固定点之前先识别进离场航线,避免被判定是固定点。⑤识别滑行线路。由于滑行线路包含滑行位置,应该在滑行位置之前先识别滑行线路,避免被判定是滑行位置。⑥分步依次识别固定点、跑道号、航路、RNAV-RNP程序、滑行位置、进近种类、灯光类别、机型、飞行情报区、字母呼号。⑦在上述识别的过程中,建立设置一个对比模板,在识别的时候,用上述识别出的标签替换原文本。如:海南拐五六拐成都雷达已识别 下降到拐两保持 下降率大于两千,替换为:<航班号> <管制区> 雷达已识别 上升到<数字组合>保持 下降率大于<数字组合>。经过上述处理,转换的对比模板,同模板有了一定相似性。

3、匹配处理:对已处理文本,对已经编译模板的进行

完整匹配,采用正则式判定是否符合模板结构。对于满足管制员的对话,我们可以输出为1,对于满足飞行员的对话,我们可以输出为4,如果不匹配,我们可以输出为0。

5 地空通话模板文本分割算法

上面是对某句已知航空通话对象的对话模板进行的匹配,在我们实际使用时,接收到的语音文本通常是混合的,比如我们从语音记录仪提取的某个时间段的数据,在设置一定的静默时间(如150ms)的情况下,实测通过语音分割,仍然会存在语音混合的情况。此时我们需要对这种混合语音进行分割,切分成管制员和飞行员的分别对话。数据分割包括文本预处理,和模板逐一匹配。

文本预处理:混合文本的预处理同上述匹配算法文本侧的预处理类似,也包括标点归一化、删除语气词、重复词压缩。

顺序文本匹配和切割:从对话语句的最开头进行文本匹配。当匹配到一个模板时,将文本进行切割。然后对剩下的文本进行再次模板匹配,并循环进行切割。按这种方式切割,仍有可能存在一个文本可以匹配多个模板的情况,导致切割就会有多重可能,此时须逐一尝试多种组合,直到找到一种能完全匹配的组合。

6 改进方向

基于本文算法,我们更希望在设置合理模板模版以避免多种组合的发生。这就需要我们在新模板输入时,将新模板同已有模板进行比对分析,一方面新模板不能命中到已有模板,另一方面,已有模板也不能命中到新模板。

参考文献

- [1] 郑诚,唐思琪.结合标签语义和对比学习的小样本文本分类[J/OL].小型微型计算机系统,1-14[2025-12-05].<https://doi.org/10.20009/j.cnki.21-1106/TP.2025-0319>.
- [2] 赵海燕,曹杰,陈庆奎,等.多层次多标签文本分类方法[J].小型微型计算机系统,2022,43(04):673-683.DOI:10.20009/j.cnki.21-1106/TP.2021-0658.
- [3] 崔海越,李延玲.基于BERT模型的多标签文本分类[J].高师理科学刊,2025,45(08):18-24.
- [4] 孙晓瑜.基于机器学习的文本分类与标签预测算法[J].吉林大学学报(信息科学版),2025,43(04):837-843.DOI:10.19292/j.cnki.jdxp.2025.04.010.
- [5] 贾桂敏,申玉茜,周熹龙.基于层级规则树与特征融合的陆空通话规范性检验模型[J/OL].北京航空航天大学学报,1-15[2025-12-05].<https://doi.org/10.13700/j.bh.1001-5965.2025.0041>.