



(12) 发明专利

(10) 授权公告号 CN 109165275 B

(45) 授权公告日 2021.03.02

(21) 申请号 201810818470.7

(22) 申请日 2018.07.24

(65) 同一申请的已公布的文献号
申请公布号 CN 109165275 A

(43) 申请公布日 2019.01.08

(73) 专利权人 国网浙江省电力有限公司电力科学
研究院

地址 310014 浙江省杭州市下城区朝晖八
区华电弄1号

专利权人 国家电网有限公司
杭州沃瑞电力科技有限公司

(72) 发明人 罗华峰 陆承宇 宣晓华 黄晓明
杜奇伟 丁峰 阮黎翔 王松
吴栋萁 杨涛 柯人观 汪冬辉
方芳 孙文文 戚宣威 孙志达
陈明 鲍威 史立勤

(74) 专利代理机构 浙江翔隆专利事务所(普通
合伙) 33206

代理人 张建青

(51) Int.Cl.
G06F 16/33 (2019.01)
G06F 40/279 (2020.01)
G06F 40/289 (2020.01)
G06N 3/04 (2006.01)

(56) 对比文件
CN 108021552 A, 2018.05.11
CN 105224511 A, 2016.01.06
CN 107153946 A, 2017.09.12
US 5630020 A, 1997.05.13
CN 105975456 A, 2016.09.28

审查员 王思文

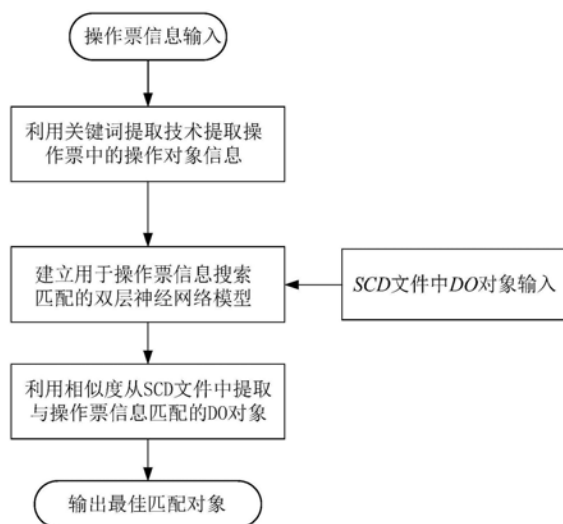
权利要求书1页 说明书4页 附图2页

(54) 发明名称

基于深度学习的智能变电站操作票信息智能搜索匹配方法

(57) 摘要

本发明公开了一种基于深度学习的智能变电站操作票信息智能搜索匹配方法。类似智能站监控信息调度联调、站内程序化操作的验证等工作,当前只能人工逐条进行,需要耗费大量的人力和时间,极大影响工程进度。本发明采用的技术方案包括:利用关键词提取技术提取操作票中的操作对象信息;建立用于操作票信息智能搜索匹配的双层神经网络模型;利用相似度计算和排序从SCD文件中提取与操作票信息匹配的DO对象。本发明能够基于自然语言处理相关技术,建立起操作票信息与SCD文件之间的关联,解决变电站操作票信息的智能识别和搜索匹配,进而服务于完整高效的变电站智能调试工作。



1. 基于深度学习的智能变电站操作票信息智能搜索匹配方法, 其特征在于, 包括以下步骤:

S1, 利用关键词提取技术提取操作票中的操作对象信息;

S2, 建立用于操作票信息智能搜索匹配的双层神经网络模型;

S3, 利用相似度计算和排序从SCD文件中提取与操作票信息匹配的DO对象;

步骤S1中, 关键词提取技术利用基于BP神经网络模型的由字构词分词法和根据分词结果利用循环神经网络模型实现操作票中操作对象的识别;

步骤S1中, 基于BP神经网络模型的由字构词分词法进行分词, 其过程如下:

选取操作票每个字前后各K个字, 连同该字本身输入共 $2*K+1$ 个字; 选择每个字的向量维度为L, 则输入神经网络模型的维度 $L*(2*K+1)$; 采用三层BP神经网络, 输入层为 $L*(2*K+1)$ 个神经元, 隐藏层为L个神经元, 输出层为4个神经元, 每个神经元对应该字的词位分别为S, B, M, E的概率, S, B, M, E分别对应字单独成词、词首、词中、词尾; 从输出词位概率值中选择概率最大的词位作为该字的词位类别;

步骤S1中, 根据分词结果利用循环神经网络模型实现操作票中操作对象的识别, 其过程如下:

将所得分词序列依次转换为词向量, 选择词的向量维度为M; 设定足够长的词数N, 不足的后面补空, 则循环神经网络模型的输入为 $M*N$ 的向量, 对应循环神经网络模型的输入层; 第二层为双LSTM; 输出层输出2维的分类情况, 输出的分类结果为0和1, 代表该词是否为操作对象, 0表示否, 1表示是;

步骤S2中, 搜索匹配采用的规则为: 判断两条信息所属的组别是否相同, 如果两条信息的组别不同, 则认定两条信息不匹配; 如果两条信息属于同一组别, 再判断两条信息的具体设备信息是否相同, 给出最终的匹配结果;

步骤S2中, 基于搜索匹配规则, 采用双层神经网络模型, 第一层区分组内和组外项: 组外为'0', 组内为'1'; 第二层区分匹配项: 组内不匹配为'0', 组内匹配为'1'; 其中每层神经网络模型采用深度BP神经网络或深度循环神经网络, 模型的输出为操作票信息与SCD文件中DO对象是否匹配的结果。

2. 根据权利要求1所述的基于深度学习的智能变电站操作票信息智能搜索匹配方法, 其特征在于, 根据每层采用神经网络的不同, 双层神经网络共有四种结构, 分别为: BP_1、BP_2, 表示模型的第一、二层均采用BP模型; BP_1、LSTM_2, 表示模型的第一层采用BP模型, 第二层采用LSTM模型; LSTM_1、BP_2, 表示模型的第一层采用LSTM模型, 第二层采用BP模型; LSTM_1、LSTM_2, 表示模型的第一、二层均采用LSTM模型。

3. 根据权利要求2所述的基于深度学习的智能变电站操作票信息智能搜索匹配方法, 其特征在于, 步骤S3中, 根据步骤S2中神经网络的输出结果是'0'、'1'的概率, 遍历判断操作票信息与SCD文件中所有DO对象是否匹配; 其中'1'的概率即为相似度或匹配度, 其大小表示操作票信息与SCD信息的匹配程度, 由此选择匹配度最大的SCD文件中的DO对象作为某条操作票信息的最佳匹配结果。

基于深度学习的智能变电站操作票信息智能搜索匹配方法

技术领域

[0001] 本发明涉及人工智能、电网智能变电领域,具体地说是一种基于深度学习的智能变电站操作票信息智能搜索匹配方法。

背景技术

[0002] 随着智能电网的发展,信息网络规模不断增大,不同应用系统之间的信息交换大大增加,信息环节对物理系统的影响问题将更加突出。电力系统应用集成是减少工作量、提高电能质量、提高电力企业经济效益的必要手段,但是电力企业在信息化过程中基本以业务部门为依托进行条块分割,各个应用系统相互独立且高度自治,导致电网建模和数据名称不一致问题严重,信息匹配已经是应用集成急需解决的问题。

[0003] 问题具体表现为:为了满足各类应用的需求,在智能站工程调试阶段需要开展大量工作来验证信息是否匹配。类似智能站监控信息调度联调、站内程序化操作的验证等工作,当前只能人工逐条进行,需要耗费大量的人力和时间,极大影响工程进度。

发明内容

[0004] 本发明所要解决的技术问题是克服上述现有技术存在的缺陷,提供一种基于深度学习的智能变电站操作票信息智能搜索匹配方法,实现自动化作业,以替代往常的人工作业模式,为实际工程生产和调试中的工作提供技术支撑。

[0005] 为解决上述技术问题,本发明采用的技术方案是:基于深度学习的智能变电站操作票信息智能搜索匹配方法,其包括以下步骤:

[0006] S1,利用关键词提取技术提取操作票中的操作对象信息;

[0007] S2,建立用于操作票信息智能搜索匹配的双层神经网络模型;

[0008] S3,利用相似度计算和排序从SCD文件中提取与操作票信息匹配的DO对象。

[0009] 作为上述技术方案的补充,步骤S1中,关键词提取技术利用基于BP神经网络模型的由字构词分词法和根据分词结果利用循环神经网络模型实现操作票中操作对象的识别。

[0010] 作为上述技术方案的补充,步骤S1中,基于BP神经网络模型的由字构词分词法进行分词,其过程如下:

[0011] 选取操作票每个字前后各K个字,连同该字本身输入共 $2*K+1$ 个字;选择每个字的向量维度为L,则输入神经网络模型的维度 $L*(2*K+1)$;采用三层BP神经网络,输入层为 $L*(2*K+1)$ 个神经元,隐藏层为L个神经元,输出层为4个神经元,每个神经元对应该字的词位分别为S,B,M,E的概率,S,B,M,E分别对应字单独成词、词首、词中、词尾;从输出词位概率值中选择概率最大的词位作为该字的词位类别。

[0012] 为了利用操作票每个字的上下文信息,本发明采用基于知识进行分词方法中的BP神经网络进行分词。该方法对操作票每个字的上下文信息记忆能力比较好,拥有不错的容错效果;同时该方法采用非线性并行处理方法,并且在分词中可以自动处理和学习得到的知识和信息,解决了操作票信息较多带来的速度问题。

[0013] 作为上述技术方案的补充,步骤S1中,根据分词结果利用循环神经网络模型实现操作票中操作对象的识别,其过程如下:

[0014] 将所得分词序列依次转换为词向量,选择词的向量维度为M;设定足够长的词数N,不足的后面补空,则循环神经网络模型的输入为M*N的向量,对应循环神经网络模型的输入层;第二层为双LSTM(Long Short Term Memory Network,长短时记忆网络);输出层输出2维的分类情况,输出的分类结果为0和1,代表该词是否为操作对象,0表示否,1表示是。

[0015] 为了解决操作票中每个词的时序敏感及操作票内容长度不确定问题,本发明采用了循环神经网络(RNN)模型。

[0016] 作为上述技术方案的补充,步骤S2中,搜索匹配采用的规则为:判断两条信息所属的组别是否相同,如果两条信息的组别不同,则认定两条信息不匹配;如果两条信息属于同一组别,再判断两条信息的具体设备信息是否相同,给出最终的匹配结果。

[0017] 由于操作票中线路或设备信息的通常格式为“间隔+具体设备信息”。相比不同间隔,同一间隔下设备的信息之间的相似程度显然要高。因此,本发明将间隔抽象为组别。

[0018] 作为上述技术方案的补充,步骤S2中,基于搜索匹配规则,采用双层神经网络模型,第一层区分组内和组外项:组外为‘0’,组内为‘1’;第二层区分匹配项:组内不匹配为‘0’,组内匹配为‘1’;其中每层神经网络模型采用深度BP神经网络或深度循环神经网络,模型的输出为操作票信息与SCD文件中D0对象是否匹配的结果。

[0019] 作为上述技术方案的补充,根据每层采用神经网络的不同,双层神经网络共有四种结构,分别为:BP_1、BP_2,表示模型的第一、二层均采用BP模型;BP_1、LSTM_2,表示模型的第一层采用BP模型,第二层采用LSTM模型;LSTM_1、BP_2,表示模型的第一层采用LSTM模型,第二层采用BP模型;LSTM_1、LSTM_2,表示模型的第一、二层均采用LSTM模型。

[0020] 作为上述技术方案的补充,步骤S3中,根据步骤S2中神经网络的输出结果是‘0’、‘1’的概率,遍历判断操作票信息与SCD文件中所有D0对象是否匹配;其中‘1’的概率即为相似度或匹配度,其大小表示操作票信息与SCD信息的匹配程度,由此选择匹配度最大的SCD文件中的D0对象作为某条操作票信息的最佳匹配结果。

[0021] 与现有技术相比,本发明具有的有益效果是:

[0022] (1) 利用BP神经网络实现了中文分词。

[0023] (2) 利用循环神经网络完成了操作票关键字提取。

[0024] (3) 利用双层深度学习模型(即双层神经网络模型)实现了操作票信息的搜索匹配。

[0025] 本发明能够基于自然语言处理相关技术,建立起操作票信息与SCD文件之间的关联,解决变电站操作票信息的智能识别和搜索匹配,进而服务于完整高效的变电站智能调试工作。

附图说明

[0026] 图1为本发明基于深度学习的智能变电站操作票信息智能搜索匹配方法的流程图;

[0027] 图2为本发明操作票信息与SCD文件中的D0对象搜索匹配流程图。

具体实施方式

[0028] 附图仅用于示例性说明,不能理解为对本发明的限制;对于本领域技术人员来说,附图中某些公知结构及其说明可能省略是可以理解的。附图中描述位置关系仅用于示例性说明,不能理解为对本发明的限制。

[0029] 下面将结合本发明中的附图,对本发明技术方案进行详细介绍。本发明技术方案的实施主要包括以下步骤(流程图见附图1):

[0030] S1,利用关键词提取技术提取操作票中的操作对象信息

[0031] (1)利用BP神经网络模型进行中文分词。

[0032] 取操作票中每个字前后各3个字,前后3个字连同该字本身输入共7个字。选择每个字的向量维度为100,则输入神经网络模型的维度 $100*(2*3+1)=700$ 。采用三层BP神经网络,输入层为700个神经元,隐藏层为100个神经元,输出层为4个神经元,每个神经元对应该字的词位分别为S,B,M,E(S,B,M,E分别为字单独成词(S)、词首(B)、词中(M)、词尾(E))的概率。

[0033] 从输出词位概率值中选择概率最大的词位作为该字的词位类别,对操作票的每个字进行依次进行同样处理便可以完成对整个操作票的分词。例如操作票“合上仙永线50121闸刀”经过分词后得到“合上仙永线50121闸刀”。

[0034] (2)根据分词结果利用循环神经网络模型实现操作票中操作对象信息识别。

[0035] 将步骤(1)所得分词序列依次转换为词向量,选择词的向量维度为10。足够长的词数30(不足的后面补空),作为固定输入的维度,则神经网络模型的输入为 $30*10$ 的向量,对应循环神经网络模型的输入层;第二层为双LSTM,隐藏层为128个神经元;输出层输出以0,1表示的2维分类情况(0和1分别对应该词是否为设备或线路,1表示是,0表示否)。

[0036] 经过操作后,从操作票分词结果中提取了以“仙永线50122闸刀”表示的线路及设备信息。

[0037] S2,建立用于操作票信息智能搜索匹配的双层神经网络模型;

[0038] (1)输入SCD文件中的DO对象,作为信息1。

[0039] (2)确定搜索匹配规则。将操作票提取后的信息作为信息2。两条信息的匹配规则为:首先判断两条信息所属的组别是否相同,如果两条信息的组别不同,则认定两条信息不匹配;如果两条信息属于同一组别,再判断两条信息的具体设备信息是否相同,给出最终的匹配结果(搜索规则流程图见附图2)。

[0040] 以操作票指令为“合上仙永线50121闸刀”例,搜索匹配结果即为表1所示。

[0041] 表1操作票信息搜索匹配SCD示例

操作票指令	关键字提取	SCD 匹配 DO 对象	关键字匹配度	匹配度
[0042] 合上仙永线 50121 闸刀	仙永线 50121 闸刀	仙永线 5012 开 关. 50121 闸刀合位	0.97300804	1
		仙永线 5013 开 关. 50131 闸刀合位	0.72973388	1

[0043] S3,利用相似度计算和排序从SCD文件中提取与操作票信息匹配的DO对象

[0044] 经过匹配度计算结果排序比较得出,搜索匹配的最佳结果为“仙永线5012开关.50121闸刀合闸”。

[0045] 显然,本发明的上述实施例仅仅是为清楚地说明本发明所作的举例,而并非是对本发明的实施方式的限定。对于所属领域的普通技术人员来说,在上述说明的基础上还可以做出其它不同形式的变化或变动。这里无需也无法对所有的实施方式予以穷举。凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等,均应包含在本发明权利要求的保护范围之内。

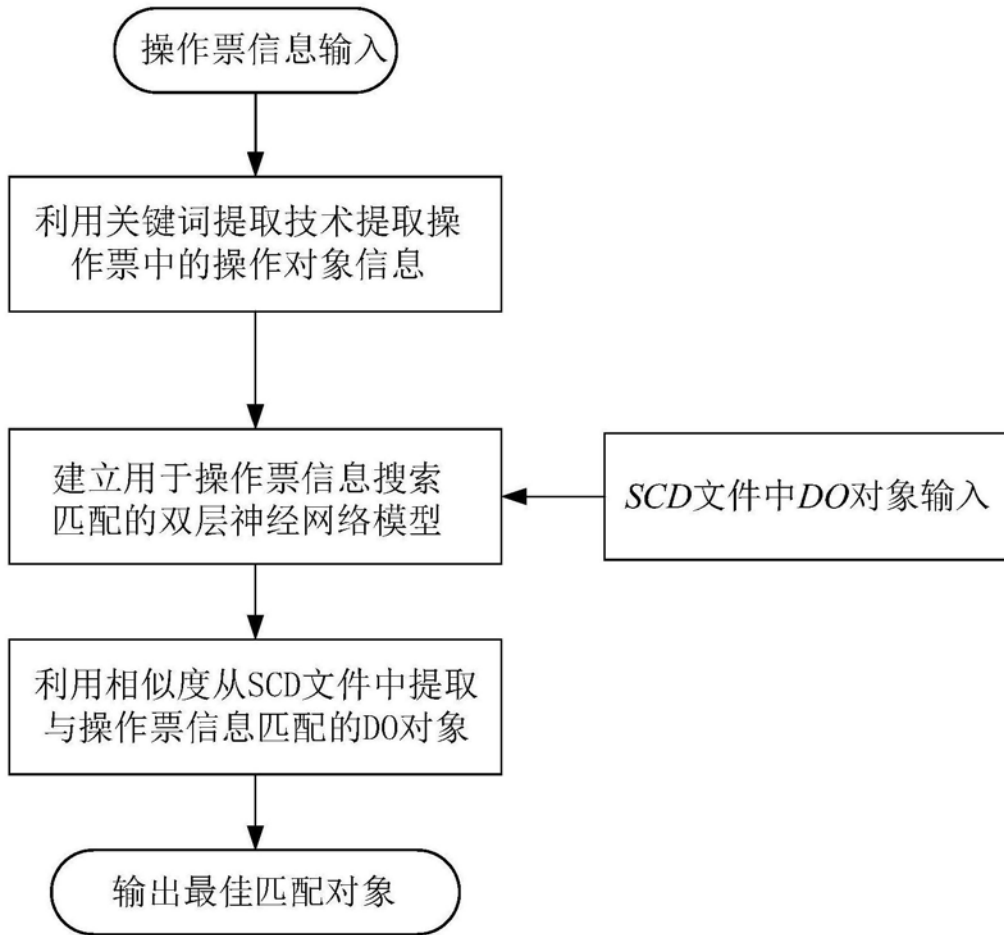


图1

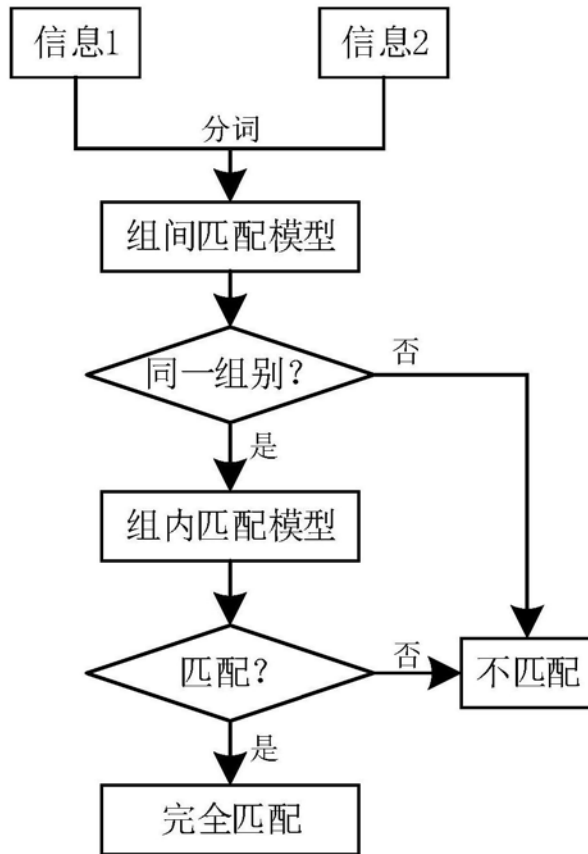


图2