



(12)发明专利申请

(10)申请公布号 CN 110968703 A

(43)申请公布日 2020.04.07

(21)申请号 201911238633.5

(51)Int.Cl.

(22)申请日 2019.12.06

G06F 16/36(2019.01)

G06N 3/04(2006.01)

(71)申请人 国网天津市电力公司电力科学研究院

G06Q 50/06(2012.01)

地址 300384 天津市西青区海泰华科四路8号

申请人 国网天津市电力公司
国家电网有限公司

(72)发明人 刘浩宇 李野 李刚 吕伟嘉
张兆杰 卢静雅 翟术然 乔亚男
陈娟 许迪 赵紫敬 董得龙
孙虹 杨光 季浩 何泽昊 顾强
赵宝国 许小亮

(74)专利代理机构 天津盛理知识产权代理有限公司 12209

代理人 王来佳

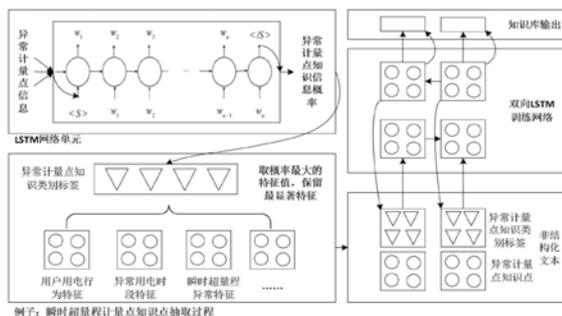
权利要求书2页 说明书7页 附图1页

(54)发明名称

基于LSTM端到端抽取算法的异常计量点知识库构建方法及系统

(57)摘要

本发明涉及一种基于LSTM端到端抽取算法的异常计量点知识库构建方法,其特征在于:包括如下步骤:(1)在用电信息采集系统抽取异常计量点信息;(2)基于异常计量点信息,通过LSTM网络单元模型计算异常计量点知识信息概率,形成多维知识标签;(3)基于双向LSTM训练网络对所形成的多维知识标签进行抽取,针对概率最大的知识特征值,建立异常计量点知识识别标签,输出多维异常计量点知识库。本发明通过知识库的匹配,识别计量异常,从而甄别出问题电能表,以便对电能表的维护、服役年限等做出准确指导。



1. 一种基于LSTM端到端抽取算法的异常计量点知识库构建方法,其特征在于:包括如下步骤:

(1) 在用电信息采集系统抽取异常计量点信息;

(2) 基于异常计量点信息,通过LSTM网络单元模型计算异常计量点知识信息概率,形成多维知识标签;

(3) 基于双向LSTM训练网络对所形成的多维知识标签进行抽取,针对概率最大的知识特征值,建立异常计量点知识识别标签,输出多维异常计量点知识库。

2. 根据权利要求1所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建方法,其特征在于:所述的异常计量点信息包括瞬时超量程计量点信息、稳定超量程计量点信息、疑似窃电计量点信息、接线错误计量点信息、采集异常计量点信息、时钟错误计量点信息、互感器过载计量点信息、户变关系错误计量点信息、互感器轻载计量点信息以及采集器异常信息。

3. 根据权利要求2所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建方法,其特征在于:所述的瞬时超量程计量点信息包括现场核查信息及用电用户档案信息,所述的稳定超量程计量点信息包括现场核查信息以及用电用户档案信息,所述的疑似窃电计量点信息包括现场的核查信息、运行误差计算信息、电压电流信息及功率因数信息,所述的接线错误计量点信息包括潮流反向、相序异常、反向电量等信息,所述的采集异常计量点信息包括日冻结数据、异常事件信息、高频采集96点电压电流数据信息,所述的时钟错误计量点信息包括日冻结数据、异常事件信息,所述的互感器过载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的户变关系错误计量点信息包括负荷环境、用电曲线、地理位置等信息,所述的互感器轻载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的采集器异常信息包括日冻结数据、96点电压电流数据信息、电表档案信息。

4. 根据权利要求1所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建方法,其特征在于:所述LSTM网络单元模型为

$$LP(S) = -\sum_{i=1}^n \log P(w_j)$$

其中, $P(w_j)$ 表示第j个知识点的预测概率。

5. 根据权利要求1所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建方法,其特征在于:所述多维异常计量点知识库包括瞬时超量程计量点知识库、稳定超量程计量点知识库、疑似窃电计量点知识库、接线错误计量点知识库、采集异常计量点知识库、时钟错误计量点知识库、互感器过载计量点知识库、户变关系错误计量点知识库、互感器轻载计量点知识库以及采集器异常知识库。

6. 一种基于LSTM端到端抽取算法的异常计量点知识库构建系统,其特征在于:包括异常计量点信息抽取模块,用于在用电信息采集系统中抽取异常计量点信息;异常计量点知识信息概率计算模块,用于基于异常计量点信息,通过LSTM网络单元模型计算异常计量点知识信息概率,形成多维知识标签;

多维异常计量点知识库构建模块,用于基于双向LSTM训练网络对所形成的多维知识标

签进行抽取,针对概率最大的知识特征值,建立异常计量点知识识别标签,输出多维异常计量点知识库。

7. 根据权利要求6所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建系统,其特征在于:所述的异常计量点信息包括瞬时超量程计量点信息、稳定超量程计量点信息、疑似窃电计量点信息、接线错误计量点信息、采集异常计量点信息、时钟错误计量点信息、互感器过载计量点信息、户变关系错误计量点信息、互感器轻载计量点信息以及采集器异常信息。

8. 根据权利要求7所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建系统,其特征在于:所述的瞬时超量程计量点信息包括现场核查信息及用电用户档案信息,所述的稳定超量程计量点信息包括现场核查信息以及用电用户档案信息,所述的疑似窃电计量点信息包括现场的核查信息、运行误差计算信息、电压电流信息及功率因数信息,所述的接线错误计量点信息包括潮流反向、相序异常、反向电量等信息,所述的采集异常计量点信息包括日冻结数据、异常事件信息、高频采集96点电压电流数据信息,所述的时钟错误计量点信息包括日冻结数据、异常事件信息,所述的互感器过载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的户变关系错误计量点信息包括负荷环境、用电曲线、地理位置等信息,所述的互感器轻载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的采集器异常信息包括日冻结数据、96点电压电流数据信息、电表档案信息。

9. 根据权利要求6所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建系统,其特征在于:所述LSTM网络单元模型为

$$LP(S) = -\sum_{j=1}^n \log P(w_j)$$

其中, $P(w_j)$ 表示第j个知识点的预测概率。

10. 根据权利要求6所述的一种基于LSTM端到端抽取算法的异常计量点知识库构建系统,其特征在于:所述多维异常计量点知识库包括瞬时超量程计量点知识库、稳定超量程计量点知识库、疑似窃电计量点知识库、接线错误计量点知识库、采集异常计量点知识库、时钟错误计量点知识库、互感器过载计量点知识库、户变关系错误计量点知识库、互感器轻载计量点知识库以及采集器异常知识库。

基于LSTM端到端抽取算法的异常计量点知识库构建方法及系统

技术领域

[0001] 本发明属于电能表计量误差分析领域,特别是一种基于LSTM端到端抽取算法的异常计量点知识库构建方法及系统。

背景技术

[0002] 2009年以来,国家电网公司大力建设用电信息采集系统,目前已实现全域4.5亿只电表运行。系统经过多年运营,积累了海量的用电数据。通过数据分析,挖掘出有效的用电信息,如电能表运行误差,用户的用电行为模式等,既可以开发出海量数据的潜力,又可以大幅降低运营成本,为电网公司提供决策支持。

[0003] 然而,用电信息采集系统所采集的海量数据中,是从现实世界中采集到的大量的各种各样的数据,原始数据质量受到多样性、不确定性和复杂性的影响,使得采集到的实际数据比较凌乱,存在着缺失、异常等现象,很多情况下不符合数据挖掘工具进行知识获取的规范要求。传统的知识库三元组抽取模型受限于样本规模,在长尾关系上效果难以满足实际应用的需求以及低压台区管理精细化的要求。

[0004] 所以需要对相关异常的情况建立知识库,以便对电能表的维护、服役年限等做出指导,以提供干净、简洁、准确的数据,使挖掘过程更有效、更容易,提高挖掘效率和准确性。

发明内容

[0005] 本发明的目的在于克服现有技术的不足,提供一种基于LSTM端到端抽取算法的异常计量点知识库构建方法及系统。

[0006] 本发明解决其技术问题是采取以下技术方案实现的:

[0007] 一种基于LSTM端到端抽取算法的异常计量点知识库构建方法,包括如下步骤:

[0008] (1) 在用电信息采集系统抽取异常计量点信息;

[0009] (2) 基于异常计量点信息,通过LSTM网络单元模型计算异常计量点知识信息概率,形成多维知识标签;

[0010] (3) 基于双向LSTM训练网络对所形成的多维知识标签进行抽取,针对概率最大的知识特征值,建立异常计量点知识识别标签,输出多维异常计量点知识库。

[0011] 而且,所述的异常计量点信息包括瞬时超量程计量点信息、稳定超量程计量点信息、疑似窃电计量点信息、接线错误计量点信息、采集异常计量点信息、时钟错误计量点信息、互感器过载计量点信息、户变关系错误计量点信息、互感器轻载计量点信息以及采集器异常信息。

[0012] 而且,所述的瞬时超量程计量点信息包括现场核查信息及用电用户档案信息,所述的稳定超量程计量点信息包括现场核查信息以及用电用户档案信息,所述的疑似窃电计量点信息包括现场的核查信息、运行误差计算信息、电压电流信息及功率因数信息,所述的接线错误计量点信息包括潮流反向、相序异常、反向电量等信息,所述的采集异常计量点信

息包括日冻结数据、异常事件信息、高频采集96点电压电流数据信息,所述的时钟错误计量点信息包括日冻结数据、异常事件信息,所述的互感器过载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的户变关系错误计量点信息包括负荷环境、用电曲线、地理位置等信息,所述的互感器轻载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的采集器异常信息包括日冻结数据、96点电压电流数据信息、电表档案信息。

[0013] 而且,所述LSTM网络单元模型为

$$[0014] \quad LP(S) = -\sum_{i=1}^n \log P(w_j)$$

[0015] 其中, $P(w_j)$ 表示第j个知识点的预测概率。

[0016] 而且,所述多维异常计量点知识库包括瞬时超量程计量点知识库、稳定超量程计量点知识库、疑似窃电计量点知识库、接线错误计量点知识库、采集异常计量点知识库、时钟错误计量点知识库、互感器过载计量点知识库、户变关系错误计量点知识库、互感器轻载计量点知识库以及采集器异常知识库。

[0017] 一种基于LSTM端到端抽取算法的异常计量点知识库构建系统,包括

[0018] 异常计量点信息抽取模块,用于在用电信息采集系统中抽取异常计量点信息;

[0019] 异常计量点知识信息概率计算模块,用于基于异常计量点信息,通过LSTM网络单元模型计算异常计量点知识信息概率,形成多维知识标签;

[0020] 多维异常计量点知识库构建模块,用于基于双向LSTM训练网络对所形成的多维知识标签进行抽取,针对概率最大的知识特征值,建立异常计量点知识识别标签,输出多维异常计量点知识库。

[0021] 而且,所述的异常计量点信息包括瞬时超量程计量点信息、稳定超量程计量点信息、疑似窃电计量点信息、接线错误计量点信息、采集异常计量点信息、时钟错误计量点信息、互感器过载计量点信息、户变关系错误计量点信息、互感器轻载计量点信息以及采集器异常信息。

[0022] 而且,所述的瞬时超量程计量点信息包括现场核查信息及用电用户档案信息,所述的稳定超量程计量点信息包括现场核查信息以及用电用户档案信息,所述的疑似窃电计量点信息包括现场的核查信息、运行误差计算信息、电压电流信息及功率因数信息,所述的接线错误计量点信息包括潮流反向、相序异常、反向电量等信息,所述的采集异常计量点信息包括日冻结数据、异常事件信息、高频采集96点电压电流数据信息,所述的时钟错误计量点信息包括日冻结数据、异常事件信息,所述的互感器过载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的户变关系错误计量点信息包括负荷环境、用电曲线、地理位置等信息,所述的互感器轻载计量点信息包括96点电压电流数据及日冻结数据、电能表规格数据,所述的采集器异常信息包括日冻结数据、96点电压电流数据信息、电表档案信息。

[0023] 而且,所述LSTM网络单元模型为

$$[0024] \quad LP(S) = -\sum_{i=1}^n \log P(w_j)$$

[0025] 其中, $P(w_j)$ 表示第j个知识点的预测概率。

[0026] 而且,所述多维异常计量点知识库包括瞬时超量程计量点知识库、稳定超量程计量点知识库、疑似窃电计量点知识库、接线错误计量点知识库、采集异常计量点知识库、时钟错误计量点知识库、互感器过载计量点知识库、户变关系错误计量点知识库、互感器轻载计量点知识库以及采集器异常知识库。

[0027] 本发明的优点和积极效果是:

[0028] 本基于LSTM端到端抽取算法的异常计量点知识库构建系统,通过对传统的三元组知识抽模型进行改进,基于LSTM端到端抽取算法将长尾关系的异常计量信息进行精准提取,有效利用头部关系的丰富样本中所包含的信息,形成瞬时超量程、稳定超量程、疑似窃电、接线错误、采集异常、时钟错误、互感器过载、户变关系错误、互感器轻载、采集器异常等十个方面的知识库,有效识别计算电能表计量点的异常,助力对电能表的维护、服役年限等做出准确指导。

附图说明

[0029] 图1是本发明的原理方框图;

[0030] 图2是本发明LSTM知识概率计算模型的示意图(其中“<S>”和“</S>”分别表示语句的开始和结束)。

具体实施方式

[0031] 以下结合附图对本发明的实施例做进一步详述:

[0032] 步骤一:异常计量点信息收集

[0033] 在用电信息采集系统抽取异常计量点信息。

| 系统来源 | 异常计量点类别 | 信息类别 |
|--------------------|-------------|--------------------------------|
| [0034] 用电信息采集系统 | 瞬时超量程计量点信息 | 现场核查信息、用电用户档案信息 |
| | 稳定超量程计量点信息 | 现场核查信息、用电用户档案信息 |
| | 疑似窃电计量点信息 | 现场的核查信息、运行误差计算信息、电压电流信息、功率因数信息 |
| | 接线错误计量点信息 | 潮流反向、相序异常、反向电量等信息 |
| | 采集异常计量点信息 | 日冻结数据、异常事件信息、高频采集96点电压电流数据信息 |
| | 时钟错误计量点信息 | 日冻结数据、异常事件信息 |
| | 互感器过载计量点信息 | 96点电压电流数据及日冻结数据、电表规格数据 |
| | 户变关系错误计量点信息 | 负荷环境、用电曲线、地理位置等信息 |
| | 互感器轻载计量点信息 | 96点电压电流数据及日冻结数据、电表规格数据 |
| | 采集器异常信息 | 日冻结数据、96点电压电流数据信息、电表档案信息 |

[0035] 步骤二:基于异常计量点信息,通过LSTM网络单元模型计算异常计量点知识信息概率,形成多维知识标签。

[0036] LSTM语言模型使用之前的词序列来预测当前异常计量点知识的概率。将输入的异常计量点信息映射到固定维度的预训练的词向量后(“_NE_”的词向量初始为零向量,但训练过程中对词向量也进行更新),输入LSTM模型中。将每个LSTM单元的输出送入softmax分类器中,最后输出当前位置词表中每个信息的知识点分布:

$$[0037] \quad LP(S) = -\sum_{j=1}^n \log P(w_j)$$

[0038] 其中, $P(w_j)$ 表示问句中第j个知识点的预测概率。 $LP(S)$ 越小,表示语句S符合要求的概率越高。

[0039] 更进一步的,针对概率最大的知识特征值,建立异常计量点知识识别标签,输出多

维异常计量知识点。

[0040] 步骤三:基于双向LSTM训练网络进行知识标签抽取,形成多维异常计量点知识库

[0041] 通过双向LSTM训练网络对长尾关系的异常计量点进行知识抽取,有效利用头部关系的丰富样本中所包含的信息,从非结构化的文本中抽取三元组,利用弱监督标注的思路为每个关系自动构造标注数据,形成异常计量点知识库。

[0042] 1. 构建瞬时超量程计量点知识库

[0043] 基于电能表瞬时超量程异常诊断模型、工单反馈的现场核查信息、用电用户档案信息建立瞬时超量程专家知识库,根据诊断模型的计算数据以及瞬时超量程用电规律及特性,分析用户用电行为,诊断存在疑似瞬时超量程异常的用户,并给出用电用户疑似瞬时超量程的用电时段。同时,根据工单反馈的现场信息,不断完善、丰富、优化知识库。

[0044] 2. 构建稳定超量程计量点知识库

[0045] 基于电能表稳定超量程异常诊断模型、工单反馈的现场核查信息、用电用户档案信息建立稳定超量程专家知识库。根据诊断模型计算出的数据以及稳定超量程的用电特性及规律,分析用户用电行为,诊断存在疑似稳定超量程的异常用户。

[0046] 一种可行的判定方式为电流平均值大于 I_{max} ,且电能表计算误差绝对值超过限值,则判定为稳定超量程异常类型。

[0047] 同时,根据工单反馈的现场信息,不断完善、优化知识库。

[0048] 3. 构建疑似窃电计量点知识库构建

[0049] 基于现场的核查信息、运行误差计算信息、电压电流信息、功率因数信息等建立知识库。疑似窃电的知识库基于多种信息的综合分析,如:

[0050] (1) 电量异常:

[0051] 当一个用户出现长期未0值,负增长、零增长或异常增长的情况,有可能是窃电所致;

[0052] (2) 线损异常:

[0053] 台区线损超出阈值,或与去年同期线损比较,联合判定台区下是否可能存在窃电;

[0054] (3) 三相不平衡分析:

[0055] 三相不平衡率的波动变化也可以表征为窃电现象;

[0056] (4) 异常事件分析:

[0057] 基于获取的计量点的异常事件,如电能表开盖、计量门开闭,这些窃电相关的异常事件上报有存在窃电的可能。

[0058] 根据以上条件,结合误差分析结果,建立疑似窃电专家知识库。给疑似窃电结论提供诊断依据。

[0059] 4. 构建接线错误计量点知识库

[0060] 通过电气方面的理论推导,得出不同接线情况下,对电表计量的影响,同时,记录接线错误计量点上报的异常事件,如潮流反向、相序异常、反向电量异常等,后期可用于异常事件的关联分析,提高错接线判别准确率。将相关内容整理成接线错误计量点知识库,通过接线错误异常诊断模型给出的输入,诊断是否存在接线错误。

[0061] 5. 构建采集异常计量点知识库

[0062] 基于电能表的日冻结数据、异常事件信息、高频采集96点电压电流数据信息,构建

采集异常计量点专家知识库。采集异常的计量点会出现用电信息数据无法上报的现象,反应到日冻结、电压电流数据里,会出现某些数据点的数据为空值,对电表进行数据点为空值的统计分析,计算电表的采集成功率,以此发现采集异常的计量点。此外还可以借助异常事件提供的采集异常信息,进行综合判断,以提高诊断成功率。

[0063] 给出诊断结果后,可以根据现场核查反馈的结果,不断完善和优化采集异常的诊断逻辑。

[0064] 6. 构建时钟错误计量点知识库

[0065] 基于电能表的日冻结数据、异常事件信息,建立时钟错误计量点专家知识库。知识库一方面可以利用发生时钟错误时,电能表上报的异常来诊断是否发生时钟错误,另一方面,当存在时钟错误时,日冻结数据可能会发生偏移,可以尝试平移某块表的用电量,检测是否对线损有较大改善,以此诊断时钟错误。

[0066] 知识库存储出现问题的计量点,计量错误信息,以便误差分析模块的调用。

[0067] 7. 构建互感器过载计量点知识库

[0068] 基于电能表的高频采集96点电压电流数据及日冻结数据、电能表规格数据构建互感器过载计量点知识库。基于日冻结数据和96点电流数据,估算日平均工作电流,平均电流大于电能表最大电流,则判定为互感器过载。

[0069] 现场核查工单反馈异常结果进行归类入库,以便后期自动化对比判别。

[0070] 8. 构建户变关系错误计量点知识库

[0071] 台区户变关系错乱导致线损异常,影响误差计算和分析,不利于台区的数据治理工作。记录户变关系错误的计量点的负荷环境、用电曲线、地理位置等信息,建立户变关系计量点知识库,可在台区误差分析时引入相关判别指标,对数据首先进行预处理,大大提高模型误差计算精度。同时,根据户变关系错误计量点知识库,可以对类似计量点的户变关系诊断和治理提供对比依据。

[0072] 9. 构建互感器轻载计量点知识库

[0073] 基于电能表的高频采集96点电压电流数据及日冻结数据、电能表规格数据构建互感器轻载计量点知识库。基于日冻结数据和96点电流数据,估算日平均工作电流,设定轻载阈值,平均电流小于电能表额定电流设定阈值,则判定为互感器轻载。

[0074] 现场核查工单反馈异常结果进行归类入库,以便后期自动化对比判别。

[0075] 10. 构建采集器异常知识库

[0076] 基于电能表的日冻结数据、高频采集96点电压电流数据信息、电表档案信息构建采集器异常专家知识库。采集器异常会导致批量的电表用电信息数据无法上报的现象,反应到日冻结、电压电流数据里,会出现多块表的数据点的数据为空值,对台区下邻近电表进行数据点为空值的统计分析,计算该区域下电表的采集成功率,以此发现采集器异常。

[0077] 给出诊断结果后,可以根据现场核查反馈的结果,不断完善和优化采集器异常的诊断逻辑。

[0078] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请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,本申请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

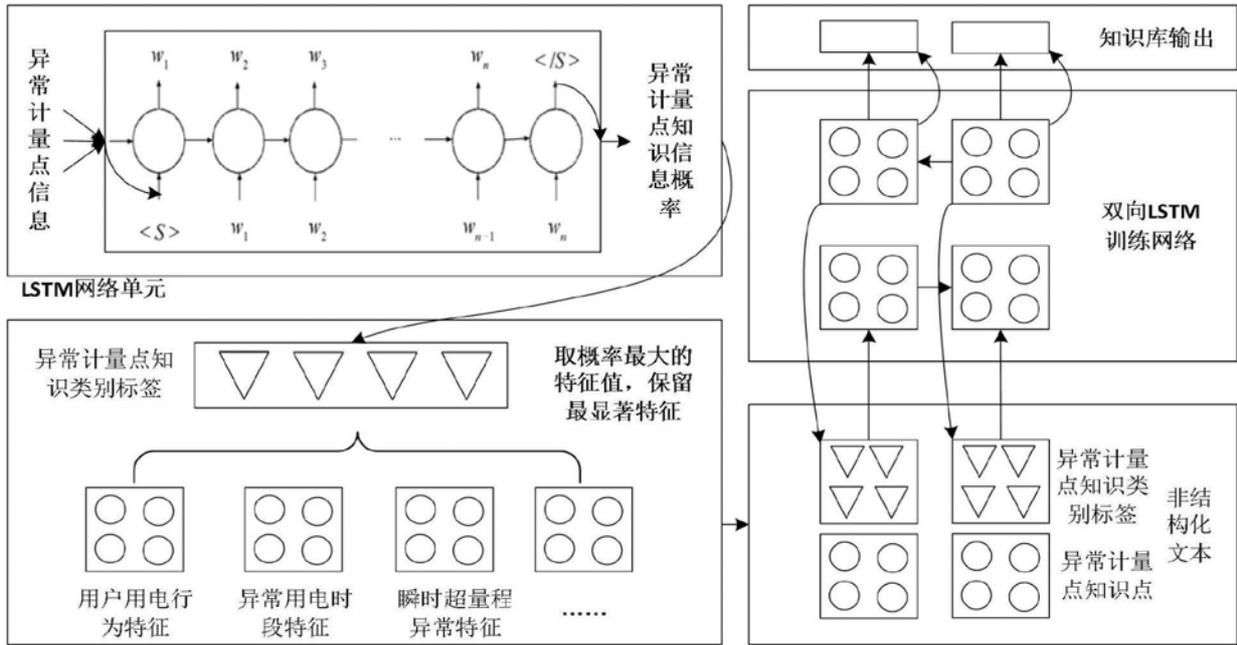
品的形式。

[0079] 本申请是参照根据本申请实施例的方法、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通用计算机、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。

[0080] 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。

[0081] 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。

[0082] 最后应当说明的是:以上实施例仅用以说明本发明的技术方案而非对其限制,尽管参照上述实施例对本发明进行了详细的说明,所属领域的普通技术人员应当理解:依然可以对本发明的具体实施方式进行修改或者等同替换,而未脱离本发明精神和范围的任何修改或者等同替换,其均应涵盖在本发明的权利要求保护范围之内。



例子：瞬时超量程计量点知识点抽取过程

图1

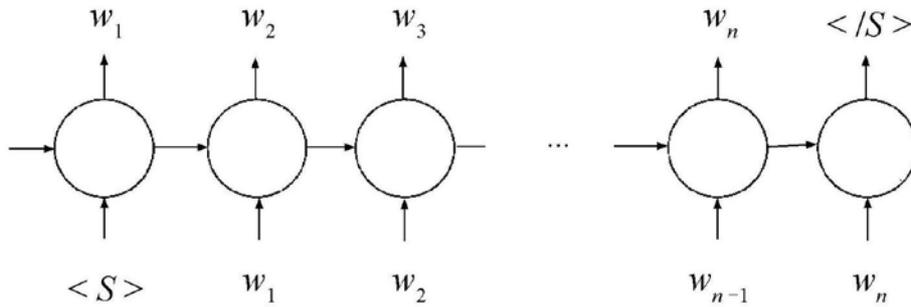


图2