



(12)发明专利申请

(10)申请公布号 CN 111553543 A

(43)申请公布日 2020.08.18

(21)申请号 202010419020.8

(22)申请日 2020.05.18

(71)申请人 润联软件系统(深圳)有限公司
地址 518000 广东省深圳市南山区桃源街
道桃源社区高发西路20号方大广场2
号楼2301

(72)发明人 刘雨桐 石强 熊娇 王国勋

(74)专利代理机构 深圳市精英专利事务所
44242

代理人 武志峰

(51)Int.Cl.
G06Q 10/04(2012.01)
G06Q 50/06(2012.01)
G06N 3/04(2006.01)

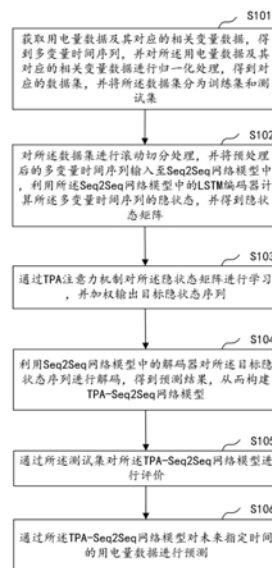
权利要求书3页 说明书11页 附图4页

(54)发明名称

基于TPA-Seq2Seq的电力负荷预测方法及相
关组件

(57)摘要

本发明公开了基于TPA-Seq2Seq的电力负荷
预测方法及相关组件。所述方法包括:获取用电
量数据及其对应的相关变量数据,得到多变量时
间序列,并对用电量数据及其对应的相关变量数
据进行归一化处理,得到对应的数据集,将数据
集分为训练集和测试集;对数据集进行滚动切分
处理,并将预处理后的多变量时间序列输入至
Seq2Seq网络模型中,利用Seq2Seq网络模型中的
LSTM编码器计算多变量时间序列的隐状态,得到
隐状态矩阵;通过TPA注意力机制对所述隐状态
矩阵进行学习,并加权输出目标隐状态序列;利
用Seq2Seq网络模型中的解码器对目标隐状态序
列进行解码,得到预测结果,从而构建TPA-
Seq2Seq网络模型;通过测试集对所述TPA-
Seq2Seq网络模型进行评价。本发明提高了电力
负荷预测精度,使预测结果更准确。



CN 111553543 A

1. 一种基于TPA-Seq2Seq的电力负荷预测方法,其特征在于,包括:

获取用电量数据及其对应的相关变量数据,得到多变量时间序列,并对所述用电量数据及其对应的相关变量数据进行归一化处理,得到对应的数据集,并将所述数据集分为训练集和测试集;

对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq网络模型中,利用所述Seq2Seq网络模型中的LSTM编码器计算所述多变量时间序列的隐状态,并得到隐状态矩阵;

通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列;

利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型;

通过所述测试集对所述TPA-Seq2Seq网络模型进行评价;

通过所述TPA-Seq2Seq网络模型对未来指定时间的用电量数据进行预测。

2. 根据权利要求1所述的基于TPA-Seq2Seq的电力负荷预测方法,其特征在于,所述对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq模型中,包括:

根据预设滑窗的窗口长度对所述数据集进行滚动切分,得到k组多变量时间序列,并使每组多变量时间序列包含w个时刻的样本数据,且每个时刻的样本数据包含n个特征变量。

3. 根据权利要求1所述的基于TPA-Seq2Seq的电力负荷预测方法,其特征在于,所述隐状态矩阵为:

$H = \{h_{t-w}, h_{t-w+1}, \dots, h_{t-1}\}$, 其中,每一隐状态的维度为m,w为滑窗窗口长度,t为最后时刻, h_t 即为最后时刻的隐状态。

4. 根据权利要求2所述的基于TPA-Seq2Seq的电力负荷预测方法,其特征在于,所述通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列,包括:

按下述公式对所述隐状态的相关变量进行卷积计算,得到目标矩阵 H^C :

$$H_{i,j}^C = \sum_{l=1}^w H_{i,(t-w-1+l)} \times C_{j,T-w+1}$$

式中, $H_{i,j}$ 表示所述隐状态矩阵H第i个行向量和第j个卷积核作用的结果值,T为TPA注意力机制过程中所覆盖的范围,l表示所述滑窗的窗口长度w中的每个时刻;

按下述公式计算所述目标矩阵 H^C 第i个行向量与最后时刻的隐状态 h_t 的相关性,并得到相关性评分函数:

$$f(H_i^C, h_t) = (H_i^C)^T W_a h_t$$

式中,f为相关性评分函数, H_i^C 为目标矩阵的第i行, h_t 为最后时刻的隐状态, $W_a \in \mathbb{R}^{m \times k}$, W_a 为待训练学习的参数矩阵, $\mathbb{R}^{m \times k}$ 为m行k列的实数矩阵;

按下述公式利用sigmoid函数对所述相关性评分函数进行归一化处理,得到注意力权重 α_i :

$$\alpha_i = \text{sigmoid}(f(H_i^C, h_t))$$

按下述公式利用所述注意力权重 α_i 对所述目标矩阵 H^c 的每一行进行加权求和,得到上下文向量 v_t :

$$v_t = \sum_{i=1}^m \alpha_i H_i^c$$

式中, α_i 为注意力权重, H_i^c 为目标矩阵的第*i*行, m 为隐状态的维度;

按下述公式将所述上下文向量 v_t 与所述LSTM编码器中输出的最后时刻的隐状态 h_t 进行融合,得到所述目标隐状态序列 h'_t :

$$h'_t = W_h h_t + W_v v_t$$

式中, $h_t, h'_t \in R^m, W_h \in R^{m \times m}, W_v \in R^{m \times k}, W_h$ 和 W_v 均为待训练学习的参数矩阵, R^m 为*m*行的实数矩阵, $R^{m \times m}$ 为*m*行*m*列的实数矩阵, $R^{m \times k}$ 为*m*行*k*列的实数矩阵。

5. 根据权利要求1所述的基于TPA-Seq2Seq的电力负荷预测方法,其特征在于,所述利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型,包括:

将所述目标隐状态序列作为LSTM网络的初始隐藏层状态并输入至所述LSTM网络中,并根据下列公式计算*t*个隐藏层状态:

$$h''_t = \sigma(W h''_{t-1} + b)$$

式中, h''_t 为第*t*个神经元输出的隐藏层状态, h''_{t-1} 为第*t-1*个神经元输出的隐藏层状态, W 为权重矩阵, b 为偏置项, σ 为神经元激活函数;

根据下列公式计算每个隐藏层状态输出的预测结果:

$$y''_t = \sigma(v h''_t + c)$$

式中, y''_t 为第*t*个神经元输出的预测结果, V 为权重矩阵, c 为偏置项, σ 为神经元激活函数。

6. 根据权利要求1所述的基于TPA-Seq2Seq的电力负荷预测方法,其特征在于,还包括:通过损失函数计算所述TPA-Seq2Seq网络模型的预测值和对应的真实值之间偏差的均方误差:

$$MSE = \frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}$$

式中, y_i 为第*i*个数据的真实值, y'_i 为第*i*个数据在所述TPA-Seq2Seq网络模型输出的预测值, n 为样本个数。

7. 根据权利要求1所述的基于TPA-Seq2Seq的电力负荷预测方法,其特征在于,所述通过所述测试集对所述TPA-Seq2Seq网络模型进行评价,包括:

通过MAPE指标对所述TPA-Seq2Seq网络模型进行评价,计算公式为:

$$MAPE = \frac{\sum_{i=1}^n (|y_i - y'_i| / y_i)}{n}$$

式中, y_i 为测试集中第*i*个数据的真实值, y'_i 为第*i*个数据在所述TPA-Seq2Seq网络模型输出的预测值, n 为测试集样本个数。

8. 一种基于TPA-Seq2Seq的电力负荷预测装置,其特征在于,包括:

获取单元,用于获取用电量数据及其对应的相关变量数据,得到多变量时间序列,并对

所述用电量数据及其对应的相关变量数据进行归一化处理,得到对应的数据集,并将所述数据集分为训练集和测试集;

编码单元,用于对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq网络模型中,利用所述Seq2Seq网络模型中的LSTM编码器计算所述多变量时间序列的隐状态,并得到隐状态矩阵;

学习单元,用于通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列;

解码单元,用于利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型;

评价单元,用于通过所述测试集对所述TPA-Seq2Seq网络模型进行评价;

预测单元,用于通过所述TPA-Seq2Seq网络模型对未来指定时间的用电量数据进行预测。

9. 一种计算机设备,其特征在于,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现如权利要求1至7任一项所述的基于TPA-Seq2Seq的电力负荷预测方法。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有计算机程序,所述计算机程序被处理器执行时实现如权利要求1至7任一项基于TPA-Seq2Seq的所述的电力负荷预测方法。

基于TPA-Seq2Seq的电力负荷预测方法及相关组件

技术领域

[0001] 本发明涉及电力预测领域,具体涉及基于TPA-Seq2Seq电力负荷预测方法及相关组件。

背景技术

[0002] 在日常生活中,时间序列数据无处不在,例如,家庭用电量、道路占用率、货币兑换率、太阳能发电量等,都可以看作是时间序列数据。在电力负荷预测领域内,大多数情况下收集到的数据为多变量时间序列数据,而这些数据通常由当地电力公司跟踪,并且不同的多变量时间序列之间存在复杂的动态相互依赖关系,从而导致难以对多变量时间序列进行捕获和分析。

[0003] 传统的用于时间序列预测的非线性自回归模型很难捕捉到一段较长时间内的数据间的时间相关性,也比较难选择相应的驱动数据来对多变量时间序列数据进行预测。传统的线性模型具体包括:自回归模型(AR,Auto-Regressive model,是一种统计上一种处理时间序列的方法)、移动平均模型(MA,Moving Average model)、自回归移动平均模型(ARMA,Auto-Regressive Moving Average Model)、自回归合成移动平均模型(ARIMA,Auto-Regressive Integrated Moving Average)等,这些模型针对小规模预测取得了很好的效果,但是不能满足用电量数据存在的突变性和多依赖性等非线性要求。

[0004] 随着神经网络、支持向量机等技术的不断提出,出现了基于机器学习算法的预测模型,例如:人工神经网络、最小二乘支持向量机等算法模型。以上算法存在的问题在于缺少对时序数据时间相关性的考虑,从而导致预测精度有限。

发明内容

[0005] 本发明实施例提供了一种基于TPA-Seq2Seq的电力负荷预测方法、装置、计算机设备及存储介质,旨在提高基于多变量时间序列数据的电力负荷预测精度。

[0006] 第一方面,本发明实施例提供了一种基于TPA-Seq2Seq的电力负荷预测方法,包括:

[0007] 获取用电量数据及其对应的相关变量数据,得到多变量时间序列,并对所述用电量数据及其对应的相关变量数据进行归一化处理,得到对应的数据集,并将所述数据集分为训练集和测试集;

[0008] 对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq网络模型中,利用所述Seq2Seq网络模型中的LSTM编码器计算所述多变量时间序列的隐状态,并得到隐状态矩阵;

[0009] 通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列;

[0010] 利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型;

[0011] 通过所述测试集对所述TPA-Seq2Seq网络模型进行评价;

[0012] 通过所述TPA-Seq2Seq网络模型对未来指定时间的用电量数据进行预测。

[0013] 进一步的,所述对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq模型中,包括:

[0014] 根据预设滑窗的窗口长度对所述数据集进行滚动切分,得到k组多变量时间序列,并使每组多变量时间序列包含w个时刻的样本数据,且每个时刻的样本数据包含n个特征变量。

[0015] 进一步的,所述隐状态矩阵为:

[0016] $H = \{h_{t-w}, h_{t-w+1}, \dots, h_{t-1}\}$,其中,每一隐状态的维度为m,w为滑窗窗口长度,t为最后时刻, h_t 即为最后时刻的隐状态。

[0017] 进一步的,所述通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列,包括:

[0018] 按下述公式对所述隐状态的相关变量进行卷积计算,得到目标矩阵 H^C :

$$[0019] \quad H_{i,j}^C = \sum_{l=1}^w H_{i,(t-w-1+l)} \times C_{j,T-w+l}$$

[0020] 式中, $H_{i,j}$ 表示所述隐状态矩阵H第i个行向量和第j个卷积核作用的结果值,T为TPA注意力机制过程中所覆盖的范围,l表示所述滑窗的窗口长度w中的每个时刻;

[0021] 按下述公式计算所述目标矩阵 H^C 第i个行向量与最后时刻的隐状态 h_t 的相关性,并得到相关性评分函数:

$$[0022] \quad f(H_i^C, h_t) = (H_i^C)^T W_a h_t$$

[0023] 式中,f为相关性评分函数, H_i^C 为目标矩阵的第i行, h_t 为最后时刻的隐状态, $W_a \in R^{m \times k}$, W_a 为待训练学习的参数矩阵, $R^{m \times k}$ 为m行k列的实数矩阵;

[0024] 按下述公式利用sigmoid函数对所述相关性评分函数进行归一化处理,得到注意力权重 α_i :

$$[0025] \quad \alpha_i = \text{sigmoid}(f(H_i^C, h_t))$$

[0026] 按下述公式利用所述注意力权重 α_i 对所述目标矩阵 H^C 的每一行进行加权求和,得到上下文向量 v_t :

$$[0027] \quad v_t = \sum_{i=1}^m \alpha_i H_i^C$$

[0028] 式中, α_i 为注意力权重, H_i^C 为目标矩阵的第i行,m为隐状态的维度;

[0029] 按下述公式将所述上下文向量 v_t 与所述LSTM编码器中输出的最后时刻的隐状态 h_t 进行融合,得到所述目标隐状态序列 h'_t :

$$[0030] \quad h'_t = W_h h_t + W_v v_t$$

[0031] 式中, $h_t, h'_t \in R^m$, $W_h \in R^{m \times m}$, $W_v \in R^{m \times k}$, W_h 和 W_v 均为待训练学习的参数矩阵, R^m 为m行的实数矩阵, $R^{m \times m}$ 为m行m列的实数矩阵, $R^{m \times k}$ 为m行k列的实数矩阵。

[0032] 进一步的,所述利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解

码,得到预测结果,从而构建TPA-Seq2Seq网络模型,包括:

[0033] 将所述目标隐状态序列作为LSTM网络的初始隐藏层状态并输入至所述LSTM网络中,并根据下列公式计算t个隐藏层状态:

$$[0034] \quad h''_t = \sigma(W h''_{t-1} + b)$$

[0035] 式中, h''_t 为第t个神经元输出的隐藏层状态, h''_{t-1} 为第t-1个神经元输出的隐藏层状态,W为权重矩阵,b为偏置项, σ 为神经元激活函数;

[0036] 根据下列公式计算每个隐藏层状态输出的预测结果:

$$[0037] \quad y''_t = \sigma(v h''_t + c)$$

[0038] 式中, y''_t 为第t个神经元输出的预测结果,V为权重矩阵,c为偏置项, σ 为神经元激活函数。

[0039] 进一步的,还包括:

[0040] 通过损失函数计算所述TPA-Seq2Seq网络模型的预测值和对应的真实值之间偏差的均方误差:

$$[0041] \quad \text{MSE} = \frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}$$

[0042] 式中, y_i 为第i个数据的真实值, y'_i 为第i个数据在所述TPA-Seq2Seq网络模型输出的预测值,n为样本个数。

[0043] 进一步的,所述通过所述测试集对所述TPA-Seq2Seq网络模型进行评价,包括:

[0044] 通过MAPE指标对所述TPA-Seq2Seq网络模型进行评价,计算公式为:

$$[0045] \quad \text{MAPE} = \frac{\sum_{i=1}^n (|y_i - y'_i| / y_i)}{n}$$

[0046] 式中, y_i 为测试集中第i个数据的真实值, y'_i 为第i个数据在所述TPA-Seq2Seq网络模型输出的预测值,n为测试集样本个数。

[0047] 第二方面,本发明实施例提供了一种基于TPA-Seq2Seq的电力负荷预测装置,包括:

[0048] 获取单元,用于获取用电量数据及其对应的相关变量数据,得到多变量时间序列,并对所述用电量数据及其对应的相关变量数据进行归一化处理,得到对应的数据集,并将所述数据集分为训练集和测试集;

[0049] 编码单元,用于对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq网络模型中,利用所述Seq2Seq网络模型中的LSTM编码器计算所述多变量时间序列的隐状态,并得到隐状态矩阵;

[0050] 学习单元,用于通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列;

[0051] 解码单元,用于利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型;

[0052] 评价单元,用于通过所述测试集对所述TPA-Seq2Seq网络模型进行评价;

[0053] 预测单元,用于通过所述TPA-Seq2Seq网络模型对未来指定时间的用电量数据进行预测。

[0054] 第三方面,本发明实施例还提供了一种计算机设备,包括存储器、处理器及存储在

所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述的电力负荷预测方法。

[0055] 第四方面,本发明实施例还提供了一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,所述计算机程序被处理器执行时实现上述的电力负荷预测方法。

[0056] 本发明实施例提供了一种基于TPA-Seq2Seq的电力负荷预测方法、装置、计算机设备及存储介质。所述方法包括:获取用电量数据及其对应的相关变量数据,得到多变量时间序列,并对所述用电量数据及其对应的相关变量数据进行归一化处理,得到对应的数据集,并将所述数据集分为训练集和测试集;对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq网络模型中,利用所述Seq2Seq网络模型中的LSTM编码器计算所述多变量时间序列的隐状态,并得到隐状态矩阵;通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列;利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型;通过所述测试集对所述TPA-Seq2Seq网络模型进行评价;通过所述TPA-Seq2Seq网络模型对未来指定时间的用电量数据进行预测。本发明实施例利用TPA-Seq2Seq网络模型对包含多变量时间序列的电力数据进行预测,充分考虑多个时序数据间的时间相关性,从而提高电力负荷预测精度,使预测结果更准确。

附图说明

[0057] 为了更清楚地说明本发明实施例技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0058] 图1为本发明实施例提供的一种基于TPA-Seq2Seq电力负荷预测方法的流程示意图;

[0059] 图2为本发明实施例提供的一种基于TPA-Seq2Seq电力负荷预测方法的子流程示意图;

[0060] 图3为本发明实施例提供的一种基于TPA-Seq2Seq电力负荷预测方法的网络结构示意图;

[0061] 图4为本发明实施例提供的一种基于TPA-Seq2Seq电力负荷预测装置的示意性框图。

具体实施方式

[0062] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0063] 应当理解,当在本说明书和所附权利要求书中使用时,术语“包括”和“包含”指示所描述特征、整体、步骤、操作、元素和/或组件的存在,但并不排除一个或多个其它特征、整体、步骤、操作、元素、组件和/或其集合的存在或添加。

[0064] 还应当理解,在此本发明说明书中所使用的术语仅仅是出于描述特定实施例的目的而并不意在限制本发明。如在本发明说明书和所附权利要求书中所使用的那样,除非上下文清楚地指明其它情况,否则单数形式的“一”、“一个”及“该”意在包括复数形式。

[0065] 还应当进一步理解,在本发明说明书和所附权利要求书中使用的术语“和/或”是指相关联列出的项中的一个或多个的任何组合以及所有可能组合,并且包括这些组合。

[0066] 下面请参见图1,图1为本发明实施例提供的一种基于TPA-Seq2Seq的电力负荷预测方法的流程示意图,具体包括:步骤S101~S106。

[0067] S101、获取用电量数据及其对应的相关变量数据,得到多变量时间序列,并对所述用电量数据及其对应的相关变量数据进行归一化处理,得到对应的数据集,并将所述数据集分为训练集和测试集;

[0068] S102、对所述数据集进行滚动切分处理,并将预处理后的多变量时间序列输入至Seq2Seq网络模型中,利用所述Seq2Seq网络模型中的LSTM编码器计算所述多变量时间序列的隐状态,并得到隐状态矩阵;

[0069] S103、通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列;

[0070] S104、利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型;

[0071] S105、通过所述测试集对所述TPA-Seq2Seq网络模型进行评价;

[0072] S106、通过所述TPA-Seq2Seq网络模型对未来指定时间的用电量数据进行预测。

[0073] 循环神经网络(RNN,Recurrent Neural Network)是一种引入循环反馈的深层神经网络,考虑到时间序列的时序相关性,在学习具有长期依赖的时序数据上表现出相对于其他神经网络更强的实用性。长短期记忆(LSTM,Long Short-Term Memory)循环神经网络是RNN的一种特殊模型,其能够学习时序数据之间的长期依赖关系,从而有效地解决了常规RNN训练过程中的梯度消失和梯度爆炸问题,并因此在时间序列预测领域得到广泛应用。但是不论是RNN网络模型,还是长短期记忆(LSTM)网络模型,它们都存在处理固定长度序列数据的限制。而Seq2Seq网络模型在将一个序列作为输入映射到另一个输出序列时,主要完成编码输入与解码输出两个基本环节,进而可以解决不定长的序列数据输入问题。因此Seq2Seq网络模型的训练过程能够更好的学习数据之间的时序关系,也能够保持更强的鲁棒性。

[0074] 另外,传统注意力机制会选择相关的时间步加权,而TPA注意力机制(Temporal Pattern Attention)则会选择相关变量进行加权,这种注意力机制(即TPA注意力机制)能够考虑到不同变量对预测变量的影响大小,侧重于选择关键变量,可对周期不明显的电力负荷数据进行多变量的精细化预测。

[0075] 如图3所示,本实施例提供的网络结构示意图具体包括数据处理过程、Seq2Seq网络模型编码过程、TPA注意力机制学习过程以及Seq2Seq网络模型解码输出预测结果。本实施例将Seq2Seq网络模型与TPA注意力机制相结合,对基于多变量时间序列的电力负荷数据进行预测,既可以解决不定长的序列数据输入问题,又可以对周期不明显的电力负荷数据进行多变量的精细化预测,从而提高电力负荷预测精度,使预测结果更准确。

[0076] 需要说明的是,在步骤S101中,所获取的用电量数据对应的相关变量数据具体可

以是指温度、天气、降水量、节假日以及季节等各自对应的数据,并且由于各变量的数量级或者量纲可能不一致,因此需要对获取的用电量数据及其对应的相关变量数据进行归一化处理。另外,本实施例中的数据集包含的用电量数据及其对应的相关变量数据均为历史数据。

[0077] 在一实施例中,所述步骤S102包括:根据预设滑窗的窗口长度对所述数据集进行滚动切分,得到k组多变量时间序列,并使每组多变量时间序列包含w个时刻的样本数据,且每个时刻的样本数据包含n个特征变量。

[0078] 本实施例中,为了匹配网络模型的输入条件,因而对数据集进行滚动切分。具体的,根据预设滑窗的窗口长度可以将数据集滚动切分为(k,w,n)形式,其中,k为经过滚动切分后的组数,w为每一组包含的样本个数,n为某时刻的样本特征个数。在这里,滑窗的窗口长度的大小和(k,w,n)中的w的大小相同。例如当滑窗的窗口长度为6,则第一组数据为时刻1到时刻7,第二组数据为时刻2到时刻8,以此类推,最终得到k组数据。通过上述滚动切分的方法,使得到的每一组数据均包含w个时刻的样本数据,每个样本数据有n个特征。

[0079] 在一实施例中,所述隐状态矩阵为: $H = \{h_{t-w}, h_{t-w+1}, \dots, h_{t-1}\}$,其中,每一隐状态的维度为m,w为滑窗窗口长度,t为最后时刻, h_t 即为最后时刻的隐状态。

[0080] 本实施例中,在将预处理后的多变量时间序列输入编码器中的LSTM网络模型后,LSTM网络模型会在每一个时刻输出一个隐状态,而这些隐状态分别为最后一个时刻的 h_t 以及之前时刻的隐状态矩阵 $H = \{h_{t-w}, h_{t-w+1}, \dots, h_{t-1}\}$,其中,每一个隐状态的维度为m。也就是说,本实施例中的隐状态矩阵由除最后时刻的隐状态以外的其余时刻的隐状态构成。

[0081] 在一实施例中,如图2所示,所述步骤S104包括:步骤S201~S205。

[0082] S201、按下述公式对所述隐状态的相关变量进行卷积计算,得到目标矩阵 H^C :

$$[0083] \quad H_{i,j}^C = \sum_{l=1}^w H_{i,(t-w-1+l)} \times C_{j,T-w+l}$$

[0084] 式中, $H_{i,j}$ 表示所述隐状态矩阵H第i个行向量和第j个卷积核作用的结果值,T为TPA注意力机制过程中所覆盖的范围,l表示所述滑窗的窗口长度w中的每个时刻;

[0085] 本步骤中,对上一步骤(即步骤S103)获得的隐状态用k个 $1 \times T$ 的卷积核做卷积。具体的,卷积核沿隐状态矩阵H的行向量计算卷积,并提取该变量在该卷积核范围内的时间模式矩阵。另外,T为TPA注意力机制过程中所覆盖的范围,通常令 $T=w$,即预设滑窗的滑窗长度。

[0086] S202、按下述公式计算所述目标矩阵 H^C 第i个行向量与最后时刻的隐状态 h_t 的相关性,并得到相关性评分函数:

$$[0087] \quad f(H_i^C, h_t) = (H_i^C)^T W_a h_t$$

[0088] 式中,f为相关性评分函数, H_i^C 为目标矩阵的第i行, h_t 为最后时刻的隐状态, $W_a \in R^{m \times k}$, W_a 为待训练学习的参数矩阵, $R^{m \times k}$ 为m行k列的实数矩阵;

[0089] S203、按下述公式利用sigmoid函数对所述相关性评分函数进行归一化处理,得到注意力权重 α_i :

[0090] $\alpha_i = \text{sigmoid}(f(H_i^C, h_t))$

[0091] 本步骤中,利用sigmoid函数对相关性评分函数所得结果进行归一化处理,将变量映射到0到1之间,从而得到便于选择多变量的注意力权重。

[0092] S204、按下述公式利用所述注意力权重 α_i 对所述目标矩阵 H^C 的每一行进行加权求和,得到上下文向量 v_t :

$$[0093] \quad v_t = \sum_{i=1}^m \alpha_i H_i^C$$

[0094] 式中, α_i 为注意力权重, H_i^C 为目标矩阵的第*i*行,*m*为隐状态的维度;

[0095] 本步骤中,上下文向量 v_t 表示对时间模式矩阵(即目标矩阵 H^C)行向量的加权求和,且其行向量代表单个变量的时间模式,因此TPA注意力机制选择相关变量加权,以便本实施例中的网络模型能够学习多个变量之间的相互依赖性。并且,这种相互依赖型不仅体现在同一时间步内,而且体现在所有历史的时间和序列中。

[0096] S205、按下述公式将所述上下文向量 v_t 与所述LSTM编码器中输出的最后时刻的隐状态 h_t 进行融合,得到所述目标隐状态序列 h'_t :

$$[0097] \quad h'_t = W_h h_t + W_v v_t$$

[0098] 式中, $h_t, h'_t \in \mathbb{R}^m, W_h \in \mathbb{R}^{m \times m}, W_v \in \mathbb{R}^{m \times k}, W_h$ 和 W_v 均为待训练学习的参数矩阵, \mathbb{R}^m 为*m*行的实数矩阵, $\mathbb{R}^{m \times m}$ 为*m*行*m*列的实数矩阵, $\mathbb{R}^{m \times k}$ 为*m*行*k*列的实数矩阵。

[0099] 本步骤中, W_a (即步骤S202中的 W_a)、 W_h 和 W_v 均是首先随机初始化,然后在所述TPA-Seq2Seq网络模型训练优化后获得的参数矩阵,也就是说, W_a 、 W_h 和 W_v 会随着所述TPA-Seq2Seq网络模型的训练优化后而对应优化,如此,可以为所述TPA-Seq2Seq网络模型找到最优权重,从而更好的将 v_t 和 h_t 融合。

[0100] 本实施例中,TPA注意力机制用于对隐状态的相关变量进行加权,生成新的隐状态序列,其主要由一维卷积和Attention机制两部分构成。在一维卷积部分中,利用卷积捕获可变信号的时间模式。通过对目标矩阵中的每一行做加权求和,得到新的隐状态序列(即目标隐状态序列)。

[0101] 在一实施例中,所述步骤S105包括:

[0102] 将所述目标隐状态序列作为LSTM网络的初始隐藏层状态并输入至所述LSTM网络中,并根据下列公式计算*t*个隐藏层状态:

$$[0103] \quad h''_t = \sigma(W h''_{t-1} + b)$$

[0104] 式中, h''_t 为第*t*个神经元输出的隐藏层状态, h''_{t-1} 为第*t-1*个神经元输出的隐藏层状态, W 为权重矩阵, b 为偏置项, σ 为神经元激活函数;

[0105] 根据下列公式计算每个隐藏层状态输出的预测结果:

$$[0106] \quad y''_t = \sigma(v h''_t + c)$$

[0107] 式中, y''_t 为第*t*个神经元输出的预测结果, V 为权重矩阵, c 为偏置项, σ 为神经元激活函数。

[0108] 本实施例中,将目标隐状态序列输入至Seq2Seq网络模型的解码器进行解码预测。所谓解码,就是将之前生成的固定向量再转化成输出序列。具体的,将目标隐状态序列 h'_t

作为解码器中的LSTM网络模型的初始隐藏状态,并输入至解码器中的LSTM网络模型中,后续只接受上一个神经元的隐藏层状态,而不接收其他的输入。例如,从初始隐藏状态开始计算,公式如下:

$$[0109] \quad h''_1 = \sigma(wh''_1 + b)$$

[0110] 式中, h''_1 为第一个神经元输出的隐藏层状态,其中W为权重矩阵,b为偏置项, σ 为神经元激活函数;

$$[0111] \quad y''_1 = \sigma(vh''_1 + c)$$

[0112] 式中, y''_1 为第一个神经元输出的预测结果,其中V为权重矩阵,c为偏置项, σ 为神经元激活函数;

[0113] 然后将第一个神经元隐藏层状态 h''_1 输入到下一个神经元,得到第二个神经元隐藏层状态 h''_2 和第二个神经元输出的预测结果 y''_2 ,以此类推,第t个神经元的隐藏层状态和预测结果即为:

$$[0114] \quad h''_t = \sigma(Wh''_{t-1} + b)$$

$$[0115] \quad y''_t = \sigma(vh''_t + c)$$

[0116] 在一实施例中,所述基于TPA-Seq2Seq的电力负荷预测方法还包括:通过损失函数计算所述TPA-Seq2Seq网络模型的预测值和对应的真实值之间偏差的均方误差:

$$[0117] \quad \text{MSE} = \frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}$$

[0118] 式中, y_i 为第i个数据的真实值, y'_i 为第i个数据在所述TPA-Seq2Seq网络模型输出的预测值,n为样本个数。

[0119] 本实施例中,根据损失函数计算出的均方误差对所述TPA-Seq2Seq网络模型进行优化,也就是对所述TPA-Seq2Seq网络模型中的参数进行优化,从而使预测结果更加准确。在一具体实施例中,选用自适应学习率的Adam优化算法对所述TPA-Seq2Seq网络模型进行优化,加快收敛速度。Adam优化算法是一种可以替代传统随机梯度下降过程的一阶优化算法,它能基于训练数据的迭代更新神经网络权重。

[0120] 在一实施例中,所述步骤S106包括:通过MAPE指标对所述TPA-Seq2Seq网络模型进行评价,计算公式为:

$$[0121] \quad \text{MAPE} = \frac{\sum_{i=1}^n (|y_i - y'_i|/y_i)}{n}$$

[0122] 式中, y_i 为测试集中第i个数据的真实值, y'_i 为第i个数据在所述TPA-Seq2Seq网络模型输出的预测值,n为测试集样本个数。

[0123] 本实施例中,利用MAPE指标对经过训练优化的TPA-Seq2Seq网络模型进行评价,确保最后得到的预测结果准确可靠。

[0124] 图4为本发明实施例提供的一种基于TPA-Seq2Seq的电力负荷预测装置400,所述装置400包括:

[0125] 获取单元401,用于获取用电量数据及其对应的相关变量数据,得到多变量时间序列,并对所述用电量数据及其对应的相关变量数据进行归一化处理,得到对应的数据集,并将所述数据集分为训练集和测试集;

[0126] 编码单元402,用于对所述数据集进行滚动切分处理,并将预处理后的多变量时间

序列输入至Seq2Seq网络模型中,利用所述Seq2Seq网络模型中的LSTM编码器计算所述多变量时间序列的隐状态,并得到隐状态矩阵;

[0127] 学习单元403,用于通过TPA注意力机制对所述隐状态矩阵进行学习,并加权输出目标隐状态序列;

[0128] 解码单元404,用于利用Seq2Seq网络模型中的解码器对所述目标隐状态序列进行解码,得到预测结果,从而构建TPA-Seq2Seq网络模型;

[0129] 评价单元405,用于通过所述测试集对所述TPA-Seq2Seq网络模型进行评价;

[0130] 预测单元406,用于通过所述TPA-Seq2Seq网络模型对未来指定时间的用电量数据进行预测。

[0131] 在一实施例中,所述编码单元402包括:

[0132] 滚动切分单元,用于根据预设滑窗的窗口长度对所述数据集进行滚动切分,得到k组多变量时间序列,并使每组多变量时间序列包含w个时刻的样本数据,且每个时刻的样本数据包含n个特征变量。

[0133] 在一实施例中,所述隐状态矩阵为: $H = \{h_{t-w}, h_{t-w+1}, \dots, h_{t-1}\}$,其中,每一隐状态的维度为m,w为滑窗窗口长度,t为最后时刻, h_t 即为最后时刻的隐状态。

[0134] 在一实施例中,所述学习单元403包括:

[0135] 卷积计算单元,用于按下述公式对所述隐状态的相关变量进行卷积计算,得到目标矩阵 H^C :

$$[0136] \quad H_{ij}^C = \sum_{l=1}^w H_{i,(t-w-1+l)} \times C_{j,T-w+1}$$

[0137] 式中, $H_{i,j}$ 表示所述隐状态矩阵H第i个行向量和第j个卷积核作用的结果值,T为TPA注意力机制过程中所覆盖的范围,l表示所述滑窗的窗口长度w中的每个时刻;

[0138] 相关性计算单元,用于按下述公式计算所述目标矩阵 H^C 第i个行向量与最后时刻的隐状态 h_t 的相关性,并得到相关性评分函数:

$$[0139] \quad f(H_i^C, h_t) = (H_i^C)^T W_a h_t$$

[0140] 式中,f为相关性评分函数, H_i^C 为目标矩阵的第i行, h_t 为最后时刻的隐状态, $W_a \in R^{m \times k}$, W_a 为待训练学习的参数矩阵, $R^{m \times k}$ 为m行k列的实数矩阵;

[0141] 归一化处理单元,用于按下述公式利用sigmoid函数对所述相关性评分函数进行归一化处理,得到注意力权重 α_i :

$$[0142] \quad \alpha_i = \text{sigmoid}(f(H_i^C, h_t))$$

[0143] 加权求和单元,用于按下述公式利用所述注意力权重 α_i 对所述目标矩阵 H^C 的每一行进行加权求和,得到上下文向量 v_t :

$$[0144] \quad v_t = \sum_{i=1}^m \alpha_i H_i^C$$

[0145] 式中, α_i 为注意力权重, \mathbf{H}_i^C 为目标矩阵的第 i 行, m 为隐状态的维度;

[0146] 融合单元, 用于按下述公式将所述上下文向量 \mathbf{v}_t 与所述 LSTM 编码器中输出的最后时刻的隐状态 \mathbf{h}_t 进行融合, 得到所述目标隐状态序列 \mathbf{h}'_t :

$$[0147] \quad \mathbf{h}'_t = \mathbf{W}_h \mathbf{h}_t + \mathbf{W}_v \mathbf{v}_t$$

[0148] 式中, $\mathbf{h}_t, \mathbf{h}'_t \in \mathbb{R}^m, \mathbf{W}_h \in \mathbb{R}^{m \times m}, \mathbf{W}_v \in \mathbb{R}^{m \times k}, \mathbf{W}_h$ 和 \mathbf{W}_v 均为待训练学习的参数矩阵, \mathbb{R}^m 为 m 行的实数矩阵, $\mathbb{R}^{m \times m}$ 为 m 行 m 列的实数矩阵, $\mathbb{R}^{m \times k}$ 为 m 行 k 列的实数矩阵。

[0149] 在一实施例中, 所述解码单元 404 包括:

[0150] 输入单元, 用于将所述目标隐状态序列作为 LSTM 网络的初始隐藏层状态并输入至所述 LSTM 网络中, 并根据下列公式计算 t 个隐藏层状态:

$$[0151] \quad \mathbf{h}''_t = \sigma(\mathbf{W} \mathbf{h}''_{t-1} + \mathbf{b})$$

[0152] 式中, \mathbf{h}''_t 为第 t 个神经元输出的隐藏层状态, \mathbf{h}''_{t-1} 为第 $t-1$ 个神经元输出的隐藏层状态, \mathbf{W} 为权重矩阵, \mathbf{b} 为偏置项, σ 为神经元激活函数;

[0153] 输出单元, 用于根据下列公式计算每个隐藏层状态输出的预测结果:

$$[0154] \quad \mathbf{y}''_t = \sigma(\mathbf{v} \mathbf{h}''_t + \mathbf{c})$$

[0155] 式中, \mathbf{y}''_t 为第 t 个神经元输出的预测结果, \mathbf{v} 为权重矩阵, \mathbf{c} 为偏置项, σ 为神经元激活函数。

[0156] 在一实施例中, 所述装置 400 还包括:

[0157] 均方误差单元, 用于通过损失函数计算所述 TPA-Seq2Seq 网络模型的预测值和对应的真实值之间偏差的均方误差:

$$[0158] \quad \mathbf{MSE} = \frac{\sum_{i=1}^n (\mathbf{y}_i - \mathbf{y}'_i)^2}{n}$$

[0159] 式中, \mathbf{y}_i 为第 i 个数据的真实值, \mathbf{y}'_i 为第 i 个数据在所述 TPA-Seq2Seq 网络模型输出的预测值, n 为样本个数。

[0160] 在一实施例中, 所述评价单元 405 包括:

[0161] 指标评价单元, 用于通过 MAPE 指标对所述 TPA-Seq2Seq 网络模型进行评价, 计算公式为:

$$[0162] \quad \mathbf{MAPE} = \frac{\sum_{i=1}^n (|\mathbf{y}_i - \mathbf{y}'_i| / \mathbf{y}_i)}{n}$$

[0163] 式中, \mathbf{y}_i 为测试集中第 i 个数据的真实值, \mathbf{y}'_i 为第 i 个数据在所述 TPA-Seq2Seq 网络模型输出的预测值, n 为测试集样本个数。

[0164] 由于装置部分的实施例与方法部分的实施例相互对应, 因此装置部分的实施例请参见方法部分的实施例的描述, 这里暂不赘述。

[0165] 本发明实施例还提供了一种计算机可读存储介质, 其上存有计算机程序, 该计算机程序被执行时可以实现上述实施例所提供的步骤。该存储介质可以包括: U 盘、移动硬盘、只读存储器 (Read-Only Memory, ROM)、随机存取存储器 (Random Access Memory, RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0166] 本发明实施例还提供了一种计算机设备, 可以包括存储器和处理器, 存储器中存有计算机程序, 处理器调用存储器中的计算机程序时, 可以实现上述实施例所提供的步骤。当然电子设备还可以包括各种网络接口, 电源等组件。

[0167] 说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的系统而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。应当指出,对于本技术领域的普通技术人员来说,在不脱离本申请原理的前提下,还可以对本申请进行若干改进和修饰,这些改进和修饰也落入本申请权利要求的保护范围内。

[0168] 还需要说明的是,在本说明书中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的状况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

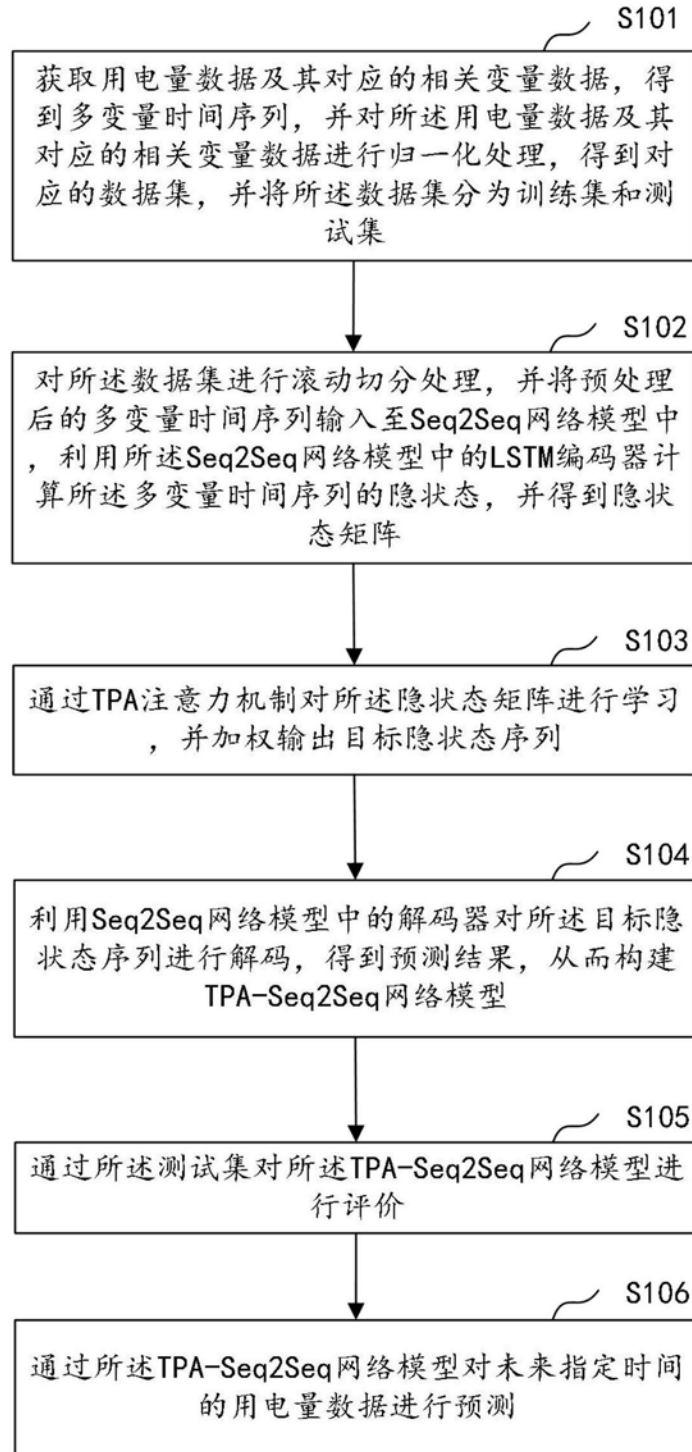


图1

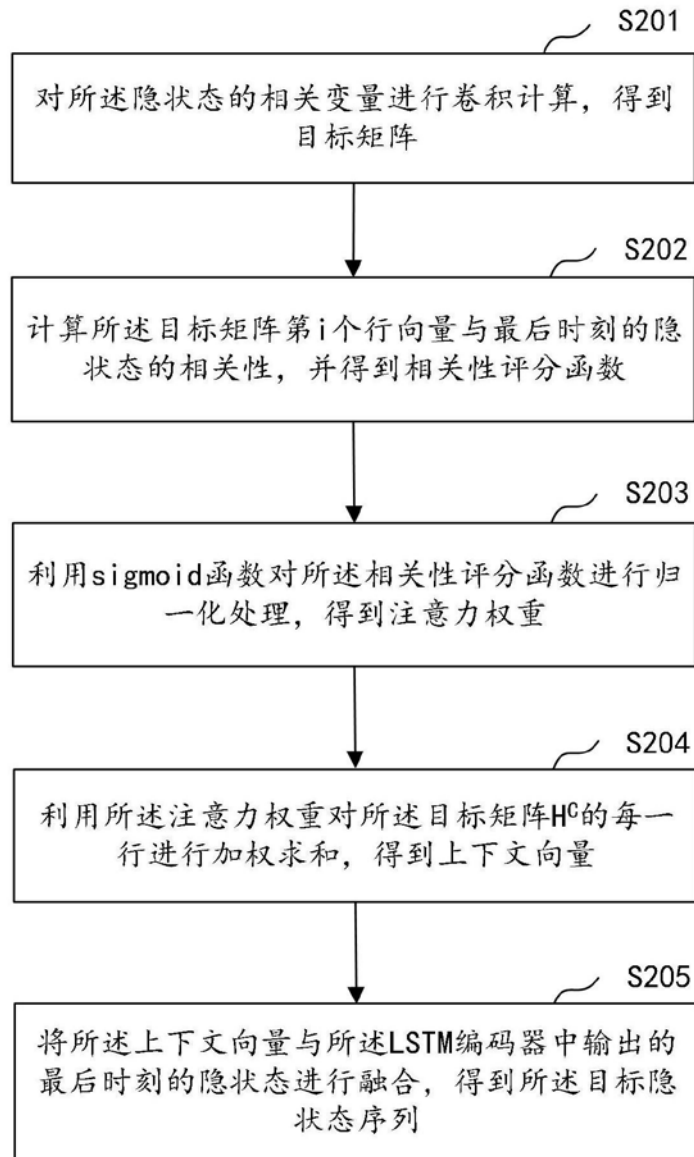


图2

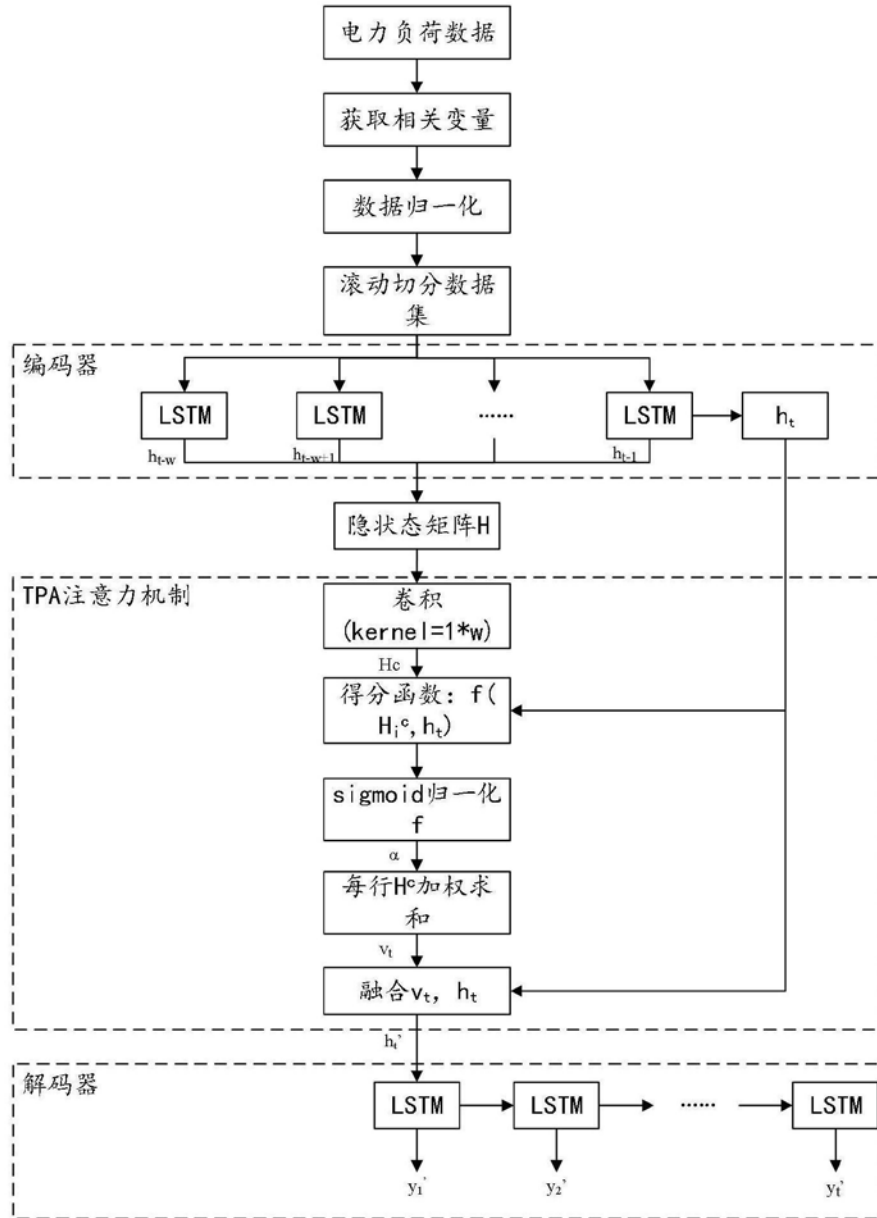


图3

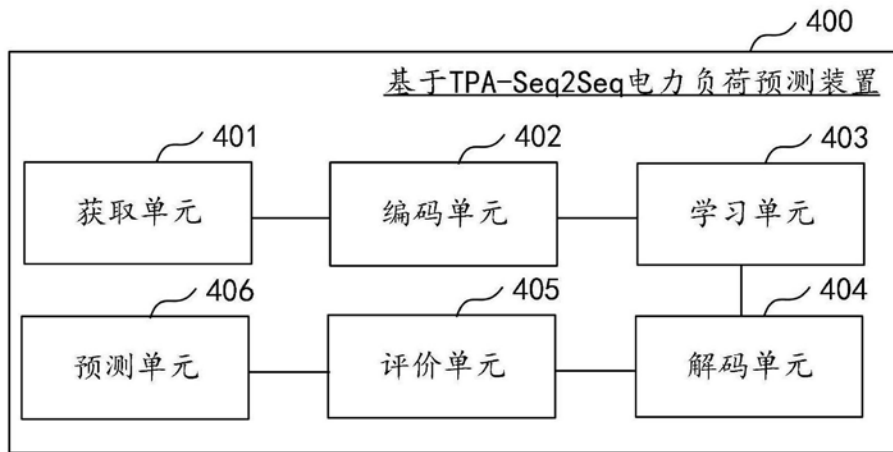


图4