

基于电-碳转换因子的工业用户用电信息分级挖掘算法

杨琦[†], 舒一飞, 梁飞, 康洁滢, 刘鹏

(国网宁夏电力有限公司营销服务中心(国网宁夏电力有限公司计量中心), 宁夏 银川 750011)

摘要:以降低碳排放为目的,研究了基于电-碳转换因子的工业用户用电信息分级挖掘算法,提升工业用户用电信息分级挖掘的全面性。利用潮流追踪法,结合工业用户用电信息,计算工业用户碳排放量;通过双向长短期循环记忆循环神经网络,结合工业用户用电信息,得到工业用户用电量;依据碳排放量与用电量,获取工业用户的电-碳转换因子,用于反映用电量与碳排放量之间的关联程度;通过模糊C均值聚类算法,分级挖掘电-碳转换因子,将具有相似电-碳转换因子的工业用户归为同一等级,完成工业用户用电信息分级挖掘。实验证明:该算法可有效计算工业用户碳排放量与用电量,得到电-碳转换因子;该算法可有效分级挖掘工业用户用电信息,确定工业用户的碳转换因子等级。

关键词: 电-碳转换因子;工业用户;用电信息;分级挖掘;潮流追踪法;模糊C均值聚类

中图分类号: TM714

文献标识码: A

Hierarchical Mining Algorithm for Industrial Users' Electricity Consumption Information Based on Electric-carbon Conversion Factor

YANG Qi[†], SHU Yifei, LIANG Fei, KANG Jieying, LIU Peng

(State Grid Ningxia Marketing Service Center(State Grid Ningxia Metrology Center), Yinchuan, Ningxia 750011, China)

Abstract: In order to reduce carbon emission, a hierarchical mining algorithm for industrial users' electricity consumption information based on electric-carbon conversion factor is studied to improve the comprehensiveness of industrial users' electricity consumption information hierarchical mining. The power flow tracking method is used to calculate the carbon emission of industrial users by combining the electricity consumption information. The power consumption of industrial users is obtained by using the bidirectional long and short term cyclic memory cyclic neural network combined with the power consumption information of industrial users. According to carbon emission and electricity consumption, the electric-carbon conversion factor of industrial users is obtained to reflect the correlation degree between electricity consumption and carbon emission. By using the fuzzy C-means clustering algorithm, the industrial users with similar electric-carbon conversion factors are classified into the same level to complete the hierarchical mining of industrial users' electricity consumption information. Experimental results show that the algorithm can effectively calculate the carbon emission and electricity consumption of industrial users, and obtain the electric-carbon conversion factor. This algorithm can effectively mine the power consumption information of industrial users and determine the carbon conversion factor level of industrial users.

Key words: electric-carbon conversion factor; industrial users; electricity consumption information; hierarchical mining; trend tracking method; fuzzy C-means clustering

工业用户用电信息是反映能源消耗和碳排放的关键数据,为了更好地满足工业用户的需求,提高能源利用效率^[1],降低碳排放,需要对工业用户用电信息进行深入的分级挖掘^[2]。分级挖掘算法是一种基于数据挖掘技术的算法^[3],通过对不同级别的数据进行分类、聚类、关联规则挖掘等操作,提取出有价值的信息和知识^[4]。在工业用户用电信息分级挖掘中,这种算法可以应用于不同级别的用电信息,以获取更全面、更深入的用电信息和知识。例如,汪颖等^[5]首先利用去噪自编码器网络对工业用户用电信息进行特征降维,实现工业用户用电信息的分级挖掘。但是如果输入数据存在较多的噪声或异常值,可能会影响特征降维的效果。魏伟等^[6]结合了多视角数据和社区发现技术,为能源管理和节能减排提供有力支持。但该算法中多视角数据需要从多个来源获取,获取难度较大。Yang等^[7]通过定义一种特殊的核函数,结合了数据点的密度和距离信息,实现用电信息的分级挖掘。但该算法在处理大规模数据集时会面临较高的计算复杂度,会导致算法在处理超大规模工业用户用电信息时效率较低。Nguyen等^[8]通过引入加权效用项集的概念,结合频繁项集挖掘技术,对工业用户的用电信息进行深入分析和挖掘。但该算法涉及复杂的数学计算和数据挖掘技术,需要结合领域知识和专家经验对挖掘结果进行解释和分析。

电-碳转换因子作为一种将电力消费量与碳排放量相联系的系数,为工业用户用电信息的分级挖掘提供了新的视角^[9],电-碳转换因子能够综合考虑用电量、设备效率、生产工艺等多种因素对碳排放的影响,这使得在分级挖掘过程中,能够更全面地评估不同工业用户的能源消耗和碳排放情况,避免单一指标的片面性。为此,研究基于电-碳转换因子的工业用户用电信息分级挖掘算法,为工业用户节能减排和可持续发展提供有力支持。

1 工业用户用电信息分级挖掘算法

1.1 基于用电信息的工业用户碳排放量计算

通过潮流追踪法,依据工业用户用电信息,计算工业用户的碳排放量。在电力系统中,一个工业用户通常被视为一个节点^[10]。

令第 i 个工业用户节点的总注入功率 P_i 为:

$$P_i = \sum_{j \in \mathcal{Y}_i^n} |P_{ij}| + P_{Gi} \quad (1)$$

其中,第 i 个工业用户节点的发电机注入有功功率是 P_{Gi} ,在工业用户用电信息内提取 P_{Gi} ; i 的进线集合是 $\mathcal{Y}_i^n$,在工业用户用电信息内提取 $\mathcal{Y}_i^n$;

第 j 个工业用户节点流入 i 的功率是 P_{ij} ,在工业用户用电信息内提取 P_{ij} 。

在无损电力网络内,存在 $|P_{ij}| = |P_{ji}|$,令 $H_{ij} = \frac{|P_{ji}|}{P_j}$,将其代入式(1)可得:

$$P_i - \sum_{j \in \mathcal{Y}_i^n} H_{ij} P_j = P_{Gi} \quad (2)$$

将式(2)变更成矩阵形式为:

$$\mathbf{K}_u \mathbf{P} = \mathbf{P}_G \quad (3)$$

其中,发电机功率列向量是 $\mathbf{P}_G$; n 个工业用户节点的总注入功率列向量是 $\mathbf{P}$; $n \times n$ 阶回溯矩阵是 $\mathbf{K}_u$, $\mathbf{K}_u$ 内元素如下:

$$\mathbf{K}_{u,ij} = \begin{cases} 1, & j = i \\ -H_{ij} = -\frac{|P_{ji}|}{P_j}, & j \in \mathcal{Y}_i^n \\ 0, & \text{其他} \end{cases} \quad (4)$$

若 $\mathbf{K}_u^{-1}$ 存在,那么式(4)可变更成:

$$\mathbf{P} = \mathbf{K}_u^{-1} \mathbf{P}_G \quad (5)$$

针对工业用户负荷节点,存在:

$$\mathbf{P}_L = \text{diag}(\mathbf{B}) \mathbf{P} \quad (6)$$

其中,工业用户负荷系数行向量是 $\mathbf{B}$, $\mathbf{B}$ 内的元素如下:

$$B_i = \begin{cases} \frac{P_{Li}}{P_i}, & i \in \mathcal{Y}_L \\ 0, & \text{其他} \end{cases} \quad (7)$$

其中,工业用户负荷节点集合是 α_L ; 第 i 个工业用户负荷节点的负荷是 P_{Li} 。因此, $\mathbf{P}_L$ 的计算公式变更为:

$$\mathbf{P}_L = \mathbf{B} \mathbf{K}_u^{-1} \mathbf{P}_G \quad (8)$$

利用式(8)能够获取电力系统内各工业用户负荷节点的发电来源^[11]。因此,各工业用户的用电碳排放量计算公式如下:

$$E_{Li} = \mathbf{B} \mathbf{K}_u^{-1} \text{diag}(\theta_{Gi}) P_{Gi} \quad (9)$$

其中,第 i 个工业用户的碳排放因子是 θ_{Gi} 。

1.2 基于用电信息的工业用户用电量计算

通过双向长短期循环记忆循环神经网络(Bi-LSTM),依据历史工业用户用电信息,计算当下工业用户的用电量^[12]。在 LSTM 的输入门内,输入历史工业用户用电信息 x_t ,输入门的输出公式如下:

$$q_t = \varphi(w_q[h_{t-1}, x_t] + b_q) \quad (10)$$

其中,激活函数是 $\varphi(\cdot)$; 输入门的权值与偏置是 w_q, b_q ; $t-1$ 时刻的隐藏状态是 h_{t-1} ,代表 LSTM 网络对输入工业用户用电信息序列的理解和表示。

遗忘门的输出公式如下:

$$f_t = \varphi(w_f[h_{t-1}, x_t] + b_f) \quad (11)$$

其中,遗忘门的权值与偏置是 w_f, b_f 。

通过更新细胞状态 s_t , Bi-LSTM 网络可以逐步学习到输入 x_t 中的长期依赖关系和模式,并将其用于计算当下用电量^[13]。 s_t 的更新公式如下:

$$s_t = f_t \cdot s_{t-1} + q_t \cdot [\varphi(w_s[h_{t-1}, x_t] + b_s)] \quad (12)$$

其中, s_t 的权值与偏置是 w_s 、 b_s 。

输出门的计算公式如下:

$$A_t = \varphi(w_A[h_{t-1}, x_t] + b_A) \quad (13)$$

其中,输出门的权值与偏置是 w_A 、 b_A 。

工业用户当下用电量计算结果为:

$$h_t = A_t \cdot \tanh(s_t) \quad (14)$$

Bi-LSTM 网络内包含两个 LSTM^[14], 利用 Bi-LSTM 网络计算当下工业用户用电量时,通过结合前向 LSTM 与后向 LSTM 的工业用户用电量计算结果 $\vec{h}_t$ 与 $\overleftarrow{h}_t$, 得到最终的用电量计算结果 $\hat{h}_t$ 。

1.3 工业用户电-碳转换因子计算

电-碳转换因子 Z 是用于描述电力消费与碳排放量之间关系的参数。它反映了每单位电力消费所产生的碳排放量^[15]。第 i 个工业用户电-碳转换因子 z_i 的计算公式如下:

$$z_i = \frac{E_{Li}}{\hat{h}_{ii}} \quad (15)$$

其中,第 i 个工业用户的用电量是 $\hat{h}_{ii}$ 。

1.4 基于电-碳转换因子的工业用户分级挖掘

利用模糊 C 均值聚类算法,对具有相似电-碳转换因子的工业用户进行分级挖掘,将其归为同一等级,完成工业用户用电信息分级挖掘。工业用户用电信息等级共包含三级,分别是高碳转换因子等级:这一等级的用户,其电-碳转换因子较高,意味着他们的用电行为对碳排放的影响较大,这些用户的能源利用效率较低。中碳转换因子等级:这一等级的用户,其电-碳转换因子适中,处于中等水平,这些用户的能源利用效率一般。低碳转换因子等级:这一等级的用户,其电-碳转换因子较低,意味着他们的用电行为对碳排放的影响较小,这些用户的能源利用效率较高。

令工业用户电-碳转换因子数据集是 $Z = \{z_1, z_2, \dots, z_n\}$, 模糊 C 均值分级挖掘的目的是将 Z 分成 C 个等级。令第 i 个工业用户电-碳转换因子 z_i 隶属于 C_l 等级的隶属度是 μ_{il} , μ_{il} 需符合 $\sum_{l=1}^C \mu_{il} = 1$ 条件。

利用模糊 C 均值聚类算法,分级挖掘工业用户用电信息的目标函数为:

$$J_m(U, C) = \sum_{l=1}^C \sum_{i=1}^n m \mu_{il} d_{il}^2 \quad (16)$$

其中,模糊系数是 m ; 隶属度矩阵是 $U =$

$$\begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\ \vdots & \vdots & & \vdots \\ \mu_{C1} & \mu_{C2} & \cdots & \mu_{Cn} \end{bmatrix}; z_i \text{ 至 } C_l \text{ 的距离是 } d_{il}; \text{ 聚类}$$

中心向量矩阵是 C 。

在 $J_m(U, C)$ 最小情况下,便可获取最佳的工业用户用电信息分级挖掘结果,最佳聚类目标函数为:

$$\min [J_m(U, C)] = \min \left(\sum_{l=1}^C \sum_{i=1}^n m \mu_{il} d_{il}^2 \right) \quad (17)$$

在 $\sum_{l=1}^C \mu_{il} = 1$ 条件下,求解 $J_m(U, C)$ 的最小值,通过拉格朗日乘法,塑造函数:

$$F = \sum_{l=1}^C \sum_{i=1}^n m \mu_{il} d_{il}^2 + \sum_{l=1}^C \sum_{i=1}^n \alpha_i \mu_{il} - 1 \quad (18)$$

其中,拉格朗日乘子是 α_i 。

依据式(18)计算 μ_{il} 的偏导获取:

$$\frac{\partial F}{\partial \mu_{il}} = m(m-1) \mu_{il} d_{il}^2 + \alpha_i = 0 \quad (19)$$

由于 $\sum_{l=1}^C \mu_{il} = 1$, 那么便可获取:

$$\mu_{il} = \frac{1}{\sum_{\tau=1}^C \left(\frac{d_{i\tau}}{d_{il}} \right)^{\frac{2}{m-1}}} \quad (20)$$

在式(17)内代入式(20),便可获取具有相似电-碳转换因子的工业用户等级,完成工业用户用电信息分级挖掘。

2 实验分析

以某工业园区内的历史工业用户用电信息数据集为实验对象,该数据集内共包含三个碳转换因子等级的工业用户,分别是高碳转换因子等级、中碳转换因子等级、低碳转换因子等级。该工业园区内共包含 4 个发电机,12 个工业用户负荷节点,该工业园区的电力系统现场拓扑结构如图 1 所示。

图 1 中工业园区电力系统内使用的发电机如图 2 所示。

该工业园区内工业用户的某日部分历史用电信息如表 1 所示。

利用本文算法计算该工业园区内,各工业用户的碳排放量,以某日 12:00 的碳排放量为例,计算结果如图 3 所示。

分析图 3 可知,本文算法可有效依据工业用户历史用电信息,计算各工业用户的碳排放量,从图 3 中可以看出,编号为 15 的工业用户的碳排放量最小在 42 t 左右,编号为 5 的工业用户的碳排放量最大,在 150 t 左右。

利用本文算法依据工业用户历史用电信息,计算该日各工业用户的用电量,以编号为 4、5、6 的工业用户为例,用电量计算结果如图 4 所示。

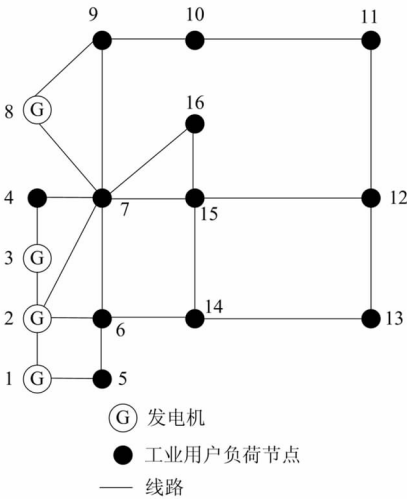


图 1 该工业园区的电力系统现场拓扑结构图

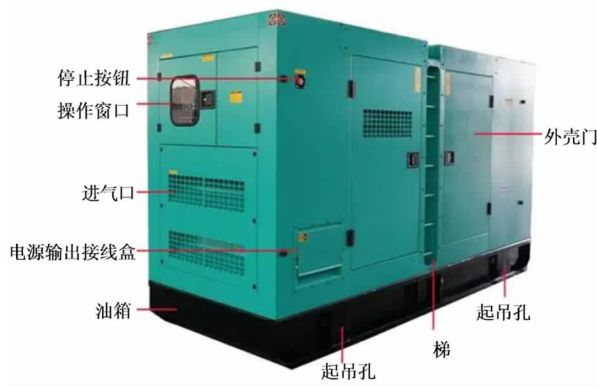


图 2 工业园区电力系统使用的发电机

表 1 工业用户的某日部分历史用电信息

用户编号	电压/V	电流/A	功率因数	设备运行时间/h	生产工艺
4	220	50	0.93	10	塑料制品加工
5	220	50	0.92	24	化工生产
6	220	50	0.95	24	机械加工
7	220	50	0.95	24	机械加工
9	220	50	0.94	24	电子制造
10	220	50	0.99	24	机械加工
11	220	50	0.98	12	食品加工
12	220	50	0.96	24	机械加工
13	220	50	0.97	12	食品加工
14	220	50	0.98	24	化工生产
15	220	50	0.96	10	塑料制品加工
16	220	50	0.98	24	机械加工

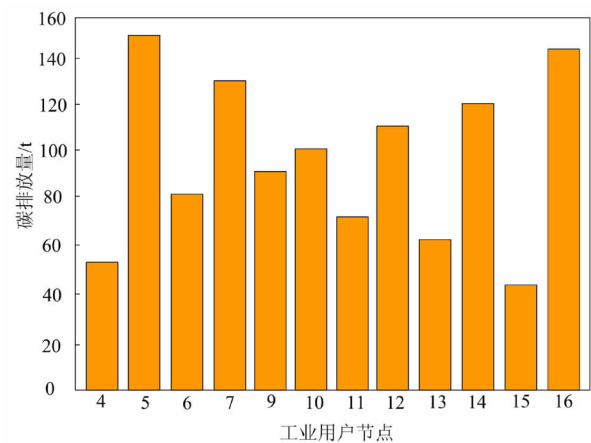


图 3 各工业用户的碳排放量计算结果

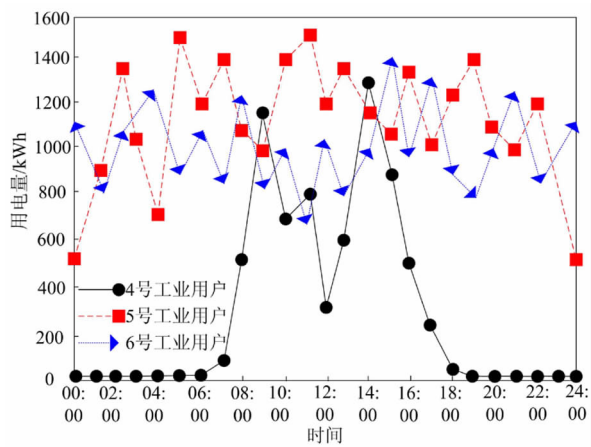


图 4 工业用户用电量计算结果

分析图 4 可知,本文算法可有效计算各工业用户一天内各时刻的用电量,其中 4 号工业用户在 7:00 至 18:00 之间(不包含 7:00 与 18:00),各时刻用电量相对较高,其余时刻的用电量非常小,原因是该工业用户每天的设备运行时间为 10 h,因此,在 7:00 至 18:00 之间的 8 个时刻用电量较高,与实际情况相符,4 号工业用户的最高用电量在 1300 kWh 左右。5 号与 6 号工业用户一天内各时刻的用电量均较高,原因是这两个工业用户每条的设备运行时间均为 24 h,与实际情况相符,5 号工业用户的最高用电量在 1500 kWh 左右,6 号工业用户的最高用电量在 1400 kWh 左右。

利用本文算法计算该工业园区内各工业用户某日的电-碳转换因子,计算结果如图 5 所示。

分析图 5 可知,本文算法可有效计算某日各工业用户的电-碳转换因子,从图 5 中可以看出,7 号工业用户的电-碳转换因子最小,在 100 g/kWh 左右,说明该工业用户的用电行为对碳排放的影响较小,其能源利用效率较高。11 号工业用户的电-碳转换因子最大,在 750 g/kWh 左右,说明该工业用户的用电行为对碳排放的影响较大,其能源利用效

率较低。

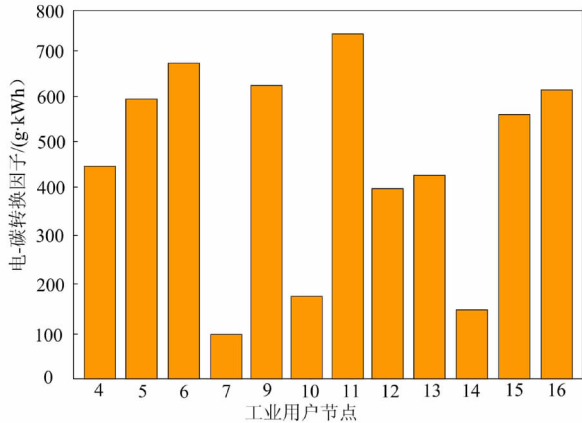


图 5 某日各工业用户的电-碳转换因子计算结果

利用本文算法对该工业园区内的工业用户用电信息进行分级挖掘,分级挖掘结果如图 6 所示。

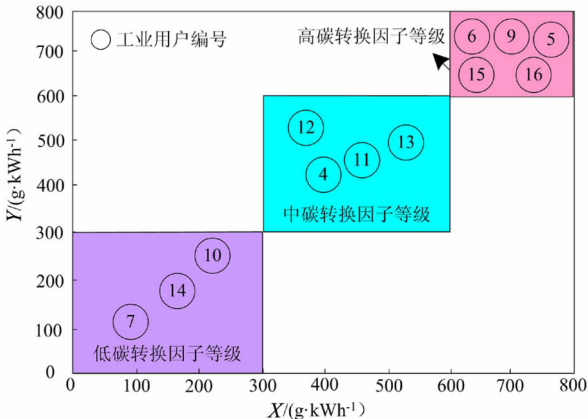


图 6 工业用户用电信息分级挖掘结果

分析图 6 可知,本文算法可有效分级挖掘工业用户用电信息,确定各工业用户的碳转换因子等级,这种分级挖掘的方法有助于更好地了解不同工业用户的碳排放情况,从而为制定相应的减排措施提供依据。其中,编号为 7、10 与 14 的工业用户属于低碳转换因子等级,说明这三个工业用户的用电行为对碳排放的影响较小,其能源利用效率较高,这可能是因为这些工业用户采用了先进的节能技术和设备,或者采用了清洁能源,从而降低了碳排放。编号为 4、11、12 与 13 的工业用户属于中碳转换因子等级,说明这四个工业用户的用电行为对碳排放的影响适中,其能源利用效率一般,这些工业用户可能需要进一步优化其生产工艺和设备,提高能源利用效率,以降低碳排放。编号为 5、6、9、15 与 16 的工业用户属于高碳转换因子等级,说明这五个工业用户的用电行为对碳排放的影响较大,其能源利用效率较低,这些工业用户需要采取更加积极的措施来降低碳排放。

3 结 论

研究了基于电-碳转换因子的工业用户用电信息分级挖掘算法,能够有效确定各工业用户的碳转换因子等级,为后续的碳排放管理和减排措施提供重要依据。通过分级挖掘,可以更清晰地了解不同工业用户的碳排放情况,从而有针对性地制定相应的减排措施,对于推动工业领域的绿色低碳发展具有重要意义。对于低碳转换因子等级的工业用户,其能源利用效率较高,但仍需继续保持并优化节能技术和设备。对于中碳转换因子等级的工业用户,其能源利用效率一般,需要进一步优化生产工艺和设备,提高能源利用效率。对于高碳转换因子等级的工业用户,其能源利用效率较低,需要采取更加积极的措施来降低碳排放,如采用更清洁的能源、优化生产工艺、提高设备效率等。

参考文献

[1] 陈静,郑垂铤,李桂敏,等. 考虑行业关联度的工业用户用电异常识别研究[J]. 仪器仪表学报,2023,44(4):72—81.

[2] 余昊杨,武昕,郭一凡,等. 基于事件感知的钢厂工业用户非侵入式负荷检测分解算法[J]. 南方电网技术,2022,16(11):29—36.

[3] 许良财,邵振国,陈飞雄. 基于 haar 小波编码和改进 K-means 算法聚合的用户负荷典型区间场景挖掘[J]. 电力自动化设备,2022,42(6):154—160.

[4] 陈逸涵,李扬,沈运帷. 基于负荷控制潜力量化模型的工业用户群体画像方法[J]. 电力自动化设备,2021,41(8):208—216.

[5] 汪颖,杨维,肖先勇,等. 基于去噪自编码器网络特征降维与改进小批优化 K 均值算法的海量用户用电行为聚类及分析[J]. 电力自动化设备,2022,42(6):146—153.

[6] 魏伟,韩颖,刘怡君,等. 基于多视角数据与社区发现的典型用电负荷模式挖掘研究[J]. 郑州大学学报(理学版),2023,55(4):68—74.

[7] YANG Q F, GAO W Y, HAN G, et al. HCDC: a novel hierarchical clustering algorithm based on density-distance cores for data sets with varying density[J]. Information Systems, 2023, 114(3):1.1—1.14.

[8] NGUYEN H, LE T, NGUYEN M, et al. Mining frequent weighted utility itemsets in hierarchical quantitative databases [J]. Knowledge-Based Systems, 2022, 237 (2): 107709. 1—107709. 13.

[9] 李扬帆,张凌浩,雷勇,等. 基于时间卷积网络和门控循环单元的短期用电量预测方法[J]. 水电能源科学,2021,39(8):198—201,173.

[10] 吴文培,宋亚林,魏上斐. 基于改进 Prophet 模型的用电量预测研究[J]. 计算机仿真,2021,38(11):473—478.

[11] 吴翔宇,荀超,肖芬,等. 基于 RF 变量选择与 LSTM 回归的长期用电量预测模型[J]. 电气传动,2023,53(5):71—76.

[12] 王新刚,朱彬若,顾臻. 基于综合能源计量大数据的中长期用电量预测[J]. 中国电力,2021,54(10):211—216.

[13] 张鸿宇,凡海波,方辉,等. 考虑负荷偏离度的零碳排放综合能源系统优化调度[J]. 电力系统及其自动化学报,2022,34(7):1—10.

[14] 张宁,李姚旺,黄俊辉,等. 电力系统全环节碳计量方法与碳表系统[J]. 电力系统自动化,2023,47(9):1—12.

[15] 袁泉,唐成亮,徐雲鹏. 基于长度约束的蝙蝠高效项集挖掘算法[J]. 计算机应用,2023,43(5):1473—1480.