

DOI: 10.19783/j.cnki.pspc.211484

基于云模型的配电网运行画像数字孪生构建方法

杜晓东¹, 曾四鸣¹, 刘科研², 詹惠瑜²

(1. 国网河北省电力有限公司电力科学研究院, 河北 石家庄 050021;

2. 中国电力科学研究院有限公司配电技术中心, 北京 100192)

摘要: 随着新型电力系统的形成, 配电网运行的不确定性和复杂性日益加大。运用数字孪生技术建立虚拟空间对物理电网的映射和交互, 是掌握配电网复杂运行特性的有效手段。提出了基于数字孪生的配电网运行画像构建方法。首先, 建立反映配电台区和线路运行状态的指标体系, 运用客观统计排序策略将特征指标转化为模糊标签模板。然后, 以日为单位提取运行指标构成采集样本, 作为云模型的一个云滴, 运用逆向云发生器形成运行指标的模糊表征。最后, 用运行指标与标签模板的模糊距离进行加权求和, 得到运行指标的打分值, 形成物理意义清晰、评价依据客观的配电网运行画像, 为全面刻画新型数字化配电网运行特征提供了技术手段。

关键词: 数字孪生; 配电网; 运行画像; 云模型; 模糊距离

Construction of portraits for a distribution network with digital twins based on a cloud model

DU Xiaodong¹, ZENG Siming¹, LIU Keyan², ZHAN Huiyu²

(1. Electric Power Research Institute of State Grid Hebei Electric Power Co., Ltd., Shijiazhuang 050021, China;

2. Power Distribution Technology Center, China Electric Power Research Institute Co., Ltd., Beijing 100192, China)

Abstract: As new power systems develop, the uncertainty and complexity of distribution network operation are increasing. Using digital twin technology to establish the mapping and interaction between virtual spaces and physical power grids is an effective way to reveal the complex operation of distribution networks. A method of constructing distribution network operation portraits based on digital twins is proposed. First, an index that reflects the operations of supply districts and feeders is established, and a statistical ranking strategy is used to transform the characteristic index into a fuzzy label template. Then, the day operation index is extracted as a unit of acquired samples and as a cloud drop of the cloud model. Fuzzy expressions of the index are obtained using a reverse cloud generator. The fuzzy distances between the index and the labels are weighted and summed to obtain the score of the operation. A distribution network operation portrait with clear physical meaning and objective evaluation basis is formed. The distribution network operation portrait also provides a technical means for comprehensively depicting the operational characteristics of the new digital distribution network.

This work is supported by the Key Research and Development Program of Hebei Province (No. 21312102D).

Key words: digital twins; distribution networks; operation portraits; cloud model; fuzzy distance

0 引言

电力系统是实现“双碳”目标的主力军, 大力发展风能、太阳能等新能源是解决碳排放问题的关键。以新能源为主体的新型电力系统面临新能源不可控性和柔性负荷的交互作用, 运行时变性和不确

定性日益突出; 电能的生产、传输、存储和消费关系更加复杂, 信息流和能量流高度耦合、实时互动。传统基于机理模型的分析 and 优化控制方法已难以满足新型电力系统规划设计、监测分析和运行优化的需要^[1-2]。依靠现代电子信息技术大力推进电力系统数字化升级是实现新型电力系统的重要保障。

数字孪生(Digital Twins, DT)是一个集成多物理量、多时间尺度、多业务流程的在线交互仿真过程^[3], 旨在挖掘海量数据资源, 在数字空间构建虚

基金项目: 河北省重点研发计划项目资助(21312102D); 国网河北省电力有限公司科技项目资助(kj2020-084)

拟模型和数字虚体与物理实体的映射关系, 通过数据的实时交互和驱动实现物理实体的镜像^[4]。电力系统高可靠性需求和高建设成本等技术特点及日益复杂的运行状况使其难以通过物理实验获得全面实时的认知, 运用数字孪生技术实现虚拟数字空间与真实物理空间的协同互动, 是构建新型电力系统的必然选择。文献[5]分析了建设电力系统数字孪生技术的必要性及DT在电力系统可能的应用场景, 对DT技术在电力和能源系统的应用前景进行了综述。文献[6-10]则从DT技术在电力系统的具体应用开展研究。文献[6]提出基于DT的多物理量实时感知实现光伏发电功率的超短期预测技术; 文献[7]提出应用DT在线预测风电波动并构建混合储能数字孪生体, 实现风电波动的优化平抑; 文献[8-9]研究了数字孪生驱动的微电网多智能体优化控制策略, 微电网多智能体按DT实时优化方案对微电网实施控制; 文献[10]提出了数字孪生在线分析系统, 用于数字化电网调度。随着新型电网的形成, 分布式新能源高比例渗透到配电网, 电动汽车等新型用电设备极大改变配电网负荷特性, 配电侧源、网、荷运行的不确定性和复杂性日益加大, 开展配电网数字孪生应用技术十分必要。

用户用电行为会极大影响配电网的运行行为, 用户用电行为分析在电网负荷预测、用电管理、需求响应、运行态势分析和安全预警等方面具有重要作用^[11-12]。由于用户行为随季节和外界条件的变化而改变, 运用DT对用户运行数据进行实时感知, 动态捕捉用户的行为特性, 是DT在电力系统的典型应用前景之一^[5]。现有文献尚未开展这方面的研究。用户画像是一种体现用户行为模式的标签化描述, 是分析用户行为、挖掘用户价值、实施精细管理的有利工具^[13-14]。文献[13]提出了基于行为感知的用户画像技术, 通过感知用户行为, 挖掘用户特性, 构建用户画像。文献[14]通过挖掘互联网用户行为的关注特性, 提取用户行业标签, 完善了互联网用户画像系统。用户画像技术也被广泛应用于电力系统领域。文献[15]提出用户用电模式的模糊C-均值聚类算法, 实现综合能源系统中不同类型用户的多属性用电行为模式画像。文献[16-17]提出配网用户群体用电行为效用模型, 用于综合能源系统用能决策。文献[11-12]提出用电特征的优选准则并基于该准则提出用电特征优选策略。文献[18-20]基于用户行为的具体样本作为分类数据, 运用聚类方法建立用户行为画像。由于用户的具体行为样本具有随机性, 不能从普适性上反映用户行为的一般画像; 同时, 刻画用户画像的标签划分通常具有模糊性, 在

聚类方法中应考虑这种模糊特性。

单一用户的用电行为因个体容量小、特性分散且群体数量庞大而不利于反映配电网的运行特性, 以台区为单元进行画像更具实用性。本文运用数字孪生的强大感知能力、镜像能力和信息平台, 以孪生配电网为画像对象和数据平台, 建立了包含台区和配电线路的运行画像方法, 为全面刻画配电网运行特征提供了技术手段。首先建立了反映配电台区和线路运行行为的特征量并形成模糊标签模板。以日周期为样本单元, 用孪生电网峰、谷等特征时段运行指标的不同时间域大数据输入逆向云发生器, 产生随机样本云滴群的云模型参数。采用运行指标与标签模板的模糊距离进行加权求和, 得到运行指标的打分值, 形成物理意义清晰、刻画尺度精细、评价依据客观的配电网运行画像。

1 基于数字孪生的配电网画像原理

1.1 数字孪生配电网

DT的概念于2003年由美国学者Michael Grieves提出, 工程中DT的早期应用集中于航空航天领域。2011年3月, 美国国家航空航天局和美国空军研究实验室相继利用DT技术在数字空间建立作业飞行器的虚拟模型, 并通过传感器技术实现两者的状态同步, 对飞行器的运行情况、健康状态、载荷能力等进行评估。

数字孪生配电网是配电网模型日渐复杂、数据呈现井喷趋势以及DT技术发展完善等多方背景共同作用下的新兴产物。相比于侧重实时操控实体的信息物理系统(Cyber-Physical Systems, CPS)或经典模型驱动的仿真软件, 数字孪生电网更侧重于数据驱动的实时态势感知(Real-Time Situation Awareness)和超实时虚拟推演(Ultra-Real-Time Virtual Test), 旨在为电力系统的运管调控决策提供参考。数字孪生电网在数字空间创造一个与物理实体电网匹配对应的数字电网, 通过全息模拟、动态监控、实时诊断、精准预测反映物理实体电网在现实环境中的状态, 进而推动电网全要素数字化和虚拟化、全状态实时化和可视化、电网运行管理协同化和智能化, 实现物理电网与数字电网协同交互、平行运转。

数字孪生配电网从物理电网量测感知各类电气量、状态量样本序列。以配网机理模型为基础, 依托大数据分析平台, 以数据云计算服务形式为配电网运行画像云模型提供数字化载体; 以微服务的形式提供配网监测、预警、控制、运维等服务; 以数字化系统用户界面形式在配电网各类三维模型或系统图中展示画像结果。

1.2 基于数字孪生的配电网画像实现机理

基于数字孪生技术构建配电网运行画像的总体实现思路如图 1 所示。主要包括以下环节。

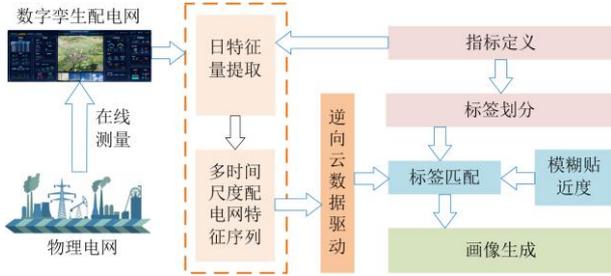


图 1 数字孪生配电网运行画像实现路线图

Fig. 1 Digital twin distribution network operation implementation roadmap

(1) 配电网运行状态的感知。数据中台按一定的采样间隔实时采集各台区日负荷状态，形成台区日负荷状态量的量测序列，用矩阵 P_T 表示，如式(1)所示。

$$P_T = \begin{bmatrix} P_{1,1}^T & P_{1,2}^T & \cdots & P_{1,t}^T & \cdots \\ P_{2,1}^T & P_{2,2}^T & \cdots & P_{2,t}^T & \cdots \\ \vdots & \vdots & & \vdots & \\ P_{k,1}^T & P_{k,2}^T & \cdots & P_{k,t}^T & \cdots \end{bmatrix} \quad (1)$$

式中： k 为台区编号； t 为数据采样时段，若采样间隔为 15 min，则日采样序列为 96 点； $P_{k,t}^T$ 为第 k 个台区在 t 时段采集的负荷数据。

孪生电网与数据中台交互数据，在物理电网运行数据驱动下进行潮流分析，得到采样时刻各配电网的潮流状况数据序列，用矩阵 P_L 表示，如式(2)所示。

$$P_L = \begin{bmatrix} P_{1,1}^L & P_{1,2}^L & \cdots & P_{1,t}^L & \cdots \\ P_{2,1}^L & P_{2,2}^L & \cdots & P_{2,t}^L & \cdots \\ \vdots & \vdots & & \vdots & \\ P_{k,1}^L & P_{k,2}^L & \cdots & P_{k,t}^L & \cdots \end{bmatrix} \quad (2)$$

式中， $P_{k,t}^L$ 为在物理电网运行数据驱动下，孪生电网第 k 个台区在 t 时段通过潮流分析得到的负荷数据。

(2) 定义反映台区和线路状态的标签体系。画像技术通过使用标签描述事物的一系列特征，用高度概括的概念化、易理解、可量化的标签体系来刻画对象。因此，首先需建立反映台区用电行为和线路运行状态的特征指标，并对指标值按一定的物理意义进行区间量化，从而使各量化区间对应一个概念，即标签。例如，若以台区用电水平为特征指标，可将台区的日用电量分为高、中、低三个等级的标签。显然，标签通常具有模糊属性，是用概念化的语言进

行描述。

(3) 基于云模型的配电网运行特征概念化归纳。根据式(1)、式(2)表示的日采样序列并根据配电网的指标定义可计算日运行特征指标，每天可形成一个指标样本。同时，考虑台区的日运行行为受人的行为及天气、气温等状况的影响，具有很大的随机性和分散性^[21-23]，因而基于单一的日样本对配电网进行画像不能反映对象的宏观特征。本文采用云模型理论作为数据驱动手段，基于一定规模的日样本生成配电网的定性概念模型。运用逆向云发生器将历史样本集合转化为样本的期望、熵和超熵，从而将样本序列转化为具有模糊特性和统计特性的概念化表达。可见，逆向云模型的输出参数反映配电网一般运行过程的宏观行为和特点，单一的日随机样本对结果的影响很小。

(4) 配电网画像的生成。如(3)所述，标签体系中各标签的定义具有模糊性，是用模糊语言表示的模糊概念。历史样本特征序列通过逆向云转化，也形成模糊集合。构建配电网画像，就是根据配电网各元件的模糊特征，识别这个模糊对象属于哪类模糊标签的问题。本文利用模糊概念的贴适度方法，将模糊对象对标上文定义的模糊标签，从而实现配电网各元件运行行为的画像生成。

上述分析可见，配电网画像生成过程就是配电网运行特征的抽象和标签匹配过程。配电网画像生成过程中特征的抽象和处理层次结构如图 2 所示。

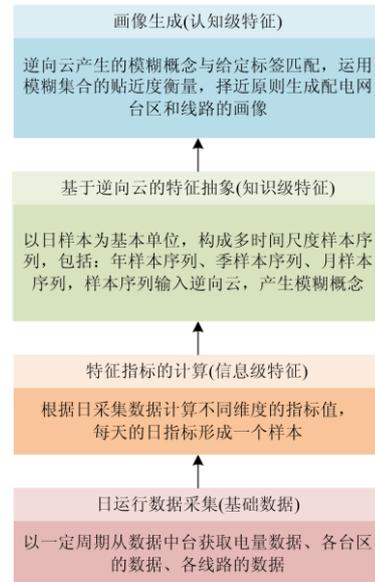


图 2 配电网运行特征的抽象过程

Fig. 2 Abstract process of distribution network operation characteristics

2 配电网标签模板的选取及模糊化

2.1 配电网特征量的构建

2.1.1 台区用电特征指标构建

台区用电特征以日为周期进行分析, 用一组反映日用电曲线特征的向量表示。每日可形成一个样本, 建立一组特征向量。特征指标的选择既要保证指标的完备性, 又要尽量降低指标的冗余性。指标的完备性就是要保证构建的一组指标能够完全区分各种负荷波形。同时, 指标应具有独立性和优选性^[11-12], 以降低指标之间的冗余。结合文献[24-26], 本文选取日负荷率、日峰谷差率、峰期负荷率、谷期负荷率、平期负荷率作为台区用电特性指标。同时, 近年来电网的尖峰负荷越来越大, 本文定义台区的尖峰电力贡献率和尖峰电量贡献率指标, 用来反映台区对配电网尖峰负荷的贡献, 各指标定义见表 1。其中, $P(t)$ 为 t 时刻的功率值, P_N 、 P_{\max} 、 P_{\min} 、 P_{pk} 和 P_{ave} 分别为台区日负荷额定、最大、最小、尖峰和平均功率, T 为全天时长, t_1 、 t_3 、 t_5 和 t_7 分别为日负荷峰期、谷期、平期和尖峰的开始时间, t_2 、 t_4 、 t_6 和 t_8 分别为日负荷峰期、谷期、平期和尖峰的结束时间。日负荷率反映台区全天的负荷用电总量, 体现了台区对报装容量的利用情况; 日峰谷差率反映日用电峰谷波动特性; 峰、谷、平负荷率反映台区在峰谷平不同时间段的用电情况; 尖峰负荷电力贡献率用尖峰时段内台区的最大功率减去台区日平均功率, 反映台区对配电网形成尖峰负荷的功率贡献; 尖峰负荷电量贡献率反映台区对尖峰电量的贡献。配电网用电特性受地区、具体电网和季节的影响, 用电规律具有差异性。相应的峰、谷、平时段需根据配电网的用电规律确定, 例如河北电网夏季的低谷时段定为 00:00—08:00; 高峰时段定为 12:00—16:00、17:00—18:00; 其余时段为平时段。

表 1 负荷特征指标

Table 1 Load characteristic index

特征指标	指标符号	指标定义
日负荷率	x_1	$\int_0^T P(t)dt / (P_N T)$
日峰谷差率	x_2	$(P_{\max} - P_{\min}) / P_N$
峰期负荷率	x_3	$\int_{t_1}^{t_2} P(t)dt / [P_N (t_2 - t_1)]$
谷期负荷率	x_4	$\int_{t_3}^{t_4} P(t)dt / [P_N (t_4 - t_3)]$
平期负荷率	x_5	$\int_{t_5}^{t_6} P(t)dt / [P_N (t_6 - t_5)]$
尖峰电力贡献率	x_6	$(P_{pk} - P_{ave}) / P_N$
尖峰电量贡献率	x_7	$\int_{t_7}^{t_8} [P(t) - P_{ave}]dt / [P_N (t_8 - t_7)]$

2.1.2 线路载荷特征指标构建

配电线路的载荷状态、利用率等指标是考核配电网建设运营成效和反映线路重要度的关键^[27-28]。配电线路运行指标除与自身结构及运行方式有关外, 主要受台区群体用电行为的影响。为提高孪生配电网的态势分析能力, 使孪生配电网从台区到线路全面反映配电网运行特性, 本文在对配电网台区画像的同时, 也建立了配电线路的画像。

考虑配电线路的功能属性, 本文选取日平均载荷、日最大载荷、日最小载荷、日电量传送率、日载荷超限率作为配电线路的特征指标, 各指标定义见表 2。日平均载荷反映线路的平均载荷状态, 体现线路的利用率; 日最大、最小载荷反映线路载荷的极值状态, 对了解线路的运行边界有帮助; 日电量传送率是指线路日传送的有功电能量, 可用来体现线路的运行绩效; 日载荷超限时间是指线路载荷超过某一给定安全阈值的时间与全天时长的比值, 用来统计配电网在不确定随机运行过程中超过安全域的状况。线路特征量的数据来源是基于与物理配电网同步运行的孪生配电网运行分析而生成的。

表 2 线路载荷特征指标

Table 2 Line load characteristic index

特征指标	指标符号	指标定义
日平均载荷	x_8	$\int_0^T S(t)dt / (S_N T)$
日最大载荷	x_9	S_{\max} / S_N
日最小载荷	x_{10}	S_{\min} / S_N
日电量传送率	x_{11}	$\int_0^T P(t)dt / (S_N T)$
日载荷超限时间	x_{12}	KT_0 / T

表 2 中, $S(t)$ 和 $P(t)$ 分别为 t 时刻线路的视在功率和有功功率值, S_N 、 S_{\max} 和 S_{\min} 分别为线路的额定、最大和最小视在功率, K 为线路载荷超过安全阈值的次数, T_0 为配网数据采集时间间隔。

2.2 特征指标的标签化和模糊化

表 1 和表 2 定义的特征指标从不同角度反映配电台区和线路的运行特征, 并提供了计算式, 可以通过孪生配电网日运行数据计算出指标向量值。但特征指标值仅能够从数域空间给出量值, 缺乏语义理解, 没有用概念形式定义指标值高低或优劣的含义。对指标不同取值赋予概念化语义, 称为标签; 对指标定义一组标尺性参照标签, 称为标签模板。

表 1 和表 2 建立的特征指标具有不同属性, 其中日负荷率、谷期负荷率、平期负荷率为效益型指标, 指标值越大越好。日峰谷差率、峰期负荷率、

尖峰电力贡献率、尖峰电量贡献率为成本型指标，指标值越小越好。将每个指标按量值分别划分为“优、良、中、差”4个等级，形成该指标的一组标签模板。显然，这些标签具有模糊属性，需建立标签的模糊隶属度。常用隶属度函数有矩形、三角形、梯形等分段线性函数以及抛物线、正态等非线性函数。为方便隶属度函数的建立，本文采用三角形分布。对于效益型指标，每组标签的隶属度分布如图3所示；对于成本型指标，每组标签的隶属度分布如图4所示。其中 x 为指标值， b_1 — b_4 为效益型指标隶属度函数的分段点； c_1 — c_4 为成本型指标隶属度函数的分段点。图3所示效益型指标的梯形分布隶属度函数表达式如式(3)一式(6)。成本型指标的梯形隶属度函数表达式与式(3)一式(6)类似，不再赘述。

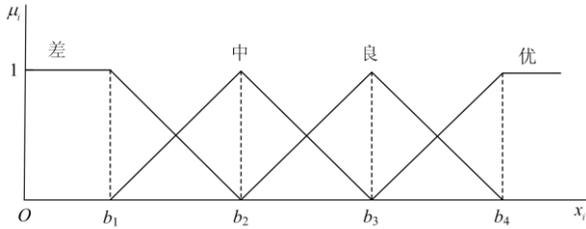


图3 效益型指标的隶属度分布

Fig. 3 Membership distribution of benefit type index

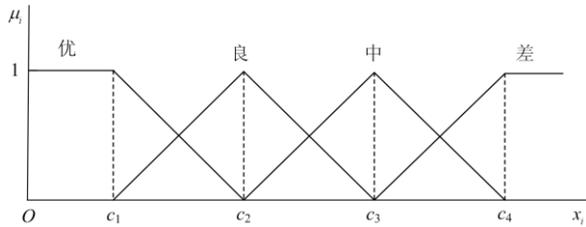


图4 成本型指标的隶属度分布

Fig. 4 Membership distribution of cost type index

$$\mu_i(x_i) = \begin{cases} 1, & x_i < b_1 \\ \frac{b_2 - x_i}{b_2 - b_1}, & b_1 \leq x_i \leq b_2 \\ 0, & x_i > b_2 \end{cases} \quad (3)$$

$$\mu_i(x_i) = \begin{cases} 0, & x_i < b_1 \\ \frac{x_i - b_1}{b_2 - b_1}, & b_1 \leq x_i < b_2 \\ \frac{b_3 - x_i}{b_3 - b_2}, & b_2 \leq x_i < b_3 \\ 0, & x_i \geq b_3 \end{cases} \quad (4)$$

$$\mu_i(x_i) = \begin{cases} 0, & x_i < b_2 \\ \frac{x_i - b_2}{b_3 - b_2}, & b_2 \leq x_i < b_3 \\ \frac{b_4 - x_i}{b_4 - b_3}, & b_3 \leq x_i < b_4 \\ 0, & x_i \geq b_4 \end{cases} \quad (5)$$

$$\mu_i(x_i) = \begin{cases} 0, & x_i < b_3 \\ \frac{x_i - b_3}{b_4 - b_3}, & b_3 \leq x_i \leq b_4 \\ 1, & x_i > b_4 \end{cases} \quad (6)$$

隶属函数的参数确定通常采用主观法或客观法。主观法依据专家打分确定，结果依赖于专家的经验，主观性强。客观法则通过实验或实测数据确定较符合实际的参数。本文采用客观法，基于电网运行大数据采用统计的方法确定隶属函数。电网大数据平台记录了多个电网大量运行数据，从中选取一些典型电网的历史数据作为统计样本，分别计算统计样本的指标并逐指标由小到大排序。依据排序结果按表3所示百分比确定各指标的隶属度参数 b_1 — b_4 或 c_1 — c_4 。

表3 百分比次序对应的隶属度参数

Table 3 Indicator value corresponding to percentage rank

百分比	10%	40%	70%	100%
效益型	b_1	b_2	b_3	b_4
成本型	c_1	c_2	c_3	c_4

3 基于逆向云的配电网运行不确定性建模

3.1 基于云模型的定量样本与定性概念转换原理

云模型理论是基于模糊论和概率论提出的一种数域空间中大量数据样本与概念域空间中定性概念的不确定性相互转换工具。在复杂的评价问题中有些情况下可获得对象的一定规模定量历史样本；有些情况下则无法得到对象的精确历史样本，只能用概念化模糊语言定性描述。云模型就是实现模糊定性概念与精确数值样本集合相互生成的方法。通常把被分析的样本称为云滴，由云滴集合统计特性生成的概念用云模型参数表征。云模型由3个参数组成，分别用 Ex 、 En 、 He 表示。 Ex 为云的期望，它反映被统计云滴集合的重心，由数域空间中最能代表定性概念的点构成； En 为云的熵，表示云概念的不确定度，是定性概念模糊性和随机性的综合度量，体现了定性概念亦此亦彼性的裕度； He 为云的超熵，表示云熵的不确定度，反映了概念在数域空间中不确定性的发散度。正态云模型是在正态分布和模糊隶属函数的基础上发展起来的用正态函数表示

的基本云模型, 具有普适性^[29], 是应用最广泛的云模型形式。云模型的转换算法称为云发生器(Cloud Generator, CG), 既可以用软件的方式实现, 又可以固化成硬件实现。云发生器分为正向云发生器和逆向云发生器。由给定的云模型参数产生特定数量云滴的过程称为正向云发生器, 如图 5 所示, 其中 (x_i, μ_i) 为一个云滴, x_i 是定性概念在论域中该次实现的具体数值, μ_i 为 x_i 属于这个语言值的程度。CG 的具体算法参见文献[29]。

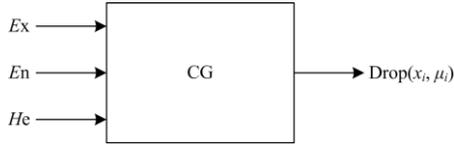


图 5 正向云发生器

Fig. 5 Forward cloud generator

对云滴集合抽象出定性概念的过程称为逆向云发生器。如图 6 所示, 逆向云发生器是实现从定量数值到定性概念的转换算法, 基于一定数量的样本数据抽象出以数字特征表示的云模型概念。

逆向云发生器的输入为 N 个云滴的定量值及每个云滴代表概念的确定度 (x_i, μ_i) , 输出为 N 个云滴表示的定性概念 A 的期望值 Ex 、熵 En 和超熵 He , 算法如下。

$$(1) \text{ 由 } x_i \text{ 计算这组数据的样本均值 } \bar{X} = \frac{1}{n} \sum_{i=1}^n x_i,$$

一阶样本绝对中心矩 $\frac{1}{n} \sum_{i=1}^n |x_i - \bar{X}|$, 样本方差

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2.$$

(2) 计算期望

$$Ex = \bar{X} \quad (7)$$

(3) 按照式(8)计算熵。

$$En = \sqrt{\frac{\pi}{2}} \times \frac{1}{n} \sum_{i=1}^n |x_i - \bar{X}| \quad (8)$$

(4) 按照式(9)计算超熵。

$$He = \sqrt{S^2 - En^2} \quad (9)$$

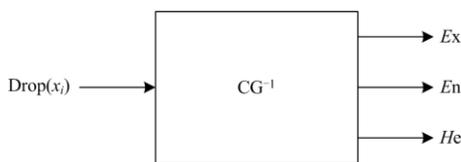


图 6 逆向云发生器

Fig. 6 Backward cloud generator

3.2 配电网运行数据驱动的云模型特征建模

配电网运行具有不确定性, 且随着季节等外部

条件的变化而变化, 用单一画像很难恰当描述配电网不同外部态势的运行特征。本文建立配电网不同时间尺度的画像模型, 用年画像表示配电网在 1 年的时间尺度上具有的一般特征, 用季或月画像表示配电网在季度或月份时间尺度上具有的一般特征, 建立配电网不同时间尺度的画像。本文从数字孪生系统获取历史样本数据, 采用云模型数据驱动的画像方法。历史样本集合可采用年度、季度、月度等多时间域样本作为数据源。不同时间尺度的画像需要利用相对应的不同时间尺度历史运行数据进行驱动, 用年数据建立年画像, 用季或月数据建立季画像和月画像。基于云模型数据驱动方法建立孪生配电网不同时间尺度特征模型的计算步骤如下。

(1) 基于孪生配电网中台数据构建样本序列。样本序列包括年度样本数据、季度样本数据和月度样本数据三类。设待分析样本序列为 $X = (X_1, X_2, \dots, X_k, \dots, X_N)$, X_k 为以日为单位的一个样本, 样本总数量为 N 。设选取的历史样本数据为 Y 年, 则年度画像的样本数 $N = 365 \times Y$, 季度画像的样本数 $N = 90 \times Y$, 月度画像的样本数 $N = 30 \times Y$, 分别利用不同数据生成年、季、月画像。从孪生配电网数据中台获取分析周期内台区日电量数据和配电线路日潮流数据, 形成样本序列基础数据。

(2) 计算样本的指标值。对样本序列的各样本按表 1、表 2 的定义分别计算台区和线路的指标 $x_1 \sim x_{12}$, 从而得到日样本 k 的指标向量 $\mathbf{X}_k = (x_{1k}, x_{2k}, \dots, x_{12k})$, 形成一个云滴。

(3) 计算云模型参数。以指标 i 的值 x_i , $i = 1, 2, \dots, 12$ 为云滴, 对样本序列 X 的 N 个云滴按式(7)~(9)计算该指标的云模型期望、熵和超熵等参数。

(4) 计算云模型的隶属度 μ_i 。用 3 个参数描述的云模型无确定的边界但有整体形状, 根据云模型参数可确定云的期望曲线如式(10)所示^[29-30]。期望曲线代表不同熵值下云滴群分布的期望, 因而用云的期望曲线可以表征云滴群属于概念 Ex_i 的隶属度, 从而得出指标 x_i 的模糊概念 (Ex_i, μ_i) 。

$$\mu_i = e^{-\frac{(x_i - Ex_i)^2}{2En_i^2}} \quad (10)$$

(5) 重复步骤(3)和(4), 得到基于 N 维大数据的各特征指标模糊概念 $(Ex_1, \mu_1) \sim (Ex_{12}, \mu_{12})$ 。

4 配电网运行行为的画像方法构建

本文第 2 节对每个特征指标建立了 4 个标签模板, 为配电网标签属性的评判提供了参照体系。在

3.2 节中基于数字孪生配电网运行数据将日指标作为云滴, 利用逆向云发生器建立了数据驱动的定性概念生成方法, 归纳形成了反映配电网实际运行水平的统计性、模糊性描述。本节研究云模型模糊指标的标签属性匹配方法, 从而实现配电网画像。

设建立的“优、良、中、差”标签模板用模糊幂集 $\{A_i\}$, $i=1,2,3,4$ 表示, 逆向云生成的运行指标用模糊集合 B 表示。则问题转化为研究模糊集合 B 与模糊集合 A_i 的归属关系。如前所述, 在建立标签的隶属函数时采用统计排序方式, 这样就可以用标签代表的名次对标签打分。用 10 分制表示, 设 A_1 为 10 分; A_2 为 7 分; A_3 为 4 分, A_4 为 1 分, 分值代表了对象在该指标下的运行质量。本文用模糊距离进行加权打分。评判两个模糊集合的模糊距离常用 Hamming 距离计算。Hamming 距离的表达式如式(11)所示。

$$D(A_i, B) = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} |A_i(x) - B(x)| dx \quad (11)$$

分别按式(11)计算 B 与 A_i 的距离 $D(A_i, B)$, 并计算相邻模糊集 A_i 与 A_j 之间的距离 $D(A_i, A_j)$ 。设 A_i 的分值为 $G(A_i)$, B 位于 A_i 与 A_j 之间, 则 B 的加权分数为

$$G(B) = G(A_i) \frac{D'(A_j, B)}{D(A_i, A_j)} + G(A_j) \frac{D'(A_i, B)}{D(A_i, A_j)} \quad (12)$$

式中: $D'(A_i, B)$ 和 $D'(A_j, B)$ 分别为距离 $D(A_i, B)$ 和 $D(A_j, B)$ 在模糊集 A_i 与 A_j 组成的空间上的投影, 分别按照式(13)和式(14)计算。

$$D'(A_i, B) = \frac{D^2(A_i, A_j) + D^2(A_i, B) - D^2(A_j, B)}{2D(A_i, A_j)} \quad (13)$$

$$D'(A_j, B) = \frac{D^2(A_i, A_j) + D^2(A_j, B) - D^2(A_i, B)}{2D(A_i, A_j)} \quad (14)$$

根据式(12)可以得出待评判概念 B 的分值信息。该分值的物理含义是表征对象运行水平在一般统计意义的排名情况, 也能够定性表示对象运行性能的“优、良、中、差”水平, 从而清晰地刻画了对象对于该项指标的画像。依据式(12)对所有指标计算分值, 得出对象在各个指标维度下的画像。

配电网的运行具有随机性和不确定性, 因而建立的物理画像也具有分散性。运行不确定性越强, 物理画像的离散程度就越大。仅从物理特性维度建立的配电网运行画像无法描绘画像的不确定特性。

因此, 本文在建立物理画像的同时, 也给出配电网运行的不确定性指标, 从不确定性的维度反映配电网运行画像的分散性。配电网运行的云模型参数不但能够表明运行特征, 也可以描述配电网运行的不确定性, 本文用方差来表示台区运行的不确定性。根据云模型的概念, 方差计算如式(15)所示。

$$DX = En^2 + He^2 \quad (15)$$

5 算例分析

采用 IEEE 33 节点电网为算例, 以台区画像为例进行分析。IEEE 33 配电网结构如图 7 所示, 对应线路参数见表 4。配电台区额定容量均为 400 kVA, 线路最大输送容量相等, 均为 6 MVA。选取某地区 2009 年 1 月 1 日至 2014 年 12 月 31 日共 2 190 天的负荷数据, 每 15 min 一个采样点, 一天 96 点。从中随机抽取 33 组日负荷数据作为 33 个台区的典型负荷曲线。例如台区 1 的日负荷曲线如图 8 所示。对 33 个台区的日负荷数据计算各自的特征指标, 每个台区得到 7 组指标数据。台区 1 的指标数据如表 5 所示, 限于篇幅其余台区未列出。对每个台区的 7 个指标分别从 -15%~15% 按等概率抽取一个随机数。每个台区抽取 3 000 组, 形成 3 000 个样本或云滴。对 33 个台区共 99 000 个样本进行排序, 按表 3 排序, 得到效益型指标和成本型指标的隶属度参数, 如表 5 所示。

表 4 IEEE33 节点系统线路参数

节点 i	节点 j	支路阻抗/p.u.	节点 i	节点 j	支路阻抗/p.u.
0	1	0.0922+j0.047	16	17	0.7320+j0.5740
1	2	0.4930+j0.2511	1	18	0.1640+j0.1565
2	3	0.3660+j0.1864	18	19	1.5042+j1.3554
3	4	0.3811+j0.1941	19	20	0.4095+j0.4784
4	5	0.8190+j0.7070	20	21	0.7089+j0.9373
5	6	0.1872+j0.6188	2	22	0.4512+j0.3083
6	7	0.7144+j0.2351	22	23	0.8980+j0.7091
7	8	1.0300+j0.7400	23	24	0.8960+j0.7011
8	9	1.0440+j0.7400	5	25	0.2030+j0.1034
9	10	0.1966+j0.065	25	26	0.2842+j0.1447
10	11	0.3744+j0.1238	26	27	1.0590+j0.9337
11	12	1.4680+j1.1550	27	28	0.8042+j0.7006
12	13	0.5416+j0.7129	28	29	0.5075+j0.2585
13	14	0.5910+j0.5260	29	30	0.9744+j0.9630
14	15	0.7463+j0.5450	30	31	0.3105+j0.3619
15	16	1.2890+j1.7210	31	32	0.3410+j0.5302

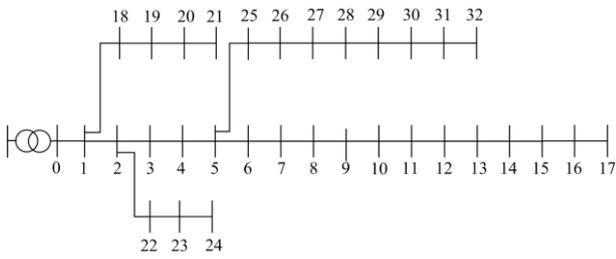


图 7 IEEE33 配电网结构图

Fig. 7 IEEE33 distribution network structure diagram

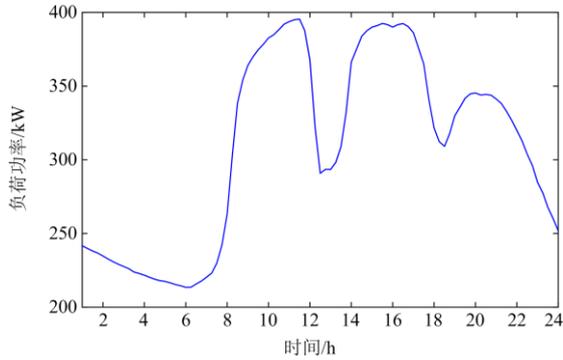


图 8 IEEE33 台区 1 的日负荷曲线

Fig. 8 Daily load curve of station area 1

表 5 台区 1 效益型指标和成本型指标的隶属度参数
Table 5 Membership parameters of benefit index and cost index in station area 1

指标名称	优	良	中	差
x_1	0.844 2	0.612 2	0.534 4	0.400 6
x_2	0.189 3	0.299 3	0.352 2	0.812 5
x_3	0.475 4	0.641 8	0.734 9	0.984 4
x_4	0.736 4	0.512 1	0.440 4	0.337 9
x_5	0.826 1	0.597 6	0.521 4	0.393 7
x_6	0.090 9	0.137 8	0.164 6	0.265 7
x_7	0.384 8	0.515 2	0.595 9	0.804 5

将每个台区随机生成的 3 000 组样本的特征指标输入逆向云发生器, 得出指标的云模型参数。求得台区 1 的各项指标云模型参数如表 6 所示。

表 6 台区 1 各项指标云模型参数

Table 6 Cloud model parameters of various indicators in station area 1

指标名称	Ex	En	He
x_1	0.551 5	0.116 9	0.041 4
x_2	0.308 5	0.084 0	0.018 8
x_3	0.657 9	0.142 8	0.051 3
x_4	0.462 4	0.100 8	0.031 0
x_5	0.537 9	0.113 2	0.041 3
x_6	0.144 2	0.039 3	0.003 0
x_7	0.534 2	0.119 3	0.043 2

将以上数据代入式(11)、式(12), 得到台区 1 各指标的得分值, 如图 9 所示。同理, 可以得出其他各台区的各指标得分情况, 从而形成各台区的运行画像。

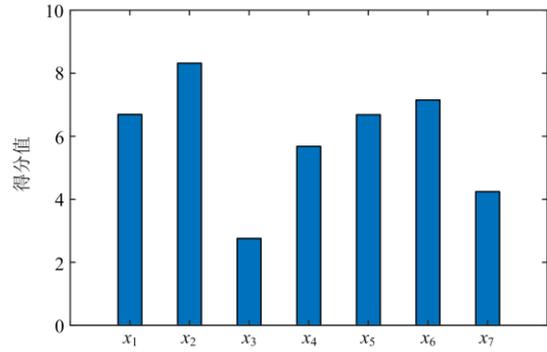


图 9 台区 1 各项指标的得分值

Fig. 9 Score value of each index of station area 1

可见, 运用本文方法能够有效构建配电网运行画像; 由于采用逆向云进行数据驱动的运行建模, 运行画像可有效刻画分析时域内配电网的一般运行规律; 以标签模板为基准, 对各指标画像采用打分的方式表征, 能够使画像分辨粒度更加细致。同时, 依据样本排序百分比对标签模板构建隶属度, 使画像的物理意义清晰, 评判更加客观和可操作。

6 结论

本文提出一种基于云模型的数字孪生配电网运行画像数据驱动构建方法, 得出以下主要结论。

(1) 应用逆向云发生器对孪生配电网运行大数据进行概念化建模, 能够有效提取配电网不同时间域宏观运行特征。

(2) 构建一种将运行指标与标签模板进行模式匹配并加权打分的配电网运行画像表征方式, 具有刻画尺度精细、评价依据客观的特点。

(3) 标签模板的隶属度模型参数运用配电网运行数据采用基于统计的客观法获得, 物理意义清晰、结果客观实用。

参考文献

- [1] 余晓丹, 徐宪东, 陈硕翼, 等. 综合能源系统与能源互联网简述[J]. 电工技术学报, 2016, 31(1): 1-13.
YU Xiaodan, XU Xiandong, CHEN Shuoyi, et al. A brief review on integrated energy system and energy internet[J]. Transactions of China Electrotechnical Society, 2016, 31(1): 1-13.
- [2] ERDIWANSYAH, MAHIDIN, HUSIN H, et al. A critical review of the integration of renewable energy sources

- with various technologies[J]. Protection and Control of Modern Power Systems, 2021, 6(1): 37-54.
- [3] BOSCHERT S, ROSEN R. Digital twin-the simulation aspect[J]. Mechatronic Futures, 2016, 1: 59-74.
- [4] GOPAL C, MANJU K, MOHAMED E. Digital twin technology[M]. CRC Press, 2021.
- [5] 贺兴, 艾芊, 朱天怡, 等. 数字孪生在电力系统应用中的机遇和挑战[J]. 电网技术, 2020, 44(6): 2009-2019.
HE Xing, AI Qian, ZHU Tianyi, et al. Opportunities and challenges of the digital twin in power system applications[J]. Power System Technology, 2020, 44(6): 2009-2019.
- [6] 崔佳豪, 毕利. 基于混合神经网络的光伏电量预测模型的研究[J]. 电力系统保护与控制, 2021, 49(13): 142-149.
CUI Jiahao, BI Li. Research on photovoltaic power forecasting model based on hybrid neural network[J]. Power System Protection and Control, 2021, 49(13): 142-149.
- [7] PARGMANN H, EUHAUSEN D, FABER R. Intelligent big data processing for wind farm monitoring and analysis based on cloud-technologies and digital twins: a quantitative approach[C] // IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), April 20-22, 2018, Chengdu, China: 233-237.
- [8] LOPEZ J, RUBIO J E, ALCARAZ C. Digital twins for intelligent authorization in the B5G-enabled smart grid[J]. IEEE Wireless Communications, 2021, 28(2): 48-55.
- [9] 王环, 曾国强, 戴瑜兴. 基于 APEO 的分布式电源改进型下垂优化控制策略研究[J]. 电力系统保护与控制, 2020, 48(2): 68-75.
WANG Huan, ZENG Guoqiang, DAI Yuxing. Research on modified droop control of distributed generation units by adaptive population-based extremal optimization[J]. Power System Protection and Control, 2021, 45(7): 2483-2491.
- [10] ZHOU M, YAN J, FENG D. Digital twin framework and its application to power grid online analysis[J]. Journal of Power and Energy Systems, 2019, 5(3): 391-398.
- [11] 陆俊, 朱炎平, 彭文昊, 等. 智能用电用户行为分析特征优选策略[J]. 电力系统自动化, 2017, 41(5): 58-63, 83.
LU Jun, ZHU Yanping, PENG Wenhao, et al. Feature selection strategy for electricity consumption behavior analysis in smart grid[J]. Automation of Electric Power Systems, 2017, 41(5): 58-63, 83.
- [12] 龚钢军, 陈志敏, 陆俊, 等. 智能用电用户行为分析的聚类优选策略[J]. 电力系统自动化, 2018, 42(2): 58-63.
GONG Gangjun, CHEN Zhimin, LU Jun, et al. Clustering optimization strategy for electricity consumption behavior analysis in smart grid[J]. Automation of Electric Power Systems, 2018, 42(2): 58-63.
- [13] 尤明辉, 殷亚凤, 谢磊, 等. 基于行为感知的用户画像技术[J]. 浙江大学学报(工学版), 2021, 55(4): 608-614, 638.
YOU Minghui, YIN Yafeng, XIE Lei, et al. User profiling based on activity sensing[J]. Journal of Zhejiang University (Engineering Science), 2021, 55(4): 608-614, 638.
- [14] 邓春燕, 郭强, 林青轩, 等. 一种基于用户关注行为的标签预测方法研究[J]. 上海理工大学学报, 2021, 43(3): 313-318.
DENG Chunyan, GUO Qiang, LIN Qingxuan, et al. Label prediction based on user's attention behaviors[J]. Journal of University of Shanghai for Science and Technology, 2021, 43(3): 313-318.
- [15] LU S, JIANG H, LIN G, et al. Research on creating multi-attribute power consumption behavior portraits for massive users[C] // 2018 8th International Conference on Power and Energy Systems (ICPES), December 21-22, 2018, Colombo, Sri Lanka: 54-59.
- [16] 代璐, 田立勃, 王海亮, 等. 计及用户综合用能行为和政策导向的广域综合能源系统联合规划方法[J]. 电力系统保护与控制, 2021, 49(9): 57-71.
DAI Lu, TIAN Libo, WANG Hailiang, et al. Integrated planning of supply and demand side considering the behavior analysis of energy users and the policy-oriented impact[J]. Power System Protection and Control, 2021, 49(9): 57-71.
- [17] 黄悦华, 王艺洁, 杨楠, 等. 基于演化博弈的用户综合用能行为决策方法研究[J]. 电力系统保护与控制, 2020, 48(23): 21-29.
HUANG Yuehua, WANG Yijie, YANG Nan, et al. Research on a decision method of a user comprehensive energy use behavior based on an evolutionary game[J]. Power System Protection and Control, 2020, 48(23): 21-29.
- [18] WANG Y, CHEN Q, KANG Chongqing, et al. Clustering of electricity consumption behavior dynamics toward big data applications[J]. IEEE Transactions on Smart Grid, 2016, 7(5): 2437-2447.
- [19] 时珊珊, 田英杰, 金瑞杨, 等. 基于聚类的电网低压台区用电画像方法研究[J]. 电工电能新技术, 2020, 39(8): 66-74.
SHI Shanshan, TIAN Yingjie, JIN Ruiyang, et al. Research on power consumption portrait method of low-voltage substation areas based on clustering[J]. Advanced Technology of Electrical Engineering and Energy, 2020, 39(8): 66-74.

- [20] 王利利, 张琳娟, 许长清, 等. 能源互联网背景下园区用户画像及成熟度评价模型研究[J]. 中国电力, 2020, 53(8): 19-28.
WANG Lili, ZHANG Linjuan, XU Changqing, et al. Research on park users portrait and maturity evaluation model under the background of energy internet[J]. Electric Power, 2020, 53(8): 19-28.
- [21] 陈涛, 吕松, 任廷林, 等. 基于最小二乘支持向量机的周用电量预测方法[J]. 华电技术, 2020, 42(1): 35-40.
CHEN Tao, LÜ Song, REN Tinglin, et al. Prediction method for weekly electricity consumption based on LSSVM algorithm[J]. Huadian Technology, 2020, 42(1): 35-40.
- [22] 常俊晓, 金之榆, 卢姬, 等. 基于集成聚类和 XGBoost 的短期光伏发电功率预测[J]. 浙江电力, 2021, 40(10): 102-107.
CHANG Xiaojun, JIN Zhiyu, LU Ji, et al. Short-term photovoltaic power prediction based on ensemble clustering and XGBoost[J]. Zhejiang Electric Power, 2021, 40(10): 102-107.
- [23] 严勤, 邓高峰, 胡涛, 等. 基于深度循环神经网络的异常用电检测方法[J]. 中国测试, 2021, 47(7): 99-104.
YAN Qin, DENG Gaofeng, HU Tao, et al. Abnormal electricity detection method based on deep recurrent neural network[J]. China Measurement & Test, 2021, 47(7): 99-104.
- [24] 宋军英, 何聪, 李欣然, 等. 基于特征指标降维及熵权法的日负荷曲线聚类方法[J]. 电力系统自动化, 2019, 43(20): 65-72.
SONG Junying, HE Cong, LI Xinran, et al. Daily load curve clustering method based on feature index dimension reduction and entropy weight method[J]. Automation of Electric Power Systems, 2019, 43(20): 65-72.
- [25] 赵会茹, 张士营, 赵一航, 等. 基于自适应噪声完备经验模态分解-样本熵-长短期记忆神经网络和核密度估计的短期电力负荷区间预测[J]. 现代电力, 2021, 38(2): 138-146.
ZHAO Huiru, ZHANG Shiyang, ZHAO Yihang, et al. Short-term power load interval prediction based on CEEMDAN-SE-LSTM and KDE[J]. Modern Electric Power, 2021, 38(2): 138-146.
- [26] 李欣然, 姜学皎, 钱军, 等. 基于用户日负荷曲线的用电行业分类与综合方法[J]. 电力系统自动化, 2010, 34(10): 56-61.
LI Xinran, JIANG Xuejiao, QIAN Jun, et al. A classifying and synthesizing method of power consumer industry based on the daily load profile[J]. Automation of Electric Power Systems, 2010, 34(10): 56-61.
- [27] 曹昆, 李国昌, 王艳松, 等. 基于大数据和随机矩阵理论的变电站状态评估[J]. 中国测试, 2020, 46(8): 149-157.
CAO Kun, LI Guochang, WANG Yansong, et al. Substation state evaluation based on big data technology and random matrix theory[J]. China Measurement & Test, 2020, 46(8): 149-157.
- [28] 王源涛, 潘武略, 马伟, 等. 基于数据中台和物元可拓理论的二次设备状态评估[J]. 浙江电力, 2021, 40(12): 111-116.
WANG Yuantao, PAN Wulue, MA Wei, et al. State assessment of power secondary equipment based on data center and matter-element extension theory[J]. Zhejiang Electric Power, 2021, 40(12): 111-116.
- [29] 李德毅, 刘常昱. 论正态云模型的普适性[J]. 中国工程科学, 2004, 6(8): 28-34.
LI Deyi, LIU Changyu. Study on the universality of the normal cloud model[J]. Strategic Study of CAE, 2004, 6(8): 28-34.
- [30] 施杰, 张安勤. 基于多参数代价敏感系数学习及数据驱动模型的电力能耗预测[J]. 华电技术, 2021, 43(8): 54-60.
SHI Jie, ZHANG Anqin. Power consumption prediction based on multi-parameter cost-sensitive coefficient learning and data-driven model[J]. Huadian Technology, 2021, 43(8): 54-60.

收稿日期: 2021-11-03; 修回日期: 2021-12-28

作者简介:

杜晓东(1989—), 男, 硕士, 工程师, 主要研究方向为电力系统仿真、数字孪生、线损分析; E-mail: sjzduxd@163.com

曾四鸣(1970—), 男, 硕士, 高级工程师, 主要研究方向为电力物联网、数字孪生电网、电力大数据; E-mail: sjzengsm@yeah.net

刘科研(1978—), 男, 博士, 教授级高工, 硕士生导师, 主要研究方向为配电网运行分析、智能有源配电网故障诊断、配电网大数据分析、配电网仿真。E-mail: liukeyan@epri.sgcc.com.cn

(编辑 张爱琴)