

• 泛在电力物联网 •

DOI:10.15961/j.jsuese.201901044



本刊网刊

电压暂降时空多粒度属性分析与知识发现方法

肖先勇, 胡文曦*, 王 杨, 汪 颖, 张文海

(四川大学 电气工程学院, 四川 成都 610065)

摘 要:随着高端制造业的发展,以电能质量为代表的电力扰动已对敏感用户造成极大的经济损失,为了提升电网安全稳定水平并优化营商环境,针对电压暂降这一典型电力扰动事件,提出电压暂降时空多粒度属性分析与知识发现方法。由于传统电力扰动分析方法依赖机理分析与精确建模,难以处理包含多不确定因素的复杂问题,从复杂问题认知规律的角度出发,提出“数据—特征—指标—信息—知识”所构成的递进认知架构,作为解决复杂问题的一般思路。在此基础上,通过电压暂降时空多粒度属性分析对不同时空尺度下的暂降相关属性进行拓展,克服了传统方法仅从单一粒度分析问题造成暂降信息缺失的问题。由于不同时空粒度下的暂降属性变化反映了其他不确定性因素对电压暂降的影响,通过粒度约简挖掘监测数据背后蕴藏的电压暂降影响程度与电网结构属性之间的关联关系,推导发现电压暂降传播规律。通过仿真和实测数据对本文方法的有效性和可靠性进行了验证,本文方法可适用于包含多不确定因素的复杂问题,有助于突破电力扰动相关的诸多技术瓶颈。

关键词:电力扰动;电压暂降;时空多粒度;知识发现;传播规律

中图分类号:TM721

文献标志码:A

文章编号:2096-3246(2020)04-0025-08

Spatial–Temporal Multi-granular Attribute Analysis and Knowledge Discovery Method for Voltage Sag

XIAO Xianyong, HU Wenxi*, WANG Yang, WANG Ying, ZHANG Wenhai

(College of Electrical Eng., Sichuan Univ., Chengdu 610065, China)

Abstract: In order to mitigate voltage sag related problems based on power quality monitoring data, it is meaningful to improve the efficiency of power quality data analysis. Due to depending on accurate models of voltage sag, the traditional methods are inadequate for complex problems with multiple uncertainty factors. Therefore, the spatial–temporal multi-granular attribute analysis of voltage sag data and a related knowledge discovery method were proposed in this paper. Inspired by the cognitive hierarchy of complex problems, a framework consisted of “data—characteristic—index—information—knowledge” was proposed as a general technical route for voltage sag related problems. Based on the framework, to solve the problem of information loss caused by single granular, sag information in different granular was extended by voltage sag spatial–temporal multi-granular analysis. The relationship between power system structure attribute and voltage sag was discovered by granular reduction. Then, knowledge about voltage sag severity and propagation was derived. The synthetic and measured data were used to validate the effectiveness of the proposed method. Results showed that the proposed method can describe and resolve complex problems with many uncertainty factors.

Key words: power disturbance; voltage sag; spatial–temporal multi-granularity; knowledge discovery; propagation law

随着数字化成为现代电力系统的重要特征^[1–2],数据从处理对象转变为一种基础资源^[3]。以电压暂

降为代表的泛在电力扰动由于其普遍性和严重性而成为国内外工业界和学术界关注的焦点。监测数

收稿日期:2019–10–28

基金项目:国家自然科学基金项目(51807126)

作者简介:肖先勇(1968—),男,教授,博士。研究方向:电能质量与优质供电。E-mail: xiaoxianyong@163.com

*通信联系人 E-mail: 408803186@qq.com

网络出版时间:2020–07–11 13:46:54

网络出版地址:https://kns.cnki.net/kcms/detail/51.1773.TB.20200709.2056.001.html

据蕴含着反映传播规律影响机理等丰富信息^[4],借力泛在电力物联网,以电力扰动数据为驱动,研究其多粒度属性和知识发现方法有助于解决包含多不确定影响因素的电力扰动复杂技术难题^[5-6],具有重要意义。

数据分析理论与知识发现是以大数据分析 with 人工智能为核心技术的泛在电力物联网的基础^[7]。传统电压暂降监测数据分析方法主要根据不同应用目标各自独立分析电能质量问题,这些研究围绕电能质量监测^[8]、特征刻画^[9]、扰动识别^[10]、综合评估^[11]等,研究成果难以互相组织,造成分析效率较低的缺陷。此外,传统分析方法依赖精确模型求解相关问题,随着现代电力系统源、网、荷侧电力电子装置的使用和用电设备的日趋集成化,对扰动产生、传播、影响等各环节的精确建模也更加困难。传统“硬计算”已逐渐难以解决包含多不确定影响因素的复杂问题,迫切需要结合“软计算”的相关方法去挖掘和发现潜在规律和知识^[12]。

电压暂降是影响最为严重的电力扰动之一,为认知和改善暂降问题,现有大量文献围绕暂降源分类、暂降严重程度评估等问题开展研究^[13]。虽然已有方法分析了暂降类型经过变压器传播后发生的变化^[14],但并未定量分析不同电网结构对暂降传播规律的影响。因此,针对电压暂降传播等涉及故障分布、电网结构属性等多不确定因素的复杂问题需要新的解决思路和手段。

为解决上述问题,本文针对电压暂降多粒度属性分析与知识发现方法展开研究。首先,从复杂问题认知思路的角度出发,提出“数据—特征—指标—信息—知识”所构成的递进认知架构,并将知识发现分解为粒子的构建与粒子的使用两个过程。其次,从特征和指标两个层面构建粒子,并从不同时间尺度和不同空间尺度对粒子进行扩展,得到电压暂降时空多粒度属性。在此基础上对这些粒子进行使用,提出以多粒度信息系统和属性约简为核心的知识发现方法。最后,通过实测和仿真数据对本文方法进行验证,实现影响严重程度和暂降传播规律等不同层面知识规则的发现。

1 电压暂降数据分析框架

1.1 双向驱动的DCIHK架构

为了理解和解决电力扰动问题,对其进行监测是有效途径^[8]。为此,在源、网、荷侧均已建立起电力系统泛在电能质量监测网络,记录到的电力扰动数据日益增加。针对泛在电力扰动数据的分析方法、理论和技术展开研究,可广泛应用于源、网、荷侧电力

扰动问题的诊断、评估、治理等措施^[4]。复杂系统理论中指出人类认知客观事物的基础由粒化、组织和推理构成^[15]。其中,粒化将复杂问题在知识空间中分解为一系列的粒子,组织将各粒子结合为整体,推理则涉及到原因与结果间的联系。因此,基于客观事物的认知规律^[16],本文提出电压暂降数据的递进认知架构如图1所示。整个过程可看作是:通过监测数据刻画事件特征,进一步计算量化事件影响程度的指标,最后提取指标中蕴含的相关信息并最终推理为知识的过程,即DCIHK(data, characteristic, index, information and knowledge)架构。

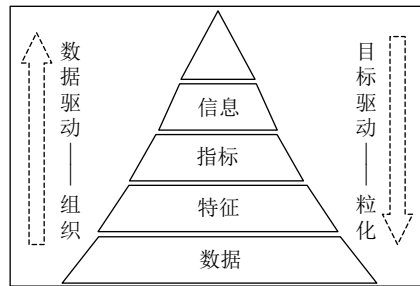


图 1 递进认知的DCIHK架构

Fig. 1 Progressive cognitive DCIHK framework

架构中不同层次概念可以看作不同粒度的知识颗粒,整个架构由数据和目标双向驱动。数据驱动体现在各认知层次的逐层递进计算和组织,目标驱动体现在计算过程中对参数的选择。

1.2 基于目标驱动的问题空间粒化

粒化准则与具体目标密切相关,因此问题空间的粒化是目标驱动的过程。信息系统或知识表达系统可表示为 $S = \langle U, A \rangle$,其中,非空有限对象集 $U = \{x_1, x_2, \dots, x_n\}$ 为论域, A 为属性的非空有限集合。由于知识源于人类的分类能力,对 U 进行划分则是对关于 U 的知识进行粒化处理。论域 U 的一个划分 $X = \{X_i | 1 \leq i \leq m\}$ 提供了一种粒度描述,因此划分中的每个 X_i 都是一个知识颗粒,称作 U 中的第 i 个基本概念。对于每个属性子集 $B \subseteq A$,定义一个二元不可分辨关系:

$$IND(B) = \{(x, y) \in U \times U | \forall a \in B, a(x) = a(y)\} \quad (1)$$

式中, $IND(B)$ 为一个等价关系,构成论域 U 上的一个划分 $IND(B)$,称作关于 U 的知识。因此,根据不同划分原则,论域 U 可被粒化为不同的知识颗粒从而构造成知识库。

在电压暂降的相关知识发现过程中, U 中的每一个元素,即为一个暂降事件的监测记录数据, A 为从监测数据中提取的诸多属性,因此属性的构建是数据驱动的过程;此外,还需根据研究目标确定必要的属性集合,即属性筛选是目标驱动的过程。结合每一

个暂降事件的属性集,即可通过关联规则挖掘发现暂降的相关规律。

1.3 基于数据驱动的粒子组织

数据本身并无意义,但通过一定规则的组织可得到知识在不同粒层上的表达形式,即特征、指标和信息。针对电压暂降事件,通常采用电压幅值 V 和持续时间 T 等参数刻画事件特征。

$$V = \sqrt{\frac{1}{N} \sum_{i=1+k-N}^k v_i^2} \quad (2)$$

$$T = T_{\text{end}} - T_{\text{start}} \quad (3)$$

式中, N 为每周波采样次数, v_i 为瞬时电压采样值, T_{end} 和 T_{start} 分别为暂降结束时刻和暂降开始时刻。

然而,基于DCIHK架构的认知层次可知,特征只反映单一暂降事件本身的波形特点,无法揭示其随时空变化背后蕴藏的严重程度和传播规律等知识。因此需要进一步在指标层面进行时空多粒度分析。可见,时空多粒度属性是知识发现的重要环节。

2 电压暂降时空多粒度属性分析

2.1 电压暂降时间多粒度属性

无论是单次电压暂降监测的三相电压波形,还是监测周期内多次暂降事件下的暂降指标,实际上都是一组时间序列。监测装置在不同时间间隔下会对电压暂降事件进行聚合和统计^[8],因此电压暂降数据具有时间多粒度属性。

由于特征只能用于刻画事件本身,基于DCIHK架构,根据暂降特征可进一步计算得到反映暂降水平的指标,即暂降幅值属性 E_v 、频次属性 E_{sf} 、能量属性 E_{se} 和严重程度属性 E_{ss} ^[17]。其中,暂降幅值属性用于量化平均电压幅值,频次属性量化暂降发生次数,能量属性和严重程度属性可表达为:

$$E_{se} = \left(1 - \left(\frac{V_{\min}}{V_{\text{nom}}}\right)\right) \cdot T \quad (4)$$

$$E_{ss} = \frac{1 - V_{\min}}{1 - V_{\text{curve}}(T)} \quad (5)$$

式中, V_{\min} 为暂降发生过程中的最低电压, V_{nom} 为标称电压, V_{curve} 为暂降耐受曲线上持续时间 T 对应的电压幅值。这些属性仅能评估一次暂降事件的影响程度,因此,考虑监测周期内的多次暂降事件,对其进行时间多粒度拓展:

$$SI = \frac{1}{n} \sum_{i=1}^n s_i \quad (6)$$

式中: n 为某个时间段内监测到的暂降事件数量; s_i 为

第 i 个暂降事件的综合严重程度指标,可通过熵权法进行加权得到^[18]:

$$s = E_v \cdot w_1 + E_{sf} \cdot w_2 + E_{se} \cdot w_3 + E_{ss} \cdot w_4 \quad (7)$$

将监测数据序列根据某一时间粒度 γ 划分成若干个时间单元二元组 (t, SI) ,则可得到粒度 γ 下的表现形式:

$$\delta_\gamma = \{(t_1, SI_1), (t_2, SI_2), \dots, (t_i, SI_i), \dots\} \quad (8)$$

客观世界常用的秒、分、小时、日、月、年等时间单位则可定义为不同时间粒度函数。因此,可对应得到不同粒度下的暂降属性表现形式,表达了从不同角度观测暂降水平及其变化规律的特点。

2.2 电压暂降空间多粒度属性

相同故障类型和故障电阻的同一故障经不同网络拓扑传播之后,其影响范围可能不同,最终造成不同节点监测到的暂降水平也不同。因此,根据监测装置在电网中的安装位置及其连接关系,可从不同母线、电压等级和区域电网对暂降进行分析,得到不同空间粒度下差异性的暂降水平结果。这种空间造成的差异性蕴含了电网结构属性对暂降传播的影响。

为了描述电网结构属性对暂降传播的影响规律,需要面向电压暂降对结构属性进行定量评估。由于暂降是以故障点为中心呈辐射状传播,该特性使得处于网络中心或者相邻节点更多的节点更容易受到从别的节点传播的暂降。因此,构建节点规模属性 K_i ,表示节点 i 的连边数目。然而,该属性仅反映了节点规模,并未反映节点之间的紧密程度。针对暂降传播路径不唯一的特性,采用节点聚集程度 TC_i 对其进行量化:

$$TC_i = \frac{H_i}{K_i(K_i - 1)} \quad (9)$$

式中, H_i 为节点 i 所连接的边数,分母为理论上可能构成的最大边数,因此二者的比值可反映节点连接的紧密程度。

电网节点的规模程度和聚集程度均反映了电网节点的拓扑属性,但难以表征物理属性,即暂降传播受电气距离和电源等因素的影响。针对暂降传播的衰减特性,节点 i 与其他节点之间电气距离倒数之和的平均值定义为传播效率 TE_i :

$$TE_i = \frac{1}{n-1} \sum_{j=1, j \neq i}^n \frac{1}{l_{ij}} \quad (10)$$

式中, l_{ij} 为节点 i 与相邻节点 j 之间的最短电气距离。节点效率值可以表征暂降从该节点传播到网络中其他节点的平均难易程度。

针对暂降传播受电源支撑的特性,以节点距电源的最短距离作为电网节点支撑能力属性 TS_i :

$$TS_i = \min(l_{ij}), j \in \alpha \quad (11)$$

式中, α 为电源节点构成的集合。

根据电网元件参数和拓扑结构计算邻接矩阵和电气距离, 针对电网各节点可计算上述结构属性^[19]。然而, 电网中的不同节点根据及电气连接关系可构成不同范围大小的节点簇, 这些簇结构为结构属性的量化提供了不同层面的描述方式。因此, 假设电网有 m 个节点, 可进行空间多粒度拓展:

$$R^j = \frac{1}{m} \sum_{x=1}^m r^j(x) \quad (12)$$

式中, $r^j(x)$ 为第 x 个节点的第 j ($0 < j \leq 4$) 个结构属性。同样地, 可将电网所有节点的结构属性构成的序列根据空间粒度划分成若干二元组 (z, R) :

$$\varphi_\gamma^j = \{(z_1, R_1^j), (z_2, R_2^j), \dots, (z_i, R_i^j), \dots\} \quad (13)$$

式中, z 为空间变量, 粒度 γ 在空间中可根据相邻节点数量、电压等级等不同区域范围划分得到。

3 电压暂降传播规律多粒度知识发现

3.1 多粒度信息系统

传统暂降数据分析存在着粒度单一、依赖精确模型的缺陷, 因此难以挖掘传播规律等复杂知识。假设电压暂降属性集中任意属性都有 I 个粒度等级, 可将传统信息系统拓展为多粒度信息系统:

$$S = \langle U, A \rangle = \langle U, \{a_j^k | k = 1, 2, \dots, m\} \rangle \quad (14)$$

式中, a_j^k 为属性 a_j 在粒度 k 上的值, 满足:

$$a_j^{k+1}(x) = g_j^{k,k+1}(a_j^k(x)), x \in U \quad (15)$$

式中, $g_j^{k,k+1}$ 为粒度变换函数。由于本文研究时空多粒度属性, 且时间域和空间域中的粒度变换函数具体表示为式(6)、(12)。综上, 从电能质量监测装置中获取的电压暂降监测数据, 可根据DCIHK架构先后计算电压暂降特征和指标, 并结合粒度变换函数进行时空多粒度拓展, 进而得到式(14)所示的多粒度信息系统。

由于目前难以针对连续属性进行关联规则挖掘, 因此, 首先需要针对信息系统中的属性进行离散化处理, 若离散化方法不合理, 往往会丢失很多有意义的信息, 对知识发现的结果造成影响。混合高斯模型采用多个高斯分布的组合来刻画数据分布, 已被证实是一种连续变量概率密度估计的有效手段^[20]。混合高斯模型的公式为:

$$P(x) = \sum_{i=1}^T w_i P(x | \mu_i, \sigma_i) \quad (16)$$

式中, x 为随机变量, T 为总的高斯分布数量, μ 和 σ 分别为高斯分布的均值和方差, 权重 w ($w \geq 0$) 满足 $\sum_{i=1}^M w_i = 1$ 。第 i 个高斯分布的密度函数可表示为:

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (17)$$

在此基础上, 根据高斯模型的 3σ 原则, 可得到每个高斯分布的主要区间为 $[\mu - 3\sigma, \mu + 3\sigma]$, 则可看作对应概念的连续分布区间。由此, 可基于混合高斯模型将连续属性变量转化为离散变量。

3.2 基于粗糙集的粒度约简方法

约简是信息系统中知识发现的关键, 是在保证论域分类能力不变的前提下, 删除信息系统中冗余属性的过程。在信息系统 $S = \langle U, A \rangle$ 中, $A = C \cup D$ 为属性集合; C 和 D 分别为条件属性和决策属性, 在本文中分别是结构属性和暂降影响程度属性。令 $P \subseteq A$, 对于 $a \in P$, 若满足:

$$IND(P) = IND(P - \{a\}) \quad (18)$$

则称属性 a 在 P 中是不必要的, 否则是必要属性。信息系统中去掉不必要属性不会改变其分类能力; 相反地, 若去掉必要属性则一定能改变系统的分类能力。此外, 不同于传统粗糙集, 多粒度信息系统的关键在于如何选择最优粒度使其能够有效挖掘出必要信息。粒度选择的本质是在给定的多个粒度空间中求取一个无冗余信息的粒空间集合。多粒度信息系统中 $A = \{A_1, A_2, \dots, A_m\}$, A_1, A_2, \dots, A_m 为属性 A 在不同粒度上拓展后得到 m 簇子属性。对 $\forall X \subseteq U$ 可定义 X 多粒度粗糙集的下近似集 $\underline{\sum_{i=1}^m A_i(X)}$ 和上近似集 $\overline{\sum_{i=1}^m A_i(X)}$ 分别为:

$$\underline{\sum_{i=1}^m A_i(X)} = \{x \in U | [x]_{A_1} \subseteq X \wedge \dots \wedge [x]_{A_m} \subseteq X\} \quad (19)$$

$$\overline{\sum_{i=1}^m A_i(X)} = \sim \underline{\sum_{i=1}^m A_i(\sim X)} \quad (20)$$

针对信息系统 $S = \langle U, A, V, f \rangle$, 有 $A' \subseteq A$, $U/D = \{Y_1, Y_2, \dots, Y_r\}$, 则多粒度粗糙集下近似分布为:

$$\mu_{A'}(U, D) = \left\{ \underline{\sum_{i=1}^m A_i(Y_1)}, \underline{\sum_{i=1}^m A_i(Y_2)}, \dots, \underline{\sum_{i=1}^m A_i(Y_r)} \right\} \quad (21)$$

若 $\mu_{A'}(U, D) = \mu_A(U, D)$, 则称 A' 是 A 粒度下的近似分布一致集; 若对 $\forall A' \subset A$, 都有 $\mu_{A'}(U, D) \neq \mu_A(U, D)$, 则 A' 是 A 粒度下的近似分布约简。可以证明, 粒度下近似分布一致集保证了每个目标决策的下近似保持不变,

即可保持分类不改变^[21]。因此,以此为判据,可使得约简后的多粒度空间具有和原始粒度空间相同的目标决策能力。

对于给定粒度 $A' \subseteq A, X \subseteq U/D$,若

$$\sum_{i=1}^m A_i^p(X) \subset \sum_{i=1, A_i \in A'}^m A_i^p(X) \quad (22)$$

则称粒度 A' 是关于 X 重要的粒度。反之,若

$$\sum_{i=1}^m A_i^p(X) = \sum_{i=1, A_i \in A'}^m A_i^p(X) \quad (23)$$

则称粒度 A' 是关于 X 不重要的粒度,或者说粒度 A' 是关于 X 冗余的。即通过删除粒度空间 A' 后决策信息系统的分类能力并不发生改变。综上,多粒度信息系统为复杂问题提供了多个层面的描述角度,而粒度约简可帮助问题分析人员选择最优的粒度,最终可挖掘出if→then语句格式的关联规则。整个暂降传播规律知识发现的流程图如图2所示。

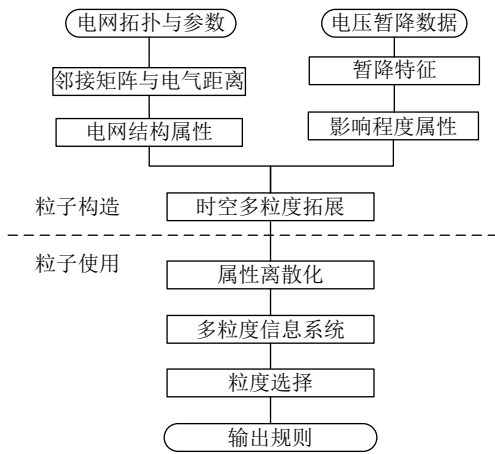


图2 暂降传播规律知识发现流程图

Fig. 2 Flowchart of knowledge discovery for voltage sag propagation

4 算例分析

4.1 电压暂降影响程度多粒度评估

本节基于蒙特卡洛法采用电网故障随机模型针对IEEE 30节点系统进行仿真^[22],电网拓扑结构如图3所示,仿真软件为MATLAB。电网故障随机模型假设故障随机分布在系统中任意一条线路上,其与节点的距离服从 $[0, 1]$ 均匀分布,故障电阻服从 $[5 \Omega, 1 \Omega]$ 正态分布,持续时间服从 $[0.06 \text{ s}, 0.01 \text{ s}]$ 正态分布。

重复1 000次蒙特卡洛仿真,可得到每次故障下所有节点的电压幅值和持续时间特征。在此基础上,可基于第2.1节计算得到暂降幅值、频次、能量和严重程度属性。受限于篇幅,仅给出部分仿真结果如表1所示。

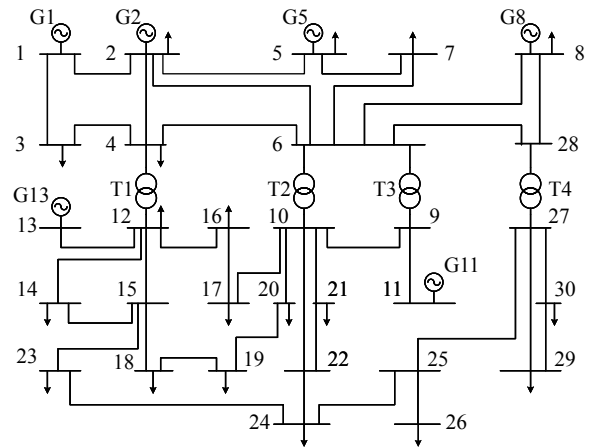


图3 IEEE 30节点测试系统

Fig. 3 IEEE 30-bus test system

表1 电压暂降节点指标

Tab. 1 Voltage sag site indices

节点编号	E_v	E_{sf}	E_{se}	E_{ss}
1	0.819 5	31	0.018 4	0.360 9
2	0.765 3	52	0.023 5	0.469 4
3	0.717 4	115	0.026 0	0.565 2
...
14	0.637 4	178	0.033 0	0.725 2
15	0.610 7	179	0.035 3	0.778 7
16	0.649 3	178	0.032 7	0.701 4
...
28	0.714 9	151	0.026 3	0.570 4
29	0.669 8	251	0.030 7	0.660 4
30	0.668 9	257	0.030 9	0.662 3

由于不同属性体现了从不同角度对电压暂降影响程度的量化,难以综合对比各节点的影响程度大小。因此,对各属性进行加权可得到综合影响程度,并根据影响大小对电网所有节点进行排序。采用不同加权方法下的排序结果如图4所示。

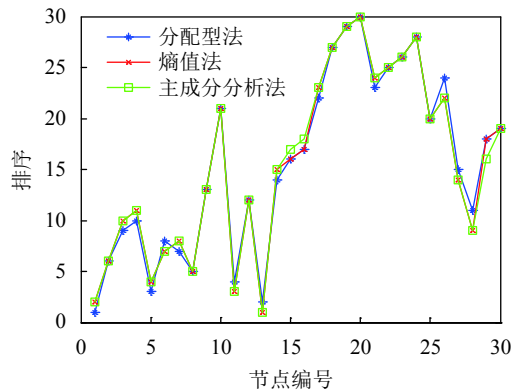


图4 不同加权方法下的电压暂降综合评估排序结果

Fig. 4 Evaluation results for voltage sags based on different weighting methods

由图4可以看出,不同加权方法所得结果趋于一致,证明加权方法对评估结果的影响较小,并不会影响后续知识发现结果。此外,由于受电源支撑作用较强,节点1、2、5、8、11、13的综合指标最小;而其余节点随着与电源距离增大、邻接节点数量增加等因素影响,暂降影响程度逐渐增大。因此,所采用的暂降严重程度等属性可以准确量化不同节点受暂降影响的程度大小。

以某优质供电园区中监测装置所记录的实测数据对电压暂降影响程度多粒度评估方法进行验证。电网的简单示意图如图5所示,图5中,M1~M6为电能质量监测装置。

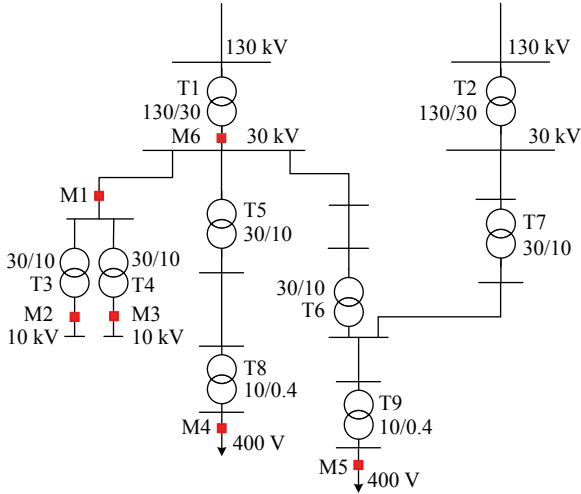


图5 供电园区电网简化示意图

Fig. 5 Simplified diagram of a power grid

根据M6监测装置于2017年7月—2018年6月一年监测到的暂降数据,分别以不同时间粒度计算暂降影响程度,结果如图6所示。由图6可以看出:当时间粒度选择为1周时,可以发现第5周、23周、27周、38周等时间段内暂降水平有明显上升;当时间粒度选择为4周时,可以看到暂降影响程度更长时间的变化规律,即第2月和第8月更加严重。这有助于电力系统分析和决策人员寻找M6点不同时间区间背后对应的故障数据和气象数据,进而为暂降治理提供决策支持。

4.2 电压暂降传播规律挖掘

将指标进行归一化和同趋势化,指标越大表示暂降严重程度越高或电网结构影响力越大。基于高斯云模型进行离散化,并采用1~5表示属性大小,得到多粒度决策系统如表2所示。表2中: $C_1 \sim C_4$ 为电网结构属性; D 为暂降影响程度; x 、 y 和 z 分别为3种空间粒度,即单一节点构成的区域、节点与相邻节点构成的区域、节点与相邻节点及间接邻节点构成的区域。例如: x_1 这行表示1号节点的4种结构属性值分别为1、3、2、5,该节点对应的暂降影响程度值为1; y_{30} 这行

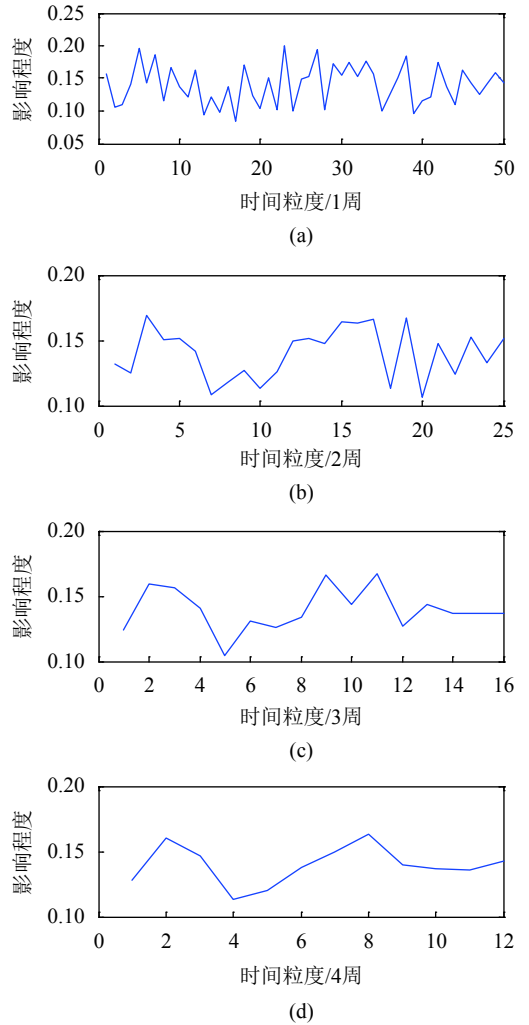


图6 时间多粒度评估结果

Fig. 6 Multi-temporal granular evaluation results

表示节点30及其相邻节点构成的区域的各结构属性的总体平均值。由此可知,表2所示的多粒度决策系统,是DCIHK架构中从监测数据中提取特征并计算为指标的过程。

为减少暂降不确定性的影响,引入置信度对决策规则的可靠性进行评价:

$$\beta(X_i \rightarrow Y_j) = \frac{N_{(C_i \cap D_j)}}{N_{D_j}} \quad (24)$$

式中,分母表示 D_j 类型暂降事件数量,分子表示这些事件中条件属性为 C_i 的数量,置信度 $\beta(X_i \rightarrow Y_j)$ 表示通过规则 X_i 推测出事件 Y_j 的可信程度。为了从指标中发现暂降传播规律的信息,采用第3.2节的粒度约简方法,得到若干关联规则,如表3所示。表3中:每一行为一条关联规则,if连接的规则前件表示规则的条件,then连接的规则后件表示当满足该条规则条件时对应推导出的结果。这些关联规则反映了电网各节点不同的结构属性所造成的电压暂降影响程度的差异,即电压暂降的传播规律。

表2 多粒度决策表

Tab. 2 Multi-granular decision table

编号	C_1	C_2	C_3	C_4	D
x1	1	3	2	5	1
x2	1	3	2	5	1
...
x11	1	3	3	5	1
x12	2	2	3	5	2
...
x21	2	3	2	2	4
x22	2	3	2	2	4
...
x30	2	2	3	5	2
y1	1	3	1	5	1
...
y11	1	3	2	5	1
y12	1	1	3	4	2
...
y21	2	3	2	2	4
y22	3	1	4	1	4
...
y30	1	3	2	4	2
z1	1	3	2	5	1
...
z11	1	3	2	5	1
z12	1	1	3	4	1
...
z21	3	1	4	1	4
z22	2	3	2	2	4
...
z29	1	3	2	4	2
z30	1	3	2	4	2

表3 电压暂降传播规律关联规则

Tab. 3 Association rules for voltage sag propagation

编号	规则	置信度 $\beta/\%$
1	If $C_1=1, C_2=3, C_3=2, C_4=5$ then $D=1$	66.7
2	If $C_1=1, C_2=1, C_3=3, C_4=4$ then $D=1$	33.3
3	If $C_1=1, C_2=3, C_3=3, C_4=4$ then $D=2$	63.6
4	If $C_1=2, C_2=2, C_3=3, C_4=5$ then $D=2$	36.4
5	If $C_1=2, C_2=2, C_3=4, C_4=4$ then $D=3$	80.0
6	If $C_1=2, C_2=3, C_3=2, C_4=2$ then $D=4$	61.3
7	If $C_1=3, C_2=1, C_3=4, C_4=1$ then $D=4$	22.5
8	If $C_1=4, C_2=2, C_3=2, C_4=2$ then $D=4$	16.1
9	If $C_1=5, C_2=3, C_3=3, C_4=1$ then $D=5$	100.0

以表3中第1条规则为例,表示当电网节点的规模为1,聚集程度为3,传播效率为2,支撑能力为5时,该节点的暂降影响程度为1。即当该节点的相邻节点数较少,距离其他节点的电气距离较远,且有电源节点支撑时,该节点的暂降水平较轻。可见,关联规则与电压暂降的认知相符合,且置信度越高表明该条规则越可信。粒度的选择说明了仅采用节点本身描述电网结构属性与暂降之间的关联规则不如考虑相邻节点的区域结构属性更为有效。粒度约简在不需要主观经验的前提下选择了最优的粒度表达方式,符合数据驱动的特点。关联规则不仅揭示了电网结构属性对电压暂降传播的影响,还为传播路径抑制等暂降解决方案给出了决策支持。

5 结论

为提升监测数据分析能力以解决电压暂降复杂问题,本文提出电压暂降多粒度属性分析与知识发现方法。相对于传统方法,本文方法存在以下特点:

1)将数据、特征、指标和信息看作知识的不同粒度表现形式,从复杂问题认知层次出发,提出双向驱动的DCIHK架构,作为知识发现的一般思路流程。

2)对电压暂降幅值、频次、能量和严重程度等属性进行时间多粒度拓展。既能捕获短时间内暂降水平升高的准确时间,又有助于分析暂降水平长期变化规律。

3)对节点规模、聚集程度、传播效率和支撑能力等电网结构属性进行空间多粒度拓展,有助于从节点、区域等不同层面分析电压暂降传播规律。

基于关联规则挖掘的知识发现方法不依赖于精确模型,适用于电压暂降传播等包含多不确定因素的复杂问题。填补了目前尚无电网结构对暂降影响定量分析方法的空白。

随着电网发、输、配、用电各环节相关技术的复杂化,以电压暂降为代表的电力扰动问题也随之变得复杂。传统基于精确模型进行分析的方法已越来越难以描述和解决相关复杂难题。未来研究中,本文提出的多粒度分析方法可通过关联规则挖掘的方式进一步分析电力扰动的诸多难题。

参考文献:

- [1] Zhou Xiaoxin, Chen Shuyong, Lu Zongxiang, et al. Technology features of the new generation power system in China[J]. *Proceedings of the CSEE*, 2018, 38(7): 1893-1904. [周孝信, 陈树勇, 鲁宗相等. 能源转型中我国新一代电力系统的技术特征[J]. *中国电机工程学报*, 2018, 38(7): 1893-1904.]
- [2] Xue Yusheng, Lai Yening. Integration of macro energy thinking and big data thinking part one: Big data and power

- big data[J].*Automation of Electric Power Systems*,2016,40(1):1–8.[薛禹胜,赖业宁.大能源思维与大数据思维的融合(一)大数据与电力大数据[J].*电力系统自动化*,2016,40(1):1–8.]
- [3] Wu Xindong,Zhu Xingquan,Wu Gongqing,et al.Data mining with big data[J].*IEEE Transactions on Knowledge and Data Engineering*,2014,26(1):97–107.
- [4] Xu Wilsun,Yong Jing.Power disturbance data analytics—New application of power quality monitoring data[J].*Proceedings of the CSEE*,2013,33(19):93–101.[徐文远,雍静.电力扰动数据分析学——电能质量监测数据的新应用[J].*中国电机工程学报*,2013,33(19):93–101.]
- [5] Elphick S,Gosbell V,Smith V,et al.Methods for harmonic analysis and reporting in future grid applications[J].*IEEE Transactions on Power Delivery*,2017,32(2):989–995.
- [6] Guder M,Salor O,Cadirci I,et al.Data mining framework for power quality event characterization of iron and steel plants[J].*IEEE Transactions on Industry Applications*,2015,51(4):3521–3531.
- [7] Huang T E,Guo Qinglai,Sun Hongbin,et al.A deep learning approach for power system knowledge discovery based on multitask learning[J].*IET Generation,Transmission & Distribution*,2019,13(5):733–740.
- [8] Electromagnetic compatibility (EMC).Testing and measurement techniques—Power quality measurement methods: BS EN 61000–4–30:2015[S/OL].[2020–06–28]. London:BSI, 2015.<https://webstore.iec.ch/publication/21844>.
- [9] CIGRE/CIREU/UIE Joint Working Group C4.110.Voltage dip immunity of equipment and installations[R/OL].[2020–06–28].<http://www.uie.org/voltage-dip-immunity-equipment-and-installations>.
- [10] Qu Hezuo,Liu Heng,Li Xiaoming,et al.Feature combination optimization for multi-disturbance classification of power quality[J].*Electric Power Automation Equipment*,2017,37(3):146–152.[瞿合祚,刘恒,李晓明,等.一种电能质量多扰动分类中特征组合优化方法[J].*电力自动化设备*,2017,37(3):146–152.]
- [11] Wang Jidong,Pang Wenjie.Application of Fisher discriminant analysis in steady-state power quality evaluation of grid-connected photovoltaic system[J].*Electric Power Automation Equipment*,2017,37(3):50–54.[王继东,庞文杰.Fisher判别分类法在光伏并网系统稳态电能质量评估中的应用[J].*电力自动化设备*,2017,37(3):50–54.]
- [12] 孙林,徐久成.粒计算的不确定性分析与知识获取方法[M].北京:科学出版社,2018.
- [13] Bollen M H J,Gu I Y H.Signal processing of power quality disturbances[M].Hoboken:John Wiley & Sons,Inc.,2006. DOI:10.1002/0471931314
- [14] Wang Ying,Yu Huaxi,Xiao Xianyong,et al.New method for calculating voltage dip/swell types based on six-dimensional vectors and Euclidean distance[J].*IET Generation,Transmission & Distribution*,2019,13(18):4206–4213.
- [15] Zadeh L A.Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic[J].*Fuzzy Sets and Systems*,1997,90(2):111–127.
- [16] Ackoff R L.From data to wisdom[J].*Journal of Applied Systems Analysis*,1989,16:3–9.
- [17] IEEE.IEEE Guide for voltage sag indices:IEEE Standard 1564—2014[S/OL].[2020–06–28].New York:IEEE,2014.<https://ieeexplore.ieee.org/document/6842577>.
- [18] Lin Deqing,Gu Wei,Wang Yuankai,et al.Synthetic evaluation of power quality based on dynamic time warping spatial distance measurement[J].*Power System Technology*,2013,37(2):562–567.[林德清,顾伟,王元凯,等.基于动态时间弯曲空间距离度量的电能质量综合评估[J].*电网技术*,2013,37(2):562–567.]
- [19] Lei Jiazhi,Gong Qingwu.Comprehensive prediction method for failure rate of transmission line based on multi-dimensional cloud model[J].*IET Generation,Transmission & Distribution*,2019,13(9):1672–1678.
- [20] Xu Da,Cai Manchun,Chen Yue.Shortest path of urban traffic based on the improved Floyd algorithm[J].*Electronic Science and Technology*,2017,30(7):17–20.[徐达,蔡满春,陈悦.基于改进Floyd算法的城市交通网络最短路径规划[J].*电子科技*,2017,30(7):17–20.]
- [21] Sang Yanli,Qian Yuhua.A granular space reduction approach to pessimistic multi-granulation rough sets[J].*Pattern Recognition and Artificial Intelligence*,2012,25(3):361–366.[桑妍丽,钱宇华.一种悲观多粒度粗糙集中的粒度约简算法[J].*模式识别与人工智能*,2012,25(3):361–366.]
- [22] University of Washington Electrical Engineering.300 bus power flow test case[EB/OL].(1993–08–15)[2020–06–28].http://www.ee.washington.edu/research/pstca/pf30/pg_tca30bus.htm.

(编辑 赵 婧)

引用格式: Xiao Xianyong,Hu Wenxi,Wang Yang,et al.Spatial-temporal multi-granular attribute analysis and knowledge discovery method for voltage sag[J].*Advanced Engineering Sciences*,2020,52(4):25–32.[肖先勇,胡文曦,王杨,等.电压暂降时空多粒度属性分析与知识发现方法[J].*工程科学与技术*,2020,52(4):25–32.]