

吴宇, 张剑, 谢涵, 等. 基于采集指标细化分析的停电投诉风险识别方法[J]. 智能计算机与应用, 2025, 15(10): 136-140.
DOI:10.20169/j.issn.2095-2163.251020

基于采集指标细化分析的停电投诉风险识别方法

吴宇¹, 张剑², 谢涵¹, 杨晓薇¹, 陈顺利², 吴林峰¹

(1 国网重庆市电力公司营销服务中心, 重庆 401100; 2 重庆广汇供电服务有限公司客户服务分公司, 重庆 401100)

摘要: 为提高敏感、重要用户停电上报的准确性和及时性, 提出一种基于采集指标细化分析的停电风险识别方法。该方法首先基于时间序列分解算法提取各类用户的停电投诉风险影响因子; 然后, 基于各用户类别和停电投诉风险影响因子, 采用模糊 C 均值算法进行用户分类的细化; 其次, 采用 BP 神经网络构建停电投诉风险识别模型, 通过对比停电识别概率值与真实停电情况, 不断优化模型识别准确率; 最后, 采用加权投票算法计算不同用户的停电投诉概率。实验表明, 本文算法可实现台区用户的停电风险及时识别, 提高用户满意度和电网安全性, 助力提升电力公司供电质量和服务水平。

关键词: 时间序列分解; 模糊 C 均值; BP 神经网络; 加权投票法; 停电投诉风险识别

中图分类号: TM933

文献标志码: A

文章编号: 2095-2163(2025)10-0136-05

Power grid outage risk identification method based on refined analysis of collection indicators

WU Yu¹, ZHANG Jian², XIE Han¹, YANG Xiaowei¹, CHEN Shunli², WU Linfeng¹

(1 State Grid Chongqing Electric Power Company Marketing Service Center, Chongqing 401100, China;

2 Customer Service Branch of Chongqing Guanghui Power Supply Service Co., Ltd., Chongqing 401100, China)

Abstract: In order to improve the accuracy and timeliness of power outage reporting for sensitive and important users, a power outage risk identification method based on detailed analysis of collection indicators is proposed. First of all, this method extracts the power outage complaint risk influencing factors of various users based on the time series decomposition algorithm. Subsequently, based on each user category and power outage complaint risk influencing factors, the fuzzy C-means algorithm is used to refine the user classification. Then, BP neural network is used to build a power outage complaint risk identification model, and by comparing the power outage identification probability value with the real power outage situation, the model identification accuracy is continuously optimized. Finally, a weighted voting algorithm is used to calculate the power outage complaint probabilities of different users. Experiments show that the algorithm proposed in this paper can realize timely identification of power outage risks for users in distribution stations, improve user satisfaction and grid security, and help improve power supply quality and service levels of power companies.

Key words: time series decomposition; fuzzy C-means; BP neural network; weighted voting; identification of power outage complaint risk

0 引言

供电可靠性直接影响社会生产生活, 尤其近年来, 随着电力市场化改革和电能替代工作推进, 对供电可靠性提出了更高要求, 如何及时且准确地识别停电故障, 防止配电网频繁出现停电事件, 已然成为电力企业提升自身服务水平的关注点。当前, 鉴于配网部分区域存在结构薄弱、运行检修不及时、计量设备故障老化等原因, 产生事故停电、修复较难等问

题。当前对用户停电分析识别及预警存在时效性差的问题, 尤其部分企业用户对供电可靠性要求高, 难以实现对用户停电投诉风险的高效识别, 导致停电投诉事件仍不时发生。故亟需开展停电高投诉风险分析及识别研究, 达到投诉超前预控、意见数量压降的目标, 不断提高公司整体服务水平^[1-2]。

当前, 部分研究机构及专家已开展该方面的研究工作。针对低压配电网对非实时故障监控和停电故障预警的客观需求问题, 文献[3]利用受限玻尔

兹曼机和深度信念网络训练历史故障数据,并深度挖掘变压器、线路和用户的历史停电故障数据特征,从而进行停电故障预警,再利用 Dempster-Shafer 证据理论对不同系统数据的预警结果进行融合,得到综合预警结果,实现台区故障类型的实时预警和精准定位。为了降低电网故障发生率,准确及时应对配电网停电风险,文献[4]基于设施设备、运行维护与网络架构水平建立预警指标体系,并利用模糊层次法计算指标的模糊综合度,及时做出停电预警并能为配电网安全运行提供技术保障。文献[5]基于物联网技术构建了电网停电故障监控及预警模型,从而解决智能电网对监控和停电预测的客观需求。该模型利用小波变换进行电网故障信息分析,并基于 GIS-MapX 技术进行故障节点地图绘制,可直观确定故障节点位置,最后通过模拟故障实验,证明了该系统具备良好的可靠性与实用性。

虽然上述方法均在一定程度上实现了用户停电故障监测和预警,但由于未结合投诉记录数据加以分析,导致停电投诉风险识别效果并不理想。针对这一问题,本文提出了一种基于采集指标细化分析的停电风险识别方法。本文所提方法的贡献如下:

(1)针对重点、敏感用户停电原因复杂、影响因

图 1 整体流程图

Fig. 1 System structure diagram

本文首先基于重点和敏感用户历史用电数据,采用时间序列分解算法提取各类用户的停电投诉风险影响因子;随后,基于各用户类别和停电投诉风险影

素较多的问题,本文基于重点用户的用电采集指标数据分析,结合停电投诉风险等级标签,采用时序分解结合权值分配,提取用户的停电投诉风险影响因子,从而提高后续重点用户停电投诉风险识别准确率。

(2)针对已有用户类别分类未考虑不同用户的用电行为特性,导致建模识别精准度低的问题,本文采用模糊 C 均值算法进行用户采集指标细化分类,提高面向停电投诉风险用户分类的类内相似度。

(3)针对当前无差别用户停电分析时效性差、准确率低的问题,本文利用 BP 神经网络算法对重点用户进行停电投诉风险识别,该方法实现了敏感客户的停电投诉风险的精准研判,可通过预警提前做好停电应急预案,减少停电故障导致的社会经济损失,大大提高电网公司的服务水平。

1 整体过程描述

为了提高停电风险识别的精准研判,本文构建了一种基于采集指标细化分析的停电风险识别方法,通过分析用户的停电影响场景,建立停电投诉风险用户标签库,采用神经网络算法构建停电风险识别模型,实现重点敏感用户的停电精准识别,模型整体流程如图 1 所示。

响因子,采用模糊 C 均值算法进行用户分类的细化,输出细化用户类别;然后,基于停电事件样本集,采用 BP 神经网络构建停电风险识别模型,通过对比停电

识别概率值与真实停电情况计算识别误差,从而更新网络的权值和阈值,不断优化模型识别准确率;最后,采用加权投票算法计算不同用户的停电投诉概率,当重点敏感用户停电投诉概率超过所在分类的设定阈值时,则给运维工作人员发送预警信息。

2 基于采集指标细化分析的停电投诉风险识别

本文从用电信息采集系统及营销业务应用系统的电力系统数据库中获取各类停电事件样本数据^[6],包含用户用电类别、用户号、用户名称、停电时间、电量值、电压值、电流值、采集成功率等数据。从 95598 系统获取停电投诉事件记录数据,包含用户名称、用户号、用户用电类别、用户社会影响度、是否重点保障类用户、投诉时间段、供电区域和重大事件场景等数据,整理汇总后将其进行数据预处理,包含对终端、电能表上送的停电事件进行有效性辨识、异常字符甄别处理、非数字字符处理、异常户号识别处理等,最终得到停电事件样本集,在此基础上从中提取具有投诉记录的用户,作为停电事件敏感用户样本数据。

根据停电事件样本集中样本数据的用户号,从用电信息采集系统中获取其正常供电下的用电采集数据,包含电压、电量、电流、采集成功率等,将其构建成用电正常样本集,将正常样本集和停电样本集混合,构成停电事件样本集,并按照 7:3 比例分为训练集和测试集。

将停电样本数据和停电投诉记录数据中的用户用电类别、用户社会影响度、是否重点保障类用户、是否存在停电投诉记录四类因素作为重点敏感用户的标记,生成重点敏感用户的停电投诉风险等级标签,完成重点敏感用户标签库的构建^[7]。

2.1 提取停电投诉风险影响因子

根据获取到的用户的用电采集数据、停电投诉风险标签,采用时序分解结合权值分配,提取重点敏感用户的停电投诉风险影响因子。

基于重点敏感用户的历史日用电量数据,采用时间序列分解算法提取用户随时间变化的日用电量指标值,通过计算不同时段的用电量指标值,作为不同用户的停电投诉风险时间影响因子。

基于重点敏感用户的停电投诉记录及重大事件场景,设定用户随重大事件场景变化产生的投诉率作为停电投诉风险场景影响因子。将重大事件场景分为 4 类:一级重大事件、二级重大事件、三级重大事件,是否恶劣天气,设非重大事件场景时停电故障引起的投诉率为平均投诉率 λ ,则重大事件场景时

停电故障引起的投诉率为:

$$\lambda_{\theta} = \omega_{\theta} \times \lambda \quad (1)$$

其中, ω_{θ} 表示重大事件权重因子。设非重大事件场景时 $\omega_{\theta} = 1$,一级重大事件时 $\omega_{\theta} = 1.2$,二级重大事件时 $\omega_{\theta} = 1.5$,三级重大事件时 $\omega_{\theta} = 2$ 。

基于重点敏感用户的停电记录及供电区域数据,设定不同区域用户产生投诉事件的停电投诉风险区域影响因子。不同供电区域数据主要受配电设备故障率的影响,设备故障率分成 3 个阶段,供电区域内配电设备存在超期服役老化时,该供电区域停电故障率会陡然升高,本文采用 Weibull 分布来拟合设备超期服役老化故障引起的投诉率,作为停电投诉风险区域影响因子 c :

$$c = \begin{cases} \alpha_1, & T < t_1 \\ 1, & t_1 \leq T \leq t_2 \\ \alpha_2, & T > t_2 \end{cases} \quad (2)$$

其中, T 表示运行时间; t_1 表示区域 B 偶发故障期开始时间; t_2 表示区域 C 耗损故障期开始时间,此处, $t_1 = 2, t_2 = 8$ 。投诉率 α_1 和 α_2 可以根据不同设备故障的概率统计求得。其中,配电设备主要考虑柱上负荷开关、变压器、架空线、柱上断路器、电缆、柱上隔离开关等 6 类设备^[8-13]。

2.2 重点敏感用户细化分类

基于现有的用户用电数据,按用户行为将现有用户分为居民生活用电、一般工商业用电、大工业用电、农业生产用电四大类。输入 4 类用户用电样本集 H 和各用户的停电投诉风险影响因子 R ,并初始化 $n \times k$ 隶属度矩阵 $U(0)$, $\mu_j(h_i)$ 为其中的第 i 行第 j 列的隶属度,利用模糊 C 均值进行用户细化分类。

文中设定待聚类用户用电数据样本集为 $H = \{h_1, h_2, \dots, h_n\}$, $m_j(j = 1, 2, \dots, k)$ 表示聚类中心, k 表示类别数目, $\mu_j(c_i)$ 表示第 i 个样本对应第 j 类的隶属度函数,则基于隶属度函数的聚类损失函数可表示为:

$$J = \sum_{j=1}^k \sum_{i=1}^n [\mu_j(h_i)]^b \|h_i - m_j\|^2 \cdot R_i \quad (3)$$

其中, $\sum_j \mu_j(h_i) = 1$, b 表示加权指数,取值为 2。

在聚类过程中,不断地进行聚类中心 $m_j(j = 1, 2, \dots, k)$ 和隶属度矩阵 $U(\mu_j(h_i))$ 的更新迭代,令聚类损失函数 J 对 m_j 和 $\mu_j(h_i)$ 的偏导为 0,求得聚类损失函数的极小值的必要条件如下:

$$m_j = \frac{\sum_{i=1}^n [\mu_j(h_i)]^b x_i}{\sum_{i=1}^n [\mu_j(h_i)]^b} \quad (4)$$

$$\mu_j(h_i) = \frac{\|h_i - m_j\|^{-2/(b-1)}}{\sum_{s=1}^k \|h_i - m_s\|^{-2/(b-1)}} \quad (5)$$

当满足终止条件 $\max_{ij} \{ \|\mu_j(h_i)\|' - \mu_j(h_i) \| \} < \varepsilon$, 则停止更新, 输出细化用户类别; 否则继续重复上述过程。

2.3 构建停电投诉风险识别模型

基于停电事件样本集的训练样本数据 $X = (x_1, x_2, \dots, x_\eta)$, 利用最大最小值标准化算法, 对用户用电采集数据进行归一化处理, 设定最小值为 0, 最大值为 1, 中间值用以下公式计算:

$$x_z = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (6)$$

其中, x_{\max} 表示训练样本的最大值, x_{\min} 表示最小值。

文中设定输入层用户用电样本数据为 $X' = (x'_1, x'_2, \dots, x'_\eta)$, 随机初始化输入层与隐含层连接权值 ω_{ph} , 输出层与隐含层连接权重 ω_{hg} , 同时假定隐含层阈值 α_h 和输出层的阈值 θ_g , 此时隐含层输出数据为:

$$Z_z = f\left(\sum_{p=1}^{\eta} \omega_{pz} \cdot d_p + \alpha_z\right), \quad z = 1, 2, \dots, \zeta \quad (7)$$

其中, f 表示隐含层激励函数; p 表示输入层的节点数; z 表示隐含层的节点数; η 表示样本个数。

将用户的第 t 时间段的电压、电量、电流、采集成功率、气象数据、节假日标记作为 BP 神经网络输入层的输入数据, 输出层输出用户第 $t+1$ 时间段的停电识别概率 Y_g 为:

$$Y_g = \sum_{z=1}^{\zeta} Z_z \omega_{zg} + \theta_g, \quad g = 1, 2, \dots, \beta \quad (8)$$

将停电识别概率值与真实停电情况进行对比, 误差函数 E 的计算为:

$$E_g = Y_g - O_g, \quad g = 1, 2, \dots, \beta \quad (9)$$

其中, O_g 表示实际是否停电。

根据识别误差 E 更新网络的连接权值和阈值:

$$\omega_{ph} = \omega_{ph} + \psi H_h (1 - H_h) d_p \sum_{g=1}^{\beta} \omega_{hg} E_g \quad (10)$$

$$\omega_{hg} = \omega_{hg} + \psi H_h E_g \quad (11)$$

$$\alpha_h = \alpha_h + \psi H_h (1 - H_h) \sum_{g=1}^{\beta} \omega_{hg} E_g \quad (12)$$

$$\theta_g = \theta_g + \psi E_g \quad (13)$$

其中, ψ 表示权重参数。

通过不断反馈迭代, 不断修正网络连接权值和阈值, 直至识别误差达到要求, 便可终止模型训练, 固定参数, 输出停电投诉风险识别模型。

根据重点敏感用户的等级标签、停电投诉风险影响因子和停电识别概率, 采用加权投票算法计算不同用户的停电投诉概率, 若重点敏感用户停电投诉概率超过所在分类的设定阈值时, 则给运维工作人员发送预警信息。

加权投票算法计算如下:

$$H(x) = c \argmax_{i=1, \dots, T} \sum_j \omega_i h_i^j(x) \quad (14)$$

其中, ω_i 表示 h_i 的权重, 通常 $\omega_i \geq 0$, $\sum_{i=1}^T \omega_i = 1$ 。根据不同用户和不同场景, 设定重点敏感用户的等级标签、停电投诉风险影响因子和停电识别概率的权重, 通过加权投票得到用户的停电风险投诉概率。

3 实验结果与分析

本文中以 95598 系统为实验平台, 获取 2021 年某公司所受理意见工单总数 1 076 件。除电力短信类问题外 (占比 27.5%), 供电可靠性问题占比 19.52%, 居意见类工单第二位, 且全部为频繁停电类。2021 年, 共受理频繁停电问题 256 件。其中, 意见 196 件, 投诉 46 件, 且频繁停电类投诉占投诉总量的 64.38%。以停电工单为例, 采用 2021 年 1~12 月的用电数据, 开展本文所述停电敏感用户分析识别方法的性能测试。选取其中 25 家重点和敏感用户的用电数据作为测试数据, 数据采集时间间隔为 15 min。

将本文所提方法与采用模糊层次法、指数平滑和 GRU 三种算法进行停电投诉风险识别准确率对比。在对比测试中, 利用均方根误差 (Root Mean Squared Error, RMSE) 和平均绝对误差 (Mean Absolute Error, MAE) 作为评价指标对模型识别准确性进行衡量。RMSE 和 MAE 的结果越小, 表示具有越高的识别精度。定义如下:

$$RMSE = \sqrt{\frac{\sum_i^n (Y_i - \hat{Y}_i)^2}{n}} \times 100\% \quad (15)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (16)$$

其中, Y_i 表示停电识别概率值; \hat{Y}_i 表示停电真实值; n 表示测试样本集中的用户用电数据的数量。

本文所提方法与其他 3 种算法的结果对比见表 1。

表 1 算法性能对比

方法名称	RMSE	MAE
本文方法	9. 83	2. 16
模糊层次法	27. 54	14. 37
指数平滑	36. 19	23. 51
GRU	38. 05	26. 43

由表 1 对比结果可以看到,其他 3 种算法虽然在一定程度上也具有不错的识别性能,但是本文所提方法的识别准确率具有明显的优势。本文采用了时序分解算法来构建停电投诉风险因子,且通过细化用户分类+BP 神经网络算法构建的敏感用户停电投诉风险识别模型,从而实现重点敏感用户的停电精准识别,有效提升了客户满意度和供电公司的服务水平。

为了验证本文所提基于时序分解提取停电投诉风险影响因子的贡献,此处以消融实验进行对比分析,即在本文方法中,分别控制变量为无时序分解提取停电投诉风险影响因子(方法 2)、不利用模糊 C 均值细化聚类(方法 3)和不采用 BP 神经网络算法进行停电识别(方法 4),并将上述 3 种方法与本文方法(方法 1)的识别精度进行对比,同样选择测试集样本进行测试, RMSE 和 MAE 作为评价指标,对比结果见表 2。

表 2 停电投诉风险识别准确率对比

方法名称	RMSE	MAE
方法 1	10. 35	2. 97
方法 2	39. 28	24. 17
方法 3	45. 63	36. 58
方法 4	37. 64	29. 05

由表 2 可以看出,采用方法 2 进行识别时,由于实际供电以及用电过程中停电的诱因较多,导致停电及时上报存在一定难度且识别准确率较低,而采用时序分解可对停电影响度大的因子进行特征提取,从而有效提升停电投诉风险识别准确率;对比方法 3,本文方法利用模糊 C 均值算法进行用户类别细化,在已有的用户分类基础上,考虑用户体量、用电趋势、各个停电影响因素的影响度进行聚类,可提高类内用户在多维度的相似度,从而降低了后续建模难度;对比方法 4,本文方法所利用的 BP 神经网络可以实现对高维非线性数据的拟合,并且能够适应不断变化的环境,自动调整权重和阈值,以适应新的情况,可实现用户停电投诉风险及时精准识别,从而提高供电公司的服务水平。

4 结束语

文中提出了一种基于采集指标细化分析的停电风险识别方法,该方法首先基于时间序列分解算法构建各类用户的停电投诉风险影响因子;随后,基于用户类别和停电投诉风险影响因子,采用模糊 C 均值算法进行了用户类别的细化;然后,采用 BP 神经网络构建了停电投诉风险识别模型,通过对比停电识别概率值与真实停电情况,不断更新迭代,逐步优化了模型识别准确率;最后,利用加权投票算法获取到不同用户的停电投诉概率。本文停电投诉风险识别模型的提出,提高了敏感、重要客户停电范围的精准研判,以预警数据指导工作人员及时处理客户停电事件,降低了客户投诉风险,也有效提升了用户满意度和电网安全性。

参考文献

[1] 张英,谢晨,苟庆林,等. 配电系统中的频繁停电预警状况分析[J]. 电子技术, 2023, 52(9):260-261.

[2] 张冬,谢世朋. 配网不停电作业全流程安全监测与预警系统[J]. 计算机技术与发展,2022,32(5):176-181.

[3] 钱利宏,彭穗,郭晓燕,等. 基于多源数据融合与深度信念网络的台区停电故障预警方法研究[J]. 广东电力,2022,35(12):32-40.

[4] 卢颖,任庭昊,覃禹铭. 基于模糊层次法的配电网停电风险预警系统[J]. 自动化技术与应用,2023,42(3):154-157.

[5] 宁波,苏波,李江鹏,等. 基于物联网的大电网停电故障监控及预警系统设计[J]. 电子设计工程,2020,28(21):70-74.

[6] 徐铭铭,牛荣泽,谢芮芮,等. 多源信息融合的配电网重复多发性停电在线监测与预警技术[J]. 山东科学,2020,33(4):117-123.

[7] 徐晓春,王栋,李佑伟,等. 基于多源信息融合的配电网停电风险预警方法研究[J]. 机械与电子,2020,38(2):3-6.

[8] 谭文,杨雪梅,叶敏等. 基于大数据的配电网停电分析与管理[J]. 四川电力技术,2018,41(5):5-9.

[9] 徐铭铭,曹文思,姚森,等. 基于模糊层次分析法的配电网重复多发性停电风险评估[J]. 电力自动化设备, 2018, 38(10):25-31.

[10] 赵文清,祝玲玉,高树国,等. 基于多源信息融合的电力变压器故障诊断方法研究[J]. 电力信息与通信技术, 2018, 16(10):30-35.

[11] 任玮. 基于深度信念网络的网络流量预测模型[J]. 山西电子技术, 2016(1): 62-64.

[12] WU Xiaomin, WANG Dianhong, CAO Weihua, et al. A genetic algorithm support vector machine and DS evidence theory based fault diagnostic model for transmission line[J]. IEEE Transactions on Power Systems, 2019, 34(6):4186-4194.

[13] YUAN Chuanlai, LIAO Yongyi, KONG Lingshuang, et al. Fault diagnosis method of distribution network based on time sequence hierarchical fuzzy petri nets[J]. Electric Power Systems Research, 2021, 191:106870.