

基于 KNN 算法的配电网单相接地故障选线研究

陈 霄, 居 荣

(南京师范大学南瑞电气与自动化学院, 江苏 南京 210023)

[摘要] 小电流接地系统单相接地故障选线是配电网领域的一个难题, 针对传统的采用单一判据的故障选线方案适用性差、选线精度低的问题, 提出了一种基于 K 近邻(K-nearest-neighbor, KNN)算法的多源信息融合的单相接地故障选线方法, 通过对故障数据处理选取故障特征量, 结合 KNN 算法进行故障线路选线。算例仿真研究表明, 该选线方法与逻辑回归算法、BP 神经网络算法相比, 在获得较高的准确率的同时可缩短选线时间, 具有较好的应用前景。

[关键词] 小波分析, K 近邻(K-nearest-neighbor, KNN)算法, 故障选线, 小电流接地系统

[中图分类号] TM773 [文献标志码] A [文章编号] 1672-1292(2020)03-0027-05

Research on Single Phase to Ground Fault Line Selection of Distribution Network Based on KNN Algorithm

Chen Xiao, Ju Rong

(School of NARI Electrical and Automation, Nanjing Normal University, Nanjing 210023, China)

Abstract: Single phase to ground fault line selection is a difficult problem in the field of distribution network. In view of the poor applicability and low line selection accuracy of traditional fault line selection schemes that use a single criterion, this paper proposes a single phase to ground fault line selection method based on K -nearest-neighbor(KNN) algorithm combined with multi-source information fusion, which selects the fault feature quantity for the fault data processing and combines the KNN algorithm to select the fault line. The simulation study of the calculation example shows that the line selection method can obtain higher accuracy and significantly reduce the line selection time compared with the logistic regression algorithm(LoR) and the back propagation neural network algorithm(BP), and that it has a better application prospect.

Key words: wavelet analysis, K -nearest-neighbor(KNN) algorithm, fault line selection, small current grounding system

配电网作为连接电力用户的最后一环,在电力系统中起着重要作用。目前我国中压配电网一般采用小电流接地方式运行^[1],该运行方式下发生单相接地故障的情况约为总故障情况的 80%以上^[2-3],发生单相接地故障后按照惯例可以在故障存在的情况下运行 1~2 h,但是处理不及时会对设备的绝缘造成极大的损坏^[4]。国内外关于配电网的故障选线方法已经做了大量的研究,依据选用故障特征量的数量不同可以分为单一故障特征量选线方法^[5-6]和多源信息融合选线方法^[7-8]。考虑到故障工况的复杂度,单一故障信号选线方法在不同工况下适应程度不同,容易造成误判。多源信息融合的选线方法综合多种故障特征量,为小电流接地系统单相接地故障选线提供了行之有效的方法。文献[9]利用 D-S 证据论的方法针对不同的选线方法构造有效域、可行域和不可行域,针对三种不同接地方式确定各条关键线路的基本信度分配值,依据基本信度分配值以及不确定度的大小做出选线决策,有效地提升了选线准确率,但在复杂工况下故障线路与非故障线路的边界仍旧难以界定。文献[10]在现有选线的基础上将神经网络与多判据选线方案结合的同时,引入人工蜂群算法,将机器学习算法与群体智能算法融合,有效地防止网络陷入局部最优,提升了选线的准确率,但是存在训练速度较慢的问题。为了在提升选线准确率的同时获得较快的选线速度,本文提出一种基于 K 近邻(K-nearest neighbor, KNN)算法的小电流接地系统单相接地故障选线的方法,利用小波变换^[11]分析故障暂态

收稿日期:2020-05-16.

通讯作者:居荣,教授,研究方向:继电保护技术、新能源控制技术及微电网控制。E-mail:jurong@njnu.edu.cn

数据,提取故障暂态特征量作为输入量,对 KNN 算法进行训练,形成故障选线模型,经仿真验证,该方法在保证选线准确率的同时可以大幅提升选线速度.同时,将本文的方法与 BP 神经网络算法与逻辑回归分类算法对比分析,体现了本文提出的基于 KNN 算法的故障选线模型的优越性.

1 基本理论分析

KNN 算法可以解决分类问题也可以解决回归问题.其核心思想是未标记样本的类别由距离最小的 K 个近邻样本投票决定^[12],将已知的样本数据集和将来可能出现的待分类样本用向量的形式表示,计算待分类样本和数据集中的每个样本的距离,选取距离最小的 K 个样本,待标记的样本所属的类别就由 K 个距离最近的样本投票产生.

样本间距离的计算方法的构建,与样本的向量表示方法有关,当采用样本的向量表示方法时,需要考虑其是否便于样本间距离的计算,样本的向量表示与样本间距离的计算方法相辅相成,常见的距离计算方法有:欧式距离、余弦距离、汉明距离、曼哈顿距离等,本文中采用欧式距离的方法计算样本间距离,

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}. \quad (1)$$

式中, d 为两个故障特征量样本间的相对距离; p_i, q_i 分别为不同故障特征量样本的相应点数据.

K 值是一个自定义的变量,是 KNN 算法中一个非常重要的参数. K 的选取会对待分类样本的分类结果产生影响,会影响算法的偏差与方差. K 值越小越容易发生过拟合,增大方差; K 值增大会增大学习的近似误差,影响分类结果的正确性,本文中采用五折交叉验证法选择最优的 K 值.

2 选线判据

2.1 零序有功分量

实际配电系统中线路存在一定的电阻,零序电流中存在一些不能被完全补偿的有功分量,系统故障后,配电网中故障线路与非故障线路的零序有功分量存在较为明显的差别.故障线路的零序有功功率大于非故障线路且方向相反,可以根据零序有功最大值的原理进行故障选线.

零序有功功率 P 为

$$P = \frac{1}{T} \int_t^{t+1} (V(t) \times I(t)) dt. \quad (2)$$

式中, $V(t)$ 表示 t 时刻的零序电压, $I(t)$ 表示 t 时刻的零序电流,在软件仿真中,可以视为对瞬时功率求和.

2.2 高次谐波综合量

当小电流接地系统的中性点接地方式为经消弧线圈接地时,发生单相接地故障后,系统中会存在大量谐波,各次谐波受消弧线圈的影响不同,消弧线圈的电抗值依据系统总的零序基波电容整定,其值满足

$$X_L = \omega_n L = \eta \frac{1}{\omega_n \sum C_i}. \quad (3)$$

式中, ω_n 为基波角频率; C_i 为第 i 条线路的等效零序对地电容; η 为补偿系数.

由式(3)可知,消弧线圈和分布电容对 n 次谐波分量所表现出的电抗呈反比, n 次谐波下系统容抗将减少到基波下的 $1/n$,消弧线圈电抗变为基波下的 n 倍.因此,高次谐波分量受消弧线圈影响较小,本文选择五次谐波、七次谐波、九次谐波幅值和作为选线判据,幅值之和最大的线路为故障线路.

2.3 暂态能量

本文利用小波分析处理故障后每条线路的暂态零序电流,小波变换具有线性的性质,经过小波分解得到的各频段的小波系数具有能量量纲,可利用小波系数表征各线路各频段能量的大小.通过小波分解将线路的暂态零序电流信号分解到各个频段,低频信号容易被消弧线圈补偿,因此选取暂态零序电流信号的高频段部分进行暂态能量的计算.

系统正常运行时,各线路零序电流信号暂态能量为零.系统发生单相接地故障后,故障线路的暂态能量最大且方向相反,计算故障后各线路的零序电流信号的暂态能量,暂态能量最大的一条为故障线路,暂态能量计算公式为

$$E = \sum_{i=1}^n \sum_{j=1}^m [C_{(k,i)}(j)]. \quad (4)$$

式中, n 为选取的频段数目; m 为某频段下的小波系数数目; k 为分解层数, $C_{(k,i)}(j)$ 为 k 层小波分解后第 i 频段下第 j 个小波分解系数。

3 故障线路选线流程

本文所提的故障选线方法流程如图 1 所示, 主要分为 4 个步骤。

(1) 故障判别. 为了正确识别各线路的故障与非故障状态, 线路的过电压应设定一个阈值. 为了消除零序基波电压及零序三次谐波电压的影响, 过电压阈值常设置为相电压的 0.15 倍, 提取各出线的零序电压信号, 当零序电压大于相电压的 0.15 倍时则为故障状态, 否则为非故障状态。

(2) 故障判据收集. 通过处理各出线的零序电流及零序电压信号, 得到各出线的零序有功分量; 采用 Fourier 变换的方法分解各条线路零序电流信号, 采集各线路的 5 次谐波、7 次谐波和 9 次谐波信号, 求取谐波信号的幅值和, 得到各线路高次谐波综合量; 利用小波变换对故障后零序电流处理得到小波分解系数, 进而得到故障后的暂态能量. 将提取的 9 个故障特征量作为输入量, 形成故障选线模型训练样本矩阵 $Y = [p_i, s_i, e_i]$ 。

(3) 样本数据处理. 由于数据大小差异较大, 计算样本间欧式距离时, 数值差值最大的属性对计算结果产生较大影响, 在处理数值差异较大的特征值时, 常用的方法是将数值归一化, 可以采用“极差法”的方式进行, 其公式为

$$x_{i_{\text{new}}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}. \quad (5)$$

式中, $x_{i_{\text{new}}}$ 为归一化后的数据, $x_{i_{\max}}$ 和 $x_{i_{\min}}$ 分别为原始数据中的最大值和最小值。

(4) 辨别分析. 将采集到的故障特征量数据集分为训练样本和测试样本, 利用训练样本训练 KNN 算法, 构建故障选线模型, 再将测试样本输入模型中进行分类, 验证选线模型的泛化能力。

4 仿真分析

4.1 仿真模型

本文使用 Simulink 搭建的 10 kV 小电流接地系统的仿真模型如图 2 所示, 其仿真模型参数如表 1 所示。

模型中的消弧线圈采用过补偿的方式安装, 补偿度设置为 10%, 消弧线圈电感值计算公式如式(6)所示:

$$L = \frac{1}{3\omega^2 C_s (1+p)}. \quad (6)$$

式中, C_s 为配电网零序基波电容之和; ω 为工频角速度; p 为消弧线圈补偿度; 经计算得电感 L 为 9.5 H。

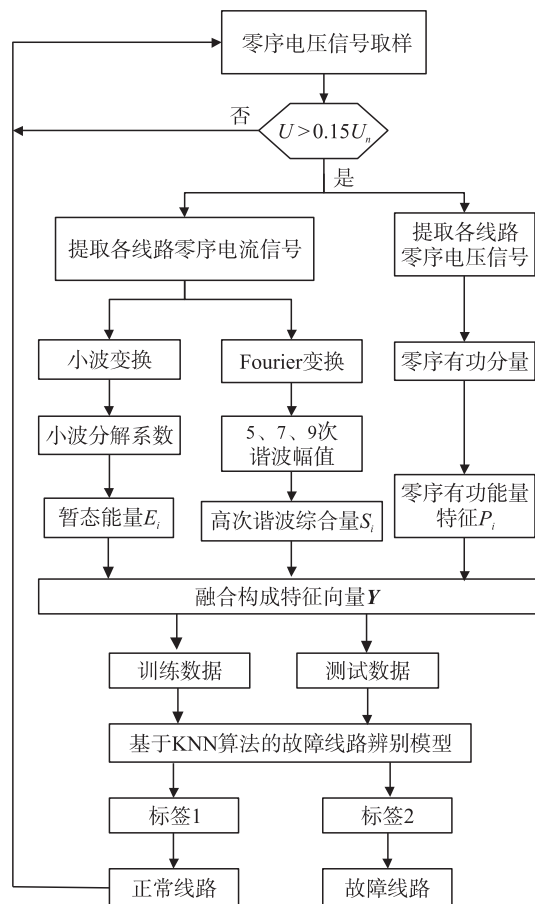


图 1 故障选线流程图

Fig. 1 Flow chart of fault line selection

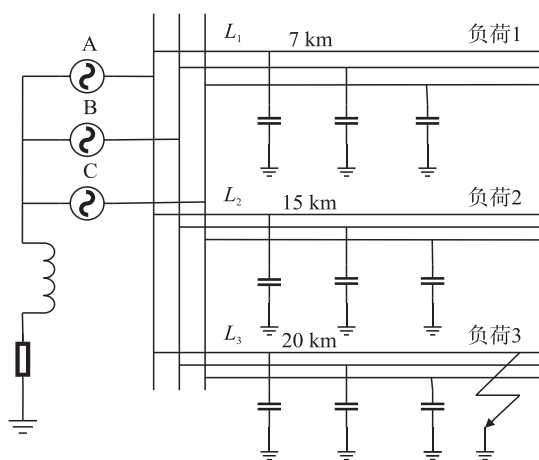


图 2 10 kV 配电网仿真模型

Fig. 2 Simulation model of 10 kV distribution network

4.2 仿真实验

本文针对配电网中性点经消弧线圈接地的不同故障工况进行了大量仿真,设置的故障工况如下:(1)故障线路为线路 $L_1 \sim L_3$;(2)故障位置为线路长度的 10%、30%、60%、90%;(3)接地电阻 R_f 分别为 100 Ω 、500 Ω 、1 000 Ω 、1 500 Ω 、2 000 Ω ;(4)故障合闸角 θ 依次取 0°、45°、90°、120°。共计获取 240 组数据,取其中 200 组数据作为训练数据,40 组作为测试数据,对提出的 KNN 算法进行验证。

本节模拟小电流接地系统中性点经消弧线圈接地系统单相接地故障,设置单相接地故障在线路 L_3 上发生,接地点发生在距母线 2 km 处,故障相为 A 相,且为金属性单相接地故障。单相接地故障发生时间设置为 0.05 s 处,仿真总时间为 0.2 s,采样时间为 1×10^{-5} s。

各条线路的零序有功分量如图 3 所示,非故障线路零序有功分量幅值较小且相位一致,而故障线路 L_3 相较于非故障线路 L_1 、 L_2 相位相反,幅值明显增大,故线路的零序有功分量可作为故障特征量的一种。

图 4 采用傅里叶分解的方法得到故障后各线路的 5 次谐波信号,从图中可以看出,对各次谐波,故障线路 L_3 的谐波幅值大于非故障线路,与上文所述一致。

图 5 给出了故障后选用 db10 作为基函数对非故障线路 L_1 及故障线路 L_3 的零序电流信号进行 5 层小波分解后 d5 频段下小波分解系数的波形,由图可知故障后零序电流信号的暂态过程大概只会持续两个工频周期,其高频信号很快衰减到较小的值,因此本文取故障后 4 000 个采样点的值计算暂态能量。

表 1 仿真模型参数表
Table 1 Parameters of simulation model

参数名称	参数值
系统电压 U/kV	10.5
线路长度 L_1 、 L_2 、 L_3/km	7、15、20
线路正、零序电阻 r_1 、 $r_0/(\Omega/\text{km})$	0.012 73、0.386 4
线路正、零序电抗 x_1 、 $x_0/(\text{mH}/\text{km})$	0.933 7、4.126 4
线路正、零序电纳 b_1 、 $b_0/(\mu\text{F}/\text{km})$	0.012 74、0.007 75
三条出线负荷有功功率 P/kW	1e6、0.2e6、2e6
三条出线负荷无功功率 Q/kvar	0.4e6

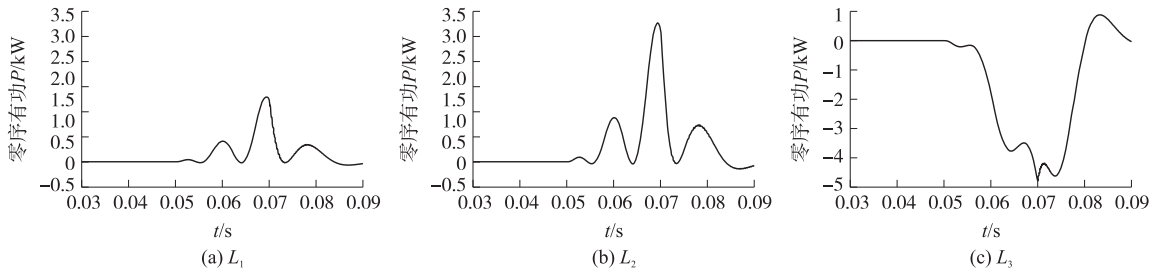


图 3 各线路零序有功分量分布
Fig. 3 Zero sequence active power distribution of each line

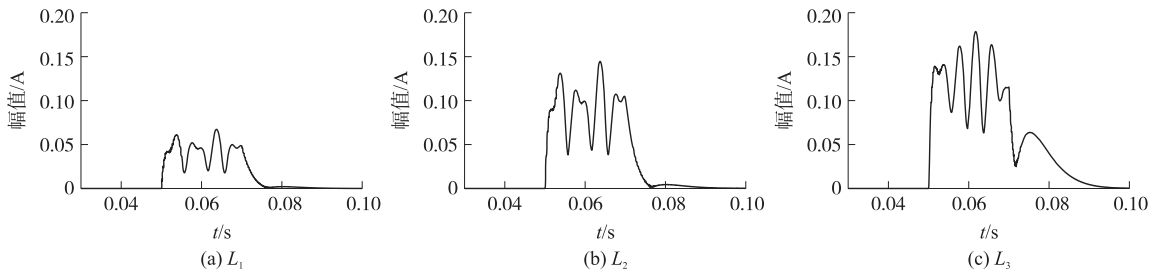


图 4 各线路 5 次谐波幅值
Fig. 4 5rd harmonic amplitude of each line

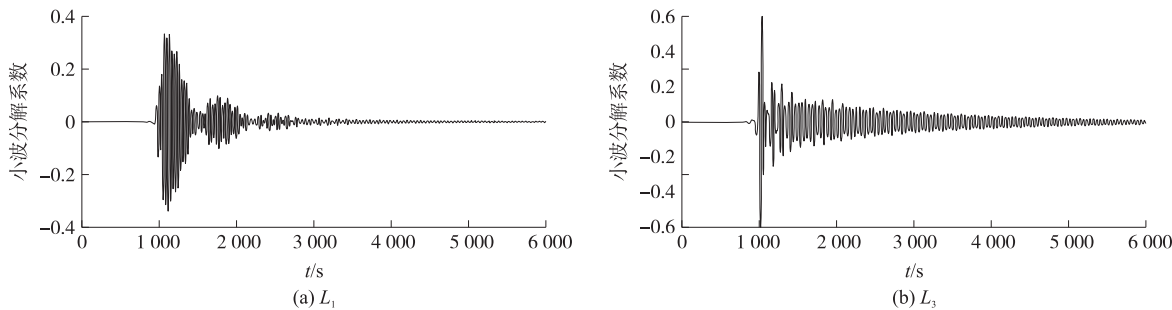


图 5 线路 L_1 、 L_3 零序电流小波分解
Fig. 5 Wavelet decomposition of zero sequence current of lines L_1 and L_3

4.3 算法性能比较

为了了解提出的 KNN 算法故障选线的准确率以及选线速度,在同一平台上,使用同样的训练集以及测试集验证 KNN 算法、BP 神经网络算法、LoR 算法的泛化能力,观察预测计算时间和预测精度。

(1)KNN 算法. 本次测试使用的 KNN 算法样本间距离计算方法采用欧式距离,经五折交叉验证后发现当 $K \geq 7$ 时,故障选线准确率大幅下降,综合考虑选线的准确率、选线速度以及噪声信号的影响,设定 K 值为 7,选线模型测试结果如图 6 所示。

(2)BP 神经网络算法. 本次对比试验使用的 BP 神经网络算法是 9-3-3 三层结构,采用的激活函数为 Sigmoid 函数,训练轮次为 1 000 次,使用的误差函数为平方和误差函数,其表达式为

$$E_p = \sum_{k=1}^m \frac{1}{2} (y_k - t_k)^2. \tag{7}$$

式中, y_k 为经神经网络预测的输出, t_k 为实际的输出. 最终的测试结果如图 7 所示。

(3)LoR 算法. 本次对比试验使用的逻辑回归算法采用主成分分析法提取其中的 5 个主要特征,防止模型过于复杂造成过拟合. 将训练集单独分出一部分作验证集,训练集与验证集的比例为 3:1,防止测试集参与模型评估,造成评估结果优于实际的情况发生. 加入 L_1 、 L_2 正则化项,提升模型的泛化能力,其测试结果如图 8 所示。

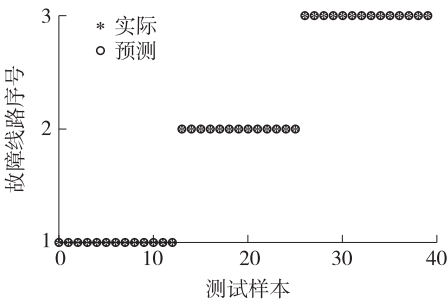


图 6 KNN 算法选线结果
Fig. 6 Line selection result of KNN

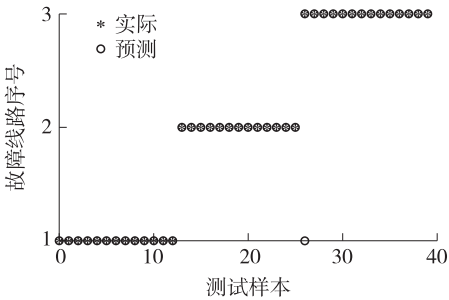


图 7 BP 神经网络算法选线结果
Fig. 7 Line selection result of BP

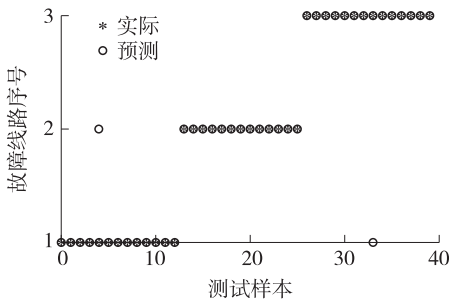


图 8 LoR 算法选线结果
Fig. 8 Line selection result of LoR

三种算法故障选线结果散点图中纵坐标 1、2、3 表示故障线路的序号,图中实际结果对应的点为仿真设定的故障线路序号,预测结果对应的点为故障特征量矩阵经选线模型处理后输出的预测故障线路的序号. 预测故障线路序号与实际故障线路序号相同则为正确,反之则为错误,最终的选线准确率为选线正确的样本数占测试总样本数的比例,三种算法性能比较如表 2 所示。

由表 2 可知相同的条件下,三种算法的选线性能各不相同,KNN 算法在取得最高的选线准确率的同时耗时最短. BP 神经网络算法由于需要迭代训练以及反向传播求梯度的缘故,导致计算复杂程度增加,耗时远高于其他两种算法. LoR 算法虽然能实现较快的选线,但是准确率却不及 BP 神经网络算法及 KNN 算法,实用性不高. 总体上 KNN 算法能在保证选线准确率的情况下大幅缩短选线时间,具有一定的优越性。

表 2 不同算法性能对比

Table 2 Performance comparison of different algorithms				
选线方法	训练样本个数	测试样本个数	选线准确率/%	收敛时间/s
KNN 算法	200	40	100	0.042
BP 神经网络算法	200	40	97.5	7.604
LoR 算法	200	40	95	0.060

5 结论

本文针对小电流接地系统单相接地故障选线提出了一种基于 KNN 算法的选线方法,通过研究分析得到以下结论:

(1)KNN 算法、BP 神经网络算法和 LoR 算法都能实现小电流接地系统单相接地故障选线,但是 KNN 算法的选线准确率更高、耗时更小。

(下转第 92 页)

- [4] 任小静,屈小娥. 我国区域生态效率与环境规制工具的选择——基于省际面板数据实证分析[J/OL]. 大连理工大学学报(社会科学版),2020,41(1):28-36.
- [5] 何念鹏,徐丽,何洪林. 生态系统质量评估方法的再思考——理想参照系和关键指标[J/OL]. 生态学报,2020(6):1-10.
- [6] 刘莎莎. 西北五省金融集聚、经济发展对生态效率的空间溢出效应研究[D]. 石河子:石河子大学,2019.
- [7] 樊霞飞. 安徽省在长三角一体化战略中的发展机遇与路径选择[J]. 山西经济管理干部学院学报,2019,27(3):68-71,85.
- [8] 宋伟东. 基于生态功能重要性和敏感性的安徽省生态安全格局构建[J]. 阜阳师范学院学报(自然科学版),2019,36(3):103-109.
- [9] 朱家明,胡榴榴,王杨,等. 基于熵值-模糊综合评价法的宜居城市排名研究[J]. 中央民族大学学报(自然科学版),2019,28(3):42-47.
- [10] 唐沛贤,肖建武. 湖南省生态-经济-社会耦合协调演化研究[J]. 金融经济,2019(24):7-9.
- [11] 李国柱,李从欣. 基于熵值法的经济增长与环境关系研究[J]. 统计与决策,2010(24):107-109.
- [12] 魏丽莉,杨颖. 西北地区绿色金融与产业结构耦合协调发展的历史演进——基于新结构经济学的视角[J]. 兰州大学学报(社会科学版),2019,47(5):24-35.
- [13] 彭建刚,徐轩. 农业产业化与普惠金融的耦合关系及协调发展——以湖南省为例[J]. 财经理论与实践,2019,40(5):19-26.
- [14] 庞庆华,李铭珍,李涵. 长江经济带金融集聚、区域创新与生态效率的空间耦合协调发展研究[J]. 工业技术经济,2019,38(2):68-76.

[责任编辑:严海琳]

(上接第31页)

(2)本文提出的配电网单相接地故障选线方法适用于35 kV及以下中性点经消弧线圈接地系统,利用多个故障特征量构成选线模型,避免了单一判据选线方法的不足,能在复杂工况下准确判断出故障线路,具有一定的实用价值。

[参考文献](References)

- [1] 张利. 中性点非有效接地系统单相接地故障定位方法的研究[D]. 北京:华北电力大学,2009.
- [2] 郭清滔,吴田. 小电流接地系统故障选线方法综述[J]. 电力系统保护与控制,2010,38(2):146-152.
- [3] 唐金锐,尹项根,张哲,等. 配电网故障自动定位技术研究综述[J]. 电力自动化设备,2013,33(5):7-13.
- [4] 刘东,张弘,王建春. 主动配电网技术研究现状综述[J]. 电力工程技术,2017,36(4):2-7,20.
- [5] 孟润泉,米建军. 五次谐波检测原理及其在矿井高压电网单相接地保护中的应用[J]. 工矿自动化,2004(3):10-13.
- [6] 牟龙华. 零序电流有功分量方向接地选线保护原理[J]. 电网技术,1999,23(9):60-62.
- [7] 舒凡娣,谢嘉晟,廖晓娇,等. 结合粒子群算法和穷举法的配电网故障诊断方法[J]. 智慧电力,2019,47(1):94-99.
- [8] 侯隽朗. 基于BP神经网络多判据小电流接地系统单相故障选线研究[D]. 太原:山西大学,2018.
- [9] 张海欣. 小电流接地系统单相接地故障多判据融合选线方法研究[D]. 秦皇岛:燕山大学,2019.
- [10] 林中鹏. 基于改进蜂群算法优化神经网络的小电流接地故障选线[D]. 青岛:山东科技大学,2018.
- [11] WEI X X, YANG D C. An adaptive fault line selection method based on wavelet packet comprehensive singular value for small current grounding system [C]//2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies. Changsha, China:IEEE,2015:1110-1114.
- [12] 杨慧敏,崔江,张卓然,等. 基于改进支持向量机的故障诊断方法[J]. 电工技术学报,2014,29(增刊1):164-169.

[责任编辑:陈 庆]