

基于 ReliefF-DDC 特征选择算法的非侵入式负荷识别研究*

邵琪¹, 包永强², 姜家辉¹, 张旭旭¹

(1.南京工程学院 电力工程学院, 江苏 南京 211167; 2.南京工程学院 信息与通信工程学院, 江苏 南京 211167)

摘要: 提取有效的负荷运行数据特征对于提高非侵入式负荷识别的精度具有重要作用。针对数据特征选择欠佳导致负荷识别准确率不高的问题, 提出了一种基于 ReliefF-DDC 特征选择算法, 用于降低特征维数减少复杂度, 改善负荷识别效果。首先, 利用 ReliefF 算法分析各特征与类别的关系计算特征权重, 筛选无关特征; 其次, 利用 DDC 算法计算特征之间与类别的互信息分析相关性, 根据特征子集评价度量删除冗余特征; 最后, 采用孪生支持向量机 (TWSVM) 作分类器进行负荷识别。实验表明, 所提出的算法在提升分类效果的同时减少了运行时间。

关键词: ReliefF; DDC; TWSVM; 特征选择; 负荷识别

中图分类号: TN911; TM714

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.200524

中文引用格式: 邵琪, 包永强, 姜家辉, 等. 基于 ReliefF-DDC 特征选择算法的非侵入式负荷识别研究[J]. 电子技术应用, 2021, 47(7): 74-77, 82.

英文引用格式: Shao Qi, Bao Yongqiang, Jiang Jiahui, et al. Research on non-intrusive load identification based on ReliefF-DDC feature selection algorithm[J]. Application of Electronic Technique, 2021, 47(7): 74-77, 82.

Research on non-intrusive load identification
based on ReliefF-DDC feature selection algorithmShao Qi¹, Bao Yongqiang², Jiang Jiahui¹, Zhang Xuxu¹

(1.School of Electrical Engineering, Nanjing Institute of Technology, Nanjing 211167, China;

2.School of Information and Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China)

Abstract: Extracting effective characteristics of load operation data plays an important role in improving the accuracy of non-intrusive load identification. In this paper, a ReliefF-DDC feature selection algorithm was proposed to reduce feature dimension, reduce complexity and improve load recognition. Firstly, ReliefF algorithm was used to analyze the relationship between each feature and category, calculate feature weight, and screen irrelevant features. Secondly, DDC algorithm is used to calculate the mutual information analysis correlation between features and categories, and redundant features are removed according to feature subset evaluation measurement. Finally, twin support vector machine (TWSVM) is used as classifier for load recognition. Experiments show that the algorithm proposed in this paper improves the classification effect and reduces the running time.

Key words: ReliefF; DDC; TWSVM; feature selection; load identification

0 引言

非侵入式负荷监测法 (Non-Intrusive Load Monitoring, NILM) 为实现智能电网和用户之间的互动提供了数据支持, 该方法在接户线入口处安装传感器, 采集总负荷的电压、电流等电气量数据进行分析, 细化系统数据, 从而辨识家用电器的类别及运行状态^[1]。相比于侵入式负荷监测法 (Intrusive Load Monitoring, ILM), NILM 具有成本低、用户接受度高、后期维护方便等优势, 但是该方法对于负荷分解算法的要求较高。特征提取和负荷识别作为 NILM 中两大关键技术^[2], 为 NILM 的发展提供了强有力

的技术支持。特征选择作为处理已提取特征的重要手段, 是目前研究的热点之一。

特征选择是在原始高维特征中遵循某个评价准则为后续任务选择一个最佳特征子集, 该理论基于少量具有代表性的特征, 不仅可以加速模型的学习过程, 而且可以提高模型的泛化能力。特征选择在图像处理、数据挖掘、机器学习等领域应用广泛^[3], 在处理包含了大量特征的高维数据时, 这些特征不可避免地包含了噪声、不相关特征以及冗余特征^[4]。在这种情况下, 有必要提取出信息量最丰富、最有价值的信息。KIRA K 和 RENDEL L A 提出了 Relief 算法, 该算法简单运行效率高, 但仅适用于二分类问题的特征选择^[5]。在此基础上, KONONENKO I 提出了 ReliefF 算法, 解决了原 Relief 算法无法对多类别

* 基金项目: 国家自然科学基金项目 (61702258); 国家重点研发计划 (2017-YFB0903504)

数据进行特征选择的问题,且对于不完整和带噪数据处理效果较好,但无法删除冗余特征^[6]。YU L 和 LIU H 使用快速相关滤波器算法(Fast Correlation-Based Filter,FCBF)算法来减少高维数据中的特征,归一化信息增益利用对称不确定性来度量特征冗余度,算法可快速消除冗余特征但对称不确定测量精度不高且无法量化特征之间与类别的依赖关系^[7]。文献[8]中 Peng Hanchuan 等人提出了最大相关-最小冗余(mRMR)算法,采用互信息作为衡量特征和类别以及特征与特征之间冗余度的准则,但同样未考虑特征之间与类别的相关性。QU G 等人提出了决策相关分析(Decision Dependency Correlation,DDC)算法,指出所选特征与决策变量密切相关^[9]。

在现有研究基础之上,针对特征选择方法存在的问题展开研究,本文提出了一种基于 ReliefF-DDC 特征选择算法。首先通过计算特征权重并按降序排列,选取权重较大的特征去除无关特征;其次计算各特征与决策变量之间的交互信息,利用决策相关分析删除冗余特征得到最终特征子集;最后将其作为孪生支持向量机(TWSVM)的输入向量进行识别。实验以非侵入式用电负荷为研究对象,将本文方法与 ReliefF、DCC 单项算法以及现有研究文献[10]和文献[11]所提算法进行识别效果比对,结果表明,本文所提出的特征选择方法可有效提高负荷识别精度。

1 特征选择

1.1 ReliefF 算法

ReliefF 算法是一种基于特征权重的高效过滤式特征选择算法,该算法根据特征属性对于各类样本的区分能力来估计特征权重衡量特征重要性^[12]。设待处理的训练集为 D , 样本 $X=\{x_1, x_2, \dots, x_d\}$, x_d 是表示第 d 维特征。在训练集 D 中随机选取一样本 R , 在与 R 同类的样本中寻找 k 个最近邻样本 $H_j(j=1, 2, \dots, k)$; 在与 R 不同类的样本中寻找 k 个最近邻样本 M_j 。若样本 R 与最近邻样本 H 在该特征上的距离小于样本 R 与不同类最近邻样本 M 的距离,此时该特征在同类和不同类最近邻样本上区分性较好,特征权重增大;相反,则表明该特征在区分同类和不同类最近邻样本效果不佳,权重则降低。其实现如下:

$$W'(d) = W(d) - \sum_{j=1}^k \text{diff}(x_d, R, H_j) / mk + \sum_{C \neq \text{class}(R)} \frac{P(C)}{1 - P(\text{class}(R))} \sum_{j=1}^k \text{diff}(x_d, R, M_j(C)) / mk \quad (1)$$

式中, $P(C)$ 表示类 C 在数据集中的先验概率分布, $M_j(C)$ 表示第 C 类第 j 个最近邻样本。其中 $\text{diff}(d, A, B)$ 表示样本 A 和样本 B 在第 d 个特征值上的区分度, m 表示循环次数。

当特征 d 的值离散:

$$\text{diff}(d, A, B) = \begin{cases} 0 & \text{value}(d, A) = \text{value}(d, B) \\ 1 & \text{value}(d, A) \neq \text{value}(d, B) \end{cases} \quad (2)$$

当特征 d 的值连续:

$$\text{diff}(d, A, B) = \frac{|\text{value}(d, A) - \text{value}(d, B)|}{\max(d) - \min(d)} \quad (3)$$

ReliefF 算法通过计算特征与各类别之间的相关性来确定“重要特征”,排除无关特征,但其计算过程中忽略了冗余特征,模型中特征之间相似度越高,会导致模型训练时间越长,造成空间浪费,同时泛化能力降低^[13],造成分类精度下降。

1.2 DDC 算法

考虑到特征之间的相关性与冗余程度对决策变量的依赖程度^[14],DDC 算法基于特征之间与决策变量之间的交互信息,通过使用一个简单的 SFS 计算方法获得特征的排序列表,再根据特征之间与决策变量的相关性分析去除冗余特征。

随机变量 X 的熵 $H(X)$ 表示该随机变量的不确定性测度。互信息是指随机变量 X 与随机变量 Y 的共享信息量,其定义如下:

$$I(X, Y) = H(X) - H(X|Y) \quad (4)$$

由式(4)可知,互信息还可看成是两个随机变量的熵的交集,使用此属性可以描述特征的相关性和冗余性。设 f 和 s 分别为两个特征, C 为类别,考虑到所选特征必须与类别有最大相关性,而与其他特征相关性需最小,因此引入了相关测度来量化特征 f 和 s 与类 C 之间的冗余程度,表示如下:

$$Q_c(f, s) = \frac{I(C; f) + I(C; s) - I(C; f, s)}{H(C)} \quad (5)$$

再者,利用 $I(C; f)$ 与 $Q_c(f, s)$ 共同构成特征子集评价准则,定义如下:

$$e(S) = \frac{I(C; f)}{H(C)} - \sum_{s \in S} Q_c(f, s) \quad (6)$$

式(6)评价准则直观指定一个子集 S , 其中与决策相关的各个特征的相互信息 $I(C; f)$ 作为对该特征子集的奖励,而特征之间的决策相关 $Q_c(f, s)$ 作为惩罚,该评价价值 $e(S)$ 越大,表明决策时该所选特征子集越优。

2 ReliefF-DDC 特征选择算法

ReliefF 算法依据特征权重来衡量属性“重要性”,通过计算样本与最近邻同类样本和异类样本在各个特征上的距离获取权重值^[15]。ReliefF 算法效率高、鲁棒性好,而且能有效处理带噪数据,降低噪声对于特征选择的影响^[16]。由于只考虑特征与类别之间的关系,ReliefF 算法只能有效去除无关特征,但无法删除冗余特征。针对在不确定性度量精度不高的情况下,所选特征可能提供错误或者不完整信息的问题,DDC 算法通过引入新测度准确量化特征之间的依赖关系或者相关性,提高了剔除冗余特征的准确性,但其去除无关特征时的能力不如 ReliefF 算法。

基于以上分析,本文结合了两个特征选择算法优点,提出了一种基于 ReliefF-DDC 特征选择算法实现特征选择分阶段处理。算法具体结构如图 1 所示。

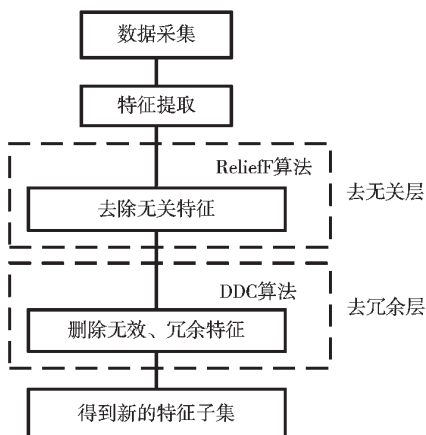


图 1 ReliefF-DDC 特征选择算法结构

为了进一步阐述本文算法,给出无关特征层和去冗余特征层的具体步骤如下:

(1)去无关特征层:利用 ReliefF 算法计算各特征权重,按降序排列后去除无关特征,得到子集 F 。

(2)去冗余特征层:将特征子集评估度量 $e(S)$ 与指定阈值 δ 的比较作为判断条件。对于每次遍历,选择同时满足两个条件的 f_j 将其放入最优子集中:

- ①与类别的相关性大于集合 F 中其他特征;
- ②与已选特征子集中所有特征的相关性最小。

ReliefF-DDC 特征选择算法的输入为原始特征数据集 D ,输出为最佳特征子集 F_{best} 。算法执行步骤如图 2

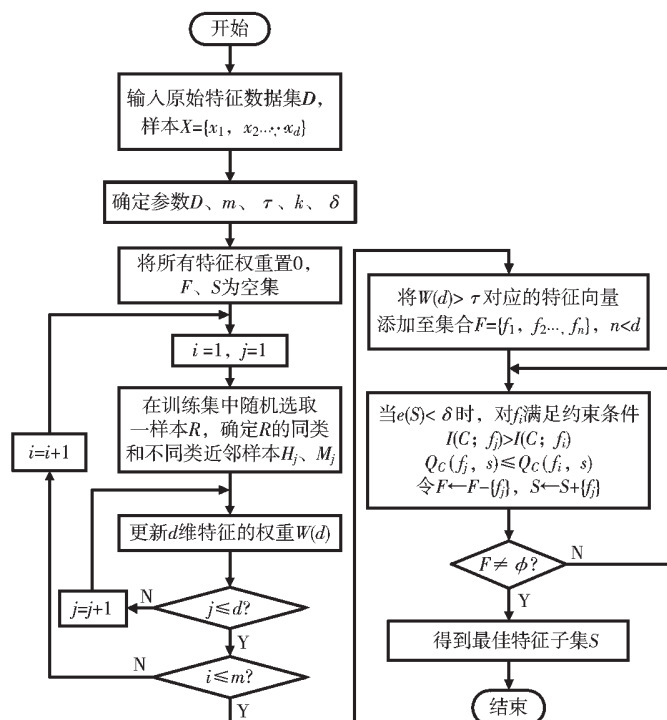


图 2 ReliefF-DDC 特征选择算法流程图

所示。

(1)输入提取的原始特征数据集 D ,确定迭代次数 m 、特征权重阈值 τ 、最近邻样本个数 k 、评价准则阈值 δ ;

(2)将所有特征权重置 0, F 为空集;

(3)令 i 从 1 循环至 m ,

①在训练集 D 中随机选取一样本 R ,在与 R 同类的样本中寻找 k 个最近邻样本 $H_j(j=1, 2, \dots, k)$,与 R 不同类的样本中寻找 k 个最近邻样本 M_j ;

②按照式(1)和式(3)更新 d 维特征的权重:

(4)输出 $W(d)$ 中大于阈值 τ 时对应的特征向量,按降序排列添加至集合 $F, F=\{f_1, f_2, \dots, f_n\}, n < d$;

(5)置 S 为空集;

(6)当 $e(S) < \delta$ 时,

①若 f_i 满足以下两条件:

(a) $I(C; f_j) > I(C; f_i) \quad \forall i \neq j, f_i \in F$

(b) $Q_c(f_j, s) \leq Q_c(f_i, s) \quad \forall i \neq j, f_i \in F, s \in S$

则令 $F \leftarrow F - \{f_j\}, S \leftarrow S + \{f_j\}$

②若 $F \neq \emptyset$,跳转至步骤(6);

(7)得到最佳特征子集 S ,结束。

3 实验与分析

3.1 实验环境的搭建

本文搭建了一套数据采集系统,包括分压模块、隔离电路模块以及 VK701H 数据采集卡等,如图 3 所示。实验模拟家用电器的运行环境,利用该数据采集系统采集了若干个典型用电设备稳定运行时的电流数据各 60 组,在 MATLAB2016a 平台上进行实验。

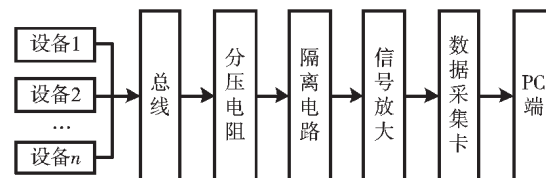


图 3 数据采集系统结构图

实验针对性地采集了 8 种家用电器的稳态电流数据用于负荷识别,采样频率为 10 kHz,负荷类型与具体参数如表 1 所示。

表 1 负荷类型与具体参数

序号	名称	型号	额定功率/W
1	节能灯	欧普 7140280	40
2	笔记本电脑	联想 INS 15-5593	65
3	电视机	海信 HZ55E3D-J	110
4	取暖器	长虹 CDN-RT01TQ	1 000
5	电热水壶	美的 MK-SH15	1 500
6	空调	海尔 KFR-35GW	1 800

为了证明本文所提出的方法在负荷特征选择上的有效性,对比实验设置 ReliefF、DCC 单项算法。为了进一

步表明本文方法优越性,另外设置现有研究文献[10]中提出的 Re-FCBF 算法和文献[11]中利用 MRMD 结合 ReliefF 算法两种方法作为对比,利用上述算法分别对特征数据进行选择。实验过程中将采集的数据 210 组作为训练样本,其余 150 组作为测试样本,采用 TWSVM 进行负荷识别,通过对比分析各特征选择方法的识别准确率来判断其算法性能。

3.2 实验结果及分析

本文对采集的各用电负荷的稳态电流数据进行分析,提取其相关时频域特征,包括电流谐波总畸变率、3 次谐波幅值、5 次谐波幅值、7 次谐波幅值以及其倒谱低频部分系数特征,共 105 维。

按第 2 节所述算法流程,首先利用 ReliefF 算法对提取的原特征数据进行处理,设定迭代次数 $m=20$,最近邻样本个数 $k=10$,特征权重阈值 $\tau=0.02$,此时得到降序排列后 38 维特征子集,如表 2 所示。

表 2 特征权重及对应维数

$W(d)$	d	$W(d)$	d	$W(d)$	d	$W(d)$	d
0.202 15	3	0.045 03	21	0.034 00	93	0.027 66	82
0.185 22	6	0.044 02	19	0.033 81	102	0.024 19	31
0.149 52	4	0.040 07	26	0.033 25	61	0.024 17	27
0.112 13	70	0.039 80	13	0.032 47	105	0.023 71	50
0.109 95	5	0.038 16	65	0.032 4	94	0.021 97	85
0.063 08	1	0.036 82	28	0.031 59	29	0.021 60	91
0.057 50	32	0.035 79	96	0.031 07	30	0.021 04	103
0.055 01	33	0.034 93	39	0.029 07	92	0.020 30	14
0.047 54	20	0.034 79	42	0.028 71	83		
0.047 26	38	0.034 26	41	0.028 28	18		

其次利用 DDC 算法去除冗余特征,置评价准则阈值 $\delta=0.9$,计算得到最终 24 维特征子集:

$$S=\{3, 6, 4, 70, 5, 1, 32, 38, 19, 26, 65, 39, 42, 102, 61, 29, 30, 83, 18, 97, 31, 50, 85, 103\} \quad (7)$$

最后将经各算法处理后得到的不同特征子集分别作为孪生支持向量机的输入特征向量进行负荷识别。各算法用时结果如表 3 所示,所得用电负荷在不同算法下的识别率如图 4 所示。

表 3 各算法运行时间

算法	ReliefF	DDC	文献[10]	文献[11]	Re-DDC
运行时间/s	20.3	17.2	15.4	15.6	11.4

从表 3 及图 4 可以看出,对于大功率负荷,文献[10]与文献[11]特征降维效果与 DDC 算法相差不大。整体而言,本文提出的方法相较于其他 4 种算法,负荷整体识别率明显提高,分别为 10.34%、7.5%、3.93%以及 4.1%,同时模型运行时间较其他算法相对减少了 8.6 s、5.8 s、4.0 s 和 4.2 s,表现出了较好的鲁棒性。

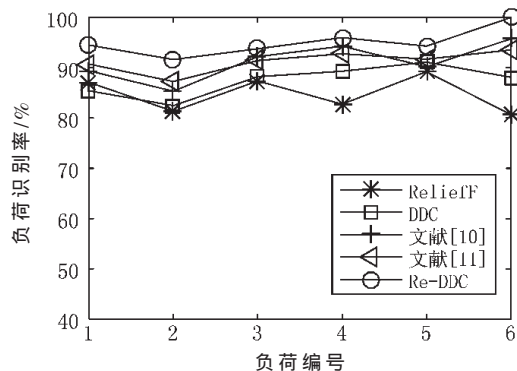


图 4 用电负荷在各算法下的识别率

4 结论

本文以非侵入式负荷监测为研究背景,针对数据特征选择优劣影响负荷识别准确率高低的问题,提出了一种基于 ReliefF-DDC 特征选择算法。算法上层对各用电负荷进行特征提取,利用 ReliefF 算法计算所选样本在各特征上到最近邻同类样本和不同类样本的距离得到各特征权重,按降序排列后依照设定的权重阈值去除无关特征;算法下层利用 DDC 算法通过计算互信息来分析特征之间与类别的依赖程度,将特征子集评价度量与设定阈值之间比较作为判断准则,从而删除冗余特征;最后利用孪生支持向量机识别分类观察识别率。实验结果表明,本文提出的方法有效提高了负荷识别率,且缩短了运行时间。

参考文献

- [1] 程祥,李林芝,吴浩,等.非侵入式负荷监测与分解研究综述[J].电网技术,2016,40(10):3108-3117.
- [2] 孙毅,崔灿,张璐,等.智能用电非侵入式负荷监测系统研究[J].电力科学与技术学报,2019,34(2):155-160.
- [3] 于宁宁,刘刚,刘森,等.基于多评价准则融合的特征选择方法[J].计算机工程与设计,2018,39(7):2075-2079.
- [4] 曹彦.基于支持向量机的特征选择及其集成方法的研究[D].郑州:郑州大学,2010.
- [5] KIRA K,RENDELL L A.A practical approach to feature selection[C].Proceedings of the Ninth International Workshop on Machine Learning(ML 1992),Aberdeen,Scotland,UK, July 1-3,1992.Morgan Kaufmann Publishers Inc.,1992.
- [6] KONONENKO I.Estimating attributes: analysis and extension of relief[C].European Conference on Machine Learning. Springer, Berlin, Heidelberg, 1994.
- [7] YU L,LIU H.Efficient feature selection via analysis of relevance and redundancy[J].Journal of Machine Learning Research,2004,5(12):1205-1224.
- [8] Peng Hanchuan,Long Fuhui,CHRIS D.Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2005,27(8):

(下转第 82 页)

参考文献

- [1] ZHANG M, ZHANG Y, VO D T. Gated neural networks for targeted sentiment analysis[C]. 2016 30th Conference on Artificial Intelligence (AAAI), 2016 : 3087–3093.
- [2] HUANG B, OU Y, CARLEY K M. Aspect level sentiment classification with attention-over-attention neural networks[C]. Lecture Notes in Computer Science, 2018 : 197–206.
- [3] VO D T, ZHANG Y. Target-dependent twitter sentiment classification with rich automatic features[C]. International Joint Conference on Artificial Intelligence (IJCAI), 2015 : 1347–1353.
- [4] TANG D, QIN B, FENG X, et al. Effective LSTMs for target-dependent sentiment classification[C]. 2016 26th International Conference on Computational Linguistics (COLING), 2016 : 3298–3307.
- [5] KIM Y. Convolutional neural networks for sentence classification[J]. ArXiv preprint : 1408.5882, 2014.
- [6] MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification[C]. 2018 10th International Conference on Knowledge and Systems Engineering (KSE), 2018 : 25–30.
- [7] FAN F, FENG Y, ZHAO D. Multi-grained attention network for aspect-level sentiment classification[C]. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018 : 3433–3442.
- [8] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv : 1706.03762, 2017.
- [9] SONG Y, WANG J, JIANG T, et al. Attentional encoder network for targeted sentiment classification[J]. arXiv : 1908.09314, 2019.

- [10] PENNINGTON J, SOCHERR, MANNING C D. GloVe : Global vectors for word representation[C]. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014 : 1532–1543.
- [11] PONTIKI M, GALANIS D, PAVLOPOULOS J, et al. SemEval 2014 Task 4 : aspect based sentiment analysis annotation guidelines[C]. 2015 9th International Workshop on Semantic Evaluation (SemEval), 2015 : 486–495.
- [12] DONG L, WEI F, TAN C, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification[C]. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 2014, 2 : 49–54.
- [13] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for aspect-level sentiment classification[C]. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2016 : 606–615.
- [14] LI C, GUO X, MEI Q. Deep memory networks for attitude identification[C]. Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (ICWSDM), 2017 : 671–680.

(收稿日期:2020-06-27)

作者简介：

杨嘉佳(1988-),男,博士,工程师,主要研究方向:社交网络、网络安全。

刘金(1990-),女,硕士,工程师,主要研究方向:社交网络、网络安全。

唐球(1985-),通信作者,男,博士,高级工程师,主要研究方向:社交网络、网络安全,E-mail:tangqiu@ncse.com.cn。



扫码下载电子文档

(上接第 77 页)

- 1226-1238.
- [9] QU G, HARIRI S, YOUSIF M. A new dependency and correlation analysis for features[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(9): 1199-1207.
- [10] 沙丽瓦尔·阿里木, 努尔布力, 吾守尔·斯拉木. 改进的 Re-FCBF 算法在入侵检测中的应用[J]. 计算机工程与设计, 2016, 37(11): 2956-2960.
- [11] 郭佳靖, 姜宏, 章翔峰, 等. 基于 MRMD 的滚动轴承损伤程度识别方法研究[J]. 组合机床与自动化加工技术, 2019(3): 98-102.
- [12] SPOLAOR N, CHERMAN E A, MONARD M C, et al. ReliefF for multi-label feature selection[C]. 2013 Brazilian Conference on Intelligent Systems, 2013: 6-11.
- [13] 杨志伟, 努尔布力, 贾雪, 等. 基于 ReliefF 的入侵特征选择方法[J]. 吉林大学学报(理学版), 2015, 53(3): 505-510.
- [14] DEISY C, SUBBULAKSHMI B, BASKAR S, et al. Efficient

dimensionality reduction approaches for feature selection[C]. International Conference on Computational Intelligence and Multimedia Applications.IEEE, 2008.

- [15] 黄晓娟.面向特征选择的Relief 算法研究[D].苏州:苏州大学,2018.
- [16] QAMAR A M, GAUSSIER E. Similarity learning in nearest neighbor and RELIEF algorithm[C]. Ninth International Conference on Machine Learning & Applications. IEEE, 2011.

(收稿日期:2020-06-19)

作者简介：

邵琪(1995-),女,硕士研究生,主要研究方向:信号处理、非侵入式负荷监测。

包永强(1978-),男,博士,教授,主要研究方向:信号处理、非侵入式负荷监测等。

姜家辉(1995-),男,硕士研究生,主要研究方向:信号处理、非侵入式负荷监测。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所