

文章编号: 2095-2163(2021)09-0165-05

中图分类号: TP301.6

文献标志码: A

基于特征加权的 SFO-K_means 用电行为研究

唐 辉, 刘晓波, 韩祥民, 邱 知, 徐邦贤

(贵州大学 电气工程学院, 贵阳 550025)

摘 要: 在传统的 K_means 算法中, 初始聚类中心大多采用任意选取或者凭借经验, 使算法的准确性易受选取结果的影响。针对这些不足, 提出了利用剑鱼算法全局寻优特性, 来改进 K_means 初始聚类中心的选取。UCI 数据对比显示, 改进算法在平均迭代次数和准确率方面优于传统 K_means 算法。在用户用电行为分析上, 利用提取的 5 个降维特征指标, 对每半小时采样一次的高维日负荷曲线进行降维处理; 引入 Critic 法来确定指标的权重系数, 并利用皮尔逊相关系数与熵权, 分别衡量指标之间的冲突性与指标内部的对比强度; 采用特征加权的 SFO-K_means 算法进行聚类分析。算例结果表明, 该法在聚类质量上有一定的优越性, 聚类结果能反应出用户的用电行为习惯。

关键词: 用电行为; K_means; 剑鱼算法; SFO-K_means; Critic 法; 熵权

Research on electricity consumption behavior based on feature weighted SFO-K_means

TANG Hui, LIU XiaoBo, HAN XiangMin, QIU Zhi, XU Bangxian

(College of Electrical Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] In the traditional K_means algorithm, the initial clustering centers are mostly selected arbitrarily or with experience, and the accuracy of the algorithm is easily affected by the selection results. In view of these shortcomings, this paper proposes to improve the selection of K_means initial clustering centers by using the global optimization characteristics of the Sailfish Optimizer(SFO). The comparison of UCI data shows that the improved algorithm is better than the traditional K_means algorithm in the average number of iterations and accuracy. In the analysis of users' electricity consumption behavior, firstly, the extracted five dimensionality reduction characteristic indexes are used to reduce the dimensionality of the high-dimensional daily load curve which is sampled every half hour. Then the Critic method is introduced to determine the weight coefficient of the indexes, and the Pearson correlation coefficient and entropy weight are used to measure the conflict between the indexes and the contrast strength within the indexes.

[Key words] electricity consumption behavior; K_means; SFO algorithm; SFO-K_means; Critic; Entropy weight

0 引 言

准确的电力客户用电行为分析, 有助于提高电力需求侧管理, 制定合理的电价和削峰填谷^[1-2]。利用数据挖掘技术, 从海量用电数据中挖掘出用户用电规律, 充分利用电力用户的特点和用电规律, 掌握电力用户的用电需求情况, 为电力系统的安全运行和调度提供理论支持, 充分利用电力用户的价值, 提高售电侧和用电侧的经济效益^[3-4]。

近年来, 相关学者已在电网用户行为分析方面展开了大量研究。文献[5]使用 K_means 算法, 对 290 户居民的用电数据进行了挖掘, 分析了不同类别用户的用电行为特征; 文献[6]在通过对用电数

据清洗、归一化、基于轮廓系数的自动聚类评价的基础上, 采用基于 K_means 聚类优化算法, 对电力用户负荷曲线进行聚类分析; 文献[7]通过调查问卷形式对居民用电情况进行统计, 并采用模糊综合评价方法, 对居民智能用电行为进行分析; 文献[8]提出了一种将 BP 神经网络和模糊聚类技术相结合的方法, 将不同的用户负荷曲线进行分类等等。但上述方法中并未考虑初始聚类中心选取对 K_means 聚类结果的影响, 同时忽略了负荷数据高维特性对 K_means 聚类结果的影响。

针对上述问题, 本文提出了一种改进的 K_means 方法。把具有全局搜索能力强且易于实施的剑鱼算法引入到 K_means 算法中, 用于改进 K_

基金项目: 国家自然科学基金(51867005)。

作者简介: 唐 辉(1993-), 男, 硕士研究生, 主要研究方向: 电力系统及其自动化; 刘晓波(1964-), 女, 博士, 副教授, 主要研究方向: 高电压技术、电力系统规划; 韩祥民(1995-), 男, 硕士研究生, 主要研究方向: 电气工程; 邱 知(1995-), 男, 硕士研究生, 主要研究方向: 电气工程; 徐邦贤(1994-), 男, 硕士研究生, 主要研究方向: 电气工程。

通讯作者: 刘晓波 Email: 799797284@qq.com

收稿日期: 2021-06-18

means 算法易陷入局部最优和对初始聚类中心敏感的缺陷。通过在 UCI 上的公开数据集测试结果表明,这种改进是有效的。在用户行为分析上,首先利用从日负荷曲线中提取的负荷特征指标,对高维日负荷曲线进行降维后,使用皮尔逊相关指数和熵权来衡量指标之间的冲突性和指标内部的对比强度,并用 Critic 法确定指标的权重系数,使用加权欧式距离代替 K_means 算法中样本点到聚类中心点的欧氏距离,最后用基于特征加权的 SFO-K_means 算法来分析用户用电行为。算例表明选取的日负荷特征指标能合理反映日负荷曲线形态特征,聚类质量较好,达到了合理分析用户用电行为习惯的目的。

1 SFO-K_means 算法

1.1 K_means 算法

K_means 算法实现过程如下:

Step 1: 从 S 个数据 $\{x_i\}_{i=1}^S$ 中任意选取 k 个初始聚类中心 c_1, c_2, \dots, c_k 。

Step 2: 计算所有样本点到每个聚类中心点的欧式距离,样本离哪个聚类中心点最近,就将其划分到这个聚类中心点。

$$c_j(i) \leftarrow \operatorname{argmin}_i \|x_i\|^2, i = 1, 2, \dots, S, \quad (1)$$

式中: $c_j(i)$ 表示 x_i 与第 j 个聚类中心点最近。

Step 3: 根据式(2)重新计算 k 个聚类中心。

$$c_j(i) = \frac{1}{S_j} \sum_{x_i \in c_j} x_i, j = 1, 2, \dots, k, \quad (2)$$

Step 4: 重复 Step2、Step3 步骤,当聚类中心点不再发生改变时停止。

该方法根据各点到所属簇的聚类中心的距离之和为标准,来衡量聚类效果的优劣,其准则函数为:

$$E = \sum_{i=1}^k \sum_{c_i} |x - m_i|^2 \quad (3)$$

式中, E 为误差平方和; x 为数据集中的点; m_i 为每类数据 c_i 的中心。 E 值越小,聚类效果越好。

1.2 剑鱼算法 (The Sailfish Optimizer, SFO)

剑鱼算法 (The Sailfish Optimizer, SFO)^[9] 的灵感来自捕食沙丁鱼的剑鱼,具有寻优能力强,收敛速度快的特点。该算法首先对剑鱼和沙丁鱼分别初始化,其中剑鱼种群位置用 X_{SF} 表示,沙丁鱼种群位置用 X_S 表示。初始化后,沙丁鱼中适度值最好的种群位置用 $X_{injuredS}$ 表示,剑鱼中适度值最好的种群位置用 $X_{eliteSF}$ 表示。剑鱼位置更新用式(4)表示:

$$X_{newSF}^i = X_{eliteSF}^i - \lambda_i * [\operatorname{rand}(0, 1) *$$

$$\left(\frac{X_{eliteSF}^i + X_{injuredS}^i}{2} \right) - X_{oldSF}^i] \quad (4)$$

其中: $X_{injuredS}^i$ 表示迭代次数为 i 时,沙丁鱼的最佳位置; $X_{eliteSF}^i$ 表示迭代次数为 i 时,剑鱼的最佳位置; X_{oldSF}^i 表示在迭代次数为 i 时,剑鱼的位置。 λ_i 的系数定义如下:

$$\lambda_i = 2 * \operatorname{rand}(0, 1) * PD - PD \quad (5)$$

$$PD = 1 - \frac{N_{SF}}{N_{SF} + N_S} \quad (6)$$

其中, N_S, N_{SF} 分别代表沙丁鱼和剑鱼的数量。沙丁鱼的位置更新如式(7):

$$X_{newS}^i = (AP + X_{eliteSF}^i - X_{oldS}^i) * \gamma \quad (7)$$

其中, AP 代表剑鱼攻击系数,其计算方式如式(8); $X_{eliteSF}^i$ 表示在迭代次数为 i 时,剑鱼的最佳位置; X_{oldS}^i 表示在迭代次数为 i 时,沙丁鱼的位置; γ 为介于 0~1 之间的随机数。

$$AP = A * (1 - 2 * Itr * \varepsilon) \quad (8)$$

其中, A, ε 控制攻击系数 AP 的变换,使攻击系数 AP 线性变换到 0, Itr 为当前迭代次数。当 $AP > 0.5$ 时用式(8)更新沙丁鱼所有位置;当 $AP < 0.5$ 时,更新沙丁鱼部分位置。部分位置的范围定义如式(9):

$$\alpha = N_S * AP \quad (9)$$

$$\beta = d_i * AP \quad (10)$$

其中, α 代表需要更新的沙丁鱼数量, β 表示需要更新的维度数量。如果 $f(X_S^i) < f(X_{SF}^i)$, 表示沙丁鱼被剑鱼猎杀,沙丁鱼的位置将被剑鱼取代。表示如下:

$$X_{SF}^i = X_S^i \quad (11)$$

其中, X_{SF}^i, X_S^i 分别表示在第 i 次迭代时剑鱼和沙丁鱼的位置, $f(X_S^i), f(X_{SF}^i)$ 分别表示在第 i 次迭代时沙丁鱼和剑鱼的适应度。

算法流程如下:

Step 1 对剑鱼数量 N_{SF} 、沙丁鱼数量 N_S 、参数 A, ε 、最大迭代次数 T_{max} 等参数初始化。

Step 2 分别计算沙丁鱼和剑鱼的适应度值,并记录两者的最优适应度值及其位置。

Step 3 分别更新剑鱼和沙丁鱼位置。如果攻击系数 $AP < 0.5$, 计算 α, β 的值,并更新部分沙丁鱼位置,否则更新全部沙丁鱼位置。

Step 4 如果剑鱼猎杀到沙丁鱼,则沙丁鱼的位置将被剑鱼替换。

Step 5 计算所有适应度值,并记录最优适应度值及其位置。

Step 6 判断是否达到迭代条件,若达到则输出最优适应度值及其位置,否则转到 Step3。

1.3 SFO-K_means 算法

K_means 算法虽具有局部搜索较强的优势,但由于 K_means 易受初始聚类中心选择的影响,易陷入局部最优解。而采用剑鱼算法(SFO)对初始聚类中心进行全局寻优,选取适应度值最优的位置作为初始聚类中心,克服了 K_means 因随机选择初始聚类中心陷入局部最优解的不足,提高了聚类的准确性。SFO-K_means 算法步骤如下。

Step 1 输入样本数据个数 (S), 聚类个数 (k) 等参数。

Step 2 初始化剑鱼数量 N_{SF} 、沙丁鱼数量 N_S 、参数 A 、 ε 、最大迭代次数 T_{max} 等 SFO 参数。分别随机选取 k 个样本数据作为一条剑鱼和一条沙丁鱼。重复上述过程直到生成剑鱼 N_{SF} 条、沙丁鱼 N_S 条为止。

Step 3 以鱼群中每条鱼为初始聚类中心,并分别计算样本数据与鱼所包含 k 个初始聚类中心的距离。离样本数据最近的聚类中心,就将其划分到这个聚类中心点。聚类结果用式(3)计算适应度值,并记录最优适应度和其所对应的 k 个初始聚类中心点。

Step 4 由位置更新公式,更新剑鱼和沙丁鱼位置,将更新之后剑鱼与沙丁鱼的位置作为新的初始聚类中心进行聚类,由聚类结果计算适应度值并记录。

Step 5 比较所有的适应度值的大小,将最优的适应度值及其所对应的 k 个初始聚类中心点记录下来。

Step 6 当满足迭代条件时,停止迭代,否则转到 Step4。

Step 7 输出最优适应度值的鱼,将其所包含的 k 个样本点作为 K_means 的初始聚类中心。

Step 8 输出最优聚类结果。

2 日负荷特性指标选取及其权重系数配置

2.1 负荷特征指标选取

考虑到居民用电负荷数据具有高维的特性,会导致进行距离测算时的意义降低^[10],从而无法准确聚类,因此从日负荷曲线中提取 5 个特征指标,用来反映各类用户的用电特性。各指标见表 1。

其中:全天为 00:00-24:00,峰期为 08:00-11:00、18:00-21:00,平期为 06:00-08:00、11:00-18:00、21:00-22:00,谷期为 22:00-24:00、00:00-06:00;

p 代表负荷; p_{av} 代表日平均负荷; P_{max} 代表日最大负荷; p_{min} 代表日最小负荷; $p_{av.sh}$ 代表平期平均负荷; $p_{av.peak}$ 代表峰期平均负荷; $p_{av.val}$ 代表谷期平均负荷。

表 1 日负荷特征指标

Tab. 1 Daily load characteristic index

时段	指标	定义	物理意义
全天	负荷率	$a_1 = p_{av}/P_{max}$	表征全天负荷变化
	日峰谷差率	$a_2 = \frac{P_{max} - P_{min}}{P_{max}}$	表征电网调峰能力
峰期	峰期负载率	$a_3 = p_{av.peak}/p_{av}$	表征峰期负荷变化
平期	平期负载率	$a_4 = p_{av.sh}/p_{av}$	表征平期负荷变化
谷期	谷期负载率	$a_5 = p_{av.val}/p_{av}$	表征谷期负荷变化

2.2 Critic 法确定指标特征指标权重

日负荷特征指标权重的不同,会在一定程度上影响聚类的结果以及后面分析用户用电习惯的准确性。Critic 法是一种客观赋权法,用来确定各评价指标的权重系数^[11]。Critic 法主要通过两方面来求每个指标的权重,一是使用指标的对比强度,二是使用指标之间的冲突性。指标对比强度通过同一指标取值的不同来衡量;评估指标之间的冲突性一般用指标间的相关系数。本文使用熵权来衡量评估指标的对比强度,用皮尔逊相关系数来衡量评估指标之间的冲突性。

熵常用来衡量某种事物的不确定性,对于有 M 个评价指标, N 个被评价对象的问题,对于第 i 个评价指标,其熵可以按照下式计算:

$$H_i = -G \sum_{k=1}^N F_{ik} \ln F_{ik} \tag{12}$$

$$F_{ik} = \frac{u_{ik}}{\sum_{k=1}^N u_{ik}} \tag{13}$$

$$G = \frac{1}{\ln N} \tag{14}$$

如果 $F_{ik} = 0$, 则定义 $F_{ik} \ln F_{ik}$ 为 0。

式中 $i = 1, 2, \dots, M$; u_{ik} 为实测数据。

第 i 个评价指标的熵权 δ_i 按照下式计算:

$$\delta_i = \frac{1 - H_i}{M - \sum_i H_i} \tag{15}$$

式中, $\delta_i \in [0, 1]$ 且 $\sum_i \delta_i = 1$ 。

从上述计算过程可以看出,若某项评价指标熵值越小,则该评价指标内不同被评价对象的指标值差异性也越大。

皮尔逊相关系数用于评估两变量之间的相关程度,评价指标 i 和评价指标 j 的皮尔逊相关系数计算

方法如下:

$$\gamma_{ij} = \frac{\sum_{k=1}^N (u_{ik} - u_{iv})(u_{jk} - u_{jv})}{\sqrt{\sum_{k=1}^N (u_{ik} - u_{iv})^2 \sum_{k=1}^N (u_{jk} - u_{jv})^2}} \quad (16)$$

式中: u_{iv} 、 u_{jv} 分别表示评价指标 i 、 j 内 N 个不同被评价对象的均值。指标 γ_{ij} 越大,表示两评价指标相关性越大。

综合皮尔逊相关系数与熵权法,用 Critic 法确定第 i 个评价指标的计算公式如下:

$$c_i = \delta_i \sum_{k=1}^N (1 - \gamma_{ik}) \quad (17)$$

$$\omega_i = \frac{c_i}{\sum_i c_i} \quad (18)$$

3 实验结果分析

3.1 SFO-K_means 算法验证

本文选用 UCI 数据库里的 Iris 和 Wine 数据集来验证本文算法 (SFO-K_means) 的收敛速度和准确率。Iris 数据集共有 150 条数据,每条数据均为 4 维,分为 3 类,每类均有 50 个数据;Wine 数据集共有 178 条数据,每条数据均为 13 维,分为 3 类,各类别数目分别为 59、71、48 条。SFO-K_means 算法的参数设置为 $A = 4$ 、 $\varepsilon = 0.001$,种群数量为 100 (剑鱼比例为 0.3),最大迭代次数为 100。将 SFO-K_means 算法与 K_means 算法在上述两个数据集上分别进行独立的 20 次重复实验,平均准确率、平均迭代次数见表 2、表 3。

表 2 平均准确率

Tab. 2 Average accuracy %

数据集	K_means 算法	SFO-K_means 算法
Iris	85.70	90
Wine	67.74	71.91

表 3 平均迭代次数

Tab. 3 Average number of iterations 次

数据集	K_means 算法	SFO-K_means 算法
Iris	5.85	2.30
Wine	6.85	4.35

从表 2 可以看出,在低维数据集 Iris 上 SFO-K_means 算法比 K_means 算法的平均准确率高出 4.3%, 高维数据集 Wine 上 SFO-K_means 算法比 K_means 算法的平均准确率高 4.17%。从表 3 可以看出,在 Iris 和 Wine 聚类实验中,SFO-K_means 算法比 K_means 算法的平均迭代次数分别减少了 3.55

次和 2.5 次。上述实验结果证明了 SFO-K_means 算法比 K_means 算法有更好的聚类效果。

3.2 居民用电行为分析

研究选取英国电网智能电表数据集,该数据集记录了伦敦 5 567 个家庭从 2011~2014 年的用电数据,采集周期为半小时。在剔除缺失数据和异常数据及归一化后,选取其中 364 个家庭于 2013 年七月某日的用电数据。利用表 1 中的 5 个特征指标,对用户数据进行降维处理,并用 Critic 法配置 5 个特征指标的权重系数 $\omega_1 \sim \omega_5$ 。权重系数取值见表 4。

表 4 特征指标权重系数

Tab. 4 Weight coefficient of characteristic index

特征指标	权重系数
a_1	$\omega_1 = 0.469$
a_2	$\omega_2 = 0.087$
a_3	$\omega_3 = 0.134$
a_4	$\omega_4 = 0.074$
a_5	$\omega_5 = 0.236$

将 K-means 算法中以欧式距离为相似性的判据改为加权欧式距离为判据。如下式:

$$d_{i,j} = \sqrt{\omega_1 |Y_{i1} - C_{j1}|^2 + \omega_2 |Y_{i2} - C_{j2}|^2 + \dots + \omega_5 |Y_{i5} - C_{j5}|^2} \quad (19)$$

式中: Y_{i1} 、 Y_{i2} ... Y_{i5} 为降维后的第 i 个样本的 5 个特征指标, C_{j1} 、 C_{j2} ... C_{j5} 为第 j 个聚类中心的 5 个特征指标,并将此法记为特征加权 K_means 算法。用 SFO 算法优化的特征加权 K_means 算法,记为特征加权 SFO-K_means 算法。

将特征加权 K-means 算法、传统 K_means 算法 (以归一化的 24 维功率值作为特征指标) 和特征加权 SFO-K_means 算法进行聚类分析 (参数为: $A = 4$, $\varepsilon = 0.001$,种群数量为 100,最大迭代次数为 100,剑鱼比例为 0.3),并采用 DBI 和 CHI 指标 (DBI 指标和 CHI 指标的定义参见文献 [12]) 评价各算法的聚类结果,结果见表 5。

表 5 各聚类算法的性能对比

Tab. 5 Performance comparison of various clustering algorithms

聚类算法	最佳聚类数	DBI 指标	CHI 指标
K_means 算法	5	1.905	109.064
特征加权 K_means 算法	5	1.311	150.769
特征加权 SFO-K_means 算法	5	1.149	160.345

从表 5 中可以看出,特征加权 SFO-K_means 算法在 DBI 指标上比其它两种算法分别下降了 0.756、0.162,在 CHI 指标上分别高出其它两种算法 51.281、9.576,由此验证了特征加权 SFO-K_means 在

聚类结果上的优越性。

3.3 用电行为模式分析

使用特征加权 SFO-K_means 算法分析居民用电行为, SFO-K_means 算法将居民分为 5 类分类结果如图 1 所示。

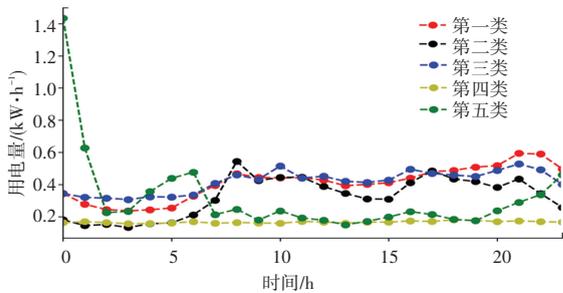


图 1 居民用电行为结果

Fig. 1 Results of residents' electricity consumption behavior

从图 1 中可以看出:第一类用户高峰用电时间段是 20:00~23:00, 00:00~6:00 用电较少, 其余时间不怎么波动;第二类用户高峰用电时间段是 7:00~11:00 和 15:00~21:00, 其余时间段较低;第三类用户全天用电都比较平缓, 波动不大;第四类用户是 5 类用户中用电最少的人群, 负荷曲线变化不大;第五类用户 00:00~6:00 用电非常活跃, 白天用电则相对很低。

4 结束语

针对 K_means 算法任意选取或凭经验选取初始聚类中心, 导致聚类效果不稳定的问题, 提出了 SFO_K-means 算法, 提高了聚类的精度和收敛速度。针对高维的用电数据, 提出了基于特征指标降维的日负荷曲线聚类方法, 从用电数据里提取出 5 种物理意义明确的特征指标, 并引入 Critic 法进行权

重配置。在用电行为分析上, 使用加权的 SFO-K_means 算法进行聚类分析。其结果表明:所提方法有一定的优越性, 能够较为合理的分析用户用电行为, 为供电公司实现用电客户分类, 制定合理的分时电价提供了一定帮助。

参考文献

- [1] 康宁宇, 李川, 曾虎, 等. 采用 FCM 聚类与改进 SVR 模型的窃电行为检测[J]. 电子测量与仪器学报, 2017, 31(12): 2023-2029.
- [2] 郝然, 艾芊, 肖斐. 基于多元大数据平台的用电行为分析构架研究[J]. 电力自动化设备, 2017, 37(8): 20-27.
- [3] 陈启鑫, 郑可迪, 康重庆, 等. 异常用电的检测方法: 评述与展望[J]. 电力系统自动化, 2018, 42(17): 189-199.
- [4] 陈鹏伟, 陶顺, 肖湘宁, 等. 短时间尺度用电行为相关性分析网络模型[J]. 电力系统自动化, 2017, 41(3): 61-69.
- [5] 王星. 基于 K-means 聚类算法的荆州市居民用电行为分析[J]. 通讯世界, 2019, 26(8): 276-277.
- [6] 冯丹彤, 钟英. 基于聚类的用户用电行为分析研究[J]. 科技经济导刊, 2020, 28(35): 77-78.
- [7] 何永秀, 王冰, 熊威, 等. 基于模糊综合评价的居民智能用电行为分析与互动机制设计[J]. 电网技术, 2012, 36(10): 247-252.
- [8] 黎祚, 周步祥, 林楠. 基于模糊聚类与改进 BP 算法的日负荷特性曲线分类与短期负荷预测[J]. 电力系统保护与控制, 2012, 40(3): 56-60.
- [9] SHADRAVAN S, NAJI H R, BARDSIRI V K. The Sailfish Optimizer: A novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems [J]. Engineering Applications of Artificial Intelligence, 2019, 80.
- [10] PARSONS L, HAQUE E, LIU H. Subspace clustering for high dimensional data: a review[J]. SIGKDD Explore, 2004, 6(1): 90-105.
- [11] DIAKOULAKI D, MAVROTAS G, PAPAYANNAKIS L. Determining objective weights in multiple criteria problems: the CRITIC method[J]. Computers & Operations Research, 1995, 22(7): 763-770.
- [12] 张斌, 庄池杰, 胡军, 等. 结合降维技术的电力负荷曲线集成聚类算法[J]. 中国电机工程学报, 2015, 35(15): 3741-3749.

(上接第 164 页)

推荐精度。实验结果表明该方法显著提高了推荐准确度的同时, 也有效缓解了数据稀疏性和冷启动问题。不足之处在于本混合推荐算法计算量大, 复杂度高。

参考文献

- [1] LIU Y, YANG H, SUN G, et al. Collaborative Filtering Recommendation Algorithm Based on Multi-relationship Social Network[J]. Ingénierie des Systèmes D Information, 2020, 25(3): 359-364.
- [2] YAN Yongjie, XIE Hui. Collaborative Filtering Recommendation Algorithm Based on User Preferences[J]. Journal of Physics: Conference Series, 2020, 1549(3).
- [3] 苑宁萍, 辛力坚, 王呼生, 等. 融合用户兴趣和信任度的协同过滤推荐算法[J]. 计算机工程与设计, 2020, 41(7): 1967-

1974.

- [4] 过金超, 杨继纲. 基于改进的 SVD 算法和二分 K-均值聚类算法的协同过滤算法[J]. 轻工学报, 2020, 35(4): 88-95.
- [5] 王光, 张杰民, 董帅含, 等. 基于内容的加权粒度序列推荐算法[J]. 计算机工程与科学, 2018, 40(3): 564.
- [6] 刘沛文, 陈华锋. 基于用户行为特征的动态权重混合推荐算法[J]. 计算机应用与软件, 2017, 34(4): 316-321.
- [7] FAN Yang. A hybrid recommendation algorithm-based intelligent business recommendation system [J]. Journal of Discrete Mathematical Sciences and Cryptography, 2018, 21(6): 1317-1322.
- [8] 李鹏飞, 吴为民. 基于混合模型推荐算法的优化[J]. 计算机科学, 2014, 41(2): 68-71, 98.
- [9] 田保军, 刘爽, 房建东. 融合主题信息和卷积神经网络的混合推荐算法[J]. 计算机应用, 2020, 40(7): 1901-1907.