

# 人工蜂与 K-means 混合算法在 VANETs 的应用\*

黄欣<sup>1</sup>, 余思东<sup>1</sup>, 赵志刚<sup>2\*\*</sup>

(1. 广西农业职业技术学院信息与机电工程系, 广西南宁 530007; 2. 广西大学计算机与电子信息学院, 广西南宁 530004)

**摘要:**针对车载自组织网(Vehicular Ad Hoc Networks, VANETs)拓扑结构经常变化导致通信链路容易断裂而通信质量不可靠的问题,将人工蜂与 K-means 混合算法应用在 VANETs 中。在成簇阶段,该混合算法利用人工蜂算法较强的全局搜索能力确定初始聚类中心,代替传统的 K-means 对初始聚类中心的选择,这样就消除了 K-means 对随机初始聚类中心的依赖。在簇头选取阶段,类内具有最小的速度方差以及到其他节点最小平均距离的车辆节点被选择为簇头。在簇的维护阶段,当最优节点即簇头有变化时,次优节点被选为临时簇头,直至更新为最优节点的簇头信息。为测试该混合算法的性能,将其和 PSO 与 K-means 混合算法、经典 K-means 算法进行实验对比,结果表明,该混合算法能够更加稳定 VANETs 通信链路,具有更高成簇质量和更高通信质量。

**关键词:** VANETs 网络拓扑 人工蜂算法 K-means 算法 混合算法

中图分类号: TP393 文献标识码: A 文章编号: 1005-9164(2020)01-0104-06

## 0 引言

车载自组织网(Vehicular Ad Hoc Networks, VANETs)是指由路边单元、车辆之间构成的一种移动自组织网络,旨在提高交通安全系数,是智能交通的重要组成部分<sup>[1]</sup>。VANETs 具有车辆高速移动的动态变化特点,容易造成其通信链路断裂而通信质量不佳。可靠而稳定的成簇算法可以实时传输数据,减少网络通信延迟,为智能交通系统奠定技术基础。

为提高 VANETs 通信的可靠性以及稳定性,如何将 VANETs 更为准确地划分成簇,使得计算成本

最小化和网络寿命周期最大化之间平衡,一些研究者将成簇算法用于 VANETs 中<sup>[2]</sup>。Gerla 等<sup>[3]</sup>针对 VANETs 成簇的稳定性,提出一种基于节点的空间位置的最小 ID 成簇算法,给所有的节点标识唯一 ID,将最小 ID 标识设置为簇头。该算法实现简单,计算成本低,但是牺牲了节点的负载均衡,不适用于高速移动的场景。Basu 等<sup>[4]</sup>基于节点移动的差异提出最小相对移动成簇算法,该算法通过比较一个节点与其他节点的移动差异的方差来选择簇头,有效地提高了成簇的稳定性,但是也不适用于高速移动的场景。Chatterjee 等<sup>[5]</sup>基于节点的移动特性、车辆流量以及节点能量提出一种按需加权的成簇算法,该算法

\* 国家自然科学基金(61363067)和广西 2019 年度中青年教师基础能力提升项目(2019KY1408)资助。

### 【作者简介】

黄欣(1983—),男,副教授,主要从事计算机网络技术研究,E-mail: Huangxin543@qq.com。

### 【\*\*通信作者】

赵志刚(1973—),男,博士,教授,主要从事计算机应用技术研究,E-mail: zzgmail2002@163.com。

### 【引用本文】

DOI:10.13656/j.cnki.gxkx.20200311.013

黄欣,余思东,赵志刚.人工蜂与 K-means 混合算法在 VANETs 的应用[J].广西科学,2020,27(1):104-109.

HUANG X, YU S D, ZHAO Z G. Application of the Hybrid Algorithm of Artificial Bee Algorithm and K-means in VANETs [J]. Guangxi Sciences, 2020, 27(1): 104-109.

较为全面地考虑了节点的特性,但是算法的计算成本也大大增加。许力文等<sup>[6]</sup>基于车辆的平均速度、位置和方向等信息,提出一种改进的 K-means 分簇算法,该算法利用改进的 K-means 算法划分成簇,再利用改进的 Floyd-Warshall 算法选择簇投,从而提高 VANETs 通信链路的稳定性。李炳圻等<sup>[7]</sup>考虑车辆的速度、位置和方向以及节点间的 QoS 提出了一种 SQ-WCA 成簇算法,实验表明该算法在 VANETs 中具有稳定性和可靠性。

K-means 算法<sup>[8]</sup>是基于划分的常见聚类算法之一,具有实现简单、较强的局部搜索能力、时间复杂度接近线性等优点,但是也存在一些不足,如对初始聚类中心依赖、容易出现“早熟”现象等。所以,有学者将群智能优化算法与 K-means 算法结合在一起,例如粒子群优化算法<sup>[9-10]</sup>、人工蜂算法(Artificial Bee Colony, ABC)<sup>[11]</sup>等。作为一种新颖的群智能优化算法,人工蜂算法是于 2005 年通过模拟蜂群的采摘蜂蜜过程而被提出,目前已经成功应用于入侵检测<sup>[12]</sup>、图像分割<sup>[13]</sup>、组合优化问题<sup>[14]</sup>等多个领域。这里提出一种基于改进人工蜂的 K-means 优化聚类算法,并将其应用到 VANETs 分簇中。通过改进人工蜂算法中蜂群的迭代搜索获取全局最优的  $K$  个聚类中心,在此基础上执行 K-means 算法划分聚类,充分利用人工蜂算法的勘探能力和 K-means 算法的开发能力,避免过早陷入局部最优,消除 K-means 算法对初始聚类中心选取的依赖性,获取近似全局最优的聚类划分,提高算法的收敛速度以及改善聚类结果的质量。

## 1 材料与方 法

### 1.1 人工蜂算法

人工蜂算法是 Karaboga 等<sup>[15]</sup>提出的一种新颖的全局优化算法,其灵感来源于蜂群的采蜜过程,蜜蜂各司其职,具有独特的正反馈机制,蜂群之间能够实现信息交互,进而快速寻找到最优解。标准的人工蜂算法<sup>[16]</sup>包括蜜源、采蜜蜂(跟随蜂)、引领蜂和侦察蜂,算法的目的是尽可能找到量多的蜜源。基本的人工蜂群算法具体步骤如下:

Step 1:初始化算法基本参数,算法的最大迭代次数 MaxIter,蜜源的最大迭代次数 Limit,以式(1)方式产生 SN 个初始蜜源(采蜂蜜) $x_{id}$ ;

$$x_{id} = x_d^{\min} + \text{rand}(0,1)(x_d^{\max} - x_d^{\min}), \quad (1)$$

其中, $i$  是蜜源的数量, $i \in (1,2,\dots,SN)$ , $d$  是蜜源的维度, $x_d^{\max}$ 、 $x_d^{\min}$  分别是解的上、下边界;

Step 2:计算各个蜜源的适应度值 fit,记录各个蜜源的最好适应度值 fit\_pbest、所有蜜源中最好适应度值 fit\_gbest 以及对应的蜜源;

Step 3:引领蜂通过式(2)的方式产生新蜜源  $x'_{id}$ 。若新蜜源的 fit 大于 fit\_pbest 则引领蜂保留新蜜源,否则不更新;

$$x'_{id} = x_{id} + r(-1,1)(x_{id} - x_{kd}), \quad (2)$$

其中, $x_{kd}$  表示相邻蜜源, $k \in (1,2,\dots,SN)$ ,

且  $k \neq i$ ;

Step 4:算法根据公式(3)计算每个蜜源的适应度值在所有蜜源的适应度值之和中所占的比重  $p_i$ :

$$p_i = \frac{\text{fit}_i}{\sum_{j=1}^{SN} \text{fit}_j}, \quad (3)$$

其中, $\text{fit}_i$  表示第  $i$  个蜜源对应的适应度值;

Step 5:跟随蜂通过轮盘赌的选择方式选择蜜源,利用公式(2)进行更新操作,若新蜜源的 fit 大于 fit\_pbest 则跟随蜂保留新蜜源,否则不更新;

Step 6:侦察蜂的转换机制,假设某个蜜源经过 limit 次迭代后适应度值仍然保持不变,说明可能陷入局部最优解,与此蜜源对应的引领蜂将放弃此蜜源并转变为侦察蜂,然后用式(1)随机生成新蜜源,侦察蜂开始对这个新蜜源进行搜索并再次转变为引领蜂;

Step 7:记录最优蜜源 gbest;

Step 8:判断是否达到算法最大迭代次数,如否,则转到 Step 3,如是,则运行结束。

### 1.2 K-means 算法

K-means 算法是根据相似性原则计算平方误差函数来作为评判标准,将一组数据  $X = \{x_1, x_2, \dots, x_n\}$  划分到  $K$  个类簇  $C = \{c_1, c_2, \dots, c_k\}$  中,满足簇间数据相似最小化,簇内数据相似最大化的要求。

算法的运行过程如下<sup>[10]</sup>:

Step 1:给定类数  $K$  并随机选取  $K$  个对象作为初始聚类中心;

Step 2:分别计算每个数据对象  $x_i$  与所有类中心的距离  $d(x_i, c_j)$ ,并根据最近邻原则将其划分;

Step 3:重新计算类内所有数据对象的均值作为各个新形成聚类的聚类中心;

Step 4:重复执行 Step 2—Step 3,直到聚类中心不发生变化,则算法结束。

### 1.3 人工蜂与 K-means 混合算法在 VANETs 分簇中的应用

人工蜂与 K-means 混合算法解决 VANETs 的

分簇问题需要解决蜜源编码以及适应度函数设计两个问题。

### 1.3.1 蜜源编码方式

根据数据集  $x_i$ , 划分簇数  $K$  和特征数  $m$ , 搜索空间解可表示为  $K \times m$  的向量; 即  $(c_{11}, c_{12}, c_{1m}, \dots, c_{K1}, c_{K2}, c_{Km})$ 。

### 1.3.2 适应度函数设计

一般而言, 将平方误差  $I$  作为评价聚类质量的标准, 但是在 ABC 中的目的是搜索花蜜量最大的蜜源, 聚类算法中  $I$  越小说明聚类效果越好, 人工蜂与 K-means 混合算法适应度函数设计如公式(4)(5)所示, 蜜源位置对应聚类中心位置:

$$I = \sum_{j=1}^K \sum_{x_i \in c_j} \|x_i - c_j\|, \quad (4)$$

$$f(x) = 1/(1 + I). \quad (5)$$

### 1.3.3 人工蜂与 K-means 混合算法原理

K-means 算法是基于划分的常见聚类算法之一, 人工蜂群算法是一种新颖的全局优化算法, 综合两者, 提出一种基于改进人工蜂的 K-means 优化聚类算法, 通过改进人工蜂算法中蜂群的迭代搜索获取全局最优的  $K$  个聚类中心, 在此基础上执行 K-means 算法划分聚类, 充分利用人工蜂算法的勘探能力和 K-means 算法的开发能力, 避免过早陷入局部最优, 消除 K-means 算法对初始聚类中心选取的依赖性, 加快收敛速度。将 ABC 的适应度函数 fit 设定为与 K-means 聚类的衡量准则函数相关的函数, 经过不断地迭代返回适应度最优的蜜源, 聚类中心点集就是使 K-means 聚类的类内距离和最小的解, 然后以这组聚类中心进行 K-means 聚类得到最优的聚类效果。

### 1.3.4 人工蜂与 K-means 混合算法伪代码

综上, 人工蜂与 K-means 混合算法实现过程如下:

Step 1: 从数据集  $X$  中随机选择  $K$  个中心点, 将其作为蜜源  $x_i$  初值;

Step 2: 执行 ABC 算法进行迭代搜索(见人工蜂算法实现流程);

Step 3: 执行 K-means 算法, 按照类输出即最终的聚类结果(见 K-means 算法实现流程)。

### 1.3.5 簇的形成

结合十字路口 VANETs 的场景, 车辆在交通规则限定下以非匀速前进, 利用人工蜂与 K-means 混

合算法应用于其分簇, 可以提高车辆之间的通信质量。簇的形成通过人工蜂算法进行, 其主要思想是代替 K-means 获取车辆节点分类的聚类点。

### 1.3.6 簇头的选择

在簇头选取阶段, 仅仅把混合算法最终输出的车辆划分的质心作为簇头, 只是依据位置来选择而未考虑到速度, 簇内计算所有 VANETs 车辆的最短距离, 将具有到其余车辆的最小平均距离且具有最小速度方差的车辆选择为簇头。参照文献[6], 利用位置以及速度方差进行簇头的选择, 根据黄金比例, 速度方差 VAR 给予  $w_1(0.618)$  的权重, 平均距离 AD 给予  $w_2(0.382)$  的权重。具有最小的合计分值(Total Score, TS)的车辆节点将被选择为簇头的候选。其计算方式是

$$TS_i = \text{count}(w_1 \text{VAR}(V_i) + w_2 \text{AD}(V_i)). \quad (6)$$

### 1.3.7 簇的维护

在簇的维护阶段, 当最优节点的簇头有变化时, 次优节点被入选临时簇头, 直至最优节点代替次优节点再次更新簇头信息。

## 1.4 实验环境及参数设置

实验环境是 MATLAB 2015Ra, Inter(R) Xeon (R) CPU E5-1620 v3 @3.5 GHz, win8 操作系统。人工蜂与 K-means 混合算法的各个参数取值如下: ABC 算法蜜源个数 SN=20, 种群规模 NP=40; 算法的最大迭代次数 MaxIter=500, 蜜源的最大迭代次数 Limit=20, 算法运行 30 次。十字路口的场景设置为每个车辆可以在 1 000 m 内运动, 速度为 60—120 km/h, 车辆节点数目是 10—100, 通过比较分析人工蜂与 K-means 混合算法、PSO 与 K-means 混合算法<sup>[17]</sup>、K-means 算法<sup>[6]</sup>, 主要从成簇的稳定性、端到端时延评价算法的性能。

## 2 结果与分析

根据公式(6), TS 越小, 簇头的稳定性越好。表 1 是其中一个簇内选择簇头的过程。

根据表 1 可知, 车辆 ID 4 具有最小的 TS 值, 因此适合作为簇头。而仅仅以距离来衡量车辆的簇头稳定性是不全面的, 如表 1 中, 车辆 ID 6 具有最小的平均距离(AD), 但与其他车辆的速度差很大, 如果选作簇头, 较大的速度差将影响簇头的持续时间和平均距离。

表1 集群簇头选择

Table 1 Cluster head selection

ID	平均距离 AD (m)	速度方差 VAR (m)	TS
1	11	68	46.23
2	14	51	36.87
3	43	64	55.98
4	16	43	32.69
5	44	87	70.57
6	9	94	61.53
7	96	19	48.41
8	57	27	38.46
9	63	56	58.67
10	44	69	59.45

3种算法的簇头节点持续时间车辆速度变化如图1所示。随着速度的增大,3种算法的簇头持续时间整体减少,这是因为速度增大,网络拓扑结构频繁变化。簇头持续时间减少,导致通信链路变差,算法的稳定性就越差。而在相同的速度下,人工蜂与K-means混合算法簇头持续时间是最长的,因为该算法不仅考虑了速度,还考虑了位置的因素,使得形成的簇比较稳定,通信链路也比较稳定。

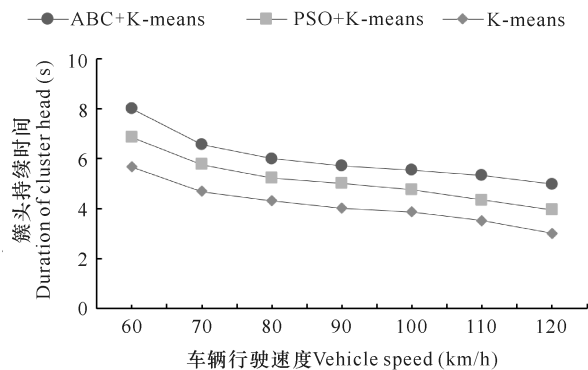


图1 簇头节点持续时间变化曲线图

Fig.1 Curve of cluster head duration

图2显示3种算法在不同速度下平均簇头变化次数变化情况。随着速度的增大,3种算法的平均簇头平均变化次数逐渐增大,与簇头持续时间呈反比,说明簇头持续时间越长,通信质量越稳定,簇头平均变化次数越少。在相同速度下,人工蜂与K-means混合算法的平均簇头变化次数是最小的,说明该算法比其他2种算法稳定。

图3是3种算法有关车辆节点与端到端时延的变化情况。随着车辆节点数目增大,会增加3种算法的端到端时延。车辆密度会影响车辆节点端到端时延。较稳定的簇头兼具有较好的信道质量,可以降低端到端时延,加快通信链路的传输。在相同的车辆节

点数目下,人工蜂-K-means混合算法具有最低的端到端时延,说明该算法对通信质量的提升程度比其他2种算法更佳。

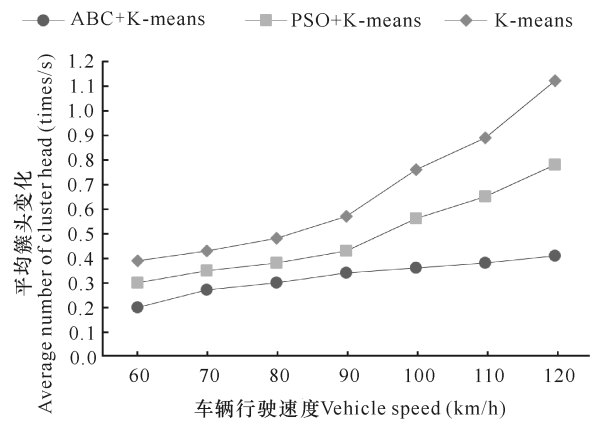


图2 平均簇头变化次数曲线图

Fig.2 Curve of average number of cluster head

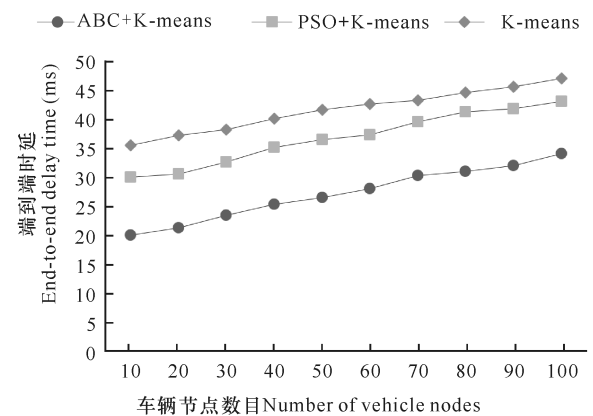


图3 端到端时延变化曲线图

Fig.3 Curve of end-to-end delay time

综上,用人工蜂与K-means混合算法求解VANETs分簇问题是可行有效的,且比PSO与K-means混合算法、K-means算法性能更佳,主要原因有:(1)K-means算法过度依赖于初始聚类中心的选择,而人工蜂与K-means混合算法利用人工蜂算法作为寻找初始聚类中心的搜索策略,再利用K-means算法进行聚类输出最终结果,消除依赖性。(2)相比PSO和K-means混合算法,人工蜂与K-means混合算法具有更强大的全局搜索能力,使得该算法更快速寻找到最优解。(3)相对于其他2种算法,人工蜂与K-means混合算法充分考虑位置和速度影响因素,更加全面地说明该算法的可靠性。

### 3 结论

VANETs的车辆节点具有动态的特性,网络拓扑结构频繁变化导致通信链路不稳定。这里将人工



蜂与 K-means 混合算法用于 VANETs 分簇中,即将场景内的车辆节点通过人工蜂算法分成不同集群,在利用 K-means 算法进行聚类结果的输出。通过改进人工蜂算法中蜂群的迭代搜索获取全局最优的  $K$  个聚类中心,在此基础上执行 K-means 算法划分聚类,充分利用人工蜂算法的勘探能力和 K-means 算法的开发能力,避免过早陷入局部最优,消除 K-means 算法对初始聚类中心选取的依赖性,加快收敛速度。分簇的稳定性影响集群通信链路的通信质量,稳定的簇头可以更好地维持分簇的形状以及簇头持续时间。在以后的工作中,应对 VANETs 进行更深入的研究,尝试将其他智能算法与 K-means 结合,并应用于 VANETs 中。

#### 参考文献

- [1] 马金忠,王文杰,田彦山.城市车联网中紧急消息的快速广播路由研究[J].西南师范大学学报:自然科学版,2019(6):84-91.
- [2] COOPER C,FRANKLIN D,ROS M,et al. A comparative survey of VANET clustering techniques [J]. IEEE Communications Surveys and Tutorials, 2017, 19(1): 657-681.
- [3] GERLA M, TSAI J T C. Multicluster, mobile, multimedia radio network [J]. Wireless Networks, 1995, 1(3): 255-265.
- [4] BASU P, KHAN N T D C. A mobility based metric for clustering in mobile ad hoc networks [C]// Proceedings 21st International Conference on Distributed Computing Systems Workshops, Mesa, AZ, USA: IEEE, 2001.
- [5] CHATTERJEE M, DAS S K, TURGUT D. WCA: A weighted clustering algorithm for mobile ad hoc networks [J]. Cluster Computing, 2002, 5(2): 193-204.
- [6] 许力文,乔丽娟,陈杰.基于 VANETs 修改的 K-means 分簇路由算法[J].计算机技术与发展,2018,28(3):15-19.
- [7] 李炳圻,梅中辉.一种适用于城市环境的 VANET 成簇算法[J].计算机技术与发展,2019,29(1):154-158.
- [8] 余秀雅,刘东平,杨军.基于 K-means++ 的无线传感网分簇算法研究[J].计算机应用研究,2017,34(1):181-185.
- [9] 谢秀华,李陶深.一种基于改进 PSO 的 K-means 优化聚类算法[J].计算机技术与发展,2014,24(2):34-38.
- [10] 徐艳. K-means 算法与群体智能算法(PSO)融合的研究与应用[D].呼和浩特市:内蒙古农业大学,2019.
- [11] 于佐军,秦欢.基于改进蜂群算法的 K-means 算法[J].控制与决策,2018,33(1):181-185.
- [12] 刘铭,黄凡玲,傅彦铭,等.改进的人工蜂群优化支持向量机算法在入侵检测中的应用[J].计算机应用与软件,2017,34(1):230-235.
- [13] 赵文昌,李忠木.融合改进人工蜂群和 K 均值聚类的图像分割[J].液晶与显示,2017,32(9):726-735.
- [14] 张懿.人工蜂群算法在投资组合问题中的优化应用[J].计算机与数字工程,2019,47(8):1869-1873,1894.
- [15] KARABOGA D, BASTURK B. On the performance of artificial bee colony (ABC) algorithm [J]. Applied Soft Computing, 2008, 8(1): 687-697.
- [16] ZHANG X, ZHANG X, GU C. A micro-artificial bee colony based multicast routing in vehicular ad hoc networks [J]. Ad Hoc Networks, 2017, 58(2): 213-221.
- [17] 汤深伟,贾瑞玉.基于改进粒子群算法的 k 均值聚类算法[J].计算机工程与应用,2019,55(18):140-145.

## Application of the Hybrid Algorithm of Artificial Bee Algorithm and K-means in VANETs

HUANG Xin<sup>1</sup>, YU Sidong<sup>1</sup>, ZHAO Zhigang<sup>2</sup>

(1. Department of Information and Electromechanical Engineering, Guangxi Agriculture Vocational and Technical College, Nanning, Guangxi, 530007, China; 2. College of Computer and Electronics Information, Guangxi University, Nanning, Guangxi, 530004, China)

**Abstract:** The VANETs topology often changes and the communication link is easy to break and the communication quality is unreliable. To solve this problem, a hybrid algorithm of artificial bee and K-means is applied to VANETs. In the clustering stage, the hybrid algorithm uses the strong global search ability of the artificial bee algorithm to determine the initial clustering center, instead of the traditional K-means selection of the initial clustering center, which eliminates K-means' reliance on random initial cluster centers. In the cluster head selection stage, the vehicle nodes with the smallest speed variance and the smallest average distance to other nodes are selected as cluster heads. In the cluster maintenance stage, when the optimal nodes, that is, cluster heads, changes, the sub-optimal nodes are selected as temporary cluster heads until the cluster head information of the optimal node is updated. In order to test the performance of the hybrid algorithm, experiments were carried out to compare PSO with K-means hybrid algorithm and classical K-means algorithm. The results show that the hybrid algorithm can stabilize the VANETs communication link more stably, have higher clustering quality and can improve communication quality.

**Key words:** VANETs, topology, artificial bee algorithm, K-means algorithm, hybrid algorithm

责任编辑:符支宏

(上接第 103 页 Continued from page 103)

## Image Encryption Scheme based on Hyper-Chaotic System and Discrete Fractional Random Transform

ZHOU Ling<sup>1</sup>, ZHOU Ying<sup>2</sup>, PAN Shumin<sup>2</sup>, CAI Jingsu<sup>1</sup>

(1. Department of Energy and Power Engineering, Guangxi Electric Polytechnic Institute, Nanning, Guangxi, 530007, China;

2. Department of Electronic Information Engineering, Information Engineering School of Nanchang University, Nanchang, Jiangxi, 330031, China)

**Abstract:** In this paper, a new image encryption scheme based on hyper-chaotic system and discrete fraction random transformation is proposed, and the principle diagram of optical device to realize the algorithm is given. In the encryption process, the chaotic sequence generated by the hyper-chaotic system is used to construct the random matrix of the discrete fraction random transformation (DFRT). The order of DFRT and the initial values of the hyper-chaotic system are used as the master key of the image encryption algorithm. Compared with the image encryption algorithm based on pure discrete fraction random transformation, this algorithm has higher complexity between the plaintext and the ciphertext without increasing the computational burden, and increases the key space and the key sensitivity. This system is a non-linear cryptosystem, which eliminates the non-security factors caused by the linear process in the traditional encryption system, and improves the ability of the encryption system to resist statistical attacks and noise attacks.

**Key words:** image encryption, hyper-chaotic system, random matrix, discrete fractional random transform, encryption algorithm

责任编辑:符支宏