



专题：工业互联网平台及安全

混沌改进鱼群算法及其在工业控制网络异常检测中的应用

吴亚萍^{1,2}, 董红召¹, 林盈盈¹, 柳菁²

(1. 浙江工业大学机械工程学院, 浙江 杭州 310032;

2. 绍兴文理学院机械与电气工程学院, 浙江 绍兴 312000)

摘要: 人工鱼群算法作为一种新型的仿生群体智能优化算法已成功地应用于各种优化问题及实际工程领域。然而, 在复杂优化问题方面, 受目标函数存在相对较大数量的局部极值的影响, 收敛速度慢、早熟等缺陷难以避免地存在。在基本人工鱼群算法中引入具有随机性、遍历性的混沌理论, 构建了一种改进算法, 以使得人工鱼群搜索能够规避可能存在的局部极值状况。在工业控制网络异常检测中进行仿真应用, 验证了算法的有效性。与基本的人工鱼群算法相比, 混沌改进人工鱼群算法可使得算法在局部极值附近长时间搜索的状况得到有效避免, 在全局收敛方面算法有更佳表现, 且搜索效率更为突出。

关键词: 混沌; Logistic 映射; 人工鱼群算法; 优化; 工业控制网络异常检测

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2020068

Improved fish swarm algorithm based on chaos and its application in abnormal detection of industrial control network

WU Yaping^{1,2}, DONG Hongzhao¹, LIN Yingying¹, LIU Jing²

1. College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310032, China

2. Mechanical and Electrical Engineering College, Shaoxing University, Shaoxing 312000, China

Abstract: Artificial fish swarm algorithm as a new type of bionic swarm intelligence optimization algorithm has been successfully used in a variety of optimization problems and practical engineering field, but when faced with the complex optimization problems, especially the multiple extreme value of peak and multimodal function optimization problems, due to the objective function has many local minima, inevitably there are defects such as premature and slow convergence speed. The random and ergodic theory was introduced into the basic artificial fish swarm algorithm, and an improved algorithm was constructed to make artificial fish swarm search avoid possible local extremum. The effectiveness of the algorithm was validated in the simulation application in industrial control network anomaly detection. Compared with the basic artificial fish school algorithm, the chaos improved artificial fish swarm algorithm can effectively avoid the long-term search of the algorithm near the local extreme value. The algorithm has better performance in terms of global convergence and the search efficiency is more prominent.

Key words: chaos, Logistic map, artificial fish swarm algorithm, optimization, abnormal detection of industrial control network

收稿日期: 2020-02-09; 修回日期: 2020-03-13

基金项目: 国家自然科学基金资助项目 (No.61773347); 浙江省教育厅科学基金资助项目 (No.LY17F030017)

Foundation Items: The National Natural Science Foundation of China (No.61773347), The Education Department Foundation of Zhejiang Province of China (No.LY17F030017)



1 引言

自李晓磊^[1]在 2003 年提出人工鱼群算法 (AFSA) 起, 此算法一直受到广泛关注。该算法模拟了鱼的觅食行为、集群行为和后端行为。通过对鱼群中的每条人工鱼进行局部优化, 在种群中突出全局最优值。该算法具有稳健性强、不依赖于初值、能并行处理、收敛性能好等优点, 被广泛应用于网络入侵检测^[2]、工业控制网络入侵检测等工程领域。但是, 由于人工鱼的个体行为属于局部随机优化行为, 不可避免地会有过早成熟和局部极端问题, 导致算法具有诸多不足, 譬如精度差、收敛速率慢等, 因而有不少对基本算法进行改进的研究^[3]和应用^[4]。近年来也有不少同类的智能算法在工业控制网络异常检测中得到应用^[5-11]。为提高检测效率, 本研究引入具有随机性、遍历性的混沌理论, 构建一种改进算法, 以使得人工鱼群搜索能够规避可能存在的局部极值状况。通过在工控网络入侵检测测试数据集^[12]上的运用, 并和同类智能算法进行性能对比, 发现改进方法可使得算法在局部极值附近长时间搜索的状况得到有效避免, 在全局收敛方面算法有更佳表现, 且搜索效率更为突出。

2 基本人工鱼群算法

2.1 算法简介

在水体中, 鱼一般会在营养丰富的地方聚集在一起。在鱼群活动中, 不存在具有协调效用的引导者, 而是以个体对群体的适应实现群体活动的目的。经对鱼类的生活习性进行长时间观察, 可归纳为觅食行为、群体行为、后端行为。

(1) 鱼类的觅食行为

鱼类游动可认为是随机的。如果鱼类发现食物, 那么会快速地游向存在食物的方向。

(2) 鱼类的群体行为

在游动时, 鱼类保持着成群聚集的习惯, 这也是为躲避有害因素及保障群体的生存而养成的习性之一。

- 对齐规则, 与相邻伙伴尽可能地保持一致的平均方向。
- 分离规则, 与相邻伙伴保持一定距离, 尽可能防止过度拥挤。
- 凝聚规则, 以相邻伙伴的中心为方向, 尽可能向此方向移动。

(3) 鱼类的后端行为

鱼类游动时, 如果有一个或多个食物被鱼找到, 那么相应也会有邻伴跟随而至。

2.2 算法描述

2.2.1 相关定义

人工鱼个体的状态可以用矢量来表示, $X = (x_1, x_2, \dots, x_n)$, $x_i (i=1, \dots, n)$ 为最优控制变量; 当前位置人工鱼的食物浓度为 $F = f(X)$; 人工鱼个体间距为 $d_{ij} = |X_j - X_i|$; Step 为人工鱼移动的最大步数; Visual 为人工鱼视场; δ 为拥挤因子; 用 N 表示参与优化的鱼类种群数量。

2.2.2 人工鱼行为描述

(1) 觅食行为

人工鱼的当前位置是随机选择一个状态 X_i , 那么以感知范围为界限, 随机性的选定状态 $X_j (d_{ij} < \text{Visual})$, 如果 X_j 的适应度优于 X_i 的适应度, 然后该人工鱼向 X_j 移动, 不然需重复以上操作, 如果经过 Try_number 次尝试仍旧难以达成觅食行为, 那么随机移动。

(2) 群体行为

人工鱼的现位置为 X_i , 人工鱼数在 Visual 内为 n_f , 这些人工鱼的中心 X_c 的适应度比 X_i 的适应度要好, 并且 $n_f < \delta N$ (N 为人工鱼规模, δ 为拥挤度, 取值为 0~1), 表明在全部人工鱼的中心 X_c 存在的食物较多 (也就是具有较好的适应度), 且拥挤度不高, 那么人工鱼以中心为方向

向前游动, 否则进行觅食行为。

(3) 后端行为

人工鱼的当前位置为 X_i , 人工鱼的数量在 Visual 内为 n_f 。其中, 具有最高适应度的人工鱼是 $X_{\max}$, 如果可满足 $X_{\max}$ 的适应度比 X_i 更高, 且 $n_f < \delta N$, 那么意味着人工鱼 $X_{\max}$ 的周边存在较多食物, 且拥挤度较低, X_i 可游向 $X_{\max}$, 否则进行觅食行为。

(4) 随机行为

人工鱼现位置为 X_i , 以视野为界限, 利用随机方式进行状态 X_j 的选择。

(5) 移动策略

围绕人工鱼实施尾部碰撞、聚类模拟。同时, 也可执行后端碰撞, 要是未取得进展, 那么执行集群行为, 要是仍旧未取得进展, 那么可执行随机行为。

(6) 公告板

鱼群中的所有人工鱼都完成了群集、尾随和觅食的过程, 评估其适应性, 并在公告板对最佳个体 X_{best} 及适应度进行记录。并使得鱼继续活动, 如此则仅以公告牌对最优值与最大的优势进行保留, 而不在鱼群中留存, 对搜索能力的提升具有积极效果。

3 混沌系统

为使得在全局收敛方面人工鱼群算法有更优性能, 并提升搜索效率, 因而把混沌系统引入算法中。混沌在大部分非线性系统中均存在, 是此类系统具有的独特现象之一。在完全确定性系统中, 混沌是系统反映出的介乎随机性与规则间的随机行为, 没有额外的随机因素。与不规则与混乱存在一定程度的不同, 尽管混乱是随机现象, 然而也具有微妙内部结构。

(1) 混沌现象的特性

- 遍历性, 即混沌能够遵循自身规律, 实现对一定范围内各类状态的遍历。

- 随机性, 即与随机变量相类似, 混沌也存在随机特性。

- 规律性, 指的是混沌现象是有一定规律性的。

鱼群系统可通过动态演化对吸入物进行传递。关键是混沌所具有的遍历性能够防止其在搜索过程中陷入局部极小值。混沌改进人工鱼群算法可使得算法在局部极值附近长时间搜索的状况得到有效避免, 使得在全局收敛方面算法有更佳表现, 且搜索效率更为突出。

(2) 以 Logistic 映射的方式, 建立起混沌模型。迭代式可表示为:

$$\left. \begin{aligned} z_{i+1} &= \mu z_i (1 - z_i) \\ z_i &\in (0, 1) \quad i = 0, 1, 2, \dots \end{aligned} \right\} \quad (1)$$

其中, z_i 为第 i 次迭代的变量值; μ 表示控制参数, 直接影响系统状态。研究发现, 如果系统在完全混沌状态, 即 $\mu = 4$, 那么系统近乎能够对 (0,1) 范围内的全部状态重复进行遍历。

(3) 混沌搜索

混沌搜索的步骤简述如下。

步骤 1 使用式 (2) 将 k 维决策变量 $x_k^{(t)}$ 映射为 $[0, 1]$ 上的混沌变量 $CX_k^{(t)}$:

$$CX_k^{(t)} = \frac{X_k^{(t)} - X_{k,\min}}{X_{k,\max} - X_{k,\min}} \quad (2)$$

步骤 2 利用式 (1) 产生 $CX_k^{(t+1)}$:

$$CX_k^{(t+1)} = 4 \times CX_k^{(t)} \times (1 - CX_k^{(t)}) \quad (3)$$

步骤 3 利用式 (4) 将混沌变量 $CX_k^{(t+1)}$ 映射回 $x_k^{(t+1)}$:

$$X_k^{(t+1)} = X_{k,\min} + CX_k^{(t+1)} (X_{k,\max} - X_{k,\min}) \quad (4)$$

步骤 4 若 $x_k^{(t+1)}$ 状态优于 $x_k^{(t)}$, 则保留 $x_k^{(t+1)}$; 否则不保留 $x_k^{(t+1)}$ 状态。

步骤 5 继续执行步骤 2~步骤 4, 直至达到预定混沌迭代次数, 在多维问题中, 每一维变量都独立搜索。



4 混沌改进 AFSA 方法

为使基本人工鱼群算法所具有的相关问题（包括收敛效率不高、易陷入局部极值、精度低等）得到妥善解决，本文选择了一种优化之后的人工鱼群算法（即混沌改进 AFSA 方法），这种方法对混沌序列所具有的遍历性与随机性进行了利用。改进的具体步骤如下。

（1）借助混沌映射形成初始鱼群

对基本算法而言，进行初始化具有随机性。多数初始鱼类种群虽能均匀分布，然而无法有效保障个体的质量，部分的溶液组也远不是最佳溶液。若是有较好的初始解组，那么不仅解的质量与效率将得到提升，还可借助混沌初始化来获取更高质量的初始解组。

给出了初值 z_0 ，由式（1）生成混沌变量 z_k 。 K 为随机整数，取值范围为 $(0, 500)$ 。在人工鱼活动范围（寻优参数取值范围） (a, b) 内，原始鱼类种群采用式 $X = a + z(b - a)$ 的方法生成。进行数次混沌迭代后，选定具有最小或者最大适配值的组，并以其作为最终初始种群。用此方法产生的人工鱼初始种群质量更加，分布更均匀。

（2）采用混沌模型优化人工鱼类种群的觅食、集群和后端行为

在觅食行为中，现状态为 X_i ，如果采取随机方式对下一状态进行选择，设为 X_j ，然后按照式（2）～式（4）进行混沌搜索，若 $X_j^{(t+1)}$ 位置优于 X_j ，则执行 $X_j^{(t+1)}$ ，否则执行 X_j 。集群和追尾行为以此类推，不再赘述。同时在执行每次行为时，所需的随机数由 Logistic 映射经过多次混沌迭代后产生。

（4）用混沌搜索代替随机行为。

由人工鱼群算法中的 4 种基本行为，可以得出如图 1 所示的行为。

由此可见，当鱼群所处位置不是很好时，也就是说，当觅食、集群和后端行为不满足时，

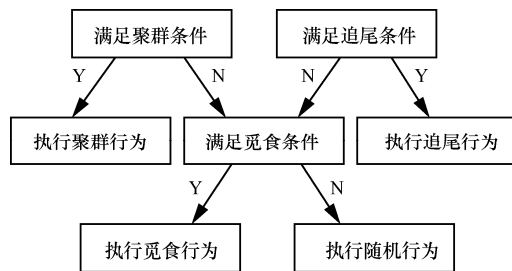


图 1 人工鱼群算法执行基本行为图解

人工鱼会进行随机行为。随机移动行为是一种相对较低级、基本的行为，它并不能使鱼群朝位置更好的方向移动，而其他 3 种行为都可以使鱼群趋向位置更优的区域。前文述及混沌搜索是一种比随机搜索更好的搜索方式，本文直接将混沌搜索取代随机搜索，以改进鱼群随机行为。

改进后的算法具体步骤如下。

（1）设置人工鱼种群大小、最大迭代数、最大尝试数、最大混沌迭代数、视场、移动步长、拥塞系数等固定参数值。

（2）以混沌映射方式，实现对人工鱼群的初始化，然后测算人工鱼个体的适应值，并对最优个体进行记录。

（3）执行集群、觅食与后端碰撞行为，并用混沌搜索代替随机行为。

（4）选择行为，选择（3）中可以达到相对最佳状态的行为，并在游动后获取新组。

（5）更新每个人工鱼个体，比较替换公告板记录。

（6）对比设置的结束条件判断是否相符，若相符则执行（7）；否则，返回（3）并重做下一次迭代搜索。

（7）计算完成并输出最优值。

5 改进后算法性能验证分析

为能够对人工鱼群算法经改进的混沌模型的效用性进行检验，本文选取两个典型测试函对混沌人工鱼群算法的可行性和精准性进行测试。此类函数有着较多局部极限，很难通过一般优化算

法找出全局最优解。

测试函数 F1: 变量范围 $[-5.12, 5.12]$, 在 $(0,0)$ 处有全局极大值 3 600。

$$f(x) = \left(\frac{3}{0.05 + x_1^2 + x_2^2} \right)^2 + (x_1^2 + x_2^2)^2 \quad (5)$$

测试函数 F2: 变量范围 $[-5, 5]$, 在 $(0, 0)$ 处的全局极小值为 0。

$$f(x) = 0.5 + \frac{(\sin \sqrt{x_1^2 + x_2^2})^2 - 0.5}{[1 + 0.001(x_1^2 + x_2^2)]^2} \quad (6)$$

测试软件为 MATLAB R2014b, 操作系统为 Windows 7。实验对每个函数进行 30 次, 且各函数优化参数设置是一致的。实验对比结果见表 1。

由表 1 可以看出, 混沌改进 AFSA 比基本 AFSA 寻优效果更好, 实验最优值更接近理论最优值。

两个测试函数通过两种优化算法寻优得到的对比结果如图 2、图 3 所示, 实验中共有 50 条人工鱼,

算法迭代 200 次。由图 2、图 3 可以看出, 经过混沌改进后的人工鱼群算法比基本算法具有更好的收敛性和搜索效率, 能有效避免陷入局部极值, 在迭代次数较少的情况下, 找寻出全局最优值。

6 仿真实验

随着工业控制系统的网络化, 工控网络异常入侵检测是一种主动的信息安全防护技术, 对检测的准确性和实时性有较高要求。本文采用改进的人工鱼群算法对异常入侵的数据分组进行特征提取和分析, 结合支持向量机分类器快速提取出最优特征, 提高检测效率。

在 Intel 2.6 GHz CPU、8 GB RAM、Windows 7 环境下, 采用 MATLAB R2014b 实现仿真实验。数据集来源于参考文献[12]提供的工业控制系统入侵检测的 4 个数据集, 具有网络流量、过程控制和过程测量特性, 数据集中的 28 个攻击来自于两个实验室级的工业控制系统, 并分为四大类:

表 1 实验对比结果

函数	基本 AFSA			混沌改进 AFSA		
	最优点	最优解	平均最优解	最优点	最优解	平均最优解
F1	-6.1291×10^{-4}	3 599.816 7	3 144.446 2	2.3715×10^{-6}	3 599.999 7	3 599.856 9
	9.4720×10^{-4}			6.6121×10^{-6}		
F2	7.6700×10^{-6}	9.7159×10^{-3}	6.9514×10^{-2}	2.8912×10^{-9}	0	2.0089×10^{-13}
	8.6130×10^{-5}			2.7872×10^{-9}		

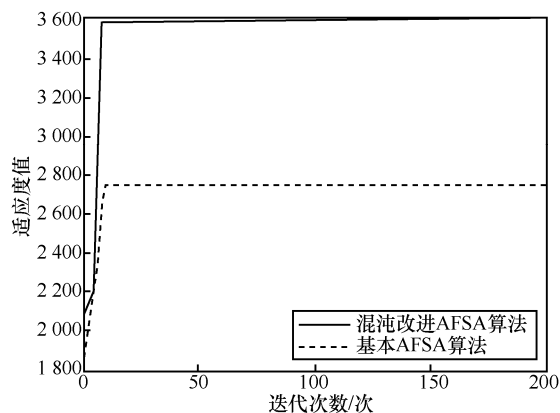


图 2 F1 函数算法收敛曲线对比

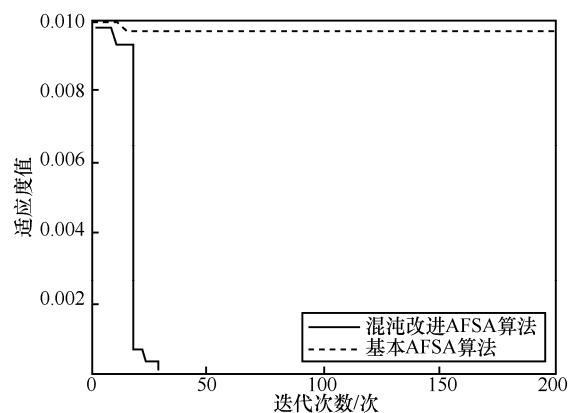


图 3 F2 函数算法收敛曲线对比



侦察 (REC)、响应注入 (MRI)、命令注入 (MCI) 和拒绝服务攻击 (DoS)。结合内存需求和处理时间, 选用的是由其中的储水罐系统数据集的 10% 实例创建的简化数据集。采用人工鱼群、遗传算法及粒子群优化算法选择特征的 SVM 模型和改进的人工鱼群 SVM 模型进行比较, 性能评价标准用检测率 (TPR) 和平均检测时间来进行对比, 其中检测率为正确检测样本数和总样本数的比值。实验结果见表 2。

表 2 检测时间对比 (单位: s)

入侵类型	REC	MRI	MCI	DoS
改进人工鱼群算法	8.76	3.27	2.89	7.35
人工鱼群算法	10.05	4.68	3.54	8.37
遗传算法	11.68	5.79	3.98	8.64
粒子群优化算法	9.56	5.02	4.02	7.96

表 2 说明改进后的人工鱼群算法特征选择加快了算法收敛速度, 减少了检测时间, 可以更好地满足工业控制系统异常入侵检测的实时性要求。

算法检测率对比结果如图 4 所示, 结果表明改进后的人工鱼群算法特征选择更加准确刻画了网络状态变化信息, 提高了检测准确率。

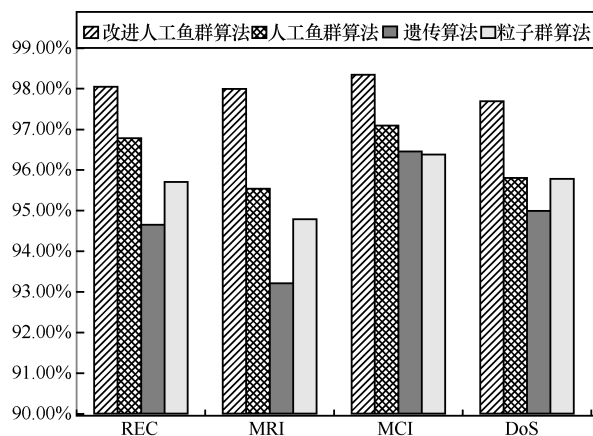


图 4 检测率对比结果

7 结束语

在人工鱼群算法中引入混沌这种具有遍历性

与随机性的系统, 使得收敛效率不高、易陷入局部极值、精度低等问题得到了妥善解决。通过对测试函数的测试和工业控制网络异常检测仿真实验, 与改进前基本算法的比较, 验证了混沌改进人工鱼群算法的有效性和准确性。改进之后的方法更具有参考性与实用性, 为改进后的人工鱼群算法更广泛地应用于工程领域提供了基础。

参考文献:

- [1] 李晓磊. 一种新型的智能优化方法——人工鱼群算法[D]. 杭州: 浙江大学, 2003.
LI X L. A new intelligent optimization method—artificial fish swarm algorithm [D]. Hangzhou: Zhejiang University, 2003.
- [2] 袁芳芳. 人工鱼群和 K 均值算法相融合的网络入侵检测[J]. 计算机仿真, 2013, 30(9): 274-277.
YUAN F F. Networks intrusion detection based on K-means clustering algorithm and artificial fish warm algorithm[J]. Computer Simulation, 2013, 30(9): 274-277.
- [3] 易新兵, 杨凯. 复合混沌-人工鱼群混合算法的改进及性能研究[J]. 计算机工程与科学, 2013, 35(8): 89-95.
YI X B, YANG K. Fusion image quality evaluation method based on human perception [J]. Computer Engineering and Science, 2013, 35(8): 89-95.
- [4] 朱文静, 白静. 一种混沌人工鱼群算法对 SVM 参数的优化及应用[J]. 微电子学与计算机, 2016, 33(3): 89-93.
ZHU W J, BAI J. Optimization and application of chaotic artificial fish swarm algorithm to SVM parameters [J]. Microelectronics and Computers[J]. Microelectronics & Computer, 2016, 33(3): 89-93.
- [5] 陈汉宇, 王华忠, 颜秉勇. 基于 CUDA 和布谷鸟算法的 KVM 在工控入侵检测中的应用[J]. 华东理工大学学报(自然科学版), 2019, 45(1): 101-109.
CHEN H Y, WANG H Z, YAN B Y. Application of CUDA and cuckoo algorithm based SVM in industrial control system intrusion detection[J]. Journal of East China University of Science and Technology, 2019, 45(1): 101-109.
- [6] 陈冬青, 张普含, 王华忠. 基于 MIKPSO—SVM 方法的工业控制系统入侵检测[J]. 清华大学学报(自然科学版), 2018, 58(4): 380-386.
CHEN D Q, ZHANG P H, WANG H Z. Intrusion detection for industrial control systems based on an improved SVM method[J]. Journal of Tsinghua University (Science & Technology), 2018, 58(4): 380-386.
- [7] 刘万军, 秦济韬, 曲海成. 基于改进单类支持向量机的工业控制网络入侵检测方法[J]. 计算机应用, 2018, 38(5): 1360-1365, 1371.

- LIU W J, QIN J T, QU H C. Intrusion detection algorithm of industrial control network based on improved one-class support vector machine[J]. Journal of Computer Applications, 2018, 38(5): 1360-1365, 1371.
- [8] 闫腾飞. 基于遗传算法优化的 OCSVM 双轮廓模型异常检测算法[J]. 计算机应用研究, 2019, 36(11): 3361-3364.
- YAN T F. Anomaly detection algorithm based on OCSVM double contour model of genetic algorithm optimization for industrial control system[J]. Application Research of Computer, 2019, 36(11): 3361-3364.
- [9] 於帮兵, 王华忠, 颜秉勇. 基于长短时记忆网络的工业控制系统入侵检测[J]. 信息与控制, 2018, 47(1): 54-59.
- YU B B, WANG H Z, YAN B Y. Intrusion detection of industrial control system based on long short term memory[J]. Information and Control, 2018, 47(1): 54-59.
- [10] 刘爱军, 杨育. 混沌模拟退火粒子群算法研究及应用[J]. 浙江大学学报(工学版), 2013, 47(10): 1722-1730.
- LIU A J, YANG Y. Research and application of chaotic simulated annealing particle swarm optimization [J]. Journal of Zhejiang University (Engineering), 2013, 47(10): 1722-1730.
- [11] 张文安, 洪榛, 朱俊威, 等. 工业控制系统网络入侵检测方法综述[J]. 控制与决策, 2019, 34(11): 2277-2288.
- ZHANG W A, HONG Z, ZHU J W, et al. A survey of network intrusion detection methods for industrial control systems[J].

Control and Decision, 2019, 34(11): 2277-2288.

- [12] THOMAS M, WEI G. Industrial control system traffic data sets for intrusion detection research[C]//Information and Communication Technology. [S.l.:s.n.], 2014.

[作者简介]



吴亚萍 (1979—), 女, 浙江工业大学机械工程学院博士生, 绍兴文理学院机械与电气工程学院实验师, 主要研究方向为智能算法优化、数据挖掘、计算机网络通信等。

董红召 (1969—), 男, 博士, 浙江工业大学机械工程学院教授, 主要研究方向为智能系统、数据挖掘等。

林盈盈 (1979—), 女, 浙江工业大学机械工程学院博士生, 主要研究方向为信息通信等。

柳菁 (1979—), 女, 绍兴文理学院机械与电气工程学院实验师, 主要研究方向为计算机应用、算法优化等。