DOI:10.16208/j.issn1000-7024.2012.02.065

入侵检测中基于遗传禁忌搜索的模糊聚类的应用

张 永，曹东侠

（兰州理工大学 计算机与通信学院，甘肃 兰州 730050）

摘 要：传统的模糊 C 均值聚类（FCM）算法须事先指出聚类数，该算法对孤立点和初始聚类敏感、易陷入局部最优，这些因素都将影响最终聚类结果的质量。针对这些缺陷，采用遗传算法和禁忌搜索的混合策略对 FCM 进行改进，该策略兼具了这两种算法的优势，改进后的算法自动生成最佳聚类数，优化初始聚类的选择，增强算法的爬山能力，有效改善了算法的性能。将改造前后的两种算法用于网络入侵检测实验，实验结果表明，改造后的算法产生的聚类质量明显优于原算法，用新算法对入侵检测建模，提高了模型的自适应性和实用性。

关键词：FCM 算法；遗传算法；禁忌搜索；混合策略；入侵检测

中图法分类号：TP309 **文献标识号**：A **文章编号**：1000-7024（2012）02-0479-05

Novel improved fuzzy clustering algorithm applied in network intrusion detection

ZHANG Yong, CAO Dong-xia

(Institute of Computer and Communications, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: When using traditional fuzzy C-Means (FCM), the number of clusters must be given beforehand. Furthermore, the algorithm is sensitive to the isolated data and the original clusters and easy to run into local critical point, and these factors have a great influence on the quality of the final clustering results. Because of the faults, the text uses a mixed search strategy which combines genetic algorithm and tabu search to improve the efficiency of FCM. The strategy possesses the advantages of these algorithms. The purpose of this mechanism is to produce the best number of clusters automatically, optimize the selection of the original cluster and advance FCM's ability of breaking away from local critical point. Experiments show that the improved algorithm, with self adaptiveness and high efficiency, gets better clustering results than the original.

Key words: FCM algorithm; genetic algorithm; tabu search; hybrid strategy; intrusion detection

0 引 言

现实生活中处于“亦此亦彼”状态的现象是大量存在的，这是事物内部复杂性的外在体现。将模糊聚类算法用于入侵检测领域是目前的一个研究热点，模糊聚类有多种形式，其中应用最广泛的是基于目标函数的模糊 C-均值聚类（fuzzy C-Means，FCM）。针对 FCM 算法事先要指定聚类数的缺陷，文献[1]通过预先给定聚类宽度值，依据欧氏距离最小策略产生聚类结果从而确定聚类数的思路取得了一定的效果，但聚类宽度的取值依然缺乏理论指导，改进后算法的自适应能力差。针对 FCM 对初始聚类敏感，易陷入局部最优的缺陷，文献[2]将全局搜索性能好的遗传算法引入 FCM 得到一定程度的改善，但该算法的爬山能力依然不是很强。鉴于此，本文采用将遗传算法与具有记忆功能且爬山能力强的禁忌搜索结合起来的混合策略对其进行改进，生成最佳聚类数下的最佳聚类结果，增加了算法的自适应性和抗初始敏感性。仿真实验的结果表明改进后的算法有更高检测率和低误报率，取得较理想的聚类结果。

1 改进后的 FCM 用于入侵检测

1.1 传统的 FCM 用于入侵检测

FCM 算法由 Dunn 和 Bezdek 于 1973 年提出，其基本思想[3-12]如下：

设训练集为 $TrD=\{x_1,x_2,\cdots,x_l\}$，其中 $x_i=(x_{i1},x_{i2},\cdots,x_{in})^T$，$i=1,2,\cdots,l$。将 TrD 划分为 c 个模糊聚类 $X_1,X_2,\cdots,X_c$，使目标函数 $J=J_m(U,V)=\sum_{k=1}^{c}\sum_{p=1}^{l}\mu_{kp}^{m}\|x_p-\nu_k\|^2$ 最小，其中 μ_{kp} 表示样本点 x_p 隶属于

收稿日期：2011-02-25；修订日期：2011-04-30

基金项目：甘肃省自然科学基金项目（0809RJZA015）

作者简介：张永（1963-），男，甘肃兰州人，教授，研究方向为人工智能、信息与网络安全；曹东侠（1983-），男，山西大同人，硕士研究生，研究方向为网络安全。E-mail：dx_cao@163.com

聚类 X_k 的程度，ν_k 为聚类 X_k 的中心列向量，$U=(\mu_{kp})_{c\times l}$ 为隶属度阵，$V=(\nu_1,\nu_2,\cdots,\nu_c)^T$ 为各聚类中心构成的 $c\times l$ 阵，m 是模糊加权指数，控制 U 的模糊程度，m 值越大 U 越模糊，通常在 $(1,5]$ 上取值，经验值为 2。U 的元素 μ_{kp} 满足

$$\sum_{k=1}^{c}\mu_{kp}=1,\ \mu_{kp}\in[0,1],\ \sum_{p=1}^{l}\mu_{kp}\in(0,l) \tag{1}$$

由于数据属性的类型既可能是字符型也可能是数字型，在定义数据间的距离时先把字符型按如下规则转化为数字型：将字符型属性值域内的值进行二进制自然编码，然后转变为对应的十进制数，这样可以认为数据只存在数值型属性。为消除数值型属性因量纲的不同对聚类结果质量的影响，需对数据作归一化预处理，此后文中涉及的数据均指归一化后的数据。具体转化规则如下

$$\overline{x_{\cdot j}}=\frac{1}{l}\sum_{i=1}^{l}x_{ij},\ s_{\cdot j}=$$

$$\sqrt{\frac{1}{l-1}\sum_{i=1}^{l}|x_{ij}-\overline{x_{\cdot j}}|^2},\ x_{ij}\leftarrow\frac{x_{ij}-\overline{x_{\cdot j}}}{s_{\cdot j}} \tag{2}$$

由式（1）及目标函数 J 构造 Lagrange 函数，由 Lagrange 乘数法易得 J 取得极小值的必要条件

$$\mu_{kp}=\left[\sum_{q=1}^{c}\left(\frac{\|x_p-\nu_k\|^2}{\|x_p-\nu_q\|^2}\right)^{\frac{1}{m-1}}\right]^{-1} \tag{3}$$

$$\nu_k=\frac{\sum_{p=1}^{l}\mu_{kp}^m x_p}{\sum_{p=1}^{l}\mu_{kp}^m} \tag{4}$$

为消除 FCM 对孤立点敏感的问题，把式（3）中的 μ_{kp} 修正为

$$\mu'_{kp}=\gamma\mu_{kp}+(1-\gamma)\mu_{kp}^2,\ \gamma\in[0,1] \tag{5}$$

聚类中心 ν_k 修正为

$$\nu_k'=\frac{\sum_{p=1}^{l}\mu_{kp}'^m x_p}{\sum_{p=1}^{l}\mu_{kp}'^m} \tag{6}$$

基于传统 FCM 算法用于入侵检测的工作流程：

输入：归一化训练数据集 TrD，聚类数 c，模糊加权指数 m，收敛终止限 δ；

输出：最佳聚类结果 U_{best} 和 V_{best}，使 J 最小。

步骤 1 置迭代次数 $t=0$，随机产生满足条件（1）的初始隶属度阵 U^t；

步骤 2 由式（5）修正 U^t，由式（6）产生聚类中心阵 V^t，计算 J^t；

步骤 3 由式（3）产生 U^{t+1}，由式（5）修正 U^{t+1}，由式（6）产生聚类中心阵 V^{t+1}，计算 J^{t+1}；

步骤 4 如果 $|J^{t+1}-J^t|<\delta$，算法终止，输出 U^{t+1} 和 V^{t+1} 为最佳聚类结果；否则 $t=t+1$，转步骤 2。

/*上述算法流程中 U^t、V^t、J^t 等符号分别表示迭代次数为 t 时的隶属度阵、聚类中心阵及目标函数值*/

对测试集 TeD 中样本 x，由式（3）计算 x 隶属于最终各聚类的程度，取最大隶属度所对应的聚类为 x 的归属聚类。

1.2 遗传算法和禁忌搜索的混合搜索策略

遗传算法（genetic algorithm，GA）[13-19]借鉴了自然界生物进化和遗传的机理，由 Holland 等人于 1975 年提出的一种仿生随机概率算法。该算法具有很高的隐并行性和全局搜索能力，广泛用于各种复杂优化问题的求解，但标准遗传算法 SGA 存在爬山能力差、易早熟的缺陷。禁忌搜索（tabu search，TS）[20-22]是对局部邻域搜索算法的扩展，由 Glove 等人于 1985 年提出的一种逐步寻优的全局搜索算法。该算法在当前最优解的邻域候选解内搜索下一个最优解时，不仅接收适配值高于当前最优解（如属禁忌对象，则采用藐视准则）的候选解，而且也接收适配值低于当前最优解且不属禁忌对象的劣候选解，克服了传统局部搜索算法中爬山能力差的缺点，增加了算法转向其它搜索区域的可能性，从而使算法能以较大概率搜索到全局最优，记忆功能是 TS 独具的特质，能有效避免迂回搜索使算法快速收敛。Glove 从理论上论证了这两种算法混合的可行性和必要性，混合策略 GATS 继承了这两种互补算法的优势。本文用 TS 对 GA 操作算子中的重组算子和变异算子进行改进。

改进后的重组算子 TSC 的工作机理是：用禁忌表记录染色体的适应度值，将父辈平均适应度值作为渴望水平，进行 TSC 操作时，若子辈染色体的适应度值超过渴望水平或虽未超过但不属禁忌对象，则接受子辈染色体，否则接受适应度值高的父辈染色体。在改进算法中由于使用禁忌表，保持了群体的多样性，有效避免了早熟现象，同时继承了 SGA 中最优个体保留法能将最优个体复制到下一代的作用。

改进后的变异算子 TSM 的工作机理是：从待变异染色体出发，以适配值函数为评价标准，通过禁忌表机制产生出更优适配值的变异个体。

1.3 基于 GATS 的 FCM 算法用于入侵检测

将遗传禁忌混合策略用于改进 FCM，不仅可自动生成最佳聚类数，而且克服了对孤立点和初始聚类的敏感性及易陷入局部最优的缺陷，取得高质量的聚类结果。在整个算法的进行过程中，适应度函数和适配值函数均取 $f(\cdot)=\frac{\mu}{\nu+J}$，$\mu,\nu>0$。

改进后的聚类算法用于入侵检测的流程：

输入：训练样本集 TrD；求解最佳聚类数所用 GATS 参数：最大进化代数 G_1，交叉概率 p_{c1}，变异概率 p_{m1}，种群规模 P_1，禁忌表 T_1，禁忌终止条件 θ_1，收敛终止限 δ_1；聚类数给定情况下所用 GATS 及 FCM 参数：最大进化数 G_2，交叉概率 p_{c2}，变异概率 p_{m2}，种群规模 P_2，禁忌表 T_2，禁忌终止条件 θ_2，收敛终止限 δ_2，模糊加权指数 m 取经验值 2；

输出：最佳聚类结果 U_{best} 和 V_{best}。

步骤 1：由 TrD 的大小 l 求出二进制染色体编码串的长度，随机产生 P_1 条染色体，表现型分别是聚类数为 $c_1, c_2, \cdots, c_{P_1}$ 的聚类，设置进化代数计数器 $t_1 = 0$；

/* 步骤 2：产生最佳聚类数 */

do {

t_1 ++；

for (i=1; i<= P_1 ; i++) {　/* i 为 P_1 条染色体编号 */

/* 步骤 3：用 GATS 求解聚类数 c_i 的最佳聚类结果 */

设置进化代数计数器 $t_2 = 1$；

for (j=1; j<= P_2 ; j++) {　/* j 为 P_2 条染色体编号，对应 c_i 的一个初始聚类 */

对聚类数 c_i，随机产生一个初始隶属度阵，由式 (6) 计算出相应的聚类中心阵，将这 c_i 个中心连成一条实数编码的长为 nc_i 且序数为 j 的染色体，计算染色体 j 的适应度值；}

/* endfor __j，产生 c_i 的 P_2 个初始隶属阵和初始聚类中心阵 */

记录 t_2 代适应度值最大的染色体的 $U_{best}^{t_2}$ 和 $V_{best}^{t_2}$；

do {

t_2 ++；

按交叉概率 p_{c2}，重组算子 TSC_2，变异概率 p_{m2}，变异算子 TSM_2 产生新一代染色体，计算 t_2 代各染色体适应度值并记录下适应度最大的个体的 $U_{best}^{t_2}$ 和 $V_{best}^{t_2}$；

} while (前后两代最大适应度值差大于 δ_2 && 进化次数 t_2 小于 G_2)；

记录聚类数为 c_i 的最佳聚类结果，其适应度值作为 c_i 在 P_1 中对应染色体的适应度值；

/* 已求得聚类数 c_i 的最佳聚类结果

} /* endfor __i，计算出聚类数 $c_1, c_2, \cdots, c_{P_1}$ 的最佳聚类结果 */

记录 t_1 代适应度最大的个体；

按交叉概率 p_{c1}，重组算子 TSC_1，变异概率 p_{m1}，变异算子 TSM_1 产生新一代染色体，计算 t_1+1 代各染色体适应度值并记录下适应度最大的个体，修正聚类数 c_i ($1 \leqslant c_i \leqslant P_1$) 的值；

} while (前后两代最大适应度值大于 δ_1 && 进化次数 t_1+1 小于 G_1)；

步骤 4：输出最佳聚类数，最佳聚类结果 U_{best} 和 V_{best}。

2　两种算法用于仿真实验

实验用数据源自 KDDCup1999，这是一个测试入侵检测性能常用的标准数据库，分训练集和测试集两部分，每一数据有 41 个特征和 1 个类别标识属性，这 41 个特征又分为 4 类属性：基本属性、内容属性、流量属性和主机流量属性，其中有 34 个数字型特征和 7 个字符型特征，详见表 1。数据共分为 5 大类：DOS、Probe、R2L、U2R 和 Normal，前 4 大类为攻击类型，后一大类为正常数据类型。每一大攻击类型又包含若干具体小的攻击种类，测试集中每一大类所含的具体攻击种类数多于相应的训练集，称测试集中每一大类所含相应训练集中的具体攻击种类为已知攻击，测试集中每一大类所含相应训练集中没有的具体攻击种类为未知攻击。由于在实际的网络流量中正常数据的数目远大于攻击数据数，实验中按正常数据与异常数据 9∶1 的比例产生 DoS＋Normal (Ⅰ)，Probe＋Normal (Ⅱ)，R2L＋Normal (Ⅲ)，U2R＋Normal (Ⅳ) 和 5 大类数据都含的 (Ⅴ) 5 个训练子集和 5 个测试子集，每个训练子集所含的样本数为 20000，每个测试子集所含样本数为 40000。各数据子集中具体攻击种类数见表 2，测试子集中与相应训练子集中的具体攻击数之差为未知攻击数。文献 [20] 的研究结论给出检测各种攻击所需的属性类型见表 3，表 3 中的序数与表 1 中的一致。传统 FCM 与本文算法在入侵检测中的性能比较见表 4、表 5，其中传统算法是对 10 个聚类数所求结果的平均值。实验环境是 Windows 操作系统，Intel processor 2.0GHz，1.00GB RAM，MATLAB7.0.

表 1　原始数据集属性列

基本属性	内容属性	流量属性	主机流量属性
1. duration	10. hot 11. num __failed __logins	23. count	32. dst __host __count
2. protocol __type	12. logged __in 13. num __compromised	24. srv __count	33. dst __host __srv __count
3. service 4. flag	14. root __shell 15. su __attempted	25. serror __rate	34. dst __host __same __srv __rate
5. src __bytes	16. num __root	26. sev __serror __rate	35. dst __host __diff __srv __rate
6. dst __bytes	17. num __file __creations	27. rerror __rate	36. dst __host __same __src __port __rate
7. land	18. num __shells	28. srv __rerror __rate	37. dst __host __srv __diff __host __rate
8. wrong __fragment	19. num __access __files	29. same __srv __rate	38. dst __host __serror __rate
9. urgent	20. num __outbound __cmds	30. diff __srv __rate	39. dst __host __srv __serror __rate
	21. is __host __login	31. srv __diff __host __rate	40. dst __host __rerror __rate
	22. is __guest __login		41. dst __host __srv __rerror __rate

表 2 各数据子集所含具体攻击类型数

	训练子集	测试子集
Ⅰ	6	10
Ⅱ	4	6
Ⅲ	8	15
Ⅳ	4	8
Ⅴ	20	36

表 3 各数据集所需的属性

数据集	所需属性类型
Ⅰ	2，5，23，34
Ⅱ	1，3，5，6，23，35
Ⅲ	1，3，5
Ⅳ	1，3，5，14，32
Ⅴ	3，5，23，32

表 4 两种不同算法的检测精度性能对比

测试样本集	本文算法		传统算法	
	已知攻击/%	未知攻击/%	已知攻击/%	未知攻击/%
Ⅰ	95.6	84.3	88.1	76.2
Ⅱ	93.7	75.1	84.0	67.4
Ⅲ	80.1	37.7	73.8	31.5
Ⅳ	86.8	40.3	78.3	38.6
Ⅴ	92.2	80.6	86.2	71.3

表 5 两种不同算法的误报率对比

测试样本集	误报率/%	
	统统算法	本文算法
Ⅰ	0.67	0.42
Ⅱ	0.59	0.33
Ⅲ	2.79	1.28
Ⅳ	2.16	1.01
Ⅴ	1.13	0.79

实验结果表明，本文算法与传统算法在检测用时相当的情况下，无论是对已知攻击还是对未知攻击的检测力度都得到了改进，检测的误报率也明显下降，说明了改进后算法的可行性和有效性。

3 结束语

FCM 算法是重要的聚类分析方法，其算法简易，适用于处理各种高维异构数据的优化聚类问题，然而该算法在运行前须给出聚类数，且对孤立点和初始聚类敏感，易陷入局部最优，这些缺陷使聚类结果通常不是很理想，制约了算法的实际应用。针对存在的缺陷，引入遗传禁忌混合策略对其进行改进，该混合策略兼具两种算法的优势，即高度的隐并行性和具有记忆功能，较强的全局和局部搜索能力，使改进后的 FCM 算法有很强的抗初值敏感性和逃逸局部最优的机制，明显改善了算法的性能。实验数据中也反映出对某些性能指标（如对未知 R2L 的检测精度）的检测还不尽人意，如何自适应地提高这些性能指标的检测精度是下一步着力要解决的问题。

参考文献：

[1] YANG De-gang. Research of the network intrusion detection based on fuzzy clustering [J]. Journal of Computer Science, 2005, 32 (1): 86-87 (in Chinese). [杨德刚. 基于模糊 C 均值聚类的网络入侵检测算法 [J]. 计算机科学, 2005, 32 (1): 86-87.]

[2] XIAO Man-sheng. Research of intrusion detectin based on GA-FCM [J]. Journal of Xiangtan Normal University (Natural Science Edition), 2008, 30 (4): 16-19 (in Chinese). [肖满生. 基于遗传模糊聚类算法的入侵检测研究 [J]. 湘潭师范学院学报 (自然科学版), 2008, 30 (4): 16-19.]

[3] ZHANG Guo-suo, ZHOU Chuang-ming, LEI Ying-jie. Improved fuzzy C-means clustering algorithm and its application to intrusion detection [J]. Journal of Computer Applications, 2009, 29 (5): 1336-1338 (in Chinese). [张国锁, 周创明, 雷英杰. 改进 FCM 聚类算法及其在入侵检测中的应用 [J]. 计算机应用, 2009, 29 (5): 1336-1338.]

[4] XIAN Ji-qing, LANG Feng-hua. Fuzzy clustering theory for analyzing intrusion detection data [J]. Journal of Chongqing University (Natural Science Edition), 2005, 28 (7): 74-76 (in Chinese). [鲜继清, 郎风华. 基于模糊聚类理论的入侵检测数据分析 [J]. 重庆大学学报 (自然科学版), 2005, 28 (7): 74-76.]

[5] LU Hu, XU Jing. Intrusion detection based on hybrid fuzzy clustering algorithm [J]. Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2008, 22 (4): 60-63 (in Chinese). [陆虎, 徐景. 基于混合聚类算法的异常检测方法 [J]. 江苏科技大学学报 (自然科学版), 2008, 22 (4): 60-63.]

[6] ZHU Weiwei, WANG Weiping, LIANG Liang. Intrusion detection method based on fuzzy cluster analysis [J]. Systems Engineering and Electronics, 2006, 28 (3): 474-477 (in Chinese). [朱卫未, 王卫平, 梁樑. 基于模糊聚类分析的入侵检测方法 [J]. 系统工程与电子技术, 2006, 28 (3): 474-477.]

[7] XIAN Jiqing, LANG Fenghua, TANG Xianlun. A novel intrusion detection method based on clonal selection clustering al-

gorithm [C]. Proceedings of International Conference on Machine Learning and Cybemetics, 2005: 3905-3910.

[8] XU Shiguo, LINGHU Dazhi. Intrusion detecting algorithm based on fuzzy cluster [J]. Journal of Liaoning Technical University (Natural Science), 2008, 27 (6): 881-884 (in Chinese). [徐世国，令狐大智. 基于模糊聚类的入侵检测算法 [J]. 辽宁工程技术大学学报（自然科学版），2008，27 (6)：881-884.]

[9] WANG Ru-shan, LI Yong-zhong, ZHANG Nian-gui, et al. Semi-supervised learning and its application to intrusion detection system [J]. Journal of Guangxi Normal University (Natural Science Edition), 2009, 27 (3): 179-182 (in Chinese). [王汝山，李永忠，张念贵，等. 半监督学习在入侵检测系统中的应用 [J]. 广西师范大学学报（自然科学版），2009，27 (3)：179-182.]

[10] Grira N, Crucianu M, Boujemaa N. Active semi-supervised fuzzy clustering [J]. Pattern Recognition, 2008, 41 (5): 1834-1844.

[11] LIU Zhi-yong, GENG Xin-qing. Text mining algorithm based on fuzzy clustering [J]. Computer Engineering, 2009, 35 (5): 44-45 (in Chinese). [刘志勇，耿新青. 基于模糊聚类的文本挖掘算法 [J]. 计算机工程，2009，35 (5)：44-45.]

[12] YANG Xiaoqiang. Algorithm for intrusion detection based on evolution semi-supervised fuzzy clustering [J]. Computer Engineering and Application, 2008, 44 (4): 33-35 (in Chinese). [杨晓强. 一种进化半监督式模糊聚类的入侵检测算法 [J]. 计算机工程与应用，2008，44 (4)：33-35.]

[13] LUO An-kun, LI Su. The apply research of genetic algorithm based on coarsegrained model in network intrusion detectin system [J]. Journal of Yunnan University (Natural Science Edition), 2006, 28 (S2): 51-55 (in Chinese). [罗安坤，李甦. 基于粗粒度模型遗传算法在网络入侵检测系统中的应用研究 [J]. 云南大学学报（自然科学版），2006，28 (S2)：51-55.]

[14] WANG Ying, DENG Ya-ping, ZENG Li-mei. Application of genetic algorithms to intrusion detection [J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science), 2006, 18 (2): 261-263 (in Chinese). [王英，邓亚平，曾立梅. 遗传算法在入侵检测中的应用 [J]. 重庆邮电学院学报（自然科学版），2006，18 (2)：261-263.]

[15] GUO Hui-ling, TANG Yong, ZHANG Dong-li. Application of genetic algorithm in rule extraction of intrusion detection [J]. Journal of Harbin Institute of Technology, 2009, 41 (1): 248-250 (in Chinese). [郭慧玲，唐勇，张冬丽. 遗传算法在入侵检测规则提取中的应用 [J]. 哈尔滨工业大学学报，2009，41 (1)：248-250.]

[16] MAO Li-min, YAO Shu-ping, HU Chang-zhen. A new hybrid attribute selection method and its application in intrusion detection [J]. Transactions of Beijing Institute of Technology, 2008, 28 (3): 218-221 (in Chinese). [毛俐旻，姚淑萍，胡昌振. 一种新型混合特征选择方法及其在入侵检测中的应用 [J]. 北京理工大学学报，2008，28 (3)：218-221.]

[17] YU Yan, HUANG Hao. An ensemble approach to intrusion detection based on improved multi-objective genetic algorithm [J]. Journal of Software, 2007, 18 (6): 1369-1378 (in Chinese). [俞研，黄皓. 基于改进多目标遗传算法的入侵检测集成方法 [J]. 软件学报，2007，18 (6)：1369-1378.]

[18] Shon T, Seo J, Moon J. SVM approach with a genetic algorithm for network intrusion detection [C]. Proc of the 20th Int'l Symp on Computer and Information Sciences. Berlin: Springer-Verlag, 2005: 224-233.

[19] FENG Li, SUN Lijuan. Application of quantum genetic algorithm to the IDS which based on artificial immune [J]. Computer Applications and Software, 2006, 23 (10): 23-24 (in Chinese). [冯莉，孙力娟. 量子遗传算法在基于人工免疫的入侵检测系统中的应用 [J]. 计算机应用与软件，2006，23 (10)：23-24.]

[20] ZHANG Hao, TAO Ran, LI Zhi-yong, et al. A research and application of feature selection based on KNN and tabu search algorithm in the intrusion detection [J]. Acta Electronica Sinica, 2009, 37 (7): 1628-1631 (in Chinese). [张昊，陶然，李志勇，等. 基于 KNN 算法及禁忌搜索算法的特征选择方法在入侵检测中的应用研究 [J]. 电子学报，2009，37 (7)：1628-1631.]

[21] LI Wen-fa, CHEN You, DUAN Mi-yi, et al. IP flow classification based on GATS-C4. 5 [J]. Computer Science, 2009, 36 (4): 68-72 (in Chinese). [李文法，陈友，段洣毅，等. 基于 GATS-C4. 5 的 IP 流分类 [J]. 计算机科学，2009，36 (4)：68-72.]

[22] CHEN You, SHEN Hua-wei, LI Yang, et al. An efficient feature selection algorithm toward building lightweight intrusion detection system [J]. Chinese Journal of Computers, 2007, 30 (8): 1398-1408 (in Chinese). [陈友，沈华伟，李洋，等. 一种高效的面向轻量级入侵检测系统的特征选择算法 [J]. 计算机学报，2007，30 (8)：1398-1408.]