基于蜂群算法的网络入侵检测模型优化①

吴建龙

(首钢工学院 机电工程系, 北京 100144)

摘 要: 基于网络入侵检测的蜂群算法优化模式是一个用于网络入侵检测开发的专用编程接口. 基于该编程接口, 在 Linux 平台上设计和实现了一个复杂的入侵检测系统. 基于网络入侵检测的蜂群算法与差分进化算法(DE)混合, 采取数据信息处理模式, 可以按照双群结构的要求, 进行数据信息独立分析, 从而能够产生数据信息交换功能. 通过分布式技术对蜂群进行空间分析, 通过空间信息搜索工具, 保证学习策略功能能够完成. 从仿真实验看提高种群解的质量. 设计了一种简单入侵检测模式的描述语言, 对入侵检测的特征数据库进行优化, 对网络异常行为进行入侵检测.

关键词: 蜂群算法; 入侵检测; 协议; 优化

Optimization of Network Intrusion Detection Model based on Artificial Bee Colony Algorithm

WU Jian-Long

(Department of mechanical and electrical engineering Shougang Institute of Technology, Beijing 100144, China)

Abstract: The Artificial Bee Colony Algorithm based on network intrusion detection is a set of application programming interface, based on which a sophisticated intrusion detection system is designed and developed on Linux platform, and based on such an algorithm combined with Differential Evolution (DE), data information exchange is thereby realized, with data processing model adopted and analysed independently under the bi-group structure rules. By analyzing bee colony with distributed technology and with the space information search tool, the study strategy function is thereby assured. The quality of population improvement can be proved through emulation experiments. A script of description language for a simple intrusion detection model is designed, with a view to optimize the detection sample database and perform the detection for network anomalous behaviors.

Key words: artificial bee colony algorithm; intrusion detection; protocol; optimizations

针对基于网络入侵检测的蜂群算法搜索速度较慢、早熟收敛、易陷入局部最优解等情况, 大批学者和专家通过学习和研究提出了许多改进的基于网络入侵检测的蜂群算法, 本文中通过对改进的基于网络入侵检测的蜂群算法的研究将它们分成三类——基于初始化参数优化的改进型、基于搜索过程优化的改进型和基于整体优化的改进型, 针对这三类改进的基于网络入侵检测的蜂群算法介绍了各类的典型算法. 用户在安全控制策略方面需要进行全面分析, 从而能够更好的提升网络入侵检测能力.

1 网络入侵检测的蜂群算法

网络入侵检测的蜂群算法(BCA)是一种基于蜜蜂行为的优化算法. 基于 Boltzmann 选择机制提出了一种改进的基于网络入侵检测的蜂群算法(BBCA)用来优化多变量函数. BBC 是对参数初始化的改进, 算法使初始群体均匀化. 采用 Boltzmann 选择机制来代替轮盘赌以防止算法过早收敛[1].

1.1 选择网络安全管理策略

通过对网络入侵检测自适应机制和机器学习模式分析, 对选择策略进行模式分析, 发挥搜索算法的信

① 基金项目:国家自然科学基金(60443004)

收稿时间:2013-07-13;收到修改稿时间:2013-08-26

© 中国科学院软件研究所 http://www.c-s-a.org.cn

息处理功能作用, 在不同的环境下按照蜂群的算法要求, 从而能够进行数据信息的决策, 让决策功能能够在信息连接过程中产生积极的作用[2]. 从当前的情况看, 需要对数据信息的估值进行优化, 从而能够产生概率信息处理模型. 保证搜索算法能够符合空间数据信息控制的要求. 算法实现过程中如果用 A 表示全体行动的集合, 那么通过价值集合模式可以对蜂群进行选择. 设 A 为全体行动的集合, V_A 为行动的估计价值集合, $\alpha_1,\alpha_2,\ldots,\alpha_v \in A$ 为当前状态下可选的行动, $\overline{v}(a_1),\overline{v}(a_2),\ldots,\overline{v}(a_n) \in A$ 分别为各行动的估计价值, 系统对蜂群算法进行解析, 从而找到入侵检测的方法, 把 Boltzmann 选择策略与网络安全形成互动机制, 逐步引入到网络入侵检测的蜂群算法中, 提出了对应的估计价值, 则当前状态下选择任一行动 $a_k \in \{a_1,a_2,\ldots,a_n\}$ 的概率为:

$$P_{rob}(a_k)=\frac{\exp(\frac{v(a_k)}{T})}{\sum_{i=1}^{N}\exp(\frac{v(a_i)}{T})} \quad (1)$$

其中, T 为退火温度. 由全概率公式可知,

$$\sum_{i=1}^{N} P_{rob}(a_k)=1,2,\ldots,n.$$

1.2 选择网络安全管理策略

初始解在整个算法搜索过程中产生起点作用, 通过对初始群体形成, 从而对个体空间分布情况进行解析, 保证算法搜索整体性能得到全面提高. 通过算法设计, 可以看出初始群体生成是随机的, 产生若干个初始群体组合, 对生产初始群体的不合理问题进行判断, 在此基础上进行算法改进, 保证群体总数的分布相对比较均匀, 通过小区间生成法的应用, 可以改进群体个体的分布均匀性, 保证群体可以含有丰富的个体, 个体可以均匀分布在不同的区间上[3]. 各个小区间组合在一起生成一个初始的个体, 这些初始个体可以有效的分布在不同的个体空间上, 群体丰富的模式可以全面体现出来, 对全局搜索提供了方便, 通过仿真实验证明, 此种方法生产的群体个体是最优良的, 加大了算法的收敛速度, 改善了算法的收敛性.

小区间在设计过程中通过等值机制进行均匀分布, 把整个区间分成 n 个小区间, n 被称为种群的规模度.

1.3 选择安全控制机制的改进

在 BCA 算法中, 需要根据概率情况选择合适的食物源, 其中采取的方式是轮盘赌的策略, 轮盘赌算法选择过程中容易导致全体多样性的下降, 因此算法会处于过早收敛的状态. 算法设计过程是个不断优化的过程, 在不同阶段需要承受不同压力, 从总体情况看, 早期进行选择产生的压力较小, 所希望较差的个体在此情况下也具有一定的生存机会, 群体在运行过程中具有多样性特点. 后期选择过程中具有较大的压力, 希望算法调整可以缩小搜索的范围, 可以提取最优解, 通过最优解对动态搜索的压力进行分解, 基于网络入侵检测的蜂群算法机制就是在此技术之上, 按照选择策略的具体要求选择搜索过程[4]. 通过自适应控制和机器学习方法的分析, 可以更好的设计搜索算法. 传统算法在处理过程中存在选择模糊问题, 难以改进搜索功能, 因此初始化过程中容易产生问题.

1.4 具体算法流程设计

在 BBCA 算法设计过程中, 可以产生 4 个控制参数类型: 其中食物源的具体个数=入侵检测中引领蜂的具体个数=网络系统中跟随蜂的具体个数(*SN*), limit 是一种限制类型参数, 系统的最大循环次数可以用 MCN 表示, 系统设计过程中需要对初始温度在 BBCA 算法中有 4 个具体控制参数进行划分: 食物源的参数具体个数=入侵检测过程中引领蜂的个数=系统运行中跟随蜂的个数(*SN*), limit 是一种限制类型参数, 产生的最大循环次数 MCN, 初始温度可以表示为 T_0.

2 自适应搜索空间的混沌网络安全入侵检测

基于网络入侵检测的蜂群算法作为一种新的随机优化算法, 算法设计过程中需要对全局群体进行最优解分析, 但是在实际操作过程中存在搜索速度慢的特点, 容易过早收敛, 导致个体多样性逐步减少, 陷入最优解难以获取的境地. 算法优化需要解决搜索速度慢的问题, 避免在过早的情况下收敛, 因此研究了基于搜索过程改进的网络入侵检测蜂群算法, 对搜索空间扩大产生重要的作用, 通过混沌入侵检测蜂群算法的使用, 可以对算法循环的过程中进行搜索, 确保在实现过程中能够构造一个搜索区域, 并且对每个值进行重新评估. 当某一个值处于局部最优的状态的时候, 可以通过混沌算法把最优解找出.

2.1 动态调整搜索空间安全策略

入侵检测过程中需要对种群进行设计, 把其空间的解设定为 n 个, 此时每个解可以看成是 d 维向量.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

在自适应搜索空间的混沌基于网络入侵检测的蜂群算法(SA-CBC)中, 搜索算法实现过程中可以对空间进行合理分布, 把 Y 的分布空间得到有效的利用, 从而能够最大可能对区间情况进行分布.

种群个体在设计过程中可以重新生成, 那么可以根据种群个体在设计过程中可以重新生成和蜂群数据信息处理情况, 从而能够按照搜索算法的要求, 进行信息比对, 确保信息能够在评估过程中产生作用, 达到搜索算法优化的目的[5]. 通过循环算法的优化, 保证循环迭代效应能够产生.

按照上述方法, 当搜索空间逐步缩小的情况下, 通常会产生两个问题, 一种是最优解没有找出来, 而处于搜索范围之外, 此种情况下也就无法获取所谓的最优解. 另一方面是个体的运动范围被缩小, 对算法在局部获取最优解的能力产生影响. 如果一些个体均匀在极值附近运动的时候, 那么会产生暂时性的停滞状态, 对局部极值的限制是非常明显的, 无法获取最优解, 因此需要针对这两个问题采取有效的措施. 算法优化过程中可以把种群的个体进行划分, 把其分成两个部分, 一部分完成动态区域搜索功能, 可以加快算法的收敛程度, 而另一部分完成在原空间的搜索任务, 保证空间边缘的解能够有效获取而不被忽视, 通过对方法实现可以看出这种方法具有可行性. 搜索空间在进行全面的调整和整合之后, 不是立即进行压缩, 而是在每次压缩之后进行迭代处理, 保证群体环境能够适应再次压缩的需求, 保证最优解能够有效获取.

2.2 关键词

混沌现象在自然界中普遍存在, 其是一种非线性现象, 从表面看具有混乱的情况, 但是从内在看起具有精致的结构, 因此具有遍历性、随机性、规律性特点, 在一定范围内, 可以根据其运行规律找到相关的遍历状态[6]. 混沌方程的确定可以得到各种算法的随机性运行模式, 可以在不重复状态下遍历各种状态, 通过混沌可以确定混沌变量, 对搜索空间优化产生重要的作用. 常用的 Logistic 映射就是一个典型的混沌系统, 算法设计过程处于混沌, 那么在遍历的过程中需要对局部情况进行最优分析, 从而能够形成一条好的搜索机制, 对算法优化产生作用. 当前需要把混沌算法和蜂群算法结合在一起, 从而能够展开算法研究, 在基于蜂群算法的网络入侵检测系统研究过程中, 需要通过循环次数的改进, 可以解决局部无法获取最优解的问题. 通过随机生产最优解模式带来以往的解. 本文研究过程中提出了 SA-CBCA 算法, 主要针对搜索停滞的状况下, 重新获取搜索的新值, 对保证最优值获取创造了条件. 通过对混沌运动的遍历情况进行分析, 可以产生混沌运动队列, 可以按照队列的位置进行位置选择, 解决搜索停滞的问题, 对提高算法收敛精度和速度具有重要作用.

2.3 选择策略的确定

在 SA-CBCA 算法设计过程中, 可以根据峰值情况对食物源进行有效选择, 选择算法实施过程中利用锦标赛的策略[7]. 锦标赛规则设计过程中主要是选择一个适应的值, 然后把此值作为一个重要的标准, 算法设计过程中通过设计一个合理的值, 可以对数据信息进行有效的约束, 避免个体对整体产生影响, 保证数据搜索信息能够产生重要的作用, 通过规避算法提高数据信息的处理能力.

3 双种群差分网络入侵检测优化实验

双种群差分基于网络入侵检测的蜂群算法(BDBCA)是全局优化的改进的基于网络入侵检测的蜂群算法. 差分进化(DE)算法从而能够产生矢量作用, 保证空间算法能够符合搜索引擎的要求, 参数设计过程中需要按照原理控制要求产生信息比对. 在基本的 DE 算法处理过程中, 需要对 NP 进行信息分解, 保证信息服务能够产生相应的 d 向量, 通过交叉选择措施确保数据信息处理符合差分要求, 每个解 $X_i=(x_{i1},x_{i2},\ldots,x_{id})(i=1,2,\ldots,NP)$ 是一个 d 维的向量. 进化过程主要由变异、交叉和选择 3 个操作组成.

(1) 变异操作: 在 DE 算法中, 根据变异个体的生成方式不同, 有多种不同的 DE 变异方式, 常见的变异方式有 DE/rand/1/bin 和 DE/best/1/bin 两种[8].

(2) 交叉操作: DE 算法在网络克隆和入侵检测过程中发挥重要的作用, 通过数据信息选择, 保证适应模型能够得到蜂群算法的支持. 从而能够产生向量控制信息. 将经过变异和交叉操作后生成的新个体 u_i 与父代个体 x_i 进行比较, 如果 u_i 的适应度优于 x_i 进入下一代, 否则, 维持 x_i 不变, 直接进入下一代. 假设在一个 D 维的目标搜索空间中, 有 N 个蜂组成一个群落, 其中第 i 个蜂表示为一个 D 维的向量

$X_i=(x_{i1},x_{i2},\cdots,x_{iD})$, $i=1,2,\cdots,N$.

第 i 个蜂的“飞行”速度也是一个维的向量, 记为

© 中国科学院软件研究所 http://www.c-s-a.org.cn

$V_i=(v_{i1},v_{i2},\cdots,v_{iD})$，$i=1,2,\cdots 3$

第 i 个蜂迄今为止搜索到的最优位置称为个体极值, 记为

$p_{best}=(p_{i1},p_{i2},\cdots,p_{iD})$，$i=1,2,\cdots,N$.

整个蜂群迄今为止搜索到的最优位置为全局极值, 记为

$$g_{best}=(p_{g1},p_{g2},\cdots,p_{gD})$$

在找到这两个最优值时, 蜂根据如下的公式(2)和(3) 来更新自己的速度和位置优化:

$$v_{id}=wv_{id}+c_1r_1(p_{id}-x_{id})+c_2r_2(p_{gd}-x_{id}) \quad (2)$$

$$x_{id}=x_{id}+v_{id} \quad (3)$$

算法优化的流程如下:

① 对蜂群进行初始化处理, 确保群体规模 N, 每个蜂的位置 x_i 和速度 v_i;

② 对每个蜂的适应值 $F_{it}[i]$进行计算;

③ 针对不同的蜂采取不同的适应度计算方法, 用它的适应度值 $F_{it}[i]$和个体极值 $p_{best}(i)$比较, 如果 $F_{it}[i]>p_{best}(i)$, 则用 $F_{it}[i]$替换掉 $p_{best}(i)$;

④ 对每个蜂, 用它的适应度值 $F_{it}[i]$和全局极值 g_{best} 比较, 如果 $F_{it}[i]>p_{best}(i)$则用 $F_{it}[i]$替 g_{best};

⑤ 根据公式(4.1), (4.2)更新蜂的速度 v_i 和位置 x_i;

⑥ 如果满足结束条件(误差足够好或到达最大循环次数)退出, 否则返回②.

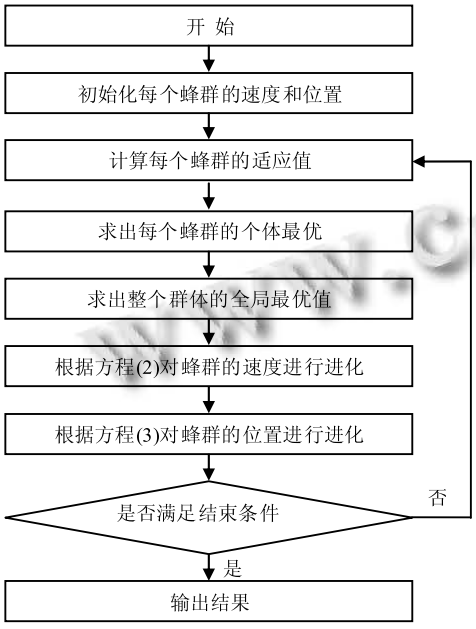

图 1　PSO 算法优化流程图

入侵检测误码率检测试验仿真如图 2 所示.

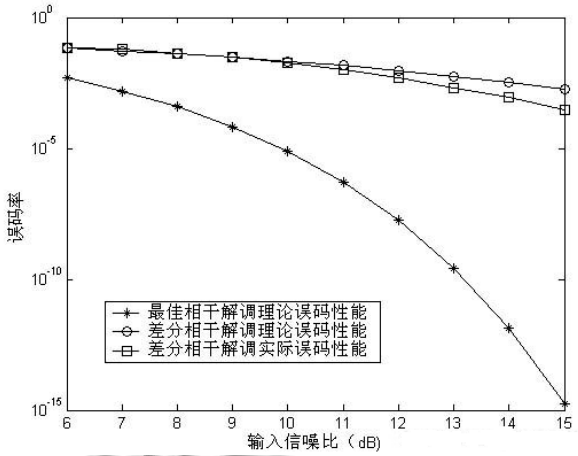

图 2　误码率检测效率提升仿真

在需优化监听程序性能时可做以下相应调整.

(1) 调整用户级缓存. 通过修改 pcap_open_live() 函数的源代码并重新编译后来调整用户级缓存.

(2) 调整函数 pcap_open_live()中的读操作等待时间值. 通常出于效率可将该值设置的比较大; 但当对响应时间要求比较高时, 应将该值改小.

被取证机信息采集器有两种设计倾向, 一种是信息采集器仅负责收集信息, 并直接传回取证机; 另一种则是信息采集器有一定的信息分析处理功能. 前者, 被取证机的负担较轻, 但是取证机的负担相对较重, 在被取证机数目较多的情况下可能引起较大入侵检测流量; 后者, 采用智能信息采集器, 有效减轻取证机负担和入侵检测流量, 但被取证机的负担较重, 可能会对被取证机的使用造成影响.

4　结语

蜂群算法对网络入侵检测产生重要的作用, 通过对算法的优化可以对网络中的数据情况进行分析, 对科学判断网络故障和入侵检测产生作用. 当网络管理人员在后台发现有入侵迹象的时候, 通过入侵检测手段可以对网络进行安全管理. 入侵检测的过程中通过主动检测和被动分析两种方法. 蜂群网络入侵检测算法可以把安全产品、自身安全级别结合在一起. 通过硬件、软件架构设计科学判断网络入侵情况. 随着网络安全的应用扩大, 对产品的易用性有了新的要求, 通过中文图形界面、自身数据库维护等手段, 提高入侵检测水平. 蜂群网络入侵检测系统在设计过程中集

(下转第 222 页)

© 中国科学院软件研究所 http://www.c-s-a.org.cn

10 Fang LY, Li ST, Nie Q, Izatt JA, Toth CA, Farsiu S. Sparsity based denoising of spectral domain optical coherence tomography images. Biomedical Optics Express, 2012, 3(5): 927–942.

11 Xu D, Huang Y, Kang JU. Assessment of robust reconstruction algorithms for compressive sensing spectral-domain optical coherence tomography. SPIE BiOS. International Society for Optics and Photonics. 2013. 1–14.

12 Zhang Z, Ganesh A, Liang X, et al. TILT: transform invariant low-rank textures. International Journal of Computer Vision, 2012, 99(1): 1–24.

13 Mardani M, Mateos G, Giannakis GB. Recovery of low-rank plus compressed sparse matrices with application to unveiling traffic anomalies. arXiv preprint arXiv. 2012, 1204. 6537.

14 Ganesh A, Min K, Wright J, et al. Principal component pursuit with reduced linear measurements. Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on. IEEE. 2012. 1281–1285.

15 Ma Y, Niyogi P, Sapiro G, et al. Dimensionality reduction via subspace and submanifold learning. Signal Processing Magazine, IEEE, 2011, 28(2): 14–126.

16 Lin Z, Chen M, Wu L, Ma Y. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices[Technical Report]. UIUC UILU-ENG-09- 2215. 2009. 1–20.

17 Liu GC, Lin ZC, Yan SC, et al. Robust recovery of subspace structures by low-rank representation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171–184.

18 Salah MB, Mitiche A, Ayed IB. Multiregion image segmentation by parametric kernel graph cuts. IEEE Trans. on Image Processing, 2011, 20(2): 545–557.

19 Gonzalez RC, Woods RE. Digital Image Processing, Reading. MA, Addison-Wesley, 1992.

(上接第 226 页)

成了这些算法优点.

随着网络系统数据流的增加, 对安全的要求越来越高, 入侵检测系统的性能要求也越来越高, 蜂群算法入侵检测算法已经运用于千兆入侵检测产品中. 网络入侵检测系统在运用过程中不仅需要对攻击情况进行分析, 还需要对网络系统的数据进行全面审计, 保证网络流量处于合理监控方位, 对网络分流产生重要的作用, 对整个网络系统优化提供解决方案.

参考文献

1 唐正军,李建华.入侵检测技术.北京:清华大学出版社,2009:8–9.

2 卿斯汉,文伟平,蒋建春,马恒太,刘雪飞.一种基于网状关联分析的网络蠕虫预警新方法.通信学报,2010,25(7).

3 Stevens WR. TCP/IP Illustrated, Volume 2: The Implement. USA: Addison Wesley, 2009. 34–36.

4 Brodleyce C. Temporal sequence learning and data reduction for anomaly detection. Proc. of the 5th Conference on Computer and Communications Security. New York. 2009. 167–170.

5 李玉波,朱自强,郭军.linux C 编程.北京:清华大学出版社,2009:254–271.

6 J. Holland. 自然界和人工系统的适应性.北京:北京科学出版社,1975:167–188.

7 陈江华.遗传算法求解 TSP 问题的研究进展.昆明理工大学学报(理工版),2009:45–47.

8 王莉.基于遗传算法的 0-1 背包问题求解.计算机仿真,2009:36–40.

© 中国科学院软件研究所 http://www.c-s-a.org.cn