文章编号:1674-2974(2022)04-0119-09 DOI:10.16339/j.cnki.hdxbzkb.2022281

基于PSO-TSA模型的网络安全态势要素识别研究

张克君[1,2],郑炜[1],于新颖[2†],王航宇[1],王志强[1]
(1. 北京电子科技学院 网络空间安全系,北京 100071;
2. 北京邮电大学 网络空间安全学院,北京 100876)

摘 要:针对网络安全态势感知技术中态势要素提取的质量与效率较低的问题,提出了融合粒子群(Particle Swarm Optimization,PSO)和模拟退火(Simulated Annealing,SA)的态势要素识别模型PSO-TSA.在位置更新模块,利用Metropolis准则对PSO算法中的个体极值和全局极值进行退火优化,增加粒子的选择性,提高态势要素提取质量.在参数优化模块,利用Metropolis准则优化PSO算法中的参数,并对参数优化过程和粒子适应度同时进行评价,避免算法陷入局部最优,提高态势要素识别效率.按照目前网络状态的实际需求,选择了37个网络安全数据字段,搭建了小型网络环境,以获取更加真实的网络安全数据集SDS-W.在开放网络安全数据集和获取的SDS-W数据集上分别进行态势要素识别实验,实验证明,PSO-TSA在时间成本保持不变甚至更少的基础上,态势要素识别的精确度平均提升了5%~7%.

关键词:网络安全态势感知;态势要素识别;粒子群算法;模拟退火算法
中图分类号:TN915.08 **文献标志码**:A

Research on Recognition of Network Security Situation Elements Based on PSO-TSA Model

ZHANG Kejun[1,2],ZHENG Wei[1],YU Xinying[2†],WANG Hangyu[1],WANG Zhiqiang[1]
(1. Department of Cyberspace Security,Beijing Electronic Science and Technology Institute,Beijing 100071,China;
2. College of Cyberspace Security,Beijing University of Posts and Telecommunications,Beijing 100876,China)

Abstract: Given the low quality and efficiency of situation element extraction in network security situation awareness techniques, this paper proposes a situation element identification model incorporating particle swarm optimization and simulated annealing (PSO-TSA). In the position update module, the Metropolis criterion is utilized to optimize the individual and global extremum in the PSO algorithm to increase the selectivity of the particles and improve the quality of the situation elements extraction. In the parameter optimization module, the parameters in the PSO algorithm are optimized using the Metropolis criterion, and the parameter optimization process and particle fitness are evaluated simultaneously to rid the local optimum and improve the efficiency of the situation element recognition. Due to the actual needs of the current network state, this paper selects 37 network security data fields and es-

* **收稿日期**:2021-11-29
基金项目:北京高校高精尖学科建设项目(20210086Z0401), Advanced Discipline Construction Project of Beijing Universities (20210086Z0401); 国家重点研发计划网络空间安全重大专项课题资助(2018YFB0803601), National Key Research and Development Program on Cyberspace Security (2018YFB0803601)
作者简介:张克君(1972—),男,山东临沂人,北京电子科技学院教授,博士
† 通信联系人,E-mail: Xinying_334@bupt.edu.cn

tablishes a small network environment to obtain a more realistic network security dataset SDS-W. This paper conducts experiments of the situation element recognition on the open cybersecurity dataset and the SDS-W, respectively. Experiments show that PSO-TSA improves the accuracy of situation element recognition by an average of 5% to 7% while the time cost remains the same or even less.

Key words: network security situation awareness; situation element recognition; Particle Swarm Optimization; Simulated Annealing

随着信息技术的不断发展,当前的网络规模具有多节点、多分支、多网段、大流量等特点,网络安全问题越来越严峻.传统的监测方法和防护手段已经无法满足新的安全需求.网络安全态势感知(Network Security Situation Awareness,NSSA)能在复杂的网络环境中实时感知网络的安全风险,安全分析人员能结合网络安全环境,快速、准确地做出判断,将风险和损失降到最低[1].网络安全态势要素识别是NSSA的基础,也是直接影响NSSA性能的关键因素之一.

粒子群算法[2]结构简单、收敛速度快,符合NSSA对时效性的要求,是目前应用最广泛的态势要素识别算法.粒子群算法在搜索最优解时通过共享粒子之间的信息,使得粒子总是在向当前最优解更新.在算法开始阶段收敛速度很快,直到所有粒子状态相似时收敛速度减慢.这会使算法在寻找到局部最优解且收敛于该位置时,粒子难以从局部最优解中跳出,从而形成粒子"早熟",导致态势要素识别的准确率降低[3].

PSO算法容易陷入局部最优解,而模拟退火算法能够接受非更好的解.同时,PSO算法在初期具有极快的收敛速度,能够弥补SA算法在收敛速度上的缺陷.因此,针对态势要素提取质量与识别效率较低的问题,本文提出融合粒子群和模拟退火的态势要素识别模型PSO-TSA.利用模拟退火中的Metropolis准则,允许粒子接受一个非更好的解,以优化个体极值(p_{best})、全局极值(g_{best})以及PSO参数设置等过程,使得粒子摆脱局部最优解,提高态势要素识别质量与效率.本文的主要工作如下:

1)在p_{best}和g_{best}更新阶段,引入退火算法中的Metropolis准则,改变p_{best}和g_{best}的接受规则,允许在一定的概率下接受一个非更好的位置,增加粒子的选择性,提高态势要素提取质量.

2)在PSO参数优化阶段,利用退火算法中的Metropolis准则将PSO算法中参数的设置作为一个统一的优化问题,在评价粒子适应度的同时,也对参数优化过程进行评价,帮助粒子跳出局部极值.

3)本文提出了NSSA数据获取集成工具,并搭建了一个小型网络环境,以获取反映网络安全状态的真实数据集SDS-W.在开放网络安全数据集和SDS-W数据集上的实验表明,PSO-TSA模型保持了较优的识别精度和效率.

1 相关工作

态势要素识别技术的研究是随着网络安全态势感知概念的提出开始的.2003年,Matheus等人提出了基于Ontology的态势感知模型,结合各个模块提出了抽象实体的思想,对网络安全态势要素提取具有一定的指导意义[4].2007年,Jin等人针对军事战场中面临的态势要素提取情况,并结合周围环境等因素,提出一种基于概念的态势要素提取技术[5],但该方法提取的数据源单一且不能应对多源攻击的情况.2014年,刘等人提出了基于时空维度分析的网络安全态势预测方法[6],在时间维度上预测未来时段内的安全态势要素集,并在空间维度上分析各安全态势要素集对网络安全态势的影响.2016年,Kaufman等人考虑了网络环境的整体结构及不同的层级之间有不同的软硬件设施和相应的通信协议,提出采用自下而上、先局部再整体的形式对网络安全状态进行描述[7].2018年,Bazrafkan等人从国家战略决策的角度提出一个国家态势感知的概念[8],以层次化结构提高了态势识别的效率,降低了错误信息的干扰.2019年,Eckhart等人提出建立系统的虚拟副本,在并行环境下对网络态势进行全面的感知[9].同年,Debatty等人提出利用网络靶场更真实地模拟安全事件[10],为决策提供更准确的支持.

国内对网络安全态势要素提取相关研究起步较晚,前期在入侵检测方面所做的相关工作为网络安

全态势要素提取技术的研究奠定了基础. 2008 年,王等人提出了一种在神经网络的基础上结合进化策略的方法[11],以优化建立神经网络的参数,然后用神经网络来提取态势要素,极大地提高了分类准确度. 2010 年,赖等人为获取反映网络整体安全态势的信息,提出了将 D-S 证据理论用于多源报警数据聚类,计算不同数据之间的相似度,融合多种设备来降低系统误报率[12]. 为了更好更快地去除冗余特征,适应网络安全领域的需要,2015 年,司等人提出了一种基于本体论的网络安全态势要素知识获取方法[13],将多源异构数据进行分类提取,依据本体构建规则建立由领域本体、应用本体和原子本体组成的网络安全态势要素知识库本体模型. 2016 年,刘等人[14]提出一种基于融合的网络安全态势感知模型,将对安全事件威胁等级和威胁要素关系的推演和多源融合算法结合,克服了态势要素获取过程中需处理网络组件间复杂隶属关系的不足. 2017 年,戚等人提出了基于信息增益的贝叶斯态势要素提取方法[15],相较于朴素贝叶斯态势要素提取方法,该方法提高了分类效果,实现了对恶意攻击的检测. 2018 年,张等人为了评估网络安全现状,提出了一种基于分布式集群的安全态势感知系统[16],该方法在要素提取的准确性和时间上都有了明显的优化. 2019 年,徐等人针对云平台的安全态势,提出一个三层安全态势指标体系[17],从而识别出精确反应云平台态势的要素. 同年,段等人提出基于RSAR的随机森林网络安全态势要素提取,有效提高了分类精确度[18]. 2020 年,赵等人利用D-S证据理论处理多源数据[19],实现了对要素更精确的识别.

1995 年,Kennedy 和 Eberhart 提出了粒子群算法[20]. 粒子群算法属于群体智能算法,具有结构简单、鲁棒性强的特点,在解决组合优化问题时有很好的表现. 近几年,针对 PSO 算法的研究主要集中在 PSO 算法的优化和与其他算法的融合方面. 最早由 Shi 和 Eberhart 提出的惯性权重线性递减的方案,在算法初期能够快速找到最优解的范围. 随着迭代次数的增多,算法进行更加精确的搜索,最终得到最优解. 文献[21]将群体灭绝的现象引入 PSO 算法,促进粒子个体进化的持续性和群体选择的多样性. 文献[22]中,每个粒子根据其自身的适应度和最优粒子选择惯性因子,使得算法具有全局收敛性,能有效地缓解早熟收敛问题. 2018 年,胡等人提出利用粒子群算法优化模拟退火降温速度过慢的问题[23],并在实际场景中获得良好的效果. 2019 年,董等人利用模拟退火解决粒子群容易陷入局部极值的问题,设计了一种新的 USV 全局路径规划算法[24],该算法在考虑收敛速度的同时能够更为准确地找到全局极值.

2　融合粒子群和模拟退火的态势要素识别模型 PSO-TSA

本文将粒子群算法和模拟退火算法应用到态势要素识别的过程中,提出了一种融合粒子群和模拟退火的态势要素识别模型 PSO-TSA.PSO-TSA 利用模拟退火中的 Metropolis 准则克服 PSO 在以往的态势要素识别过程中识别准确率不高的问题,能较好地摆脱局部最优解,快速准确地找到态势要素 .PSO-TSA 模型的总体框架如图 1 所示,模型主要由两个部分组成:p_{best} 和 g_{best} 退火更新模块、PSO 参数退火优化模块.

2.1　Metropolis 准则

假设一个 D 维的目标搜索空间,有种群数目为 N 的粒子群,其中把第 i 个粒子表示为一个 D 维向量,记为:

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

第 i 个粒子的飞行速度也是一个 D 维向量,记为:

$$V_i = \left(v_{i1}, v_{i2}, \cdots, v_{iD}\right), i = 1, 2, \cdots, N \tag{2}$$

第 i 个粒子目前找到的最佳位置称为个体极值,记为:

$$p_{\text{best}} = \left(p_{i1}, p_{i2}, \cdots, p_{iD}\right), i = 1, 2, \cdots, N \tag{3}$$

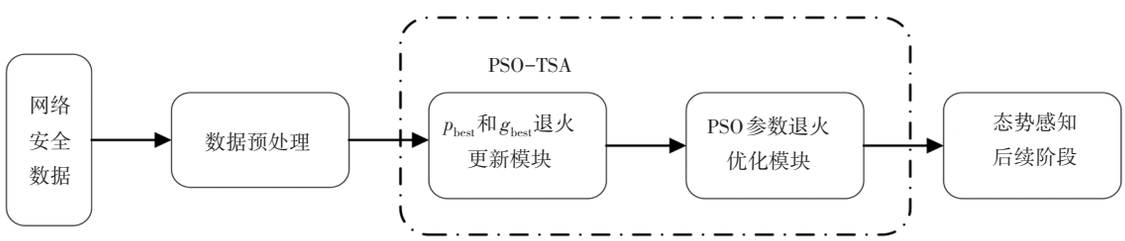

图 1　PSO-TSA 总体框架

Fig.1　Overall framework of PSO-TSA

整个粒子群目前找到的最佳位置为全局极值,记为:

$$g_{\text{best}} = \left(g_1, g_2, \cdots, g_D\right) \tag{4}$$

当粒子找到个体极值和群体极值后,粒子会更新自己的速度和位置,如公式(5)和公式(6)所示:

$$v_{id} = w \times v_{id} + c_1 \times r_1 \times \left(p_{id} - x_{id}\right) + c_2 \times r_2 \times \left(g_d - x_{id}\right) \tag{5}$$

$$x_{id} = x_{id} + v_{id} \tag{6}$$

式中:w 为惯性权重,表示在多大的程度上保留原有速度.w 越大则全局收敛能力越强;w 越小则局部收敛能力越强.c_1 和 c_2 为学习因子,r_1 和 r_2 为[0,1]内的均匀随机数.

Metropolis 准则是一种以概率接受新状态的采样法.给定一个初始状态 i 作为当前状态,记当前状态的能量为 E_i.然后通过一定的手段产生一个变化,使其进入一个新的状态 j,记新状态的能量为 E_j.若 $E_j < E_i$,则接受新状态;否则,考虑到热力学运动,这个新状态是否被接受要依赖一定的概率来判断.该过程服从正则分布,如公式(7)所示:

$$P\left\{E = E_i\right\} = \frac{1}{Z(T)}\exp\left(-\frac{E_i}{kT}\right) \tag{7}$$

式中:T 是绝对温度,k 是 Boltzmann 常数,$\exp\left(-\frac{E_i}{kT}\right)$ 称为 Boltzmann 因子,$Z(T)$ 如公式(8)所示:

$$Z(T) = \sum \exp\left(-\frac{E_i}{kT}\right) \tag{8}$$

物体处于状态 i 和状态 j 的概率应为相应 Boltzmann 因子的比值,如公式(9)所示:

$$r = \exp\left(-\frac{E_j - E_i}{kT}\right) \tag{9}$$

若 $r > 0$,则接受新状态 j,否则舍弃新状态 j.由此得到简化的 Metropolis 接受准则:若系统当前处在状态 i,由于某种变化进入状态 j.相应地,系统的能量也由 E_i 变为 E_j,那么系统接受这种状态改变的概率为:

$$p = \begin{cases} 1, E_j < E_i \\ \exp\left(-\dfrac{E_j - E_i}{T}\right), E_j \geqslant E_i \end{cases} \tag{10}$$

2.2 p_{best} 和 g_{best} 更新模块

p_{best} 和 g_{best} 更新模块在利用 Metropolis 准则更新个体粒子的最好位置 p_{best} 时,允许 P_{best} 在一定的概率下向一个非更好的位置更新.同样,在更新群体粒子的最好位置 g_{best} 时,g_{best} 也被允许在一定的概率下向一个非更好的位置更新,算法结束时输出 G_{best}.两次退火更新能增加粒子的选择性,防止 PSO 算法陷入局部最优解.p_{best} 和 g_{best} 更新模块算法步骤如下.

Step1:对算法进行初始化,包括最大迭代次数 $t_{\max}$、初始温度 T、降温系数 α、最低温度 $t_{\min}$、参数组合 $S\left(w, c_1, c_2\right)$、适应度函数 $f\left(x_i\right)$.

Step2:计算所有粒子各自的适应值 $f\left(x_i\right)$ 以及每个粒子目前的最好位置 p_{best} 的适应值 $f(p_{\text{best}})$.计算 $\Delta f = f\left(x_i\right) - f(p_{\text{best}})$,当 $\Delta f > 0$ 时,更新 $p_{\text{best}} = x$;当 $\Delta f < 0$ 时,引入 Metropolis 准则,得到 $p = \exp(-\Delta f/T)$.当 $p > 0$ 时,$p_{\text{best}} = x$.

Step3:设置两个变量 G_{best} 和 g_{best} 来记录群体粒子经历的最好位置.比较当前个体粒子最好位置的适应值 $f(x_{\text{best}})$ 和群体粒子最好位置的适应值 $f(G_{\text{best}})$.如果 $f(x_{\text{best}}) > f(G_{\text{best}})$,则 $G_{\text{best}} = g_{\text{best}} = x_{\text{best}}$.否则,计算 $\Delta f = f(x_{\text{best}}) - f(g_{\text{best}})$,当 $\Delta f > 0$ 时,更新 $g_{\text{best}} = x_{\text{best}}$:当 $\Delta f < 0$ 时,引入 Metropolis 准则,得到 $p = \exp(-\Delta f/T)$.当 $p > 0$ 时,$g_{\text{best}} = x_{\text{best}}$.

p_{best} 和 g_{best} 更新算法流程如图 2 所示.

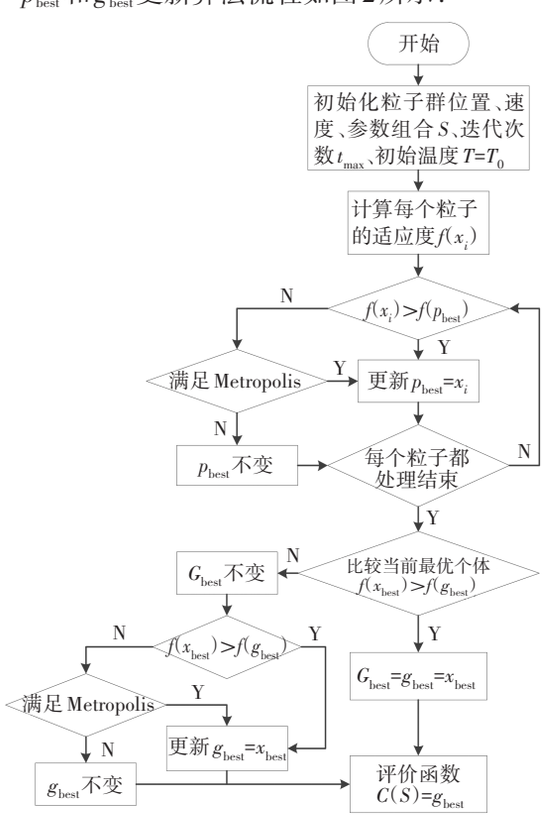

图 2 p_{best} 和 g_{best} 退火更新模块流程图

Fig.2 Flow chart of p_{best} and g_{best} annealing update module

2.3 PSO参数优化模块

PSO参数优化模块对PSO算法中的参数(惯性权重w,学习因子c_1、c_2)进行退火优化.在每次迭代中,PSO算法对粒子群的适应度和参数组合的优化评价值都可以用最优适应度函数来表示.由于Metropolis准则能接受非更好的参数组合,从而粒子可以更好地摆脱局部最优解.PSO参数优化模块算法步骤如下.

Step1:取评价函数$C(S)=g_{\text{best}}$,求解得到新的参数组合$S'(w',c_1',c_2')$,按照公式(5)、(6)和新的参数组合更新速度v_i和位置x_i,并计算适应度$f(x_i)$.

Step2:令$C(S')=\min[f(x_i),i=1,2,\cdots,m]$,其中$m$为粒子个数,$\Delta C=C(S)-C(S')$.当$\Delta C>0$时,接受$S'$,进行退火操作,并依据$S'$更新速度和位置;当$\Delta C<0$时,引入Metropolis准则,得到$p=\exp(-\Delta C/T)$.当$p>0$时,接受$S'$,进行退火操作,并依据$S'$更新速度和位置.否则拒绝$S'$的状态,$S$仍为当前状态,依据$S$更新粒子的速度和位置.

Step3:判断是否满足终止条件,若满足,则算法结束,输出最优值;否则,跳转到p_{best}和g_{best}更新模块的Step2.

PSO参数优化模块算法流程如图3所示.

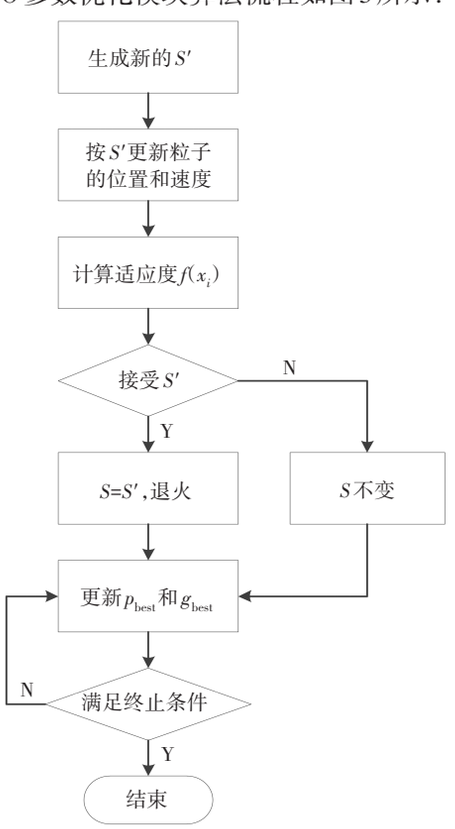

图3　PSO参数优化模块算法流程图

Fig.3　Flow chart of PSO parameter optimization module

2.4 PSO-TSA性能分析

为了验证PSO-TSA在寻找最优解时的能力、有效性和收敛性,本文选择了囊括单多峰函数的三个测试函数来分析算法的性能,分别是:

1)Rosenbrock单峰函数.

$$f(x)=\sum_{i=1}^{N-1}\left[100\left(x_{i+1}-x_i^2\right)^2+\left(x_i-1\right)^2\right],$$

$x\in(-10,10)$,当$(x_1,x_2,\cdots,x_N)=(1,1,\cdots,1)$时,有最小值$f(x)=0$.

2)Rastrigin多峰函数.

$$f(x)=\sum_{i=1}^{N}\left[x_i^2-10\cos\left(2\pi x_i\right)+10\right],$$

$x\in(-10,10)$,当$(x_1,x_2,\cdots,x_N)=(0,0,\cdots,0)$时,有最小值$f(x)=0$.

3)Griewank多峰函数.

$$f(x)=\frac{1}{4\,000}\sum_{i=1}^{N}x_i^2-\prod_{i=1}^{N}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1,$$

$x\in(-10,10)$,当$(x_1,x_2,\cdots,x_N)=(0,0,\cdots,0)$时,有最小值$f(x)=0$.

本文对标准粒子群算法(PSO)、线性惯性权重粒子群算法(LDPSO)以及本文提出的融合粒子群和模拟退火的PSO-TSA进行对比分析.在对比试验中,三个函数的维度设置为$N=10$.算法的参数设置:种群大小$m=40$,初始温度$T=1\,000$,最大迭代次数$t_{\max}=1\,500$,降温系数$\alpha=0.96$,惯性权重$w=0.6$,学习因子$c_1=c_2=2$.算法各自独立运行100次,实验结果如表1所示.

由表1可知,针对本文选取的三种测试函数,PSO-TSA的测试效果要优于其他两种传统的PSO算法.在算法速度上,对于单峰函数Rosenbrock,PSO-TSA相比于另外两种算法没有明显的优势;对于多峰函数Rastrigin和Griewank,PSO-TSA在收敛速度上的表现明显优于传统的PSO算法.在算法的收敛精度上,对于单峰函数Rosenbrock,PSO-TSA算法相较于另外两种算法有了明显的提高;对于多峰函数Rastrigin和Griewank,PSO-TSA算法在收敛速度提升的同时,精度也有了明显的提高.总的来说,PSO-TSA对于单峰函数,速度上没有明显的优势,但在收敛精度上有明显的优势;对于多峰函数,其优化后的收敛速度和精度都有很大的提升.PSO-TSA增加了

一定的算法复杂性,在维持原有收敛速度甚至提高速度的情况下,能够很好地摆脱局部最优值,提升寻找全局最优值的性能.

表1 三种算法在不同测试函数下的性能对比结果

Tab.1 Performance comparison results of the three algorithms under different test functions

测试函数	指标	PSO	LDPSO	PSO-TSA
Rosenbrock	平均迭代次数	1 495	1 499	1 423
	平均最优值	7.301 24	7.276 31	0.030 68
	平均时间	0.706 49	0.689 24	0.683 06
	收敛率	0.81	0.89	0.96
Rastrigin	平均迭代次数	1 492	1 287	894
	平均最优值	5.760 25	2.436 17	1.043 02
	平均时间	0.973 68	0.690 36	0.701 38
	收敛率	0.41	0.73	0.92
Griewank	平均迭代次数	924	1 061	437
	平均最优值	0.193 02	0.078 43	0.024 38
	平均时间	0.418 94	0.316 72	0.203 19
	收敛率	0.57	0.79	0.95

3 实验分析

3.1 实验拓扑

全面采集网络安全态势数据并充分考虑多方面信息,进而选择合适的态势数据字段组成态势指标,这是进行准确态势理解[25]的重要保证.为了数据获取的真实性,本文搭建了一个小型网络环境.实验环境拓扑如图4所示,包括七台主机,五个网络组件和两套NSSA数据获取工具,其中三台主机进行服务器模拟,包括了Web服务、FTP服务和database服务.

3.2 数据采集

网络安全态势感知旨在实现对网络安全态势的全面掌握,需要对反映网络安全状态的数据进行全面的获取.本文考虑了漏洞信息、攻击信息、流量信息三个方面,提出了NSSA数据获取集成工具,包括Nessus、Snort和Netflow.NSSA数据获取集成工具对网络的基本运行、正在面临的攻击和潜在的安全隐患进行全面的测评,为态势要素提取提供了强有力的支持.

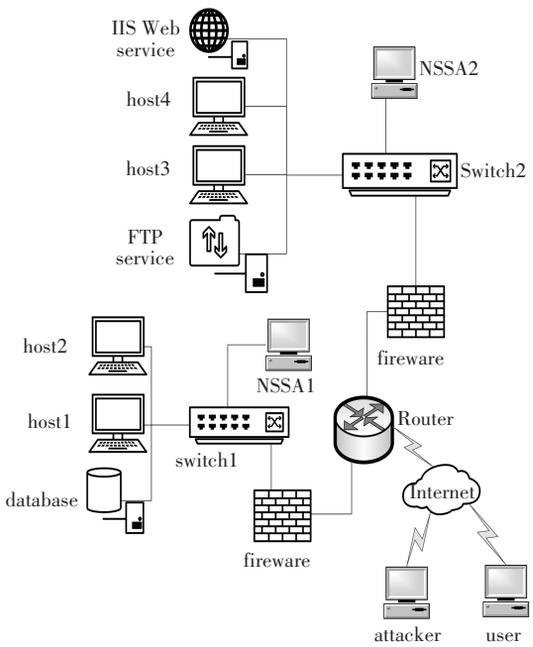

图4 实验搭建的小型网络环境拓扑图

Fig.4 Topology diagram of small network environment built by the experiment

NSSA数据获取工具对网络中节点的态势数据进行采集,剔除缺失数据,生成态势要素数据集SDS-W.SDS-W数据集包含了各类网络安全数据信息,共37个字段,数据格式如表2所示.

3.3 实验结果

本文为了全面分析PSO-TSA算法在态势要素识别上的性能,将实验分为两个部分.一部分实验采用现有的开放网络安全数据集,包括KDD99、NSL-KDD和UNSW-NB 15;另一部分实验采用本文模拟的实验环境所获取到的态势要素数据集SDS-W.本文选取五种在态势要素识别领域应用较多并且较为经典的算法与PSO-TSA算法进行比较,包括支持向量机(SVM)、分类决策树(CART)、线性惯性权重粒子群算法(LDPSO)、BP神经网络、循环神经网络(RNN).实验结果如表3和表4所示.

KDD99数据集包含了四种攻击类型和一种正常状态,训练集共494 021个,测试集共331 029个;NSL-KDD数据集包含四种攻击类型和一种正常状态,训练集共125 973个,测试集22 544个;UNSW-NB 15数据集为二分类集合,包含训练集175 341个和测试集82 332个.本文获取的SDS-W数据集为二分类集合,包含训练集148 957个和测试集51 043个.

表 2　SDS-W 数据集中的数据字段

Tab.2　Data fields in the SDS-W data set

字段名	说明	字段名	说明
Sreaddr	源IP地址	Dstddr	目的IP地址
Flow_sequence	信息流的序列计数器	dPkts	信息流中的数据包
dOctets	数据包第三层总字节数	Dstport	TCP/UDP目的端口号码
Port	IP协议	Tos	IP业务类型
Src_mask	源地址的前缀源码	Dst_mask	目的地址的前缀源码
Src_port	源IP地址端口号	Dst_port	目的IP地址端口号
Src_mac	源主机MAC地址	Dst_mac	目的主机的MAC地址
Src_OS	源主机操作系统和版本	Dst_OS	目的主机操作系统和版本
Creat_time	报警产生的时间	Ident	报警的标志符
Service_name	攻击利用的服务类型	Port_name	攻击采用的协议类型
Classification	攻击类型	Severity_S	危害等级
Starttime	扫描开始时间	Openports	端口开放数
High_count	高危漏洞数	Medium_count	中危漏洞数
Low_count	低危漏洞数	Severity_N	漏洞威胁等级

表 3　PSO-TSA 在开放网络安全数据集上的性能

Tab.3　Performance of PSO-TSA on open network security data sets

数据集	识别算法	训练准确率	训练时间/s	测试准确率	测试时间/s
KDD99	SVM	0.998 6	2 442.876	0.920 1	6.299
	CART	0.999 4	82.904	0.865 5	0.938
	LDPSO	0.999 6	65.231	0.904 9	0.893
	BP	0.999 6	173.482	0.988 1	1.451
	RNN	0.999 4	182.951	0.989 6	1.942
	PSO-TSA	0.999 6	71.705	0.984 8	0.901
NSL-KDD	SVM	0.996 3	7 630.762	0.761 3	0.657
	CART	0.992 8	27.743	0.717 2	0.031
	LDPSO	0.995 3	21.691	0.843 9	0.093
	BP	0.998 1	275.021	0.951 6	0.296
	RNN	0.998 9	309.966	0.954 2	0.301
	PSO-TSA	0.997 7	16.204	0.971 5	0.089
UNSW-NB 15	SVM	0.936 1	2 715.224	0.811 2	23.332
	CART	0.938 4	90.001	0.640 8	0.379
	LDPSO	0.954 9	50.801	0.831 5	0.221
	BP	0.964 4	161.273	0.916 4	0.294
	RNN	0.954 7	158.053	0.904 2	0.443
	PSO-TSA	0.949 2	36.195	0.931 4	0.097

表4 PSO-TSA在安全态势数据集SDS-W上的性能

Tab.4 The performance of PSO-TSA on the security situation data set SDS-W

识别算法	训练准确率	训练时间/s	测试准确率	测试时间/s
SVM	0.966 7	4 528.502	0.884 1	17.492
CART	0.959 1	118.861	0.795 8	2.173
LDPSO	0.989 7	46.986	0.904 6	1.021
BP	0.994 3	378.094	0.925 5	3.147 3
RNN	0.996 6	411.892	0.930 5	2.907 6
PSO-TSA	0.993 5	43.081	0.954 9	0.215

由表3可知,在开放网络安全数据集的实验中,相较于SVM、BP和RNN,PSO-TSA在KDD99数据集上的训练时间和测试时间上都体现了明显的优越性.从训练准确率和测试准确率来看,六种算法在KDD99数据集上的性能表现相差不大.在NSL-KDD数据集上,PSO-TSA在训练时间、测试时间和测试准确率上有了明显提升.在UNSW-NB 15数据集上,PSO-TSA在四项性能指标中都有明显的优势,尤其在大幅度缩小测试时间的基础上,对测试准确度也有了显著的改进.整体来看,PSO-TSA虽然在各个数据集上的识别准确率略逊于BP神经网络和RNN,但是在时间成本上具有明显的优势;与同类型的LDPSO相比较,在测试准确率有小幅提升的情况下,时间成本极大减少.

由表4可知,在本文搭建的小型网络环境所采集的网络安全态势数据集SDS-W上,PSO-TSA在各项性能指标上的表现比较突出.与SVM、CART、LDPSO相比,在时间成本和识别精确度上都有比较好的表现.与BP和RNN神经网络算法相比,虽然在准确率上有1%~2%的下降,但是其时间成本不到这两种算法的10%.

总的来说,PSO-TSA在开放网络安全数据集上的整体表现比较好,在个别数据集上全面优于实验中的其他五种算法.在本文搭建的模拟网络环境中,对于采集到的网络安全数据集SDS-W,PSO-TSA在识别精度保持一定的基础上,时间成本大幅降低.

4 结 语

本文提出了一种融合粒子群和模拟退火的网络安全态势要素识别模型PSO-TSA,用于对网络安全态势要素进行快速准确的识别.首先分析了粒子群算法和模拟退火中的Metropolis准则,讨论了两者结合的可能性.接着提出了以粒子群为基础,在位置更新和参数选择阶段三次引入Metropolis准则的全新融合算法.然后从测试函数的角度证明了PSO-TSA算法比其他PSO改进算法在避免陷入局部最优解的方面具有更优的效果.最后为了实验数据更加逼近实际网络环境,搭建了一个小型网络环境,用于获取全新的网络安全态势要素数据集SDS-W.从现有开放网络安全数据集和全新态势要素数据集SDS-W两方面,将PSO-TSA与几种常见的态势要素识别算法进行对比实验.实验表明,PSO-TSA算法在绝大多数的态势要素数据集中的性能表现都要优于其他识别算法,在维持甚至大幅度降低原有时间成本的基础上,显著提升了态势要素识别的精确度.

未来的工作将讨论在现有态势要素识别的基础上,如何将识别后的要素更快、更准确地进行态势理解,从而对网络安全态势有一个更直观有效的表达,为之后的决策、分析提供强有力的支持.

参考文献

[1] 刘效武,王慧强,赖积保,等. 基于多源异质融合的网络安全态势生成与评价[J]. 系统仿真学报,2010,22(6):1411-1415.
LIU X W,WANG H Q,LAI J B,*et al*. Network security situation generation and evaluation based on heterogeneous multi-sensor fusion[J]. Journal of System Simulation,2010,22(6):1411-1415.(In Chinese)

[2] 胡志刚,常健,周舟. 面向云环境中任务负载的粒子群优化调度策略[J]. 湖南大学学报(自然科学版),2019,46(8):117-123.
HU Z G,CHANG J,ZHOU Z. PSO scheduling strategy for task load in cloud computing[J]. Journal of Hunan University (Natural Sciences),2019,46(8):117-123.(In Chinese)

[3] 郭思源,刘海峰,李理,等. 基于混合粒子群算法的PSS4B参数优化研究[J]. 湖南大学学报(自然科学版),2018,45(4):112-121.
GUO S Y,LIU H F,LI L,*et al*. Research on parameter optimization of PSS4B based on hybrid particle swarm algorithm[J]. Journal of Hunan University (Natural Sciences),2018,45(4):112-121.(In Chinese)

[4] MATHEUS C J, KOKAR M M, BACLAWSKI K. A core ontology for situation awareness[C]//6th International Conference on Information Fusion. Cairns, Australia: IEEE Computer Society, 2003: 545-552.

[5] JIN W,SRIHARI R K, WU X. Mining concept associations for knowledge discovery through concept chain queries[J]. International Journal of Computer Applications in Technology, 2007, 29

(2/3/4): 243-246.

[6] 刘玉岭,冯登国,连一峰,等. 基于时空维度分析的网络安全态势预测方法[J]. 计算机研究与发展,2014,51(8):1681-1694.
LIU Y L, FENG D G, LIAN Y F, *et al*. Network situation prediction method based on spatial-time dimension analysis[J]. Journal of Computer Research and Development, 2014, 51(8):1681-1694. (In Chinese)

[7] KAUFMAN C, PERLMAN R J, SPECINER M. Network security: private communication in a public world[M]. Pearson Education India, 2016.

[8] BAZRAFKAN M H, GHARAEE H, ENAYATI A. National cyber situation awareness model[C]//2018 9th International Symposium on Telecommunications (IST). Tehran, Iran: IEEE, 2018: 216-220.

[9] ECKHART M, EKELHART A, WEIPPL E. Enhancing cyber situational awareness for cyber-physical systems through digital twins [C]//2019 24th IEEE International Conference on Emerging Technologies and Factory Automation. Zaragoza, Spain: IEEE, 2019: 1222-1225.

[10] DEBATTY T, MEES W. Building a cyber range for training CyberDefense situation awareness [C]//2019 International Conference on Military Communications and Information Systems (ICMCIS). Budva, Montenegro: IEEE, 2019: 1-6.

[11] WANG H Q, LIANG Y, YE H Z. An extraction method of situational factors for network security situational awareness[C]//2008 International Conference on Internet Computing in Science and Engineering. Harbin, China: IEEE, 2008: 317-320.

[12] 赖积保,王慧强,郑逢斌,等. 基于DSimC和EWDS的网络安全态势要素提取方法[J]. 计算机科学,2010,37(11):64-69.
LAI J B, WANG H Q, ZHENG F B, *et al*. Network security situation element extraction method based on DSimC and EWDS[J]. Computer Science, 2010, 37(11): 64-69. (In Chinese)

[13] 司成,张红旗,汪永伟,等. 基于本体的网络安全态势要素知识库模型研究[J]. 计算机科学,2015,42(5):173-177.
SI C, ZHANG H Q, WANG Y W, *et al*. Research on network security situational elements knowledge base model based on ontology[J]. Computer Science, 2015, 42(5): 173-177. (In Chinese)

[14] 刘效武,王慧强,吕宏武,等. 网络安全态势认知融合感控模型[J]. 软件学报,2016,27(8):2099-2114.
LIU X W, WANG H Q, LÜ H W, *et al*. Fusion-based cognitive awareness-control model for network security situation[J]. Journal of Software, 2016, 27(8): 2099-2114. (In Chinese)

[15] 戚犇,王梦迪. 基于信息增益的贝叶斯态势要素提取[J]. 信息网络安全,2017(9):54-57.
QI B, WANG M D. A method using information gain and naive Bayes to extract network situation information[J]. Netinfo Security, 2017(9): 54-57. (In Chinese)

[16] ZHANG P, HAN X, ZHANG D J, *et al*. A security situation awareness system based on wide & deep[C]//2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems. Nanjing, China: IEEE, 2018: 107-111.

[17] XU Y B, WANG W J. A security situational awareness method for cloud platform[C]//2019 IEEE 5th International Conference on Computer and Communications. Chengdu, China: IEEE, 2019: 1927-1934.

[18] 段詠程,王雨晴,李欣,等. 基于RSAR的随机森林网络安全态势要素提取[J]. 信息网络安全,2019(7):75-81.
DUAN Y C, WANG Y Q, LI X, *et al*. RSAR-based random forest network security situation factor extraction[J]. Netinfo Security, 2019(7): 75-81. (In Chinese)

[19] ZHAO Z W, ZHOU T T, WANG H. Quantitative evaluation model of network security situation based on D-S evidence theory [C]//2019 6th International Conference on Dependable Systems and Their Applications (DSA). Harbin, China: IEEE, 2020: 371-376.

[20] KENNEDY J, EBERHART R. Particle swarm optimization[C]// Proceedings of ICNN'95 - International Conference on Neural Networks. Perth, WA, Australia: IEEE, 1995: 1942-1948.

[21] XIE X F, ZHANG W J, YANG Z L. Hybrid particle swarm optimizer with mass extinction[C]//IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions. Chengdu, China: IEEE, 2002: 1170-1173.

[22] ZHU J R, ZHAO J B, LI X N. A new adaptive particle swarm optimization algorithm[C]//2008 International Workshop on Modelling, Simulation and Optimization. Hong Kong, China: IEEE, 2008: 456-458.

[23] HU X M, XU H J, XU J, *et al*. A novel workshop layout optimization algorithm based on SA-PSO[C]//2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference. Xi'an, China: IEEE, 2018: 2431-2435.

[24] DONG J, CHEN X, ZHANG J Q, *et al*. Global path planning algorithm for USV based on IPSO-SA[C]//2019 Chinese Control and Decision Conference (CCDC). Nanchang, China: IEEE, 2019: 2614-2619.

[25] 龚俭,臧小东,苏琪,等. 网络安全态势感知综述[J]. 软件学报,2017,28(4):1010-1026.
GONG J, ZANG X D, SU Q, *et al*. Survey of network security situation awareness[J]. Journal of Software, 2017, 28(4): 1010-1026. (In Chinese)