用于网络安全态势预测的粒子群与支持向量机算法研究

孙 卫 喜
(渭南师范学院网络安全与信息化学院 陕西 渭南 714099)

摘 要 网络安全已经越来越受到人们的重视,网络安全态势预测作为一种阻隔网络安全威胁的新兴手段受到了学者的广泛关注。针对威胁网络安全的特异性因素,提出一种改进的网络安全态势预测技术。介绍网络安全态势感知预测相关的背景;针对网络安全态势预测过程中存在的时变性与非线性等特征,在分析了支持向量机与改进粒子群算法的基础上,给出一种改进的 PSO-SVM 算法。通过相关仿真实验说明该方法的可行性与实用性。实验表明,使用该预测方法处理先前收集到的网络安全数据,明显提高了网络态势的预测精度,实现了对网络安全威胁的有效防御。

关键词 安全态势 支持向量机 粒子群算法 态势预测

中图分类号 TP301.6 **文献标识码** A DOI:10.3969/j.issn.1000-386x.2019.06.056

PSO AND SVM FOR NETWORK SECURITY SITUATION PREDICTION

Sun Weixi
(*College of Network Security and Information Technology, Weinan Normal University, Weinan 714099, Shaanxi, China*)

Abstract Network security has been paid more and more attention by people. The network security situation prediction has attracted extensive attention from scholars as an emerging means to block network security threats. This paper proposed an improved network security situation prediction technology for the specific factors that threatened network security. I introduced the background related to network security situational awareness prediction. Then, based on the characteristics of time-varying and nonlinearity in the network security situation prediction process, an improved PSO-SVM was presented based on the analysis of SVM and PSO. Furthermore, the relevant simulation experiments were given to illustrate the feasibility and practicability of the proposed method. Experiments show that using this prediction method to process the previously collected network security data significantly improves the prediction accuracy of the network situation and achieves effective defense against network security threats.

Keywords Security situation SVM PSO Situation prediction

0 引 言

现今网络已经成为国家战略部署、人民日常生活、社会经济发展的重要基础。人们在工作、学习、生活等方面分享网络带来极大便利的同时也为经常出现的网络安全问题感到困惑,而网络结构的复杂化、数据的综合化及协议的多样化等使管理人员感到茫然[1],特别是互联网环境高速进化中网络安全威胁及网络攻击手段多样化已超越了防范措施的推出速度[2]。面对网络攻击行为规模化、常态化发展的趋势,研究如何在网络攻击之前,利用有效的防护措施及时发现攻击行为并予以阻止就显得非常有意义。

传统的安全卫士、杀毒软件、防火墙的网络安全防护往往是单一的,难以主动防御网络威胁。美国的 Time Bass[3-4]针对网络安全的复杂性于 1999 年提出了网络安全态势感知的概念,网络安全态势感知 NSSA (Network Security Situation Awareness)能对影响网络安

收稿日期:2018-09-04。陕西省自然科学基础研究计划资助项目(2017JM6110);渭南师范学院自然科学类研究项目(18YKS13)。
孙卫喜,高工,主研领域:网络安全,网络应用。

全的各种因素信息进行解析、收集、综合处理,并建立数学模型,给出评估网络安全的方法,进而对网络安全进行预测。NSSA 包含对网络态势要素获取、态势信息融合量化及对未来态势预测等方面的内容。传统的网络安全技术主要存在,检测的数据流数据单一、速度慢、更新周期繁琐等问题。与传统的网络安全技术相比,NSSA 具有实时性、多样性、整体性等优点。其更注重网络安全细节的特征[5]。NSSA 作为网络安全的新技术,以网络安全发展状况为关注点,能从整体上感知网络的安全状况。它的数据支持来源于硬件设备与软件系统,更加注重统筹分析影响网络安全的各种因素,全面实时、多角度反映网络当前时刻的安全状况。采用更加科学合理的网络安全评估手段和方法对网络安全状况实时监测,及时发现并处理发现的网络安全问题,NSSA 是网络管理的发展方向。网络安全态势感知系统以网络安全态势要素为基础,网络安全态势要素以数据精炼、对象精练、态势精练三次抽象来获取,从网络系统安全中获取高质量的网络态势要素是实现网络安全主动防御的前提[6]。

网络安全的重要性无论是国家或单位都有充分的认识,纷纷构建其网络安全态势感知系统,该系统不仅需要对网络实施漏洞检查、入侵检测、防火墙等基础防御,更需要掌握全局网络系统的安全运行数据,为管理者提供整个网络变化的决策依据。

在实际操作中,NSSA 能通过实时监测的手段了解大规模复杂网络环境中影响网络安全状态的网络、用户、网络设备运行情况。在技术上,NSSA 通常通过对当前和早期检测信息表现出的网络安全状态进行分析处理,预测后续的网络安全态势[7],从而使网络管理者能及时利用可视化的网络安全预测系统,对发现的网络安全弱点、预测到的威胁,制定出相应的措施主动进行防御。本文研究的重点为网络安全态势感知中有关态势预测的部分。

1　网络安全态势感知

SA(Situation Awareness),态势感知起源于军事领域,用于对复杂结构、影响因素众多、大范围事件的整体理解及快速决策处理[8]。Endsley[9]以“在一定的范围内,感知和理解环境状况,并对后续发展趋势进行推算”最先确定了态势感知,并将态势获取、态势理解、态势预测界定为态势感知的三个方面。Loke[10]提出了“网络态势感知是网络环境中所有信息按照逻辑约束关系进行组合的结果”,Lacey 等[11]通过系统研究网络态势感知给出了网络空间态势感知框架,虽然网络态势感知目前还没有权威的定义,但网络态势感知作用很明确,即:通过感知系统观测整个网络的安全情况,依据观测到的数据对网络安全事件及时做出判断,并以可视化方式提供给管理者决策[12]。

用于进行网络安全态势感知的三大技术包括数据挖掘、态势评估以及态势预测。其中,数据挖掘中力求准确、快速、全面地找到网络威胁事件;态势评估要求更为有效客观地评价网络安全态势;态势预测则强调预测网络安全的准确性,以便网络管理者依据预测结果采取相应的措施,保护网络安全。本文的研究重点是在分析网络安全态势预测中的相关技术后,对其进行改进,并使改进后的态势预测系统精度得以提高。

国外网络安全态势感知系统架构的建立主要采用的是集成化思想,卡内基梅隆大学 SEI 2005 年在网络态势感知系统中集成了 Netflow 工具,能对潜在的、恶意的网络攻击行为进行识别与响应并做出相应的防御。King 等[13]在分类属性网络中运用深度包检测技术提供了深刻全面的态势感知结果,Varun 等在基于事件学习理论的基础上构建了分析网络攻击态势的网络攻击检测模型,Bode 等[14]构建了网络态势风险管理的贝叶斯网络模型,Friedberg 等[15]给出了网络异常行为的态势感知基于事件自动关联的事件检测 AECID。CID(Cooperative Infrastructure Defense)中 Fink 等[16]提出了分层协同管理的思路,并在 CID 之上,提出了数字蚂蚁的思路[17]。Szwed 等[18]对网络依赖关系中重要资产采用模糊认知图的方式获取并评估其危害程度,该方法存在数据来源单一、主观性、误差较大的问题。Liu 等[19]基于隐马尔可夫模型与报警观测序列,以 Viterbi 算法推导最大可能状态转移序列,实现攻击意图识别,但缺乏全网风险值量化。Ghasemigol 等[20]针对攻击发生概率的不确定性,提出了一种综合预测算法,提高了预测精度。Wu 等[21]基于大数据分析提出了电网系统安全态势感知机制,将博弈论和模糊聚类相结合,降低错误率提高预测效率。

国内研究网络安全态势感知主要包括:网络安全数据的全面获得,数据融合方法及数据之间的关联性分析,网络安全态势指标体系的建立以及网络安全态势评估。陈秀真等[22]利用网络运行情况与告警信息数据,对网络结构及主机和服务发挥的作用进行系统分析,提取影响网络态势的多个因素,给出了网络安全态势计算方法和层次化的量化评估模型,但其缺乏对网络攻击间的联系及整体性分析。李方伟等[23]给出的一种径向基函数神经网络的网络安全态势预测模型,虽然对网络安全态势预测精度有所提高,但有预测

结果不稳定与过拟合的问题存在。韦勇等[24]依据层次化的思想,对节点上的安全要素利用 D-S 证据理论做了融合,再按照节点、子网、全网层融合,最后获得网络态势值,从而在信息融合的基础上建立网络安全态势评估模型。由于其有效处理了多源安全事件,使多源信息间的互补性得到了充分利用,从而使态势感知的准确性进一步提高,也使得在网络态势感知中多源信息融合优于单源的特征得到验证。周新卫等[25]基于安全态势值与多节点网络安全态势重要影响因子值,采用构建的模型对多节点网络安全态势进行预测,实现对突变态势的有效预测,提高网络安全态势的预测精度。刘效武等[26]以数据信息从融合异质多传感器获得,再用支持向量机结合特征约简算法生成网络安全态势值,最后依据评价指标评价量化态势感知。贾焰等[27]建立态势感知模型时采用关联分析与集中处理所收集到的网络安全态势数据,再对网络安全态势用建立的指标体系进行预测,该模型可用于较大规模的网络环境。张丹等[28]利用自律反馈特性获取网络安全态势数据,建立态势评估模型时用层次分析法(AHP),再用神经网络的改进方法对态势进行预测。甘文道等[29]使用网络径向基函数神经网络更利于神经网络的结构与参数控制,使得用网络安全态势图反映网络安全更直观。黄亮亮[30]在网络安全态势预测中用 PSO(Particle Swarm Optimization)优化 RBF(Radial Basis Function)网络的方法,给出 PSO-RBF 网络模型,通过对历史数据分析并映射出未来的网络态势值。胡冠宇等[31]将云群的高维差分进化算法应用到预测网络安全态势中补充了算法的多样性及搜索精度。琚安康等[32]将现有大数据处理技术与安全事件管理需求相结合,构建了集数据存储、实时关联检测、收集整理、离线分析发现、态势呈现、威胁预警等有效地实施了全流程网络安全态势预测。

经过专家学者对网络安全态势预测大量的分析研究及验证,发现网络式非线性的状态特征使得时间序列应用于网络安全态势预测中存在许多问题。而传统的统计学方法表示的是线性特征,因此需要借助于人工智能算法如支持向量机、马尔可夫、神经网络等提高对网络安全态势的预测精准,但支持向量机参数选定的盲目性、神经网络参数的难以确定和马尔可夫算法公式推导与建模过程的繁琐也是网络安全态势预测面临新的问题[33-35]。

网络安全态势感知常用的方法包括基于神经网络的方法、支持向量机、模糊逻辑、知识推理、基于贝叶斯网络、深度学习的方法等。神经网络利用大量的实验来进行神经网络模型的建立,存在计算量大、选择基函数较为困难的问题。知识推理方法克服数学模型难以处理的情况,避免客观性受主观因素的影响,模拟人类思维过程。贝叶斯网络用有向图表示,图中每一个节点表示一种变迁状态,可从状态属性图最小割、重要性程度、可信任性等分析网络系统。网络安全态势感知方法按照其原理可分为:基于模式识别的态势研究方法、基于知识推理的态势研究方法及基于数学模型的态势研究方法。

已有文献对网络安全态势预测提供了理论基础以及思路上的借鉴。目前网络安全态势预测的技术尚不成熟,用于网络安全态势预测的技术一般为一些智能优化算法,然而这些算法较少考虑网络安全态势感知的实际情况。因此,造成了预测准确率低,时变性和非线性等因素未被考虑等问题。基于此,本文给出了一种将改进粒子群算法与支持向量机相结合的新算法来提高网络安全态势预测的精度。

2 网络安全态势预测模型的构建

2.1 网络安全态势预测

网络安全态势的预测以发生网络安全事件的数量、频率、网络受威胁程度等因素经过处理而获取反映网络态势的数据为根据。目前关于网络安全态势预测思路与框架尚未完全形成,发展较快的网络安全态势预测方法主要有:灰色预测、神经网络、时间序列预测法及支持向量机预测。

本文对上述网络安全态势预测的方法进行了分析比较[36-49],并给出了相关方法的优缺点以及适用范围,具体如表 1 所示。

表 1 网络安全态势预测方法比较

算法名称	优点	缺点	适用范围
灰色理论	参数设置不需人工干预、速度快、预测方法相对简单、便于计算	由于其忽略原始序列自身周期性、随机性的特性,使预测网络安全态势的计算精度不高、缺乏系统性	对灾难、人口、环境及信息量少的情形预测比较有效
神经网络	泛化能力、映射能力、推理能力强。可进行自组织学习、容易实现、速度快。能够以任意精度逼近任何非线性关系,数据来源多样、客观性较强	易受人为调整隐含层权值与个数的影响,出现随机性不收敛或收敛慢的问题。基于经验风险最小化原则,泛化能力有限,容易得到局部极小值,训练时间过长、评估结果的精度不足	适应于多变量、非线性等预测问题
时间序列预测	考虑了时间因素	未考虑到事态发展的因果关系	周期性往复的事件

续表 1

算法名称	优点	缺点	适用范围
支持向量机	自由度高、快速收敛、有较强的数学理论支持、适应性好、构造简便、基于结构风险最小化原则泛化能力强	参数的优化上需要结合其他算法、自由参数多支持向量数目易受样本量影响	解决非线性、不确定预测、高维数、局部最优问题

依据网络安全态势预测相关文献及上述研究分析,得出网络安全态势预测具有以下特征:

(1) 数据呈现非线性,高维度等特征。

(2) 预测结果表现出不确定性,无周期性。

鉴于此,本文选择支持向量机作为网络安全态势预测的基本模型。

然而传统的支持向量机方法存在参数确定困难以及结果可能为局部最优等问题。针对传统支持向量机存在的问题,本文引入了一种改进的粒子群算法。改进的粒子群算法与传统粒子群算法相比,使用了混沌优化函数,使得改进后的算法更加偏向全局最优,从而解决了参数确定及局部最优的问题。

综上所述,本文是在对目前网络安全态势预测方法大量的分析研究后,给出一种支持向量机与改进粒子群优化算法相结合的网络安全态势预测方法,即在 SVM 优点特征的基础上引入改进的粒子群优化算法,通过用无体积无质量的粒子作为个体且规定各粒子的行为规则,使用个体之间的协作寻优在表现出复杂特性的整个粒子群中寻找最优解,进而优化支持向量机的三个参数。该安全态势预测方法还克服了使用线性方法评估网络安全态势带来的预测精度低、描述网络目前状态与未来状态关系困难等问题,更适应网络安全态势变化时变性、非线性等特点。

2.2　改进的粒子群与支持向量机算法

本文将按照由简单到复杂的顺序行文。首先给出一般的支持向量机算法,然后,引入改进的粒子群算法确定支持向量机所需参数,最后,给出综合算法的计算步骤。

2.2.1　支持向量机算法

为解决复杂的模式识别问题,Vapnik 提出了支持向量机,取得了较大进展[50],并在统计学习理论的基础上给出了 SVM 分类器,较好地解决了线性不可分的问题[51]。后续的学者们研究出基于二叉树的多分类方法[52]、序列最小优化训练算法(SMO)[53]、多分类理论、决策导向非循环图法(DDAG)[54]、1-a. r 方法[55],近年来学者们又研究出 Class-SVM、v-SVM、C-SVM 等算法,进一步完善支持向量机的理论体系。张翔等[56]在态势评估指标的时间序列预测中采用支持向量回归预测方法。王庚等[57]采用遗传算法的染色体编码优化支持向量机参数。李洁等[58]给出的通过对态势样本集进行归一化处理和相空间重构,利用相关向量机最优的超参数与和声搜索算法搜索,采用 wilcoxon 符号秩检验验证模型预测性能之间的差异性,提升了网络安全态势预测模型的预测精度和速度。

支持向量机 SVM 训练样本通过预设函数的支持向量机训练,函数的确定是在用不断拟合方法给出重要参数的基础上获得的。SVM 泛化能力强,对复杂的非线性数据与小样本数据的建模识别能力很好。

本文所给样本集为:

$$\{\{X_1,Y_1\},\{X_2,Y_2\},\cdots,\{X_n,Y_n\}\}\quad X_i\in R^n$$

观测样本值 $Y_i\in R^n$,设回归模型为:

$$f(x)=\boldsymbol{\omega}^{\mathrm{T}}x+b \tag{1}$$

式中:$\boldsymbol{\omega}$ 为支持向量机法向量,b 为偏移量。

实现合理拟合样本集需要用到损失函数 ε,$|y_i-f(x_i)|=\max\{0,|y_i-f(x_i)|-\varepsilon\}$ 观测值与 $f(x_i)$ 回归预测值间误差的相对值不能大于 ε,因而:

$$\begin{cases}y_i-\boldsymbol{\omega}\cdot x_i-b\leqslant\varepsilon\\ \boldsymbol{\omega}\cdot x_i+b-y_i\leqslant\varepsilon\\ i=1,2,\cdots,n\end{cases} \tag{2}$$

ε 损失函数稳定性好,在不知道 SVM 训练样本分布特征的情况下误差不能直接计算。$f(x)$ 按结构风险最小化 $\frac{1}{2}\|\boldsymbol{\omega}\|^2$ 应该最小,给出数值大于或等于零的松弛因子 $\xi_i\geqslant 0$ 和 $\xi_i^*\geqslant 0,i=1,2,\cdots,n$,则 SVM 的优化目标问题为:

$$\min\left(\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}+C\sum_{i=1}^{n}(\xi+\xi_i^*)\right) \tag{3}$$

$$\text{s. t.}\begin{cases}y_i-\boldsymbol{\omega}\cdot x_i-b\leqslant\varepsilon+\xi_i\\ \boldsymbol{\omega}\cdot x_i+b-y_i\leqslant\varepsilon+\xi_i^*\\ \xi_i,\xi_i^*\geqslant 0\end{cases} \tag{4}$$

惩罚因子 $C>0$,利用拉格朗日乘子求解,即:

$$\begin{aligned}L(\boldsymbol{\omega},\xi_i,\xi_i^*)=&\frac{1}{2}\|\boldsymbol{\omega}\|^2+c\sum_{i=1}^{n}(\xi_i+\xi_i^*)-\\ &\sum_{i=1}^{n}\alpha_i(\varepsilon+\xi_i-y_i+(\boldsymbol{\omega}\cdot x)+b)-\\ &\sum_{i=1}^{n}\alpha_i(\varepsilon+\xi_i^*-y_i+(\boldsymbol{\omega}\cdot x_i)+b)-\\ &\sum_{i=1}^{n}\eta_i\xi_i+\eta_i^*\xi_i^*\end{aligned} \tag{5}$$

分别对该函数的 $\boldsymbol{\omega}$、b、ξ_i、ξ_i^*,求偏导则有:

$$\begin{cases}\dfrac{\partial L}{\partial \omega}=\sum_{i=1}^{n}\omega-\sum_{i=1}^{n}(a_{i-\eta}a_i^*)x=0\\ \dfrac{\partial L}{\partial b}=\sum_{i=1}^{n}(a_i-a_i^*)x_i=0\\ \dfrac{\partial L}{\partial \xi_i}=c-a_{i-\eta}=0\\ \dfrac{\partial L}{\partial \xi_i^*}=c-a_i^*-\eta_i^*=0\end{cases}\tag{6}$$

并满足：

$$\begin{cases}\alpha_i(\varepsilon+\xi_i-y_i+\omega\cdot x+b)=0\\ \alpha_i^*(\varepsilon+\xi_i^*-y_i+\omega\cdot x+b)=0\\ (C-a_i)\xi_i=0\\ (C-a_i^*)\xi_i^*=0\end{cases}\tag{7}$$

把式(6)代入式(4)获取优化目标函数，给出$K(x,x_i)$核函数来替换点积运算，则：

$$\min\frac{1}{2}\sum_{i=1}^{n}y_i(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)K(x_i,x_j)-\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)+\sum_{i=1}^{n}\varepsilon(\alpha_{i-\eta}\alpha_i^*)\tag{8}$$

$$\text{s.t.}\begin{cases}\sum_{i=1}^{n}(\alpha_{i-\eta}\alpha_i^*)=0\\ \alpha_{i-\eta},\alpha_i^*\in[0,C]\quad i=1,2,\cdots,n\end{cases}\tag{9}$$

于是，问题就变为二次规划问题，该问题的解：

$$f(x)=\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)K(x_i,x)+b\quad\alpha_i,\alpha_i^*=0\tag{10}$$

鉴于在支持向量机核函数中使用高斯核函数较好，对高斯核函数做如下设定：

$$f(x,x_i)=\exp\left(\frac{|x-x_i|^2}{\sigma^2}\right)\tag{11}$$

如下表达式是支持向量机预测模型：

$$f(x)=\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)\exp\left(\frac{|x-x_i|^2}{\sigma^2}\right)+b\tag{12}$$

高斯核函数宽度用σ表示。

2.2.2 改进的粒子群优化算法

使用SVM算法能从庞杂的网络安全因素中找出规律，足以说明其对网络安全态势预测的有效性，而传统确定SVM参数的方法如网络搜索法、穷举法及经验法存在耗时长、难以找到最优参数、模型的预测精度较低等问题，分析前期对SVM用于网络安全态势预测的研究，发现SVM参数的优化问题是决定预测精度的关键。支持向量机的主要参数是：① 核函数的宽度σ，非线性问题最优解的复杂度用σ确定，σ的取值关系支持向量机的泛化能力；② 惩罚因子C，是过学习还是欠学习由C的取值过大或过小决定；③ 不敏感损失函数ε，支持向量数目和计算复杂度由ε确定，其表示训练时的误差期望。支持向量机主要参数的选取决定其预测精度，本文采用粒子群结合混沌优化两种算法对SVM的三个参数进行了优化。

粒子群算法是基于群体智能的优化理论所抽象出算法，其包括三方面内容：一是用一个具备初始位置和速度的粒子表示优化问题的解；二是按其最优位置以及全局最优位置动态调整粒子飞行速度和当前所处位置，搜寻粒子最好的目标获得最优解；三是用适应度函数衡量解的优劣程度。

粒子群优化算法是一种全局优化进化算法，有着进化初期需要调整的参数少、容易实现、概念简单、快速收敛等特征。但在用PSO优化SVM的三个主要参数时发现：粒子群中当单个粒子搜索到某个局部最优解时会影响到其周围的其他寻优粒子，导致它们快速靠近该粒子，这样就会出现局部最优解及粒子早熟等问题[59]。针对PSO存在局部最优解及粒子早熟的问题，本文引入了混沌优化算法对粒子群优化算法做了改进。混沌优化算法具有全局性的优点[60]，其可以按某种规则一次性搜索一定范围内的所有情况，当粒子群出现部分收敛后，再按照粒子群变化的适应情况对粒子群最优值及情况差的粒子进行混沌变异，使粒子群优化算法避开部分最优的能力得到进一步的提高。

采用改进的PSO算法优化支持向量机中参数的思路为：将支持向量机中的三个参数作为一个组合优化(σ,C,ε)，计算利用该组合得到的模拟值$f(x)$与实际值y的标准差(MSE)，并以此为目标建立多目标规划模型，设置初始化参数并进行迭代，最终得到目标最小值。在迭代过程中，需要注意每个粒子的适应度值。若适应度值大于所给阈值，则使用一般PSO算法进行更新，若适应度值小于所给阈值，则使用混沌优化算法进行处理，再对粒子进行更新。

当模型取最小值时，得到的组合为SVM的最优参数。值得指出的是，需要预先给出这三个参数的取值范围，并作为多目标规划的约束。

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(y-f(x))^2$$

式中：y为实际值，$f(x)$为利用SVM得到的模拟值。

$$\min g(z_1,z_2,\cdots,z_i)=\min MSE$$
$$\text{s.t. }a_i\leqslant z_i\leqslant b_i\quad j=1,2,3\tag{13}$$

式中：z_1、z_2、z_3分别对应SVM中需要得到的三个参数σ、C、ε，$[a_i,b_i]$分别为其对应的上下界。

使用PSO算法进行迭代的公式为：

$$V_i^{k+1}=\mu V_i^k+d_1r_1(P_{\text{ibest}}^k-X_i^k)+d_2r_2(P_{\text{ibest}}^k-X_i^K)\tag{14}$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \quad (15)$$

式中:μ 为惯性权重系数,d_1、d_2 为加速因子,且均为非负常数。r_1 和 r_2 为[0,1]间的随机数。P_{ibest}^k为个体极值。

下面给出用使用改进的 PSO 算法计算支持向量机参数的步骤:

Step 1　给定初始值。具体包括:迭代次数,随机生成的初始粒子速度 v_0 和位置 $X_i^k = (x_{i1}^k, x_{i2}^k, x_{iM}^k)$。$k=0$,搜索标志 $r=0$,$t_i^k = t_i(0)$,$t_i^* = t_i(0)$,$a_i^r = a_i$,$b_i^r = b_i$,g^* 为一个初始化较大的正数。$t_{k+1} = \mu t_k(1-t_k)$。

Step 2　利用式(12)－式(13)计算所有粒子的初始个体极值以及全局初始极值。

Step 3　迭代和更新。将 t_i^k 映射到区间上成为 z_i^k,$z_i^k = a_i^r + (b_i^r - a_i^r)t_i^k$。利用式(12)－式(13)计算该阶段每个粒子的适应度值 $g(z_i^k)$。这里给出一个适应度变化率阈值 0.8。则分情况讨论:

(1) 粒子中适应值比率的粒子$\left(\dfrac{g^* - g(z_i^k)}{g^*} \geqslant 0.8\right)$,可利用式(14)－式(15)进行迭代,得到下一代的粒子位置和速度。

(2) 对于适应值比率低的粒子$\left(\dfrac{g^* - g(z_i^k)}{g^*} < 0.2\right)$,需要进行混沌变异。

令 $e_r^{(k)} = \dfrac{x^{(k)} - x_{\min}}{x_{\max} - x_{\min}}$,其中 $x_{\max}$ 和 $x_{\min}$ 表示 $x^{(k)}$ 位置向量的上界和下界。将其进行变异后得到:

$$x^{(k+1)} = x_{\min} + e_r^{(k)}(x_{\max} - x_{\min})$$

Step 4　进行混沌优化搜索。若 $g^* > g(z_i^k)$,则令 $g^* = g(z_i^k)$,$t_i^* = t_i^k$,否则,令 $k = k+1$,$t_i^k = \mu t_i^k(1-t_i^k)$,直到 g^* 保持不变,或达到最大迭代次数。转向 Step 5。否则转向 Step 3。

Step 5　输出最优解向量。得到支持向量机所对应的函数表达式 $f(x)$。

本文充分综合上述两种算法的优点计算 SVM 的三个参数。用混沌优化算法在参数选取时能使用普遍的参数选取法,不用考虑模型的变量维数和复杂度的特性,再利用混沌理论的规律性、遍历性、随机性等特点有效地解决了用 PSO 算法优化时出现的局部最优解及粒子早熟的问题,也就是用混沌变异算子对粒子群优化算法进行必要的改进。在粒子群进化中确定粒子是否早熟依据粒子群适应度最优变化情况,若变化不大于确定值时,则对粒子群中优胜粒子的位置与速度进行更换,再用混沌变量映射非优胜粒子,然后把替换了的优胜粒子与使用混沌优化后的非优粒子组成新种群。使用混沌优化法对此时全局最优值进行扰动,以便增加寻找全局最优解的几率,使得粒子群经过本操作后避开出现局部最优点的问题。粒子的速度与位置经混沌变量随机性初始化后,种群的遍历性及多样性得到进一步的提高。

2.3　网络安全态势预测步骤

预测步骤如图 1 所示。

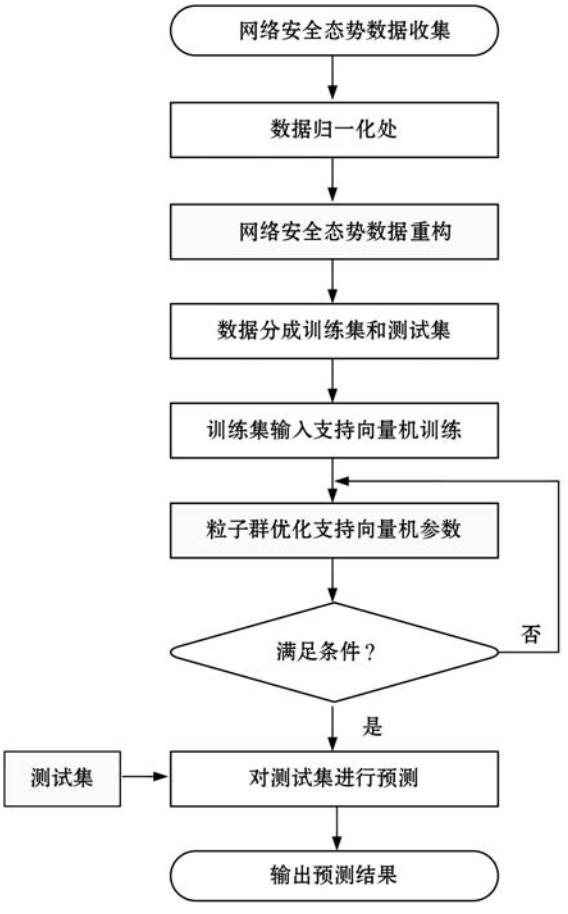

图 1　网络安全态势预测

(1) 收集、整理网络安全态势数据,量化处理网络安全监测数据。

(2) 数据归一化处理,影响网络安全态势的因素众多,有时收集到的数据差异明显,而支持向量机数据预测敏感区在(0,1)之间,故需要在(0,1)之间归类原始的网络安全态势数据。

(3) 在前两步的基础上通过确定嵌入维数和时间延迟将一维的网络安全态势数据转换为多维样本态势数据。

(4) 把获得的样本数据分为测试集与训练集两部分,把训练集数据输入 SVM 学习。

(5) SVM 主要参数的优化采用改进的粒子群优化算法,实现用最优参数建立预测模型。

(6) 对测试集使用建立的预测模型进行预测,完成反归一化预测结果等处理,再依据得到的处理数据预测网络安全态势。

3 实验分析

3.1 网络安全态势数据的选取

选取某公司2017 年3 月1 日 -4 月29 日和5 月1 日 -6 月29 日的安全测试数据,每天取样 4 回,安全测试数据按两月为一批,通过计算后每批各获得 240 个态势值,以相同的过程分别对两批数据进行实验,通过 MATLAB 2016b 进行实验。

3.2 实现预测网络安全态势的模型

累加所得的各组态势值,以获得新数据样本,实行归一化处理新数据样本。把 NSSA 时间延迟设定为 1,用试凑法得到的嵌入维数为 6,这样 SVM 就有了 1 个输出变量和 5 个输入变量,最后通过嵌入维数与延迟时间对获得的数据进行重构,生成 SVM 的测试集与训练集样本,再将生成的训练集样本数据输入到预测模型中学习。预测模型为改进的粒子群优化后的 SVM。

3.3 预测网络安全态势

对本文所给网络安全态势预测模型通用性与有效性的检验,采用上述获得的两组重构数据,以未改进的 PSO-SVM 模型与改进后的模型分别进行预测,然后比较两种模型所得的预测结果。具体操作方法采用:① 先将两组重构数据的前 200 个点作为训练样本,用于两种方法的训练及模型的构建;两组数据的后 40 个点作为测试样本,用于将两种模型的预测结果与实际值进行比较。② 分别将两组重构数据的训练样本输入 SVM 进行学习,SVM 的三个参数用改进的粒子群优化算法进行优化,获得用第一批数据时 $\sigma = 5$、$\varepsilon = 0.001$、$C = 98$;第二批数据时 $\sigma = 5$、$\varepsilon = 0.001$、$C = 76.12$。③ 用两组数据所获得的三个参数再分别将两组重构数据输入到支持向量机进行学习、训练得到新模型的预测结果。④ 图 2 为比对未经处理的原始值、本文所给方法得到的预测结果及未改进 PSO-SVM 得到的预测结果,图 3 为 PSO-SVM 改进前后的误差比较。

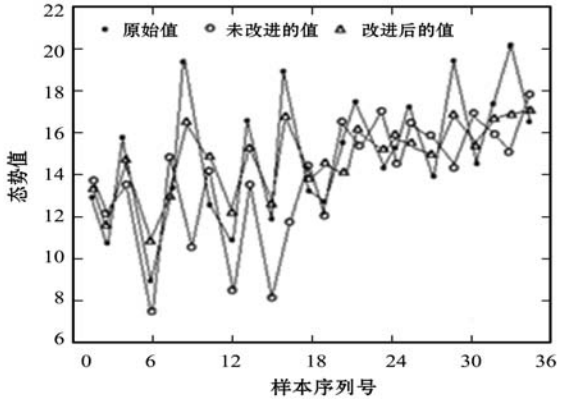

图 2 三种数据比较图

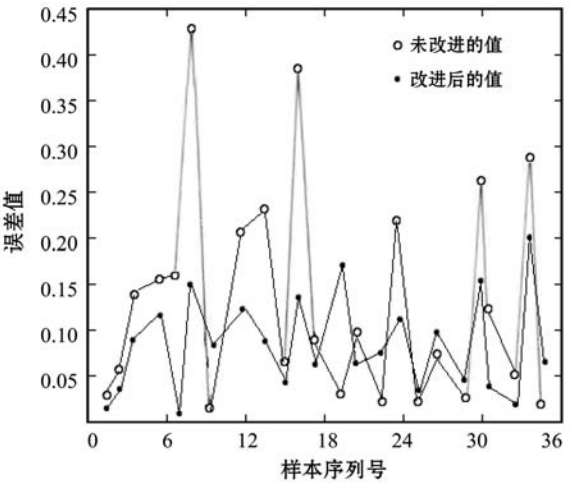

图 3 两种方法误差比较

结果表明,采用本文给出的网络安全态势预测方法,预测未来的网络安全状态精度高、误差小。

4 结 语

本文给出的采用支持向量机与改进粒子群优化算法相结合的网络安全态势预测方法,是基于实际问题展开的,理论基础深厚、可实施性强。实验表明,该方法进一步提高了网络安全预测的精确度及有效性。用本文给出的网络安全态势预测模型,能对先前网络安全态势的变化趋势做出准确、客观的评估,很好地预测后续的网络安全态势,便于指导网络管理者做出更好应对网络安全威胁的决策。

参 考 文 献

[1] 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 - 1024.

[2] 肖汉杰,桑秀丽. 相关向量机超参数优化的网络安全态势预测[J]. 计算机应用,2015,35(7): 1888 - 1891.

[3] Bass T. Multi-sensor data fusion for next generation distributed intrusion detection systems [C]//Proceedings of IRIS National Symposium on Sensor and Data Fusion, 1999.

[4] BASS T. Multi-sensor data fusion for next generation distributed intrusion detection systems[EB/OL]. [2016. 03. 10]. http://citeseerx. ist. psu. edu/viewdoc/summary? doi = 10. 1. 1. 51. 1753.

[5] 尤马彦. 网络安全态势评估技术的研究与实现[D]. 广州:广东工业大学,2012.

[6] 赵冬梅,李红. 基于并行约简的网络安全态势要素提取方法[J]. 计算机应用,2017,37(4):1008 - 1013.

[7] 李方伟,杨绍成,朱江,等. 基于模糊层次法的改进型网络安全态势评估方法[J]. 计算机应用, 2014,34(9):2622 - 2626.

[8] Durso F T, Gronlund S D. Situation Awareness[C]//Pro-

ceedings of Handbook of Applied Cognition. John Wiley & Sons,New York,1999: 283 - 314.

[9] Endsley M R. Design and evaluation for Situation Awareness enhancement[C]//Proceedings of the Human Factors Society 32nd Annual Meeting,Santa Monica,CA,1988: 97 - 101.

[10] Loke S W. Representing and reasoning with situations for context-aware pervasive computing: A logic programming perspective[J]. Knowledge Engineering Review, 2004, 19(3):213 - 233.

[11] Lacey T H, Mills R F, Raines R A, et al. A Qualia Framework for Awareness in Cyberspace[C]//Military Communications Conference. IEEE, 2007.

[12] Wei Y, Lian Y F, Feng D G. A Network Security Situtional Awareness Model Based on Information Fusion[J]. Journal of Computer Research and Development, 2009, 46(3): 353 - 362.

[13] King D, Orlando G, Kohler J. A case for trusted sensors: Encryptors with Deep Packet Inspection capabilities[C]// 2012 IEEE Military Communications Conference. IEEE,2012.

[14] Bode M A, Alese B K, Thompson A F, et al. A Bayesian Network Model for Risk Management in Cyber Situation [C]//Proceedings of the World Congress on Engineering and Computer Science 2014 Vol I, 2014.

[15] Friedberg I, Skopik F, Fiedler R. Cyber situational awareness through network anomaly detection: state of the art and new approaches[J]. e & i Elektrotechnik und Informationstechnik, 2015, 132(2):101 - 105.

[16] Fink G A, Haack J N, Maiden W M, et al. A Cooperative Cyber Defense for Securing Critical Infrastructures[J]. ACM Transactions on Information and System Security, 2005, 4(3): 186 - 205.

[17] Haack J N, Fink G A, Maiden W M, et al. Cooperative Infrastructure Defense[EB/OL]. 2014 - 05 - 06. http://www. vizsec. org/workshop2008/fink. pdf.

[18] Szwed P, Skbzynski P. A new lightweight method for security risk assessment based on fuzzy cognitive maps[J]. International Journal of Applied Mathematics and Computer Science,2014,24(1):213 - 225.

[19] Liu S, Liu Y. Network security risk assessment method based on HMM and attack graph model[C]//IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, NETWORKING and Parallel/distributed Computing. 2016:517 - 522.

[20] Ghasemigol M, Ghaemi B A, Takabi H. A comprehensive approach for network attack forecasting[J]. Computers & Security, 2016, 58: 83 - 105.

[21] Wu J, Ota K, Dong M, et al. Big Data Analysis-Based Security Situational Awareness for Smart Grid[J]. IEEE Transactions on Big Data, 2018, 4(3):408 - 417.

[22] 陈秀真,郑庆华,管晓宏,等. 层次化网络安全威胁态势量化评估方法[J]. 软件学报,2006, 17(4): 885 - 897.

[23] 李方伟,郑波,朱江,等. 一种基于 RBF 神经网络的网络安全态势预测方法[J]. 重庆邮电大学学报(自然科学版),2014,26(5):576 - 583.

[24] 韦勇,连一峰,冯登国. 基于信息融合的网络安全态势评估模型[J]. 计算机研究与发展,2009, 46(3):353 - 362.

[25] 周新卫,李小玲. 基于改进 G-K 算法的多节点网络安全态势预测模型[J]. 科学技术与工程,2018,18(25):72 - 77.

[26] 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.

[27] 贾焰,王晓伟,韩伟红,等. YHSSAS:面向大规模网络的安全态势感知系统[J]. 计算机科学,2011, 38(2):4 - 8.

[28] 张丹,郑瑞娟,吴庆涛,等. 基于自律计算的网络安全态势感知模型[J]. 计算机应用,2013, 33(2): 404 - 407.

[29] 甘文道,周城,宋波,等. 基于 RAN-RBF 神经网络的网络安全态势预测模型[J]. 计算机科学, 2016,43(11):388 - 392.

[30] 黄亮亮. 网络安全态势评估与预测方法的研究[D]. 兰州:兰州大学,2016:36 - 41.

[31] 胡冠宇,乔佩利. 基于云群的高维差分进化算法及其在网络安全态势预测上的应用[J]. 吉林大学学报,2016,46(2):568 - 577.

[32] 琚安康,郭渊博,朱泰铭. 基于开源工具集的大数据网络安全态势感知及预警架构[J]. 计算机科学,2017,44(5): 125 - 131.

[33] 白景斐,赵文仓. 基于 G-K 算法的网络安全态势预测模型[J]. 科技通报,2017,33(11):216 - 219.

[34] 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.

[35] 王坤,邱辉,杨豪璞. 基于攻击模式识别的网络安全态势评估方法[J]. 计算机应用,2016, 36(1):194 - 198.

[36] 邓聚龙. 灰色预测与决策[M]. 武汉:华中科技大学出版社,1986.

[37] 汪材印. 灰色关联分析和支持向量机相融合的网络安全态势评估[J]. 计算机应用研究,2013, 30(6): 1859 - 1862.

[38] 王丽珍,郑展,张爱新. 神经网络预测控制算法及其应用[J]. 机械工程与自动化,2007(2):98 - 100.

[39] 任伟, 蒋兴浩, 孙锬锋. 基于 RBF 神经网络的网络安全态势预测方法[J]. 计算机工程与应用, 2006, 42(31): 136 - 144.

[40] 张勇. 网络安全态势感知模型研究与系统实现[D]. 合肥: 中国科学技术大学,2010.

[41] 王晋东,沈柳青,王坤,等. 网络安全态势预测及其在智能防护中的应用[J]. 计算机应用,2010,30(6):1480 - 1488.

[42] Ko C N, Chang Y P, Wu C J. A PSO Method With Nonlinear Time-Varying Evolution for Optimal Design of Harmonic

Filters[J]. IEEE Transactions on Power Systems, 2009, 24(1): 437 - 444.
[43] 曾斌,钟萍. 网络安全态势预测方法的仿真研究[J]. 计算机仿真,2012,29(5):170 - 173.
[44] 高昆仑,刘建明,徐茹枝,等. 基于支持向量机和粒子群算法的信息网络安全态势复合预测模型[J]. 电网技术,2011,35(4):176 - 182.
[45] Snidaro L, Visentini I, Bryan K. Fusing Uncertain Knowledge and Evidence for Maritime Situational Awareness Via Markov Logic Networks[J]. Information Fusion,2015(21): 159 - 172.
[46] Endsley M R. Final Reflections:Situation Awareness Models and Measure[J]. Journal of Cognitive Engineering and Decision Making,2015,9(1): 101 - 111.
[47] 飞思科技. 神经网络理论和 matlab7 实现[M]. 北京:电子工业出版社,2016:15 - 30.
[48] 陈维鹏,敖志刚,屠义强,等. 基于 PSO 优化 LS-SVM 算法的网络空间态势预测研究[J]. 通信技术,2018,51(5): 1154 - 1160.
[49] 杨豪璞, 邱辉, 王坤. 面向多步攻击的网络安全态势评估方法[J]. 通信学报, 2017, 38(1): 187 - 198.
[50] Boser B,Guyon L,Vapnik Y. A training algorithm for optiomal margin classifer[C]//fifth annual workshop on Computational Learning Therory, ACM Press. 1992:144 - 152.
[51] Cortes C,Vapnik V. Support vector network[J]. Machine Learning,1995,20(3): 273 - 297.
[52] 安金龙,王正鸥,马振平. 一种新的支持向量机多分类分类方法[J]. 信息与控制,2004,33(3): 262 - 267.
[53] Lin Y L, Hsieh J G, Wu H K, et al. Three-parameter sequential minimal optimization for support vector machines [J]. Neurocomputing, 2011, 74(17):3467 - 3475.
[54] Yang J, Yang X, Zhang J. A Parallel Multi-Class Classification Support Vector Machine Based on Sequential Minimal Optimization[C]//First International Multi-Symposiums on Computer and Computational Sciences(IMSCCS'06). IEEE, 2006: 443 - 446.
[55] Hsu C W, Lin C J. A Comparison of Methods for Multiclass Support Vector Machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2):415 - 425.
[56] 张翔,胡昌振,刘胜航,等. 基于支持向量机的网络攻击态势预测技术研究[J]. 计算机工程. 2007,33(11):10 - 12.
[57] 王庚,张景辉,吴娜. 网络安全态势预测方法的应用研究[J]. 计算机仿真,2012,29(2):98 - 101.
[58] 李洁,张兆薇. 基于和声搜索算法和相关向量机的网络安全态势预测的方法[J]. 计算机应,2016,36(1):199 - 202.
[59] 汤永利,李伟杰,于金霞,等. 基于粒子滤波的网络安全态势预测方法研究[J]. 计算机应用与软件,2017,34(1): 293 - 297.
[60] 袁小芳, 王耀南. 基于混沌优化算法的支持向量机参数选取方法[J]. 控制与决策, 2006, 21(1):111 - 113.

(上接第 281 页)

表 7 两种算法的比较结果

算法	最优值	最差值	平均值	方差
FPA	1.331 5e+06	1.426 1e+06	1.377 2e+06	2.642 7e+04
MFPA	1.203 8e+06	1.289 5e+06	1.246 5e+06	2.486 2e+04

5 结 语

本文针对当前算法求解物流配送中心选址问题时,普遍存在求解精度低、收敛速度慢及规模较小等缺陷,提出了一种改进的花朵授粉算法求解物流配送中心选址问题,将花朵授粉算法进行全局搜索的同时融合遗传算子进行局部搜索。通过两者的结合及多组不同规模的仿真实验表明了改进的花朵授粉算法在求解物流配送中心选址问题的有效性。

参 考 文 献

[1] Weber A. Alfred Weber's theory of the location of industries [M]. The University of Chicago Press,1909.
[2] 杨茂盛,姜华. 基于重心法与离散模型的配送中心选址研究[J]. 铁道运输与经济,2007,29(7):68 - 70.
[3] 王思宇,乔辉. 层次分析法在配送中心选址中的应用研究[J]. 计算机应用与软件,2013,30(6): 222 - 224.
[4] 解丹蕊,薛惠锋,和文全,等. 基于遗传算法的西安邮政配送中心选址研究[J]. 计算机仿真,2008,25(1): 208 - 211,220.
[5] 周艳平,顾幸生. 差分进化算法研究进展[J]. 化工自动化及仪表,2007,34(3): 1 - 6.
[6] 费腾,张立毅,陈雷. 配送中心选址问题的 BFO-AFSA 算法研究[J]. 计算机工程与应用,2015,51(23): 1 - 5.
[7] 李磊,杨爱峰,唐娜. 基于多种群搜索的 PSO 的物流配送中心寻址求解[J]. 合肥工业大学(自然科学版),2017,40(2):266 - 271.
[8] 赵世安,屈迟文. 改进的布谷鸟算法求解物流配送中心选址问题[J]. 数学的实践与认识,2017,47(3):206 - 213.
[9] 徐小平,师喜婷. 关于物流配送中心供需优化选址仿真[J]. 计算机仿真,2018,35(10):345 - 349.
[10] Yang X S. Flower Pollination algorithm for global optimization[C]//International Conference on Unconventional Computing and Natural Computation. Lecture Notes in Computer Science, 2012,7445:240 - 249.
[11] 肖辉辉,万常选,段艳明. 一种改进的新型元启发式花朵授粉算法[J]. 计算机应用研究,2016,33(1):126 - 131.