

引用格式:庄海燕. 大数据分析技术的无线通信网络安全风险预测[J]. 微电子学与计算机, 2019, 36(8): 97-100. [ZHUANG H Y. Network Based on Big Data Analysis Technology[J]. Microelectronics & Computer, 2019, 36(8): 97-100.]

大数据分析技术的无线通信网络安全风险预测

庄海燕

(铁道警察学院, 河南 郑州 450053)

摘 要: 当前无线通信网络安全风险预测模型存在误差大, 预测耗时长等缺陷, 为了提高无线通信网络安全风险预测精度, 加快无线通信网络安全风险预测速度, 提出了基于大数据分析技术的无线通信网络安全风险预测模型. 首先采集无线通信网络安全风险预测历史数据, 然后采用大数据分析技术建立无线通信网络安全风险预测模型, 最后进行了无线通信网络安全风险预测仿真实验, 结果表明, 本文模型的无线通信网络安全风险预测精度高, 预测误差远远小于其它无线通信网络安全风险预测模型, 加快了无线通信网络安全风险预测速度.

关键词: 无线通信; 网络安全; 风险预测; 大数据分析

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2019)08-0097-04

Network based on big data analysis technology

ZHUANG Hai-yan

(Railway Police College, Zhengzhou 450053, China)

Abstract: The current wireless communication network security risk prediction model has some shortcomings, such as large error and long prediction time. In order to improve the accuracy of wireless communication network security risk prediction and accelerate the speed of wireless communication network security risk prediction, a wireless communication network security risk prediction model based on large data analysis technology is proposed. Firstly, the historical data of wireless communication network security risk prediction are collected, and then the security risk prediction model of wireless communication network is established by using large data analysis technology. Finally, the simulation experiment of wireless communication network security risk prediction is carried out. The results show that the security risk prediction accuracy of wireless communication network based on this model is high and the prediction accuracy is high. The measurement error is far less than other wireless communication network security risk prediction models, which accelerates the speed of wireless communication network security risk prediction.

Key words: wireless communication; network security; risk prediction; large data analysis

1 引言

无线网络系统给人们生活带来便利的同时, 一些不法分子提供了一些便利, 通过各种攻击手段, 对无线网络传输的敏感信息进行窃取, 每年攻击的频率, 窃取数据数量呈大幅增长趋势, 无线网

络安全问题随之而来^[1,2]. 风险预测是一种新的无线网络安全技术, 可以描述无线网络的整体安全状态, 成为当前无线网络安全研究的一个重要方向^[3].

无线通信网络安全风险预测模型可以分为两类: 线性预测模型和非线性预测模型. 线性预测模

收稿日期: 2018-12-20; 修回日期: 2019-02-22

基金项目: 河南省科技攻关项目(182102310809), 铁道警察学院院级教学改革项目(JY2018019); 河南省高等教育教学改革研究与实践项目(2017SJGLX521)

型目前已经处于淘汰状态,没有什么实际应用价值^[4].非线性预测模型主要包括:各种神经网络的无线通信网络安全风险预测模型以及支持向量机的无线通信网络安全风险预测模型^[5],它们从非线性角度对无线通信网络安全风险进行跟踪和建模.

最小二乘支持向量机(Least Squares Support Vector Machines, LSSVM)是一种大数据分析技术,将支持向量机的不等式约束改成等式约束,提高了求解速度及收敛精度^[6].为了解决无线通信网络安全风险预测过程存在的难题,设计大数据分析技术的无线通信网络安全风险预测模型,并与其它无线通信网络安全风险预测模型进行了具体的仿真测试,验证了本文无线通信网络安全风险预测的可行性和优越性.

2 大数据分析技术的无线通信网络安全风险预测模型

2.1 大数据分析技术

最小二乘支持向量机算法是一种新型的大数据分析技术,设训练样本集: $\{x_k, y_k\}_{k=1}^l$. x_k 和 y_k 分别代表第 k 个样本的输入向量和期望输出; l 表示训练样本数量,最小二乘支持向量机建立如下形式最优分类面函数:

$$f(x) = \omega^T \varphi(x) + b \quad (1)$$

式中, b 为阈值; $\varphi(\cdot)$ 为变换函数; ω 为超平面的权值向量.

根据结构风险最小化原则,式(1)可以转换为求解如下的最优化问题

$$\begin{aligned} \min J(\omega, \xi) &= \frac{1}{2} \omega^T \cdot \omega + \frac{1}{2} \gamma \sum_{i=1}^l \xi_i^2 \\ \text{s. t.} & \\ \begin{cases} y_i = \omega^T \cdot \varphi(x_i) + b + \xi_i \\ i = 1, 2, \dots, l \end{cases} & \end{aligned} \quad (2)$$

式中, γ 为正则化参数; ξ_i 为松弛变量.

将约束条件改为等式约束,引入拉格朗日算子: α_i , ($i = 1, \dots, l$),构造如下拉格朗日方程:

$$\begin{aligned} L(\omega, b, \xi, \alpha_k) &= \frac{1}{2} \omega^T \cdot \omega + \frac{1}{2} c \sum_{i=1}^l \xi_i^2 - \\ &\sum_{i=1}^l \alpha_i (\omega \cdot \varphi(x_i) + b + \xi_i - y_i) \end{aligned} \quad (3)$$

要使目标函数取得最小值,对 ω, b, ξ, α_k 进行求偏导数,并且使偏导数为 0,即:

$$\begin{cases} \frac{\partial L}{\partial \omega} = \omega - \sum_{i=1}^l \alpha_i \varphi(x_i) = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^l \alpha_i = 0 \\ \frac{\partial L}{\partial \xi_i} = \gamma \xi_i - \alpha_i = 0 \\ \frac{\partial L}{\partial \alpha_i} = \omega^T \cdot \varphi(x_i) + b + \xi_i - y_i = 0 \end{cases} \quad (4)$$

通过求解该二次规划问题件,建立核函数 $K(x, x_i) = \varphi(x)^T \varphi(x_i)$ 得到如下形式的最小二乘支持向量机决策函数:

$$f(x) = \sum_{i=1}^m \alpha_k K(x, x_k) + b \quad (5)$$

式中, m 表示支持向量的数量.

核函数定义如下

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right) \quad (6)$$

式中, σ 为核参数.

2.2 粒子群优化算法

由于核参数 σ 和正则化参数 γ 影响最小二乘支持向量机的性能,因此本文采用粒子群优化算法对于核参数 σ 和正则化参数 γ 进行优化.粒子群优化算法不断变换自己的飞行轨迹对问题进行求解,而飞行轨迹包括速度($V_{ik}(t+1)$)和位置($X_{ik}(t+1)$)两个部分,在第 $t+1$ 时刻和第 t 时刻,两者之间的联系

$$V_{ik}(t+1) = \omega V_{ik}(t) + c_1 r_1 (P_{ik}(t) - X_{ik}(t)) + c_2 r_2 (P_{gk}(t) - X_{ik}(t)) \quad (7)$$

$$X_{ik}(t+1) = X_{ik}(t) + V_{ik}(t+1) \quad (8)$$

2.3 大数据分析技术的无线通信网络安全风险预测步骤

(1) 对一个无线通信网络安全风险预测问题,选择报警数量、被植入后门数量、攻击数量、新增安全漏洞数量作为无线通信网络安全风险指标,收集它们一段时间的这些指标数据.

(2) 将无线通信网络安全风险等级划分为 5 种,它们具体分别为:优[0 1],良(1 2.5],中(2.5 6],差(6 8],危(8 10],通过专家根据指标数据标记出相关的无线通信网络安全风险等级,它们构成无线通信网络安全风险预测的样本集合.

(3) 从无线通信网络安全风险样本集合按一定比例随机选择一部分作为训练样本集,用于构建无线通信网络安全风险预测模型,剩余下的样本作为测试样本,以验证无线通信网络安全风险预测模型的有效性.

(4) 为了提高最小二乘支持向量机的收敛效率,加快最小二乘支持向量机学习速度,对无线通信网络安全风险等级的值进行如下归一化处理.

$$y' = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \tag{9}$$

式中, $y_{\max}$ 和 $y_{\min}$ 分别表示无线通信网络安全风险等级的最大和最小值.

(5) 采用粒子群优化算法对于核参数 σ 和正则化参数 γ 进行优化,确定最小二乘支持向量机的参数,采用最小二乘支持向量机对无线通信网络安全风险的训练样本集进行学习,建立无线通信网络安全风险预测模型.

(6) 采用验证本对无线通信网络安全风险预测模型的性能进行测试.

最小二乘支持向量机的无线通信网络安全风险预测流程如图 1 所示.

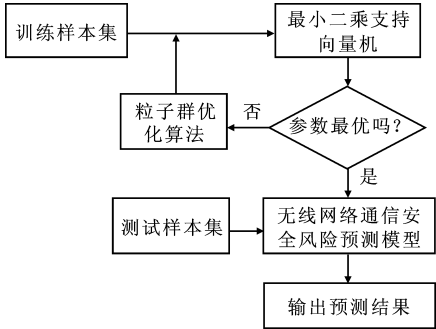


图 1 大数据分析技术的无线通信网络安全风险预测流程

3 无线通信网络安全风险预测模型的性能分析

3.1 无线通信网络安全风险实验数据

选择一个无线通信网络的一段时间安全态势历史数据作为测试对象,无线通信网络安全风险样本数据变化如图 2 所示. 为了体现实现的公平性,进行 5 次仿真实验,每一次随机从图 200 数据选择个数据作为训练样本,剩余样本 50 个作为测试样本.

3.2 结果与分析

3.2.1 安全态势的预测精度对比

采用 RBF 神经网络、支持向量机、最小二乘支持向量机分别对 200 个训练样本进行学习,并对 50 个测试样本的值进行预测,并统计每一次仿真实验的无线通信网络安全风险平均预测精度,结果如图 3 所示,对图 3 的无线通信网络安全风险平均预测精度进行对比和分析,可以得到如下结论:

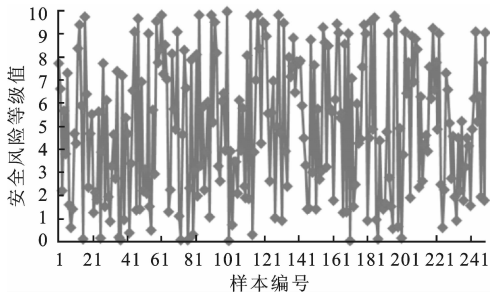


图 2 无线通信网络安全风险历史数据

(1) RBF 神经网络的无线通信网络安全风险平均预测精度比较低,而且稳定性比较差,这是因为它们基于经验风险最小化原则进行建模,建立的无线通信网络安全风险预测模型稳定不够,无法获得理想的无线通信网络安全风险预测结果.

(2) 支持向量机和最小二乘支持向量机的无线通信网络安全风险平均预测精度要高于 RBF 神经网络,主要因为它们基于结构风险最小化原则进行建模,可以更好对无线通信网络安全风险进行描述,降低了无线通信网络安全风险预测误差,同时最小二乘支持向量机的无线通信网络安全风险预测精度要高于支持向量机,这是因为克服了支持向量机的缺陷,建模效果更佳.

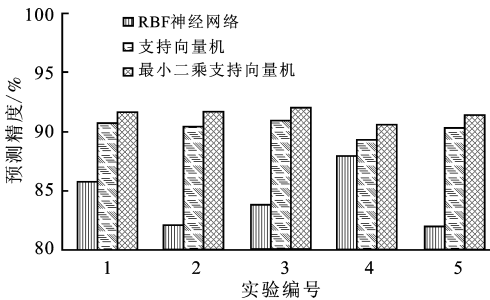


图 3 无线通信网络安全风险的预测精度

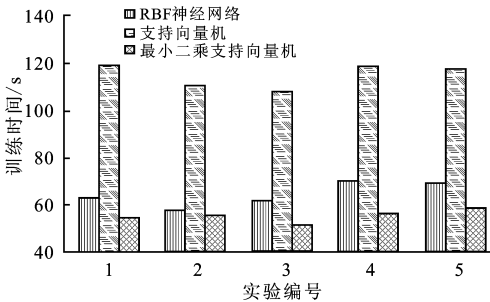


图 4 无线通信网络安全风险预测的训练时间

3.2.2 无线通信网络安全风险预测效率对比

统计各种模型的无线通信网络安全风险训练时间,具体如图 4 所示. 从图 4 可以发现,无线通信网

络安全风险建模的训练时间相差较大,其中支持向量机的训练时间最长,而最小二乘支持向量机的训练时间最短,提升了无线通信网络安全风险预测效率.

4 结束语

针对当前无线通信网络安全风险预测模型存在误差大,预测耗时长等缺陷,设计了基于大数据分析技术的无线通信网络安全风险预测模型,测试结果表明,本文模型的无线通信网络安全风险预测精度,加快了无线通信网络安全风险预测速度,具有十分广泛的应用前景.

参考文献:

[1] 沈利香,曹国. 基于不确定语言变量的网络安全风险评估方法[J]. 常州工学院学报,2018,31(1):36-43.

[2] 项立. 基于智能分类算法的光纤通信网络风险评估[J]. 激光杂志,2015,36(7):93-96.
[3] 刘建峰,陈健. 基于模糊博弈规则的网络节点入侵风险评估[J]. 计算机科学, 2018, 45(10): 138-141.
[4] 刘刚,李千目,刘凤玉,等. 面向网络实时风险预测的马尔可夫时变模型[J]. 兵工学报,2012,33(2):163-169.
[5] 熊飞,李军,钮焱. 基于藤 Copula-GARCH 模型的网络资源风险预测[J]. 湖北工业大学学报,2016, 31(5): 68-71.
[6] 郭祖华,李扬波,徐立新,等. 面向云计算的网络安全风险预测模型的研究[J]. 计算机应用研究, 2015, 32(11):3421-3425.

作者简介:

庄海燕 女,(1976-),硕士,副教授. 研究方向为电子数据取证、数据分析. E-mail:zhuanghaiyan@rpc.edu.cn