

一种基于用户行为习惯的主机安全性检测方法

余林琛, 沈江波, 何仲彻, 钟 贝, 刘 路
(中国地质大学(武汉) 计算机学院, 湖北 武汉 430074)

摘 要: 文中提出了一种基于用户行为习惯的主机安全性检测方法. 通过建立一种以攻击树 root 代表主机的拓展攻击树模型来研究用户的行为习惯对主机的影响, 并以此来描述主机安全性. 该方法实现了对不同用户主机安全性的智能评估, 为主机安全性检测提供了一种新方向.

关键词: 用户行为习惯; 主机安全性检测; 文件来源地; 拓展攻击树

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2011)09-0158-02

A Method to Detect Host's Security Based on User's Behavior Habits

YU Lin-chen, SHEN Jiang-bo, HE Zhong-che, ZHONG Bei, LIU Lu

(College of Computer, China University of Geosciences (Wu han), Wuhan 430074, China)

Abstract: A method to detect host's security based on user's behavior habits is presented in this paper. A extended attack tree model which uses attack tree's root to represent the host is built to research the influence of user's behavior habits to the host, it's also used to describe the host's security. The method realizes an intelligence assessment of host's security for different hosts, and provides a new direction for detect host's security weight.

Key words: user's behavior habits; host's security weight; files' source; extended attack tree

1 引言

随着信息技术的高速发展, 主机安全性检测成为不可忽略的问题. 常见的主机安全性判断方法^[1-2]是综合分析主机内所有可执行文件, 以同一标准进行文件恶意性判定, 并不考虑用户的行为习惯对可执行文件的恶意性的影响.

通过对主机上大量随机可执行文件分析后发现, 来源地不同的可执行文件其安全性有很大差异, 而用户行为习惯决定文件来源地. 据此, 文中提出了一种基于用户行为习惯的主机安全性检测方法.

2 主机安全性检测方法的具体实现

检测主机安全性就是综合分析来源地对主机安全性的影响. 因此, 判断主机安全性可分为以下几个步骤: 计算可执行文件权值, 判断文件恶意性; 计算文件来源地权值; 计算主机安全性权值, 判断主机安全性.

可执行文件权值影响文件来源地的权值, 文件来源地的权值又影响主机安全性权值, 体现的是一种自下而上的影响关系. 因此, 文中在攻击树^[3]模型的基础上, 建立了一种以攻击树 Root 代表主机的拓展攻击树模型, 图 1 为拓展攻击树的结构模型.

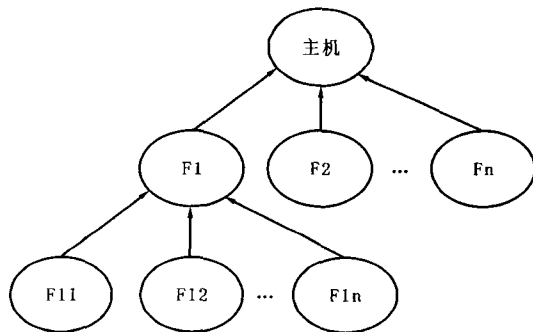


图 1 拓展攻击树的结构

由拓展攻击树的模型图可知: 该拓展攻击树的根结点为主机, 分支结点 F1、F2、Fn 表示主机上的

文件来源地,叶子结点 F11、F12、F1n 表示此类可执行文件来自文件来源地 F1.

2.1 可执行文件权值的计算

分析可执行文件的行为,以不区分来源地的方式计算可执行文件的固有权值 $P(D)$,表示一个可执行文件在不考虑文件来源地的前提下有毒的概率.

2.2 可执行文件恶意性判断

现有攻击树中计算根结点权值的算法没有体现出分支结点对根结点影响.即没有体现文件来源地对主机安全性的影响.所以,本方法引进贝叶斯算法^[4],该算法将文件来源地作为判断文件恶意性的影响因素,避免了此前出现的弊端.

根据贝叶斯公式:

$$P(D | F(i)) = \frac{P(F(i) | D) * P(D)}{P(F(i))} \tag{1}$$

式中, $P(D | F(i))$ 表示一个可执行文件来自文件来源地 $F(i)$ 的前提下是病毒的概率; $P(F(i) | D)$ 表示一个可执行文件在确定有毒的情况下是来自文件来源地 $F(i)$ 的概率; $P(D)$ 是可执行文件的固有权值; $P(F(i))$ 表示任意给以个可执行文件时来自文件来源地 $P(F(i))$ 的概率.

$$P(F(i)) = \frac{a(i)}{\sum_{k=1}^n a(k)} \tag{2}$$

式中, $a(i)$ 表示第 i 个文件来源地中可执行文件的个数, n 为文件来源地的个数.

$$P(F(i) | D) = \frac{b(i)}{\sum_{k=1}^n b(k)} \tag{3}$$

式中, $b(i)$ 表示第 i 个文件来源地中有毒文件的个数,由公式(2)(3)可得:

$$P(D | F(i)) = \frac{\sum_{k=1}^n a(k) * b(i) * p(D)}{a(i) * \sum_{k=1}^n b(k)} \tag{4}$$

比较 $P(D | P(i))$ 与给定的阈值,若大于阈值,则该文件为病毒.

2.3 文件来源地权值的计算

$$W(i)' = \frac{(a(i) - m) * W(i) + P(D)}{a(i) - m + 1} \tag{5}$$

式中, $W(i)$ 为可执行文件在被检测之前该可执行文件所在文件来源地的权值, $W(i)'$ 为可执行文件被检测之后其所在文件来源地的权值. m 为一个常数,表示来源地中 $a(i)$ 的初始值,由大量已有数据分析得出.

2.4 主机安全性权值 W 的计算

$$W' = \frac{M * W - W(i) + W(i)'}{M} \tag{6}$$

式中, M 为文件来源地的个数, W 为文件检测前主机安全性权值, W' 为文件检测之后主机安全性权值.该公式表明主机安全性权值是不断在变化的,且与文件来源地权值相关.

2.5 主机安全性判断

根据计算得到的主机安全性权值,分析出主机的安全性.

3 实验结果及其分析

由于该检测方法与用户行为习惯相关联,不同用户的主机安全性具有差异性.所以,本实验选取两个用户分别在两台主机上进行操作.规定:两个用户在获取文件时,其文件来源地的个数为 14 个,且这些文件来源地与其平时的行为习惯相吻合.实验相同时间后计算两台主机各自的权值,比较两台主机的安全性,从而判断该方法的可行性.如表 1、表 2 所示.

表 1 用户一的统计表格

$a(i)$ / 个	$b(i)$ / 个	W
314	3	0.46
129	6	
212	9	
324	48	
409	23	
264	16	
278	3	
312	6	
242	9	
174	17	
156	8	
229	4	
192	10	
171	6	

表 2 用户二的统计表格

$a(i)$ / 个	$b(i)$ / 个	W
214	8	0.38
154	3	
263	6	
143	9	
532	3	
323	5	
232	10	
382	8	
192	3	
204	11	
326	7	
319	6	
132	3	
201	12	

5 结束语

为了提高区域物流需求预测的精度,文中结合线性预测模型、周期性预测模型和非线性预测模型优点,提出一种基于 BP 神经网络的区域物流需求组合预测型.通过对上海区域物流需求数据进行验证性实验,实验结果表明,组合模型很好揭示了区域经物流需求与影响因素间的非线性映射关系,提高了区域物流需求的预测精度.组合预测模型比单一预测模型更适合于非线性特征的复杂系统,可以准确描述区域物流未来发展趋势,为区域物流需求预测提供了一种新的思路和方法.

参考文献:

- [1] Rodrigo A, Hani S. Forecasting freight transportation demand with the space-time multinomial probity model [J]. Transportation Research Part B, 2000;403-418.
- [2] 李莉,张建华,周海燕.物流产业发展与国民经济整体水平提升的相关性分析[J].中国机械工程,2003,14(10):884-887.
- [3] 过秀成,谢实海,胡斌.区域物流需求分析模型及其算法

[J].东南大学学报:自然科学版,2001,31(3):1-5.

- [4] 潘英英,宋国喜.基于灰色系统理论的广西物流需求预测[J].广西财经学院学报,2008,21(1):26-30.
- [5] 黄毅,夏国恩.基于支持向量回归机的广西物流需求预测[J].科技管理研究,2011,12(2):142-146.
- [6] 杨海光,夏国恩.基于灰色预测模型的广西物流需求预测[J].中国管理信息化,2011,14(5):39-41.
- [7] 陈以,万梅芳.RBF神经网络在物流系统中的应用[J].计算机仿真,2010,21(4):41-44.
- [8] 后锐,张毕西.基于MLP神经网络的区域物流需求预测方法及其应用[J].系统工程理论与实践,2005,12(1):43-47.
- [9] 孙晓玲,王宁,梁艳.应用BP神经网络的教学评价模型及仿真[J].计算机仿真,2010,21(11):84-86.

作者简介:

万励女,(1972-),硕士,工程师,讲师.研究方向为数据库和计算机应用.

李余琪男,(1980-),硕士研究生,讲师.研究方向为数据分析预测.

吴洁明女,(1959-),研究生,副教授.研究方向为计算机软件与理论.

(上接第159页)

由表格可以看出,用户一的主机安全性权值 W 为 0.46,用户二的主机安全性权值 W 为 0.38.经比较,用户一主机安全性比用户二的主机安全性低,这验证了用户行为习惯对主机安全性是有影响这一结论.实验证明:该检测方法可行.

4 结束语

文中通过建立拓展攻击树模型,利用贝叶斯算法判断主机安全性的方法是有效可行的,并且该方法可以作为判断未知文件恶意性.当来源地集合越完善时,该方法对主机安全性的检测越精确.

随着来源地集合的不断完善,对拓展攻击树的进一步改造以及对贝叶斯算法的优化将是接下来的

工作的重点.

参考文献:

- [1] 张国力,马军.主机安全技术研究[J].信息安全与技术,2010(8):56-58.
- [2] 夏冰,徐飞,郑秋生.一种主机安全评估系统的研究和设计[J].中原工学院学报,2009,20(5):24-29.
- [3] 刘巍伟,石勇,郭煜,等.一种基于综合行为特征的恶意代码识别方法[J].电子学报,2009,37(4):696-700.
- [4] 张波云,殷建平,蒿敬波,等.基于多重朴素贝叶斯算法的未知病毒检测[J].计算机工程,2006,32(10):18-21.

作者简介:

余林琛女,(1977-),讲师.研究方向为网络安全、对等计算.