文章编号:1673-9469(2008)01-0042-03

基于记忆原理的网络入侵检测

黄光球,李 眩
(西安建筑科技大学 管理学院,陕西 西安 710055)

摘要:以生物记忆的三种形式的形成与衰减为基础建立一种数学分析模型,并将此应用到入侵检测中去。应用短时记忆容量限制来节省系统大量的存储空间,应用长时记忆使系统能检测较长时间跨度的入侵行为。感应阀技术能灵活调整系统的灵敏度,这种系统拥有高精确的入侵检测,高效的决策过程,以及系统资源消耗低的优点。

关键词:入侵检测;记忆原理;感应阀;衰减
中图分类号:TP393 **文献标识码**:A

An intrusion detection system based on memory principle

HUANG Guang-qiu, LI Xuan
(Department of Management, Xi'an University of Architecture and Technology, Xi'an 710055, China)

Abstract: An analytical math model was established basing on three form and decline of biological memory and was used in intrusion detection system. The capacity limitation of short-term memory saved much storage space for the detection system; the long-term memory enabled the detection system to detect intrusion in a long time; the sense valve technology can regulate sensitiveness of system flexibly. IDS have advantages of the high accuracy of detection intrusion, an efficient decision-making process and low consumption of system resource.

Key words: intrusion detection; memory principle; sense valve; decline

现在,越来越多的业务通过信息网络实现,网络遭受入侵的风险性比以前急剧增多,要想完全避免入侵的发生并不太现实,所能做到的只能是尽力发现和察觉入侵及入侵企图,以便采取有效的措施来防范和制止入侵,这样的研究称为入侵检测。入侵检测系统中最为核心的问题是数据分析,它将直接决定系统的检测能力和效果。入侵检测系统(IDS)用来识别针对计算机系统和网络系统,或者更广泛意义上的信息系统的非法攻击,包括检测外界非法入侵者的恶意攻击或试探,以及内部合法用户的超越使用权限的非法行动。

1 用户行为异常度值的计算

中心极限定理认为无论研究的统计总量服从什么样的分布,样本均值的分布均接近一个正态分布。正态分布的均值等于总体分布的均值,通过分析数据的均值和标准偏差就可以了解用户的行为,可将系统正常情况下的数据样本作为描述系统行为的测度集。对测度集数据计算均值和标准偏差,设定一个置信区间。如果网络在不同时段的数据在此区间以外,则认定网络异常。

定义测度集$(x_1, x_2, \cdots, x_n)$,则均值为

$$m_n = (x_1 + x_2 + \cdots + x_n)/n = \frac{\sum_{i=1}^{n} x_i}{n}$$

数据的标准偏差为

$$\sigma_0 = ((x_1^2 + x_2^2 + \cdots + x_n^2)/n - m_n^2)^{1/2} = sqrt(\frac{\sum_{i=1}^{n} x_i}{n - m_n^2})$$

则总体均值的置信区间为 $U = [m_n - H\frac{\sigma_0}{\sqrt{n}}, m_n +$

收稿日期:2007-09-13
基金项目:陕西省教育厅专项基金资助项目(06JK258),西安建筑科技大学基础研究基金项目(JC0616)
作者简介:黄光球(1964-),男,陕西西安人,教授,博导,从事电子商务与信息安全的研究。

$H\frac{\sigma_0}{\sqrt{n}}$,其中 H 为系统所设定的感应阀值。如果新的取样值 $x_{n+1}\in U$,则该行为正常,若 $x_{n+1}\notin U$,则用户行为异常。下面计算异常用户行为的异常程度值

当 $x_{n+1}<m_n-H\frac{\sigma_0}{\sqrt{n}}$ 时,$S=m_n-H\frac{\sigma_0}{\sqrt{n}}-x_{n+1}$;

当 $x_{n+1}>m_n+H\frac{\sigma_0}{\sqrt{n}}$ 时,$S=x_{n+1}-m_n-H\frac{\sigma_0}{\sqrt{n}}$。

在计算均值和标准偏差后,不是固定的选取不变的置信区间,而是根据感应阀值去调整置信区间的宽度,灵活调整检测系统的灵敏度。

2 记忆原理分析

感应阀值直接影响短时记忆的接收能力,它是信号强度接受的控制值。其值的大小直接影响系统处理信息的活跃程度。由于新信息的加入而将旧信息挤出记忆库的现象称为遗忘。遗忘量的大小与时间有紧密联系,时间越长,信息就会变得越弱。

2.1 感应阀的控制

感应阀值的大小受 S 值控制,其用来确定 S 是否超过异常阀值,并是否将相应数据输入短时记忆的控制值。感应阀越高,则系统越迟钝;感应阀越低,意味着系统越灵敏。定义感应阀为 H,t_2 为当前时间,t_1 为上次计算阀值的时间(H 的初始值为 a)。其计算公式为

$$H=H'-\frac{b}{t_2-t_1}S \qquad (x_{n+1}\in U, H'>0)$$

$$H=H'+2^{r(t_2-t_1)}b \quad (x_{n+1}\notin U, H'<H_{max})$$

式中 b —单位调整常量;r—衰减速率,$r=\frac{-\log_2 0.5}{10}=0.1$;$H_{max}$—最大限制值。

计算中,如果 $H<0$,则令 $H=0$;如果 $H>H_{max}$,则令 $H=H_{max}$。当 $S>a$ 时,感应阀值处于降低状态,意味着系统会对异常更加敏感,而当 $S\le a$ 时,感应阀值开始恢复。

2.2 短时记忆及其衰减

当确定 $x_{n+1}\notin U$ 后,即被输入短时记忆库,否则丢弃该数据。为了保证系统的永久可载性,对短时记忆库进行了容量限制。定义短时记忆集 $J=(W_1,W_2,\cdots,W_n)$,限制长度 $L=5$,当前长度 $n<=L$,集合中记忆细胞 $W_k=(t_k,S_k,d_k)$ 为三维数据集。当数据进入短时记忆集前,先判定 n,L 的大小,d 为 S 值对异常度综合评价值的贡献度,初始值皆为 1。

如果 $n<L$,则 $W_{n+1}=(t_k,S_k,d_k)$,此时集合细胞个数变为 $n+1$ 个。

如果 $n=L$,则寻找 J 中存放时间最长即 t 最大的细胞 W_m 替换为 $W_m=(t_k,S_k,d_k)$。

短时记忆是一个随时间刷新的数据集合,其长度到达容量极限后不会再增长。这样保证系统始终保存最近的异常数据,又能保证系统不因数据量过多而崩溃。记忆衰减是一个以时间为变量的减弱过程。在最终的异常度综合评价中,必须考虑到每个记忆细胞信息的衰减。

定义衰减函数 $d=d'\frac{1}{2^{r(t_2-t_1)}}$。$r$ 为衰减速率,决定了贡献系数的半周期,取 d 的半周期为 30s 则 $r=\frac{-\log_2 0.5}{30}$,d' 为衰减前的贡献系数,初始值为 1,t_2 为当前时间,t_1 为最后更新时间。

该函数具有如下特性 $\lim_{t_2-t_1\to 0} d'\frac{1}{2^{r(t_2-t_1)}}=d'$;

$\lim_{t_2-t_1\to\infty} d'\frac{1}{2^{r(t_2-t_1)}}=0$。

时间差为 0,不产生衰减,经历无限长时间后,衰减为 0。

2.3 长时记忆及其衰减

长时记忆由一个记忆细胞 $L_m=(t,A,\omega)$ 组成,其中 t 为时间,A 为长时记忆权重(即衍生异常值),ω 为长时记忆细胞信息对综合异常度的贡献值。长时记忆会随短时记忆集中数据的刷新和增加而增长,又会随时间而缓慢减弱。记忆细胞中的信息值是衰减与迭代的结果。

$$A=A'+N(W_n)$$

A' 为原长时记忆细胞所记录值,初始值为 0。$N(W_n)$ 函数是以短时记忆细胞信息值为输入的迭加函数,$N(W_n)=\frac{c}{n}\sqrt{1+S_n^2}$。

当短时记忆集更新或增加时,c 为常数,否则 $c=0$。S_n 为短时记忆细胞所携带的值(即短时异常值),以上就体现了短时记忆对长时记忆的促进作用。

长时记忆和短时记忆细胞对综合评价值的贡献都会随时间衰减,另外长时记忆是短时记忆集中的数据迭加而来,故其贡献系数初始值须设置

高些。

$$\omega_1 = \omega'_1 \frac{1}{2^{\frac{t_2 - t_1}{r_2}}}$$

ω'_1 为原贡献系数,初始值设定为 2,t_1 为最后更新时间,t_2 为当前时间。长时记忆较短时记忆衰减半周期较长设定为 120s,$r_2 = -\frac{\log 0.5}{120} = \frac{1}{120}$。

3 基于异常累积的入侵判定

异常度综合评价值结合短时记忆和长时记忆的结果来计算。我们用 T^2 综合表明最近用户行为的异常程度,下面介绍如何从 S 值计算得来 T^2。

假设有 n 个测量值表示为 $S_i, 1 \le i \le n$,测量值 S_i 与 S_j 之间的相关性表示为 C_{ij}。T^2 则定义为 $T^2 = (S_1, S_2, \cdots, S_n) C^{-1} (S_1, S_2, \cdots, S_n)^T$。

这里 C^{-1} 是向量 $(S_1, S_2, \cdots, S_n)$ 相关矩阵的逆矩阵,而 $(S_1, S_2, \cdots, S_n)^T$ 是该向量的转置矩阵,排除变量之间的相关性,C^{-1} 则变为单位矩阵。则 T^2 值简化为 $S_1^2 + S_2^2 + \cdots + S_n^2$,即 S 值的平方和。

$$T^2 = S_1^2 + S_2^2 + \cdots + S_n^2$$

由于各 S 值对 T^2 值的影响各不相同,较早计算的 S 值对 T^2 产生的影响会比较晚计算的 S 值要小,因此在上式中引入贡献系数,来表示他们对 T^2 的影响程度。

$$T^2 = d_1 S_1^2 + d_2 S_2^2 + \cdots + d_n S_n^2$$

短时记忆集中异常数据的产生或变更意味着系统遭受入侵的可能性不断加大,这种可能性用长时记忆细胞值来衡量。为此在上式中加入长时记忆因子。

$$T^2 = d_1 S_1^2 + d_2 S_2^2 + \cdots + d_n S_n^2 + \omega A$$

$$T^2 = \sum_{i=1}^{n} d_i S_i^2 + \omega A$$

A 为长时记忆细胞值,ω 为长时记忆贡献值。至此完成了异常度综合评价值的计算,对于 T^2,可选择一个"临界值",临界值与所关心的入侵程度水平相连。当计算得出的结果超出这个临界值时,可认为系统正遭受外界的入侵。

4 结语

在计算机网络建设的现状下,基于防火墙和加密技术的安全固然很重要,但是,发展入侵检测也同样重要。应用生物记忆原理和统计分析技术使检测系统具有较高的检测效率和灵敏度。但随着网络入侵技术的不断发展,入侵的行为表现出不确定性、复杂性、多样性等特点,入侵检测面临许多有待解决的关键问题,如高效率的检测算法、入侵模式确认,高速网络中的入侵检测等一系列问题都有待进一步研究。

参考文献:

[1] 唐正军.网络入侵检测系统的设计与实现[M].北京:电子工业出版社,2002.

[2] 梁巍.记忆的形式模型[J].中国听力语言康复科学杂志,2004,(5):61-62.

[3] 石昌文.基于记忆原理的 Web 安全预警研究[D].西安:西安建筑科技大学,2006.

[4] 汪荣鑫.数理统计[M].西安:西安交通大学出版社,1986.

[5] 阎彗,曹元大.一种基于入侵统计的异常检测方法[J].计算机工程与应用,2002,(22):48-50.

[6] 肖政宏,尹浩.基于网络流量统计分析的入侵检测研究[J].微电子学与计算机,2003,23(5):76-82.

[7] PORRAS P, KEMMERER R. Penetration state transition analysis: A rule-based intrusion detection approach [C]. In: Proceedings of the 8th Annual Computer Security Applications Conference, San Antonio, Texas, 1992.

[8] DEBAR H, DACIER M, WESPI A. Towards a tasonomy of intrusion-detection systems[J]. Computer Networks. 1999, 31:805-822.

[9] SOMMER P. Intrusion detection system as evidence[J]. Computer Networks. 1999, 31:2477-2487.

[10] 李劲峰.分布式入侵检测系统研究与实现[D].北京:北京邮电大学,2001.

[11] 蒲荣富.基于正态分布的异常入侵检测系统[J].阿坝师范高等专科学校学报,2006,23(3):118-120.

[12] 朱磊,杨治良.多种记忆分类之研究[J].心理科学,2003,4(26),694-697.

(责任编辑 刘存英)

基于记忆原理的网络入侵检测

作者： 黄光球， 李晓， HUANG Guang-qiu， LI Xuan
作者单位： 西安建筑科技大学,管理学院,陕西,西安,710055
刊名： 河北工程大学学报（自然科学版） ISTIC
英文刊名： JOURNAL OF HEBEI UNIVERSITY OF ENGINEERING(NATURAL SCIENCE EDITION)
年，卷(期)： 2008,25(1)
被引用次数： 1次

参考文献(12条)

1. 唐正军 网络入侵检测系统的设计与实现 2002

2. 梁巍 记忆的形式模型[期刊论文]-中国听力语言康复科学杂志 2004(05)

3. 石昌文 基于记忆原理的Web安全预警研究 2006

4. 汪荣鑫 数理统计 1986

5. 阎彗;曹元大 一种基于入侵统计的异常检测方法[期刊论文]-计算机工程与应用 2002(22)

6. 肖政宏;尹浩 基于网络流量统计分析的入侵检测研究[期刊论文]-微电子学与计算机 2003(05)

7. PORRAS P;KEMMERER R Ponetration state transition analysis:A rule-based intrusion detection approach 1992

8. DEBAR H;DACIER M;WESPI A Towards a tasonomy of intrusion-detection systems[外文期刊] 1999(8)

9. SOMMER P Intrusion detection system as evidence[外文期刊] 1999(23-24)

10. 李劲峰 分布式入侵检测系统研究与实现[学位论文] 2001

11. 蒲荣富 基于正态分布的异常入侵检测系统[期刊论文]-阿坝师范高等专科学校学报 2006(03)

12. 朱磊;杨治良 多种记忆分类之研究[期刊论文]-心理科学 2003(04)

引证文献(1条)

1. 金民锁.孙道.朱单 边界检测在入侵模式分类与特征提取中的应用[期刊论文]-黑龙江科技学院学报 2011(2)