

基于 ACBM 的高速网络入侵检测算法研究

谈 娴 茹

(中国民航大学 计算机系, 天津 300300)

摘 要: 本文提出一种基于 ACBM 规则匹配的网络入侵算法, 通过定义入侵规则, 将入侵数据与数据库中的数据规则化, 然后根据 ACBM 算法实现入侵规则的匹配, 这样就避免了传统算法对单一数据串的逐一匹配, 大大缩短了高速网络检测所消耗的时间. 实验证明, 这种方法能够有效提高网络入侵检测的效率, 对高速网络的检测效率也有了很大的改善, 取得了满意的结果.

关键词: 规则匹配; 入侵检测; 数据采集

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2012)03-0035-04

Research on High Speed Network Intrusion Detection Algorithm Based on the ACBM Rules Matching

TAN Xian-ru

(Computer Department, Civil Aviation University of China, Tianjin 300300, China)

Abstract: This paper proposes a ACBM rules based on matching network intrusion algorithm, through defining invasion rules and get the invasion data and data in database regularization, then based on the ACBM algorithm to complete the invasion rules matching. This prevents the traditional algorithm for a single data string by the match, greatly reducing the high-speed network detection spending time. Experiments show that the method can effectively improve the efficiency of the network intrusion detection, high-speed network detection efficiency also has made a lot of improvement, and satisfactory results were obtained.

Key words: rules matching; intrusion detection; data acquisition

1 引言

随着网络化程度的日益提高, 网络资源共享带来的信息安全问题也逐渐凸显出来, 成为亟待解决的问题^[1-2]. 传统的基于 BM(R. S. Boyer and J. S. Moore)算法的入侵检测方法, 是将待检测的数据字符串逐个与数据库中的字符串一一匹配比较, 先比较两组字符串的第一个字符, 若匹配则移动到一个字符进行比较, 若不匹配则将数据库中的字符串移动到与待检测字符串第一个字符相同的位置, 继续以同样的方式匹配^[3]. 但是, 当所要检测的网络是高速网络时, 通过数据采集器会采集到达规模的数据, 这时通过传统方法检测网络入侵, 需要将数据采集器采集到的大规模数据逐一与数据库中的字符串

比较, 这样的检测就耗费了大量的时间, 不能高效将入侵威胁检测出来. 因此, 如何高效处理数据采集器采集到的大规模数据, 提高入侵检测的效率, 成为入侵检测领域的难点问题.

针对传统基于 BM 算法的入侵检测方法不能快速检测高速网络数据安全的问题, 本文提出一种基于 ACBM(Aho_Corasick R. S. Boyer and J. S. Moore)规则匹配的网络入侵算法, 通过定义入侵规则, 将入侵数据与数据库中的数据规则化, 然后根据 ACBM 算法实现入侵规则的匹配, 这样就避免了传统算法对单一数据串的逐一匹配, 大大缩短了高速网络检测所消耗的时间. 这种方法能够有效提高网络入侵检测的效率, 对高速网络的检测效率也有了很大的改善.

2 网络入侵检测原理

入侵检测是通过对网络行为、安全日志或者是审计数据或者其它的网络上可以获得的有效信息进行操作,检测出可能对网络系统的闯入或者是闯入的企图,来保证计算机信息的安全,是一种能够及时检测出计算机网络中违反信息安全行为的技术^[4]. 入侵检测系统一般分为四个部分,具体的框图如图 1 所示.

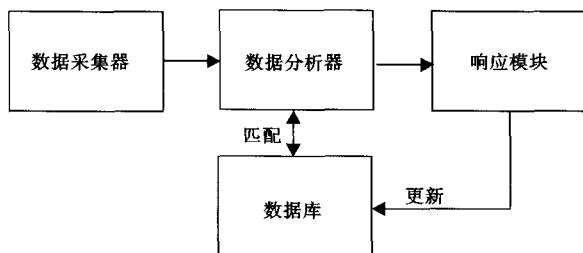


图 1 入侵检测系统框图

(1) 数据采集器. 主要完成的是从检测系统之外的计算机网络环境中收集威胁事件,并将所有的这些威胁事件转换成数据字符串的形式输入到入侵检测系统中.

(2) 数据库. 用来存储入侵数据,需要与待检测数据匹配时提取使用.

(3) 数据分析器. 用来将数据采集器输送过来的数据字符串与数据库中的数据字符串相匹配,匹配结果输送到响应模块.

(4) 响应模块. 在对攻击信息进行分析匹配后确定攻击报警,然后针对确定的攻击类型进行相应的防攻击处理.

此外,为保证数据库包含丰富的入侵信息,并保证入侵检测的有效性,高效保护网络信息的安全,需要实时对数据库的数据进行更新^[5]. 在入侵检测系统中数据的匹配是关键的一环,直接关系到入侵检测的效率和有效性.

但是,传统的基于 BM 算法的入侵检测方法,在数据采集器输送到数据分析器采集数据字符串后,是直接将输入的待检测数据字符串逐个与数据库中的字符串一一匹配比较. 具体的匹配准则是:先比较两组字符串的第一个字符,若匹配则移动到下一个字符进行比较,若不匹配则将数据库中的字符串移动到与待检测字符串第一个字符相同的位置,继续以同样的方式匹配. 当所要检测的网络是高速网络时,由于网络通信信息量较大,通过数据采集器采集到的数据量是大规模的,用传统 BM 方法检测高速

网络入侵,需要将数据采集器采集到的大规模数据逐一与数据库中的字符串比较,这样的逐一检测耗费了大量的时间,不能实现入侵的高效检测.

针对传统基于 BM 算法的入侵检测方法不能快速检测高速网络数据安全的问题,提出一种基于 ACBM 规则匹配的网络入侵算法. 通过定义入侵规则,将数据采集器采集到的入侵数据与数据库中的数据进行规则化,然后根据 ACBM 算法对入侵规则配准,实现入侵检测的匹配,将匹配结果传送到响应模块进行相应的反应处理,这样就避免了传统算法对所有待检测数据串的逐一匹配,有效地缩短了对高速网络入侵检测所耗费的时间,能够有效提高网络入侵检测的效率,对高速网络的检测效率也会有很大的改善.

3 基于 ACBM 规则匹配的入侵检测算法

计算机网络的入侵检测就是通过采用各种的技术和措施,对网络的运行情况进行监测,对网络攻击及时检测出并处理,保证网络系统的安全正常运行,从而保证网络数据的保密性和安全性,满足人们对网络安全的要求. 其中在入侵检测的系统中,最关键的环节就是数据字符串的匹配问题,其匹配的快慢直接影响着入侵检测系统的效率. 传统的基于 BM 的检测方法是将数据字符串逐一比较,这样造成在检测高速网络时,由于要处理的数据量巨大而造成匹配耗时长,造成入侵检测低效的问题. 因此文中提出一种基于 ACBM 规则匹配的算法,解决高速网络检测时数据量巨大带来的耗时长的问题.

3.1 规则的定义

由于通过网络中采集的数据字符串之间有一定的语义联系,这时就可将字符串以一定的规则进行归一化,在数据匹配时直接对规则化后的数据进行匹配,就避免了逐一对字符串的各个字符比较匹配带来的耗时过长的问题. 为了使规则能够包含数据字符串所拥有的全部信息,将规则定义为两部分:规则头和规则项.

(1) 规则头,包含数据字符串的规则行为、起始地址和端口信息. 规则的行为主要有:警告、记录事件、忽略事件、激活规则、退出等. 起始地址就是数据字符串的存储起始 IP 地址. 端口表示的是数据字符串的端口所能达到的范围.

(2) 规则项. 规则项是由常见字符串组成的. 由于入侵数据字符串处会出现语义上的重复,所以将常用的字符串列在规则项中. 比如,比较常用的字符串

有msg(在日志中提取的警告信息)、content(某个事件的可疑模式)、content-list(搜索事件模式的集合)、depth(字符串匹配的长度值)、offset(字符串匹配时的可能偏移量)、nocase(对字符串信息的写入威胁)。这样得到规则定义的结构图如图2所示。

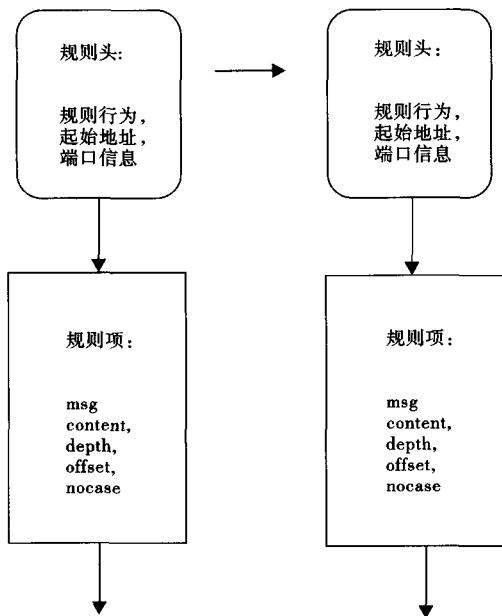


图2 规则定义结构图

3.2 基于ACBM算法的规则匹配

入侵检测系统中,通过数据采集器将网络中的可疑信息数据采集,并将采集的数据转换为字符串的形式后输送到系统的系一个单元,即数据分析器中。然后将数据分析器中的字符串与数据库中的字符串匹配处理,这一环节也是系统的关键环节,决定了入侵检测的效率。ACBM算法是对传统的单一匹配模式的BM算法的一种改进,通过将数据字符串规则化避免了对单一字符串的逐个比较。ACBM算法将所有的数据规则放在一棵树上,这棵树称为规则树,又称模式树。通过对规则树的逐层遍历实现数据规则的匹配,即完成最终的采集数据信息与数据库中信息的匹配。ACBM规则匹配的流程图如图3所示。

ACBM规则匹配的具体算法为:

(1) 构造数据库中的攻击数据规则树。从起始地址上的规则开始,挨个将数据库中的规则插入到规则树上,沿着规则树从左到右、从父节点到子节点的顺序,在最后数据规则结束的位置用标识符标示,并完成最后的根节点的转向函数的确立。

(2) 从规则树的第一个状态字符串出发,按照广度优先搜索算法开始逐个选取规则树上的字符串与数据库中的字符串相匹配,如果匹配成功,则根据

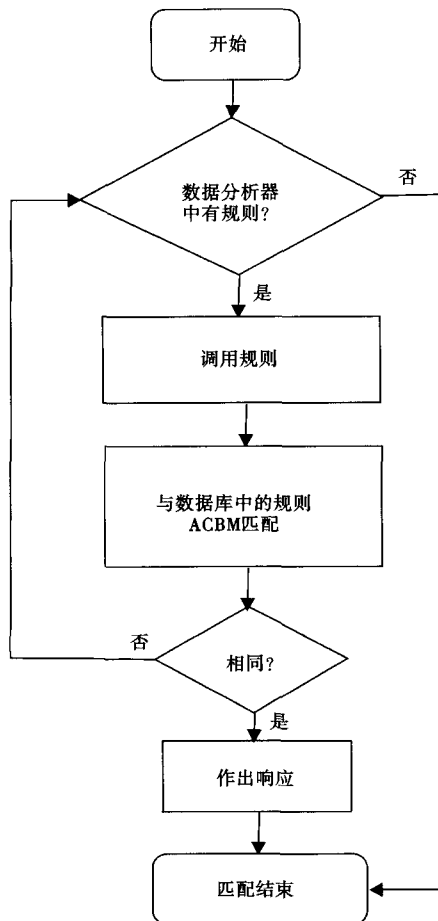


图3 ACBM规则匹配流程图

规则树上的F函数到达相应的响应模块,如果匹配不成功则转向下一个规则树状态,直至规则树的标示的结束位置标识符,即完成了对数据的匹配。下面列出一个ACBM规则树,如图4所示。

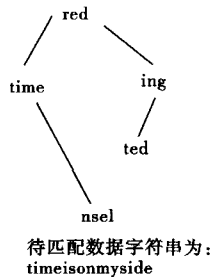


图4 ACBM匹配规则树示例

由图4可以看出,在进行ACBM规则树匹配时,需要广度优先遍历此规则树,遍历时是按照字符串构成的规则进行遍历的,而不是对数据字符串单个字符的逐一遍历,这样就避免了传统方法检测高速网络时检测效率不高的问题。当在遍历过程中匹配成功时,直接转向此规则相应的响应模块进行响应,及时保护网络信息的安全。

3.3 数据库升级

因为网络资源共享时随时会出现网络漏洞,针对每个漏洞会有很多全新的攻击威胁,为了提高入侵检测的有效性,保证对网络安全的实时保护,需要将数据库实时更新,及时找出应对各种网络攻击的新方法,来保证网络的安全.对数据库进行网络升级的流程如图 5 所示.

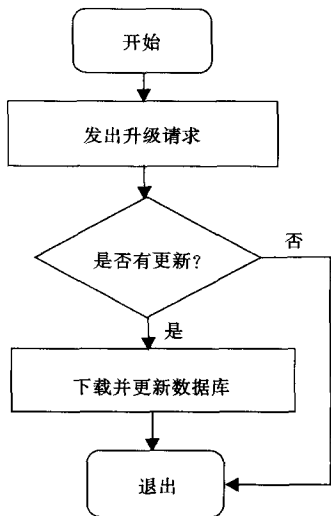


图 5 数据库更新流程图

这样通过将数据规则化,对规则化后的数据字符进行匹配,就避免了传统方法直接对数据字符串的单个字符逐一匹配,大大缩短了匹配耗时,提高了对高速网络入侵检测的效率.

4 实验结果及分析

基于 ACBM 规则匹配的高速网络入侵检测算法,通过将数据规则化后直接对数据规则进行匹配,有效提高了入侵检测效率,其具体的算法流程图如图 6 所示.

实验构建一个高速网络,分别采用传统的基于 BM 的算法和文中提出了基于 ACBM 规则匹配的算法对此高速网络的入侵情况进行检测,记录两种方法分别的检测耗时,并将两种方法耗时柱状图进行对比,耗时对比柱状图如图 7 所示.两种检测方法实验数据的对比如表 1 所示.

表 1 实验数据对比

检测方法	算法复杂度	平均检测耗时
基于 BM	$O(mn)$	41.73 s
基于 ACBM	$O(n)$	27.88 s

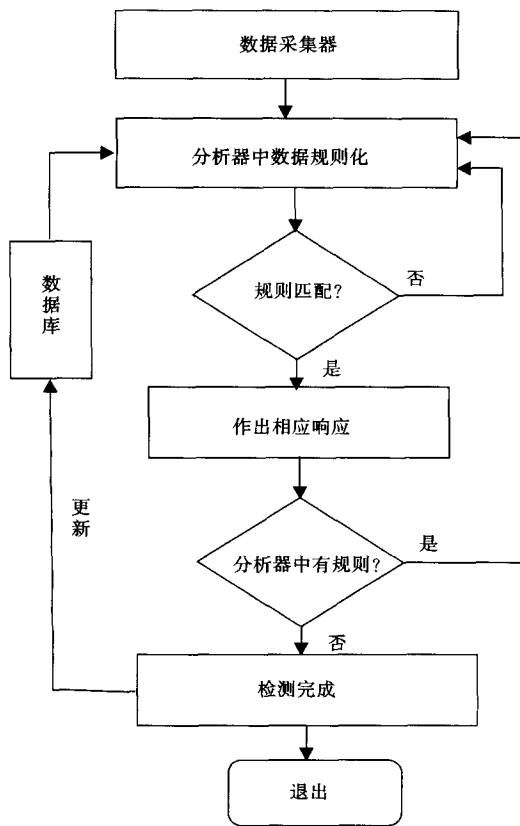


图 6 基于 ACBM 入侵检测算法流程图

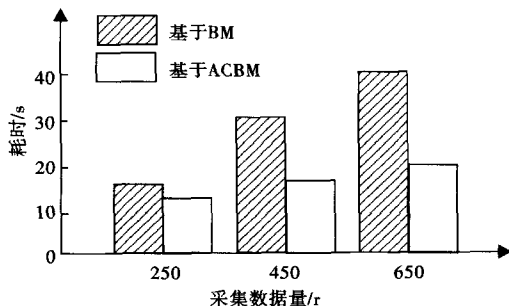


图 7 耗时对比柱状图

由表 1 可知,传统基于 BM 的算法逐一对数据字符串进行匹配,算法复杂度很高,平均检测耗时是文中改进方法的约 1.5 倍.可见基于 ACBM 的检测方法,由于将数据规则化,避免了直接对字符的逐一匹配,而是对数据规则的匹配,算法复杂度比传统方法小,检测耗时也远远优于传统方法.由图 7 的检测耗时柱状图可知,随着数据采集器采集数据量的增大,即所要检测的数据量的增大,传统方法由于工作量的加大,逐一检测数据字符串造成耗时逐渐增大.而基于 ACBM 算法随着工作量的加大其工作耗时变化不大,一直处在一个稳定的范围内.综上可知,基于 ACBM 规则匹配的检测算法,通过将数据规则

(下转第 42 页)

知,相对于 QPSO-BPNN 算法,基于 QPSO-LSSVC 算法提高了网络入侵的检测准确率,加快检测速度,结果表明 LSSVC 很好的解决了 BPNN 局部极值和收敛速度慢等缺陷,十分适合于现代网络异常检测的实时性和检测准确率高的要求。

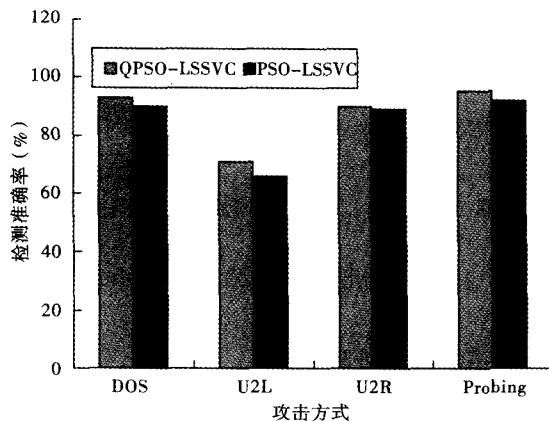


图 3 QPSO-LSSVC 和 QPSO-BPNN 结果对比

6 结束语

针对 LSSVC 在网络入侵检测中参数优化问题,文中提出采用 QPSO 对 LSSVC 参数进行优化,并通过仿真实验进行验证。仿真实验结果表明,QP-SO-LSSVM 解决传统 LSSVC 模型参数优化的难题,提高网络异常入侵检测检测准确率,降低误报率,加快了网络入侵检测速度,在网络安全应用中具有广泛的应用前景。

(上接第 38 页)

化后对规则直接匹配,避免了对数据字符串的逐一匹配,大大缩短了检测耗时,提高了检测效率,实现了高速网络中入侵的快速检测,满足了人们对网络入侵检测的实时性要求。

5 结束语

本文提出一种基于 ACBM 规则匹配的网络入侵算法。通过定义入侵规则,将入侵数据与数据库中的数据规则化,然后根据 ACBM 算法实现入侵规则的匹配,这样就避免了传统算法对单一数据串的逐一匹配,大大缩短了高速网络检测所消耗的时间。实验证明,这种方法能够有效提高网络入侵检测的效率,对高速网络的检测效率也有了很大的改善,满足了人们对检测实时性的要求,具有很高的实用价值。

参考文献:

- [1] Taeshik Shon, Jongsub Moon. A hybrid machine learning approach to network anomaly detection[J]. Information Sciences, 2007, 177 (18): 3799-3821.
- [2] Ozgur Depren, Murat Topallar, Emin Anarim, et al. An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks[J]. Expert Systems with Applications, 2005, 29(4): 713-722.
- [3] 赵俊忠. 入侵检测系统中检测技术的研究[J]. 计算机工程与应用, 2005(2): 11-13.
- [4] 卿斯汉, 蒋建春, 等. 入侵检测技术研究综述[J]. 通信学报, 2004, 25(7): 9-29.
- [5] 刘渊, 马汝辉, 林星. 基于 QPSO 小波神经网络的网络异常检测[J]. 宁波工程技术大学学报: 自然科学版, 2009, 28(2): 261-264.
- [6] 许劲松, 覃俊. 一种基于支持向量机的入侵检测模型[J]. 计算机仿真, 2004, 22(5): 43-55.
- [7] 李旭芳, 王士同. 基于 QPSO 训练支持向量机的网络入侵检测[J]. 计算机工程与设计, 2008, 29(1): 34-36.
- [8] 池越, 夏克文, 陈国鹰, 等. 一种基于 QPSO 和 WLS-SVM 智能方法[J]. 河北工业大学学报, 2009, 38(2): 9-15.

作者简介:

姚 晔 女, (1973—), 硕士, 副教授. 研究方向为计算机应用及网络。

参考文献:

- [1] 王哲, 阮永良. 异常入侵检测技术综述[J]. 电脑知识与技术, 2004(20): 56-59.
- [2] 郁伟, 胡蓓, 陈昌芳. 入侵诱骗系统中的数据捕获[J]. 网络安全技术与应用, 2007(3): 36-38.
- [3] 朱元忠. 入侵检测系统发展研究[J]. 北京工业职业技术学院学报, 2006(1): 8-11.
- [4] 杨光, 巫春玲, 程远征. 基于 RS 和 WSVM 的网络入侵检测算法研究[J]. 计算机仿真, 2011(5): 175-178.
- [5] 彭诗力, 谭汉松. 基于特征值的多模式匹配算法及硬件实现[J]. 计算机工程与应用, 2005(1): 148-150.

作者简介:

谈娟茹 女, (1972—), 硕士, 讲师. 研究方向为计算机硬件、嵌入式系统、软件工程。