基于小波分析的无线传感网实时异常检测算法

李致远,朱求志,吴永焜,唐振宇,胡华明

(江苏大学计算机科学与通信工程学院,江苏 镇江 212013)

[摘要] 异常检测技术能够检测到未知攻击,对于保障无线传感器网络安全具有重要意义. 当前的异常检测技术实时性差,误报率高且计算量大,因此,无法直接应用在无线传感器网络中. 鉴于此,提出基于小波分析的实时无线传感网异常检测(Wavelet Analysis-Based Real-time Anomaly Detection,WARAD)算法. 在整个检测过程中,WARAD 算法采用了逆向获取实时网络流量,然后通过对小尺度区间使用小波系数方差法计算 Hurst 值,从而提高异常检测的实时性、准确率,并降低求解 Hurst 值的运算复杂度. 最后,在 MeshIDE 平台上实现了基于 WARAD 算法的异常检测系统,实验结果表明此算法极大地提高了无线传感网环境下异常检测的实时性,并降低了异常检测的误报率和漏报率.

[关键词] 无线传感器网络,安全,异常检测,小波分析,Hurst 参数

[中图分类号]TP393 [文献标志码]A [文章编号]1001-4616(2014)01-0087-06

Wavelet Analysis-Based Real-Time Anomaly Detection Algorithm for Wireless Sensor Network

Li Zhiyuan,Zhu Qiuzhi,Wu Yongkun,Tang Zhenyu,Hu Huaming

(School of Computer Science and Telecommunication Engineering,Jiangsu University,Zhenjiang 212013,China)

Abstract: Anomaly detection can detect new and unknown attacks, which has great significance on the wireless sensor networks security. Nowadays, the proposed anomaly detection schemes has poor real-time, high false positive rate and the large amount of computational overhead, and hence the schemes are not suitable for wireless sensor networks. In this paper, a wavelet analysis-based real-time anomaly detection (Wavelet Analysis-based Real-time Anomaly Detection, WARAD) algorithm for wireless sensor network is proposed. Throughout the detecting process, the WARAD algorithm reversely collects the real-time network traffic, and then uses the variance of the wavelet coefficients in the small-scale interval to compute the Hurst values, which can improve the real-time and the accuracy of anomaly detection, and reduce the computational complexity of solving the Hurst values. Finally, the WARAD algorithm-based intrusion detection system is implemented on the platform of MeshIDE. The experimental results showed that the proposed algorithm greatly improved the real-time of anomaly detection for wireless sensor networks, and reduced the false positive rate and the false negative rate of anomaly detection.

Key words: wireless sensor networks, security, anomaly detection, wavelet analysis, Hurst parameter

传感器网络的无线传输和无人看守等固有特性使它极易受到各种恶意攻击. 通常,攻击者通过干扰无线信道,大量重放插入或丢弃报文等手段发起攻击,消耗传感器节点有限的资源,使网络部分或全部瘫痪. 因此,当传感器网络被用于具有重要使命的场景时,如何迅速、准确地检测出 DoS 攻击,保证网络设施的基本功能可用,对整个任务的成败至关重要.

异常检测技术能检测到一些未知攻击,是目前入侵检测的主流研究方向. 它主要是对网络正常行为进行建模,将不符合正常模型的事件定义为攻击. 当前异常检测的主要方法有概率统计、机器学习、时间序列

收稿日期:2013-07-01.

基金项目:国家自然科学基金(61202474、61103195)、江苏省自然科学基金(BK20130528)、江苏大学高级专业人才科研启动基金项目(12JDG049)、江苏大学本科生创新计划项目(2012075).

通讯联系人:李致远,博士,讲师,研究方向:无线传感器网络安全. E-mail:lizhiyuan81@126.com

分析等技术,尽管这些技术能够学习到较好的检测模型,然而由于无线传感器网络节点能量和计算能力受限等特征,传统的入侵检测方法无法直接应用在无线传感器中,实时检测的性能问题尚未很好地解决.

本文借助于网络数据流分形特性与小波分析解决无线传感网的实时异常检测问题.大量研究表明,无线传感器网络数据流具有分形特性和自相似特性.当网络数据流中存在入侵时,网络数据流的分形特性将会受到干扰,其Hurst及频谱与正常情况下相比将会有较大的区别.小波分析方法是一种窗口大小固定,而其形状、时间窗和频率窗可改变的时频局部化分析方法.这种特性使得小波分析在分形信号处理和分形参数估计中显示出其多分辨率时频分析的独特优势,但是小波系数的计算需要耗费大量的计算资源,这是影响实时检测和传感器能耗的最重要因素.本文采用逆向获取实时网络流量数据,通过小尺度计算Hurst值,以及小波系数方差变化等多种优化策略提高传感网异常流量检测的实时性,并降低了异常检测的误报率和漏报率.

1 相关工作

传统的网络异常检测主要采用统计、机器学习、时间序列分析建立主观检测模型,根据当前的行为与模型的偏离来检测异常.下面归纳出4类常见的异常检测模型:

(1)基于机器学习和数据挖掘的异常检测.包括基于隐马尔科夫模型的异常检测系统[1],基于马尔科夫链的异常检测模型[2],基于免疫遗传算法的异常检测[3]及基于BP神经网络的异常检测[4].这些方法将异常检测作为分类或聚类问题,借助机器学习的有效学习能力,构建具有一定精度的异常检测模型,具有较高的准确率,但是,不足之处是需要的样本量较大,训练时间长.近年来颇受关注的是基于支持向量机的异常检测.文献[5]利用混合的无监督聚类方法和支持向量机(Support Vector Machine,SVM)算法提高异常检测的快速性和精确性.SVM克服传统机器学习方法大样本的缺陷,根据有限的样本信息在模型的复杂性和学习能力之间寻求最佳平衡点,能获得最好的范化能力.但SVM在训练之前必须进行模型选择,要确定核函数,目前核函数的选择往往凭经验进行,此外,SVM一般只能处理二元分类问题,如果要区分多种入侵方式,就要使用多个SVM来实现,代价高并影响异常检测的实时性.

(2)基于信任的异常检测.文献[6]首先提出了层次化的信任管理模型,然后基于该信任模型提出动态的入侵检测方案.该方案利用具有较高信任度的节点来交替检测簇内节点,达到了实时甄别恶意节点和自私节点的目标.文献[7]提出了基于加权信任机制的入侵检测方法,在系统开始就给每个传感器节点分配权重,每个周期当节点发送与其他节点不同的报告时进行更新,这样当节点的权重低于某一阈值时就被检测出是恶意检点.这些方法具有低功耗、高安全性等优点,但当簇头节点入侵,或遇到Sybil攻击时检测精确度降低,而且其阈值设置会影响算法的精度,如何找到合适的阈值是等待解决的问题.

(3)基于博弈的异常检测.文献[8]在移动传感器节点和入侵者之间提出零和博弈下的Nash均衡混合策略来实时检测入侵节点的异常行为,这种方法能够帮助传感网实现可信安全的全覆盖.然而,初始博弈矩阵的设置随环境变换而变化,博弈矩阵的变化必然引起最终博弈策略的变化,因此,该类方法缺乏通用性和灵活性.

(4)基于小波分析的异常检测.文献[9]提出了一种新的基于小波包分析的网络数据流异常检测新机制,对高、中、低频异常数据流具有同样的检测能力.该类异常检测方法具有较好的实时性,适合在线异常检测,且具有很好的应用前景.然而,小波系数的计算需要耗费大量的计算资源,这是影响实时检测和传感器能耗的最重要因素.

综上所述,目前的各种异常检测技术还不能实时、准确地对各种入侵进行检测.鉴于小波分析在信号处理中具有独特的频谱多分辨率优势,基于频谱、小波分析的异常检测被证实适合实时的异常检测,因此,本文主要对基于小波分析的无线传感器网络实时快速异常检测模型进行研究.重点在于解决小波系数的快速绿色计算难题.

2 基于小波分析的实时异常检测法

Hurst参数是网络数据流所具有的分形指数.目前常通过Hurst值的变化来判断传感网是否遭受到异常网络攻击.本节提出一种实时的小波系数方差法(On_Time Wavelet Coefficient Variance,On_Time

WCV),如图 1 所示.

(1)逆向滑动窗口,小尺度计算 Hurst 策略

如图 1 所示,传统的 WCV 采用顺序工作,Phase 0 阶段是抓包时间 T_c,将 Phase 1→Phase 2→Phase 3→Phase 4 定义为小波分析时间 T_w,那么有效分析时间率 η 为 $\frac{T_w}{T_w+T_c}$. 传统 WCV 的缺陷在于按照 Phase 1 到 Phase 4 顺序采集流量的方式具有滞后性,即攻击可能已发生过,致使不能实时检测出当前的异常攻击. 而本文所提出的 On_Time WCV 是将大序列分析划分为若干个小段,使用时间窗口从当前时刻逆向抓取少量的实时流量进行分析,称之为小尺度求解 Hurst. 尽管只取了少量数据,根据小波分析时多分辨率特点,只要满足信号长度大于滤波器长度即能够完全解析出信号的特征,因此,可以较快地计算出 Hurst 值.

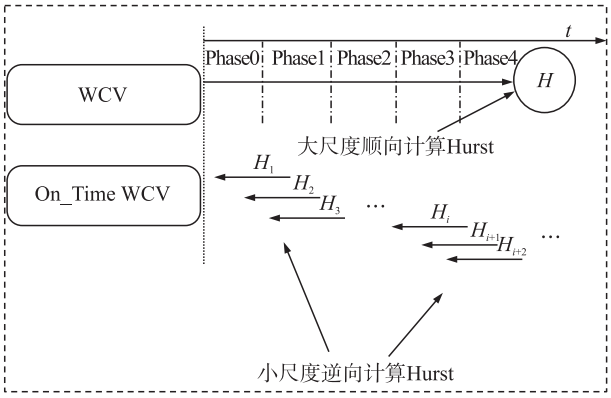

图 1　基于实时小波系数方差法的异常检测

Fig. 1　The anomaly detection based on real-time wavelet coefficient variance

考虑到网络数据包的动态变化,逆向时间窗口根据实时数据包速率进行动态调整,可以保障时间序列与小波分析数据的可靠性. 通过以上策略,可以减少抓包时间 T_c,提高 η 和分析时间 T_w 的频次. 因此,采用逆向滑动窗口、小尺度计算 Hurst 值策略对实时数据的快速获取及有效提高异常检测性能有重要的作用.

(2)采用小波方差变化及并行策略求解 Hurst 参数

①小波方差变化法求解 Hurst 值

设$\{x(t)\}$为统计自相似过程,对$\{x(t)\}$作小波变换,得到小波系数:

$$d_l^{(j)}=2^{\frac{j}{2}}\int x(t)\psi(2^jt-l)\,\mathrm{d}t. \tag{1}$$

其中,$\psi_{j,l(t)}=2^{\frac{j}{2}}\psi(2^jt-l)\,\mathrm{d}t$ 为二进正交小波,其正则度为 R. 小波系数的期望:

$$E[d_l^{(j)}]=E[x(t)]2^{\frac{j}{2}}\int\psi(2^jt-l)\,\mathrm{d}t=0. \tag{2}$$

根据相关系数的定义,任意 2 个小波变换系数 $d_l^{(j)}$,$d_{l'}^{(j')}$ 间的相关值见式(3):

$$E[d_l^{(j)},d_{l'}^{(j')}]=\iint E[x(t)\psi_{j,l}(t)x(t')\psi_{j',l'}(t')]d_td_{t'}=\int\psi_{j,l}(t)[R_x(t)*\psi_{j',l'}(t')]\,\mathrm{d}t. \tag{3}$$

对式(3)作傅立叶变换并用 Parseval 公式,得到式(4):

$$E[d_l^{(j)},d_{l'}^{(j')}]=\frac{2^{\frac{-j+j'}{2}}}{2\pi}\int_{-\infty}^{+\infty}\frac{\sigma_x^2}{|\omega|^\gamma}\hat{\psi}(2^{-j}\omega)\overline{\hat{\psi}(2^{-j'}\omega)}\,\mathrm{e}^{-i[l2^{-j}-l'2^{-j'}]\omega]}\,\mathrm{d}\omega. \tag{4}$$

由式(2)和(3)得到 $d_l^{(j)}$ 的方差见式(5):

$$\sigma^2=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{\sigma_x^2}{|\omega|^\gamma}|\hat{\psi}(\omega)|^2\mathrm{d}\omega. \tag{5}$$

由式(4)和(5)得到式(6):

$$Var[d_l^{(j)}]=\sigma^2 2^{-j\gamma}. \tag{6}$$

通过对式(6)两边取对数后,在均方差最小的条件下进行线性拟合得到以 j 为自变量,以 $\log_2^{\mathrm{var}[d_l^{(j)}]}$ 为函数的直线,斜率即为 γ. 利用 $\gamma=2H+1$ 便可求得 H 值.

②并行的小波方差变化策略求解 Hurst 参数

在 Phase 0 执行结束后,再对 Phase 1 到 Phase 4 阶段进行时间序列划分,统计得到$\{x(t)\}$. 由于 Phase 1 在 Phase 0 阶段得到一部分数据后即可启动,对 Phase 0 后续得到数据没有依赖性,因此,将抓取数据包阶段 Phase 0 与时间序列划分阶段(Phase 1 到 Phase 4)按照流水线理论进行并行处理.

(3)利用改进的小波系数方差变化法判定网络异常攻击

在用传统 WCV 方法判定攻击时,首先依据正常数据流求出 Hurst 值,建立一个正常模型,以此为判定依据,对当前抓获的数据报进行 Hurst 求解,最后进行比较判定是否发生了网络攻击. 显然,这对于弱攻击

检测是可行的,因为数据量小,Hurst 参数求解速度较快. 然而,当攻击数据量较大(强攻击)时,待完全求解出 Hurst 参数再判定,之后再采取措施,就难以满足实时性异常检测与防御的需求. 鉴于强攻击会使得 Hurst 参数发生明显变化,而 Hurst 参数是各个分解级数上小波系数方差拟和而成的,因此,方差必然会产生明显的变化,这就是能量分布变化检测攻击的原理. 因此,本文以检测到相邻级别的小波系数方差发生明显变化时判断异常发生,就无需求出全部的小波系数方差,进而达到快速检测强攻击的目的.

3　WARAD 算法实现与分析

3.1　WARAD 算法的实现

在南京邮电大学江苏省无线传感网高技术研究重点实验室研发的面向 nesC 的无线传感器网络集成开发平台软件 MeshIDE 上实现了 WARAD 算法. MeshIDE 集成开发平台为用户提供了开发 nesC 应用程序的友好接口. 在 MeshIDE 平台上,用户可以自动创建 . project 项目文件和 makefile 文件;然后,通过创建 nesC 文件生成 . nc 文件;最后通过 Make 指令实现代码编译,并通过串口将可执行程序烧制到传感器节点上.

基于 WARAD 算法的入侵检测系统的实现步骤如下,初始化界面如图 2 所示.

步骤 1:通过网络数据包接收器从网络层接收网络数据包. 为了搜集传感网内的数据包,必须将网卡设置为混杂模型.

步骤 2:从获得的数据包中,提取求解 Hurst 值所需要的特征量,在此需要 2 个特征量,一个是时间,一个是数据包大小. 将这些特征存储于特征库.

步骤 3:得到这 2 个特征后,对数据包进行时间序列划分.

步骤 4:采用本文提出的 WARAD 小波分析法求解 Hurst 值.

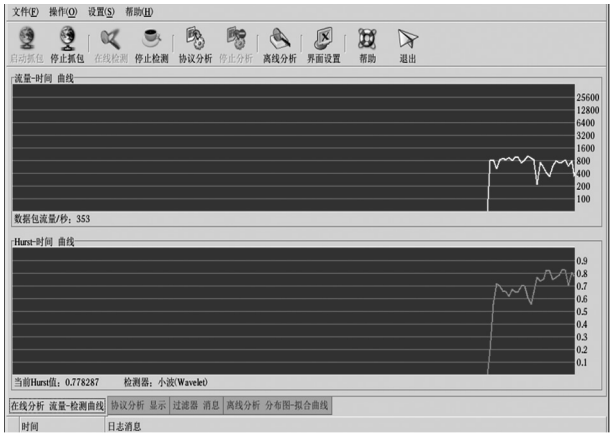

图 2　异常检测软件初始化界面

Fig. 2　The initialization interface of anomaly detection software

步骤 5:根据经验阈值判断当前的 Hurst 值是否正常,以及攻击是否发生.

采用权威数据回放的方式对系统的检测性能进行测试. 这里用 tcpreplay 工具分别以 20 倍和 50 倍速率回放绿野千传项目 2010 年产生的正常数据集,如图 3 和图 4 所示. 对于正常流量,Hurst 值都在 0.55 以上,且“Hurst-时间”曲线较平稳,说明异常检测系统对正常流量能准确稳定地识别.

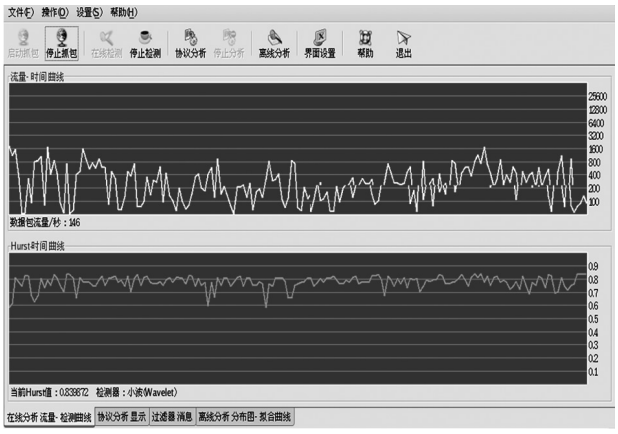

图 3　20 倍速回放正常数据集

Fig. 3　20 times playback of the normal data set

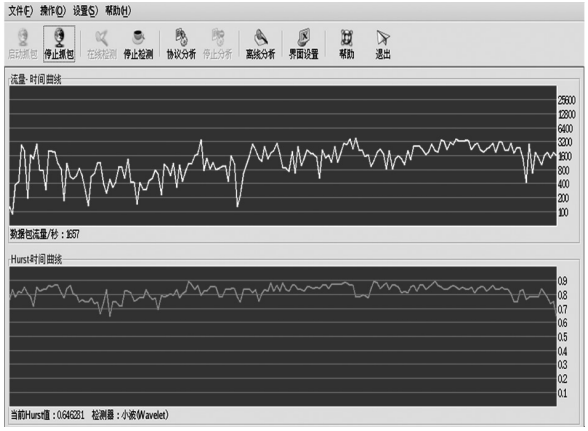

图 4　50 倍速回放正常数据集

Fig. 4　50 times playback of the normal data set

然后,添加 5 个恶意传感器节点,让它们依次分别发起 TCP SYN Flooding 和 UDP Flooding. 异常检测结果如图 5 和图 6 所示.

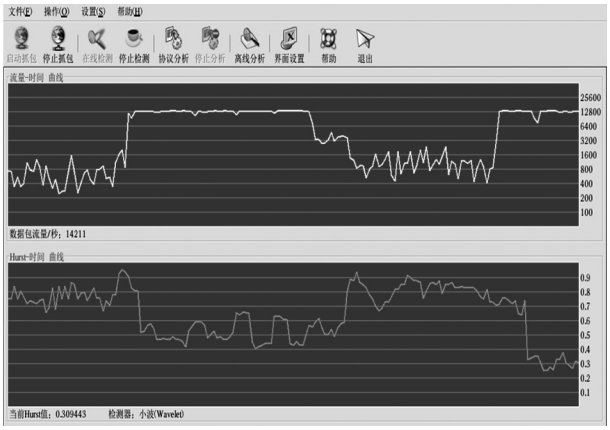

图 5　20 倍速回放正常数据集,5 个 TCP 方式攻击

Fig. 5　20 times playback of the normal data set, and 5 TCP connection attack

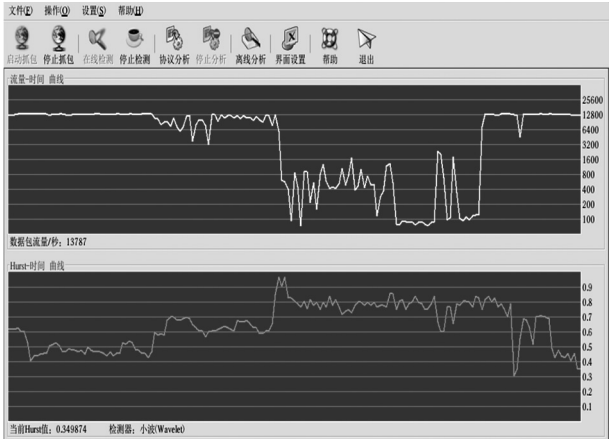

图 6　20 倍速回放正常数据集,5 个 UDP 方式攻击

Fig. 6　20 times playback of the normal data set, and 5 UDP connection attack

从实验结果看,攻击开始后,“Hurst-时间”曲线发生明显的变化,表现为 Hurst 值急剧下降,且小于 0.65;统计检测时延在 3 s 内;停止攻击后,“Hurst-时间”曲线开始上升,Hurst 值恢复到 0.65 以上,且保持相对平稳.

3.2　WARAD 算法的性能分析

本节对算法的性能进行深入地分析,包括检测准确率和误检率. 其中检测准确率表征准确检测到的恶意报文数量与全部报文数量的比值,误检率表征将检测到的正常报文误认为恶意报文的报文数量与全部检测到的报文数量比值.

(1)算法检测的准确率

在本次实验中,验证了当恶意攻击(TCP 和 UDP 方式攻击)的连接数从 1 到 15 递增时文献[2,3]所提出的算法以及 WARAD 算法的检测率和过警率. 整个实验反复进行了 10 次,取平均值作为实验结果,具体如图 7 所示.

结果表明:恶意攻击的连接数较低时,3 种算法的检测率都能够保持较高的水平,同时误检率都能控制在较低的水平. 比如连接数为 1 时,3 种算法的检测率为 100%,而误检率被控制在 2% 以下. 然而,随着恶意攻击连接数的增加,文献[2,3]的检测率迅速下降,同时误检率迅速上升. 而 WARAD 算法却始终保持着 95% 以上的检测准确率和 1.3% 以下的误检率. 这就表明,上述模型的检测率和误检率与未知攻击的强度有直接的关系. 当面对强度较高的未知攻击时,文献[2]所提出的基于攻击模式挖掘的马尔科夫入侵检测模型和文献[3]所提出的基于人工免疫技术的危险理论模型都无法准确地检测出新的恶意攻击,且其误检率也大大增加. 而本文提出的 WARAD 算法则利用了被誉为“数学显微镜”的小波多尺度分析技术,极大地提高了未知恶意攻击的检测准确率并降低了误检率.

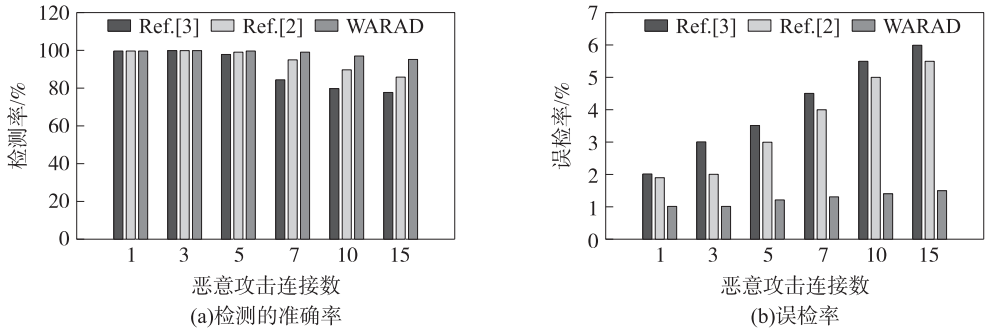

图 7　检测准确度

Fig. 7　The accuracy of anomaly detection

(2)算法检测的实时性

图 8 所示为 3 种算法检测恶意攻击的实时性能分析. 横坐标表示报文的回放速率,纵坐标表示检测到

恶意攻击所需要的时间. 回放的报文包含正常的通信和数据报文,常规攻击所产生的攻击报文,以及未知攻击所产生的攻击报文.

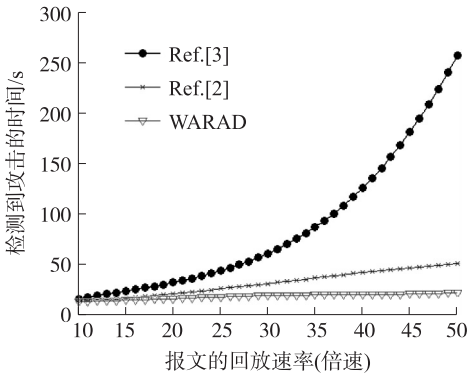

图 8　检测的实时性

Fig. 8　Real-time of anomaly detection

图 8 的实验结果表明:文献[3]检测到的攻击时间随着回放速率的增加呈指数增长,这主要是由于文章所采用的危险理论主要是针对正常报文和常规攻击报文的区分,而对于新型的异常攻击缺乏相应理论支撑和技术支持;文献[2]和 WARAD 算法检测到的攻击时间随着回放速率的增加呈线性增长,WARAD 算法的斜率更小,因此,其实时性更强. 这主要是由于文献[2]采用的未知攻击模式的数据挖掘方法需要的样本空间较大,若样本空间达不到要求,则无法检测出异常攻击. 此外,文献[2]提出的数据挖掘算法的收敛速度较慢,达不到异常检测实时性的需求. 本文提出的 WARAD 算法采用了逆向获取实时网络流量,然后通过对小尺度区间使用小波系数方差法计算 Hurst 值,从而提高异常检测的实时性、准确率,并降低求解 Hurst 值的运算复杂度.

4　结束语

本文针对目前的异常检测技术不能实时、有效地检测异常攻击的现状,提出了基于小波分析的无线传感器网络实时快速异常检测算法 WARAD. WARAD 算法的创新点在于使用逆向滑动窗口、小尺度及流水线策略实时准确计算 Hurst 值;在检测到相邻级别的小波系数方差发生明显变化时判定异常攻击发生,无需求解全部小波系数方差,进而达到快速检测强攻击的目的. 之后,在面向 nesC 的无线传感器网络集成开发平台软件 MeshIDE 上实现了 WARAD 算法. 最后,采用权威数据回放的方式对系统的检测性能进行测试. 实验结果表明 WARAD 算法可以达到实时快速检测网络异常流量的目标. WARAD 算法的主要贡献在于为传感网实时异常检测的进一步研究与应用提供了新的思路与理论及实验数据.

[参考文献]

[1] Du Y, Yang S, Zhang R H. Design of an intrusion detection system for wireless sensor networks[J]. Sensor Letters, 2011, 9(5): 2 082-2 086.

[2] Huang J Y, Liao I E, Chung Y F, et al. Shielding wireless sensor network using Markovian intrusion detection system with attack pattern mining[J]. Information Sciences, 2013, 231: 32-44.

[3] 傅蓉蓉,郑康锋,芦天亮,等. 基于危险理论的无线传感器网络入侵检测模型[J]. 通信学报,2012,33(9):31-37.

[4] Yan K Q, Wang S C, Wang S S, et al. Hybrid intrusion detection system for enhancing the security of a cluster-based wireless sensor network[C]//Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT'10). Chengdu: IEEE, 2010: 114-118.

[5] Sedjelmaci H, Feham M. Novel hybrid intrusion detection system for clustered wireless sensor network[J]. International Journal of Network Security and Its Applications, 2011, 3(4): 1-14.

[6] Bao F, Chen I R, Chang M J. Hierarchical trust management for wireless sensor networks and its applications to trust-based routing and intrusion detection[J]. IEEE Transactions on Network and Service Management, 2012, 9(2): 169-183.

[7] Srikanth H, Shroff N B, Saurabh B. Secure neighbor discovery through overhearing in static multihop wireless networks[J]. Computer Networks, 2011, 55(6): 1 229-1 241.

[8] Liu B, Olivier D, Philippe N. Dynamic coverage of mobile sensor networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(2): 301-311.

[9] Gao J, Hu G, Yao X. Anomaly detection of network traffic based on wavelet packet[C]//Proceedings of the Asia-Pacific Conference on Communications. Busan, Korea: APCC, 2006: 1-5.

[责任编辑:黄　敏]