

改进 Boyer 匹配算法在 Snort 入侵检测中的应用

马小雨¹, 刘双红²

(1. 河南工程学院 计算机学院, 河南 郑州 451191;
2. 郑州航空工业管理学院 计算机科学与技术系, 河南 郑州 450046)

摘要: 以 Snort 入侵检测系统为研究对象,探讨其规则匹配环节的适用算法,并在 Boyer 算法的基础上设计一种改进方法.此方法首先设计了一个统计数组,然后以两个相邻字符为组合执行匹配,并分为 3 种策略判断如何确定最大移动长度.实验结果表明:这种改进措施,使得最大移动长度更加合理,相比于 Boyer 方法,改进方法的字符比较次数明显降低,窗口移动次数明显降低,执行时间明显减少.

关键词: 网络安全;入侵检测;Snort 系统;Boyer 算法

中图分类号: TP 393.08 **文献标志码:** A

为了应对黑客攻击风险、协议漏洞风险、防御缺陷风险等一系列网络间的安全隐患^[1-2],信息安全、密码理论、入侵检测等领域的研究被大量开展^[3-5].Snort 入侵检测系统是一个非常成功的网络安全检测系统,具有免费开源、精准匹配的诸多优点^[6-7],且在大多数操作系统上都可以运行,这对国内普遍使用的 Windows 系统具有重要意义.从执行流程上看,Snort 入侵检测系统以字符为单位对网络信息进行解读,分为捕获数据、解码数据、预处理数据、规则匹配、判断输出等环节^[7-8].规则匹配是用于判断信息是否为安全的核心环节,在这个阶段中,很多字符匹配算法被成功应用,而 Boyer 匹配算法是一种最常见的方法^[9].Boyer 匹配算法原理简单,便于实现,但也存在重复比较等问题.为此,很多基于 Boyer 算法的改进算法被构建出来,如 BMH 算法、BMHS 算法等^[10-11].本文在 Boyer 算法的基础上进行改进,提出更适合 Snort 入侵检测系统使用的规则匹配算法.

1 改进 Boyer 匹配算法

从实现原理上看,经典的 Boyer 匹配算法属于典型的后缀匹配方法.在字符串匹配执行的过程中,设置一个标准串,将标准串和待匹配的字符串,按照由右及左的顺序执行比较.匹配过程是否结束的判断依据有 2 个:第 1 种是达到预先设定的,确定为匹配失败;第 2 种是确定为匹配成功.整个匹配过程,都需要根据两个重要的特征字符指引,一个称为“优良后缀”,一个称为“不良字符”.

在实际运用中,为了减少带匹配字符串和标准串的比较次数,Boyer 匹配算法增大了移动距离.据此,在经典 Boyer 匹配算法的基础上进行改进,以期获得更大的移动距离.

首先,根据经典的 Boyer 匹配算法获得一个移动距离 L_1 .然后,把 $T[k+1]$ 和 $T[k+2]$ 组合起来,再计算出一个移动距离 L_2 .最后,比较这两个移动距离的大小,选取最大的那个作为匹配过程执行时的移动距离.其中: T 为表示待匹配字符串; k 为一个以 i 为边界,同标准串后缀能够匹配上的最大长度; B 为标准字符串.算法有如下 4 个实际的执行过程.

1) 设置一个统计数组,用于统计各个字符在标准串中出现的次数,即

$$D_{\text{time}}(\text{Char}) = \begin{cases} 1, & B_{\text{time}}(\text{Char}) = 1, \\ 0, & B_{\text{time}}(\text{Char}) > 1. \end{cases} \tag{1}$$

式(1)中; $B_{\text{time}}(\text{Char})$ 为字符 Char 在标准串中出现的次数。如果字符 Char 在标准串 B 中只出现 1 次,统计数组纪录为 1;如果字符 Char 在标准串 B 中出现的次数大于 1,统计数据纪录为 0。

2) 从待匹配字符串中按照 $T[k+1]$ 和 $T[k+2]$ 的顺序提取相邻的字符组合,并纪录为 $\text{Char}(1,2)$,并用 $\text{Char}(1,2)$ 和标准串 B 进行匹配。

3) 若标准串 B 没有字符组合和 $\text{Char}(1,2)$ 相同,需要进一步比较 $\text{Char}(1,2)$ 和标准串的第 1 个字符 $B[1]$ 。如果 $\text{Char}(1,2)$ 不含字符 $B[1]$, L_2 的值就是 $m+2$;如果 $\text{Char}(1,2)$ 中的第 1 个字符和 $B[1]$ 相等,那么 L_2 的值是 $m+2$;如果 $\text{Char}(1,2)$ 中的第 2 个字符和 $B[1]$ 相等,那么 L_2 的值是 $m+1$ 。

4) 若标准串 B 中有字符组合和 $\text{Char}(1,2)$ 相同,且 $\text{Char}(1,2)$ 中不含字符 $B[1]$, $D_{\text{time}}(\text{Char})=1$,那么 L_2 的值是 $m+2$;若标准串 B 中有字符组合和 $\text{Char}(1,2)$ 相同,且 $\text{Char}(1,2)$ 中不含字符 $B[1]$, $D_{\text{time}}(\text{Char})=0$,那么 L_2 和 L_1 相同。 L_1 的表达式为

$$L_1 = \begin{cases} m-k, & k = \max\{k \mid B[k] = \text{Char}, \quad 1 \leq k \leq m-1\}, \\ m, & \text{其他.} \end{cases}$$

2 实验结果与分析

从执行先后顺序的逻辑关系上看,Snort 入侵检测系统从主函数 Main()开始,然后调用并执行 Detect()函数完成入侵检测。实验中,需要将各种字符匹配算法编制成函数,并用 Detect()函数调用。为了和改进算法形成对比,选择了经典的 Boyer 匹配算法,通过程序实现后封装在 Boyer()函数中。对于提出的改进算法,通过程序实现后封装在 ImBoyer()函数中。在验证实验的过程中,Boyer()函数和 ImBoyer()函数都可以供 Detect()函数调用。

实验所用的计算机配置:酷睿双核主频为 2.0 GHz 的 CPU,8 GB 的内存,500 GB 的硬盘;Windows 7.0 操作系统;Snort 入侵检测系统。实验针对 50 个随机长度的待匹配字符串,并以不同长度标准串完成匹配。50 个待匹配字符串的部分样本及全部不同长度的标准串,如表 1 所示。

随着标准串长度的改变,比较经典 Boyer 算法和提出的 ImBoyer 算法完成匹配的字符比较次数、窗口移动次数和执行时间,结果如图 1 所示。图 1 中: n_1 为字符比较次数; n_2 为窗口移动次数; t 为执行时间; l 为标准串长度; s 为数据包大小。由图 1(a)可知:提出的 ImBoyer 算法的字符比较次数明显低于 Boyer 算法;同时,随着标准串长度的增加,两种算法的字符比较次数都开始下降,但

表 1 待匹配字符串样本和标准串
Tab.1 String samples and standard string used in the experiment

| 序号 | 待匹配字符串 | 标准串 |
|----|---|---------|
| 1 | fdakliogakljagjlkagjalkgdeggasgasdgadbo | Im |
| 2 | dafdbymagpouagagigbmabvgdgdysosdgg | ImB |
| 3 | byoajsagyasgnbbhgdaygsdgsdlkstyasggs | ImBo |
| 4 | adsgsimagyaskgjasgdsaghbtagoasogdsgk | ImBoy |
| 5 | ksvbfgfdkgfbxczmgzxcnsadyegskhgiugdgi | ImBoye |
| 6 | vnxcenvxzhdsdgsdgsdzxvngsazvgsadgvifd | ImBoyer |

ImBoyer 算法下降的速度明显更快。这充分说明,所提出的 ImBoyer 算法性能更加优秀。由图 1(b)知:提出的 ImBoyer 算法的移动次数明显低于 Boyer 算法;同时,随着标准串长度的增加,两种算法的移动

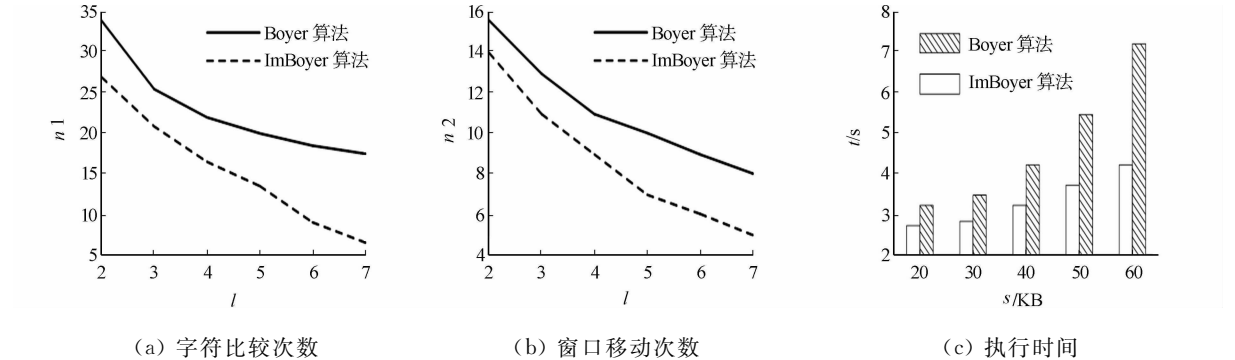


图 1 经典 Boyer 算法和 ImBoyer 算法的比较

Fig.1 Comparison between Boyer algorithm and ImBoyer algorithm

次数都开始下降,但 ImBoyer 算法下降的速度明显更快.这充分说明,所提出的 ImBoyer 算法性能更加优秀.由图 1(c)可知:提出的 ImBoyer 算法的执行时间明显低于 Boyer 算法;同时,随着数据包大小的增加,两种算法的执行时间都开始增加,但 ImBoyer 算法增加的速度明显更慢.这充分说明,所提出的 ImBoyer 算法性能更加优秀.

3 结束语

通过在规则匹配环节算法设计,对 Boyer 字符匹配算法进行了改进.通过对字符比较次数、窗口移动次数、执行时间 3 个方面进行实验研究,证实了改进工作的有效性.

参考文献:

- [1] SRIDHAR M,VAIDYA S,YAWAKJAR P. Intrusion detection using keystroke dynamics and fuzzy logic membership functions[C]//Proceedings International Conference on Technologies for Sustainable Development, Switzerland;Bridges Press,2015,27(4):444-458.
- [2] 李杰. 基于 Snort 的入侵检测系统规则解析及改进研究[J]. 电子技术与软件工程,2014,19(8):240.
- [3] PARVAT T J,CHANDRA P. Performance improvement of deep packet inspection for intrusion detection[C]//Proceedings 2014 IEEE Global Conference on Wireless Computing and Networking. [S. l.]:IEEE Press,2014:224-228.
- [4] PASTRANA S,TAPIADOR J E,ORFILA A. Defidnet: A framework for optimal allocation of cyberdefenses in intrusion detection networks[J]. Computer Networks,2015,80:66-88.
- [5] 谭笑,柯泽贤. 基于混合高斯和帧间差分的机场安全入侵检测[J]. 计算机仿真,2014,31(11):38-41.
- [6] 陈柏生,吴可沾,杨育辉. 互联网用户安全登陆平台设计[J]. 华侨大学学报(自然科学版),2011,32(6):638-640.
- [7] 储泽楠,李世扬. 基于节点生长马氏距离 K 均值和 HMM 的网络入侵检测方法设计[J]. 计算机测量与控制,2014,22(10):3406-3409.
- [8] MACDERMOTT A,SHI Q,KIFAYAT K. Collaborative intrusion detection in a federated cloud environment using the Dempster Shafer theory of evidence[C]//European Conference on Information Warfare and Security. [S. l.]:Earlybird Press,2015,195-203.
- [9] PAN Zhiwen,HARIRI S,AI-NASHIF Y. Anomaly based intrusion detection for building automation and control networks[C]//Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, Morocco:EasyChair Press,2015:72-77.
- [10] 袁其帅,刘云朋. 基于人工免疫原理的网络入侵检测系统的应用与研究[J]. 科技通报,2014,30(11):131-135.
- [11] 钱勤,张减,张坤,等. 用于入侵检测及取证的冗余数据删减技术研究[J]. 计算机科学,2014,41(11):252-258.

Application of Improved Boyer Matching Algorithm in Snort Intrusion Detection

MA Xiaoyu¹, LIU Shuanghong

(1. School of Science, Henan University of Engineering, Zhengzhou 451191, China;

2. Department of Computer Science and Application, Zhengzhou University of Aeronautics, Zhengzhou 450046, China)

Abstract: In this paper, the application of Snort intrusion detection system is studied. An improved method based on Boyer algorithm is designed. This method first designs a statistical array, then executes the matching with two adjacent characters, and divides into three strategies to determine the maximum movement length. Experimental results show this improvement makes the maximum movement length more reasonable. Compared with the Boyer method, the proposed method is significantly lower than the number of characters method, the number of windows mobile number is significantly reduced, the execution time is significantly reduced.

Keywords: network security; intrusion detection; Snort system; Boyer algorithm