文章编号:1008-0570(2009)02-3-0106-02

Snort 匹配机制的改进

Improvement of Matching Mechanism in Snort

(中国矿业大学 徐州) 张永平 徐冬阳
ZHANG Yong-ping XU Dong-yang

摘要: 基于规则的模式匹配是 Snort 检测引擎的主要机制,本文在结合协议分析和模式匹配的基础上,对 Snort 匹配机制进行了改进。首先对从网络中获取的数据包进行预先处理,利用协议分析技术对数据包进行高层应用协议分析;根据分析的结果,再利用提出的新型的模式匹配算法,对数据包中的其他相应信息进行模式匹配,从而显著地提高了 Snort 规则匹配的效率。测试表明,改进过的 Snort 在性能上得到了提高。

关键词: 入侵检测; 协议分析; 规则树; 模式匹配
中图分类号: TP393 **文献标识码:** A

Abstract: The pattern matching based on rule is the main mechanism in the detection engine of Snort. Based on the combination of protocol analysis and pattern matching, we improved the matching mechanism in snort. Firstly, we foreclose the data packet gains from the network, analyzing the data packet by using the technology of protocol analysis; based on the result, we match the other corresponding information of data packet by using a new algorithm of pattern matching; all of these can raise the efficiency of rule matching in Snort obviously. Finally a performance test is held, which shows that the improved snort has better performance.

Key words: Intrusion Detection; Protocol Analysis; Rule Tree; Pattern Matching

技术创新

引言

Snort 是一个开放源代码的轻量级的基于网络的入侵检测系统。基于规则的模式匹配是 Snort 的核心,大多数基于 Snort 的入侵检测系统所使用的模式匹配技术主要是字符串匹配技术,它的分析速度虽然比较快、误报率也比较小,但它也存在着很大的弊端。

1)Snort 只对 TCP/IP/UDP/ICMP 进行了分析,没有对高层应用协议进行协议分析。如果对高层应用协议进行分析可以显著地提高匹配效率。

2)字符串匹配主要就是算法的问题,Snort 的规则匹配多数属于字符串的匹配,因此一个好的字符串匹配算法可以大大提高检测的效率。

由于没有对高层应用进行协议分析,因此检测时的计算量大。对于一个特定网络,每秒需要比较的最大次数为 攻击特征字节数×网络数据包字节数×每秒数据包数量×攻击特征数量。假设所有攻击特征长度为 10 字节,网络数据包平均长度为 300 字节,每秒 40,000 数据包,攻击特征库中有 4000 条特征,那么每秒比较次数为 :10×300×40000×4000 = 480000000000。可以看出计算量是非常的大。

另外由于传统的模式匹配只能检测特定类型的攻击,对未知的攻击和攻击特征微小的变形不能够准确的检测出来。因此也容易产生漏报和误报,降低了检测的准确性。

本文提出了一种改进的 Snort 的匹配机制, 该机制引入高层应用协议分析,并且使用一种新的字符串模式匹配算法进行规则匹配,大大的提高了检测的效率。

1 改进办法

我们在进行入侵检测的时候,首先利用协议分析技术进行预处理,对获取的数据包进行高层协议分析 然后,再利用新的字符串模式匹配算法对数据包中相应的信息进行检测。

1.1 高层应用协议分析

协议分析将获取的数据包视为具有严格定义格式的数据流, 并将获取的数据包按照各层协议报文封装的反向顺序,层层解析出来。然后,再根据各层网络协议的定义,对各层协议的解析结果进行逐次分析。协议分析技术是利用预先定义好的关于协议字段的期望值或合理值的详细知识,来判断是否出现了恶意的网络流量 ;同时,这种解析是沿着协议栈向上解析的,因此可以使用所有当前已知的协议信息,来排除所有不属于这一个协议结构的攻击。例如,检测出第四层上的协议是 TCP,就不必再去检测第四层其它协议的攻击了,如 UDP。

协议分析具有能够读取攻击特征串及其所有可能的变形,并发掘其本质含义的能力。例如,将发现的"/./pf"、"pf"或者用"/."所做的其它变换判断为同一个攻击,也就是"/pf"攻击。这样,在攻击特征库中只需要一个特征,就能检测这一攻击所有可能的变形。

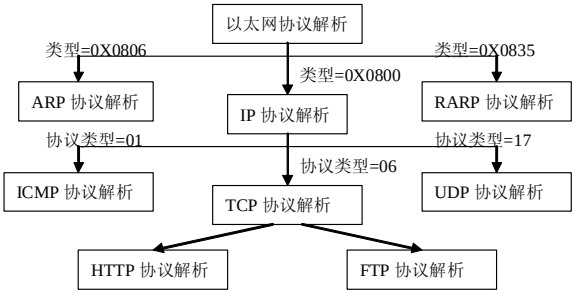

图 1 Snort 规则树结构

当前, 利用 Snort 开发的入侵检测系统只对网络层和传输层的协议进行了分析,没有对高层协议进行分析。另外 Snort 的

张永平: 副教授 硕士生导师

规则树结构过于简单，可能造成某些规则树结构下的选项链分枝比较庞大，因此在开发基于 Snort 的入侵检测系统时，要有效利用网络协议的层次性和相关协议的知识，使它快速地判断攻击特征是否存在。这样才能够使得匹配的计算量大幅度减小。Snort 规则树结构图 1 所示。

图 2 利用协议分析技术进行数据包分析

下面就根据图 1 中的 Snort 规则树结构，利用协议分析技术对图 2 中的数据包进行高层应用分析处理：

协议规范指出以太网络数据包中第 13、14 字节为第三层协议标识。我们利用这个知识开始第一步检测：

1) 过前面 12 个字节，读取 13、14 字节处的 2 字节协议标识：0800。根据协议规范可以判断这个网络数据包是 IP 包。

2)IP 协议规定数据包的第 24 字节处有一个 1 字节的第四层协议标识：06，这个数据包是 TCP 数据包。

3)TCP 协议规定在 TCP 数据包的第 37 字节处有一个 2 字节的应用层协议标识（端口号）：80，则表明该数据包是一个 HTTP 协议的数据包。

4) 根据 HTTP 协议规定，在数据包第 55 字节开始读取 URL，利用模式匹配检测入侵特征。

从上述数据包的检测过程可以看出，在入侵检测中，首先利用协议分析技术对数据包进行高层协议分析，可以避免不必要检测，大大减小模式匹配的计算量，提高匹配的精确度，减少误报率。

1.2 一种新的模式匹配算法

Snort 的检测模块中采用的是 BM 算法。

BM 算法的关键是根据给定的模式 P=P1P2…Pm，定义一个映射函数 dist: x→{1,2, …,m}，这里 x ∈ ∏ (∏是字符集)，函数 dist 也称为滑动距离函数，它给出了正文中可能出现的任意字符 x 在模式 P 中的位置，dist 函数的具体定义如下：

$$dist[x]=\begin{cases} m:若x不在P中或者x=P[M]但x\neq P[j](1\le j\le M-1) \\ m-j:其余情况，其中j=\max\{j|P[j]=x,1\le j\le M-1\} \end{cases}$$

BM 算法思想是：假如在执行正文中从位置 1 起“返前”的一段与模式匹配（自右向左）检查中，一旦发现不匹配（无论在哪个位置），立即执行由 P[M]与 T[1+d(T[1])]起始的自右向左的新一轮匹配检查。

通过分析可以看出影响模式匹配效率和速度的关键因素是：模式 P 和文本 T 在某个位置匹配失败时，如何使模式 P 向右移动的幅度最大，即尽可能多的跳过不需要比较的文本字符，从而减少模式匹配的次数和匹配过程中字符比较的次数。

这里以 BM 算法为基础，提出了一种新型的模式匹配算法。

新算法引入对模式 P 中的字符进行优先级别的分类的方法：首先将模式 P 的末尾字符 P[M]作为最优级独立出来单独处理；其次，将 P 中余下的字符按字符重复出现的次数为标准进行分类，出现一次的所有字符列为第一级，出现两次的所有字符列为第二级，依此类推，直至将所有的字符分配完毕。

字符比较时按照优先级的高低进行，优先级高的先比较（字符重复次数越少的优先级越高）。同一级别的字符，按照在 P 中从左至右出现的顺序进行比较。

进行匹配时，利用文本中匹配失败的字符 T[i]，文本中和模式最后一个对应位置字符 T[end]，以及它的下一个字符 T[end+1]共同启发模式 P 向右进行滑动。则新算法的滑动位数分两种情况：

(1) T[end]与 P[M]匹配成功时：

$$dist[x]=\begin{cases} m+1:T[end+1]不在P中出现 \\ \max\{dist[T[i]],dist[T[end+1]]\}:T[end+1]在P中出现 \end{cases}$$

(2) T[end]与 P[M]匹配失败时：

$$dist[x]=\begin{cases} m:T[end]不在P中出现 \\ m+1:T[end+1]在P中出现且T[end+1]不在P中出现 \\ \max(dist[T[i]],dist[T[end+1]]):T[end]和T[end+1]都在P中出现 \end{cases}$$

在完全随机的情况下，T 中某个位置出现的字符 x (x ∈ 字符集∏)的概率都为 q(q 是字符集∏基数的倒数)。假设在匹配检测中，T 中指针所指某个位置的字符 x 给 P 产生了滑动，滑动位数为 dist(x)，产生的数学期望(平均值)为∑。

则 BM 算法中的期望值为：

$$\sum_1=\sum_{x\in\Pi}\max\left(dist(x)\right)*q,(x是T中指针j所指字符) \quad (Ⅰ)$$

其中：$dist(x)=\left(j+d(x)\right)-end,(x是指针j所指字符)$

由于在新算法中，模式的滑动位数在 0 到 m+1 之间(0 为匹配成功)，且新算法由 T 中的三个字符共同启发模式 P 向右滑动，所以新算法的期望值为：

$$\sum_{123}=\sum_{x_1\in\Pi}\sum_{x_2\in\Pi}\sum_{x_3\in\Pi}\max\{dist(x_1),dist(x_2),dist(x_3))*q^3 \quad (Ⅱ)$$

其中，x_1,x_2,x_3 分别为：当前匹配时，文本中匹配失败的字符 T[i]，文本中和模式最后一个对应位置字符 T[end]，以及它的下一个字符 T[end+1]。

因为：

$$\sum_{123}\ge\sum_{x_1\in\Pi}\sum_{x_2\in\Pi}\sum_{x_3\in\Pi}dist(x_1)*q^3\ge\sum_{x_1\in\Pi}dist(x_1)*q$$

所以，比较(Ⅰ)和(Ⅱ)可得 $\sum_{123}\ge\sum_1$，所以新算法加大了模式滑动的幅度，减少了匹配的次数。

1.3 测试比较

测试环境为：主频 2.8GHz，内存为 1GB，操作系统为 Windows 2000，编译器为 Visual C++ 6.0，数据库系统为 SQL Server 2000，入侵检测系统采用了开源的 Snort。在这样的环境下，针对匹配机制未改进的 Snort 和改进过的 Snort 进行了测试比较。

在测试中，分别针对 Snort 日志中的部分数据进行了测试，选取了 100KB,200KB,300KB 和 400KB 的数据进行测试。测试结果如图 3 所示。

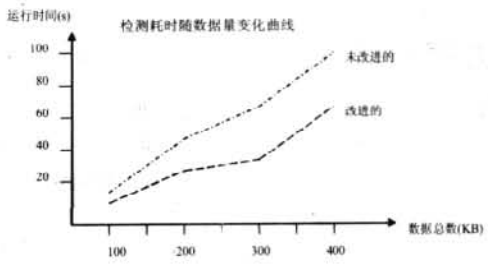

图 3 测试结果比较

由图 3 可见，随着数据量的增加，两种检测耗时均逐渐增加。曲线表明经过改进的 Snort 检测耗时比未改进的耗时少。分析原因，改进的方法采用对高层应用进行协议分析，避免了不

（下转第 80 页）

技术创新

3.3 RSA 算法的优化

基于乘同余对称特性的算法 SMM (Symmetry of Modulo Multiplication) 是利用乘同余对称特性来减少 RSA 加解密计算中乘法和求模运算量的一种快速算法。RSA 加密是对明文求幂剩余的过程: $y=\langle Me\rangle n$

$\langle.\rangle n$表示括号内的数对 n 求模。上述的 RSA 算法是将指数 e 表示成 t 位二进制数的形式,并将幂剩余变成一系列乘同余的迭代,(以 $a=g^x \bmod p$ 为例),由上面的讨论知道每一步迭代必有运算 $a=a^2 \bmod p$ 和可能有运算 $a=a\times g \bmod p$。

SMM 算法是在每步迭代计算中对乘数和被乘数进行有条件的代换。具体代换情况如下:如果 a_{i-1}表示第 $(i-1)$步迭代的结果,则在进行第 i 步迭代时,若a_{i-1}或$g \le (n-1)/2$,则保持原数不变,但是如果 a_{i-1} 或 $g>(n-1)/2$,则使用 $(n-a_{i-1})$ 或 $(n-g)$来代替 a_{i-1} 或 g。下面给出证明来说明算法的正确性:(以 $x^2 \bmod n$ 和 $xy \bmod n$ 为例)

$\because (n-x)^2 \bmod n=(n^2-2nx+x^2)\bmod n=x^2 \bmod n$. 可$(n-x)$用来代替 x

又 $\because (n-x)(n-y)\bmod n=(n^2-nx-ny+xy)\bmod n=xy \bmod n$

$\therefore$ 可分别用$(n-x)$来代替 x 和 $(n-y)$来代替 y

由于使用 SMM 方法,减少了乘法时间和求模运算量,使算法速度得到一定程度的改善。

4 结束语

本方案在研究传统通信加密和数字签名的基础上,结合对称密钥加密算法和非对称密钥加密算法优点,用混合密钥的方式既达到信息加密传送的高速性,也解决了对称密钥加密算法密钥分发的不安全性,实现了系统的要求。同时服务器密钥管理模块还解决了局域网公要基础设施 (PKI) 不健全带来的安全性问题。采用两次或者多次 hash 化来降低"碰撞"可能性消除了对消息摘要算法存在理论上被破译的问题。本系统网络配置为光纤网络,交换机为华为 24 口光交换机,型号 LS-3026FM,内置两个千兆模块 LS-GIMU。

本文作者的创新点:用混合密钥的方式实现系统要求。用密钥管理模块解决 PKI 不健全的问题。用多次 hash 化消除摘要算法存在理论上被破译的可能。

参考文献

[1]Bruce Schneier. Applied Cryptography: Protocols, Algorithms, and Source Code in C[M], Second edition. John Wiley &Sons. Inc. 1996.

[2]Atul Kahate,邱仲潘.密码学与网络安全=Cryptography and network security.北京:清华大学出版社, 2005

[3]Man Young Rhee,金名,张长富.网络安全:加密原理、算法与协议.北京:清华大学出版社,2007

[4]周书锋,孙玉真.基于混合密钥的数字签名研究[J].微计算机信息,2006,12.3:54-57.

[5]纪秀君.《王小云—破解国际两大密码的中国女教授》[N].中国教育报, 2005-03-26.

[6]李可胜.简论加密技术.计算机与信息技术.2007, 20: 83

[7]程庭, 张明惠. 一种基于 DES 和 RSA 算法的数据加密方案及实现. 河南教育学院学报(自然科学版), 2003(2):69-71

[8]陈运.基于乘同余对称特性的快速 RSA 算法的改进.电子科技大学学报,1997,

作者简介:曾一民 (1980-),男 (汉族) 重庆人,电子科技大学硕士研究生 研究方向:智能机电系统及控制软件开发等。骆德渊 (1970-),男 (汉族),副教授,中共党员,工学博士,硕士生导师;主要研究方向:精密仪器开发,自动控制技术,软件开发等。

Biography: ZENG Yi-min (1980-), Male (Han), Chongqing, Master's candidate of UESTC. Research area: Intelligent electromechanical control systems and software development

(610054 电子科技大学机械电子工程学院成都)曾一民 骆德渊

(College of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China) ZENG Yi-min LUO De-yuan

通讯地址:(610054 四川 成都电子科技大学欣苑 6 栋 202) 曾一民

(收稿日期:2009.01.13)(修稿日期:2009.02.15)

(上接第 107 页)

必要的协议检测,减少了计算量;同时,采用的新模式匹配算法,降低了匹配的次数,提高了匹配的效率。

2 结束语

协议分析技术和模式匹配技术的结合,减少了 Snort 检测时产生的漏报和误报,增加了 Snort 检测时的可靠性;大大减小模式匹配的计算量,提高匹配的精确度,显著地提高匹配效率。

本文作者创新点:提出针对 Snort 匹配机制改进的方法:结合协议分析和模式匹配技术,提出对从网络中获取的数据包先进行高层应用的协议分析;然后,根据分析的结果,利用提出的新型的模式匹配算法,对数据包中的其它相应信息进行匹配。其中对新模式匹配算法的优越性用数学期望值进行了说明。测试表明,新的改进方法提高了 Snort 规则匹配的效率。

项目经济效益:50 万元

参考文献

[1]HORSPOOL RN. Practical fast searching in strings [J]. Software Practice and Experience. 1980,10(6):501-506.

[2]赵念强,鞠时光.入侵检测系统中模式匹配算法的研究[J].微计算机信息,2005,8-3:22-24。

[3]胡大辉,刘乃琦.高效的 Snort 规则匹配机制[J].微计算机信息,2006,2-3:10-11。

作者简介:张永平 (1958-),男,汉族,中国矿业大学计算机科学与技术学院,副教授,硕士生导师,研究方向:计算机网络与信息安全,密码学。徐冬阳 (1983-),男,汉族,中国矿业大学计算机科学与技术学院硕士研究生,研究方向:计算机网络与信息安全。

Biography: ZHANG Yong-ping, 1958-, male, Han nationality, Associate professor of Computer Science and Technology Department, China University of Mining and Technology, tutor of postgraduate, study in computer network and information security, cryptography.

(221008 江苏徐州 中国矿业大学 计算机科学与技术学院) 张永平 徐冬阳

(Computer Science and Technology Department, China University of Mining and Technology, 221008, China) ZHANG Yong-ping XU Dong-yang

通讯地址:(221008 江苏省徐州市中国矿业大学计算机科学与技术学院 2006 级硕士研究生) 徐冬阳

(收稿日期:2009.01.13)(修稿日期:2009.02.15)

技术创新