高效数据包过滤器部署算法

王 杰,董鹏飞,李 振

郑州大学电气工程学院,郑州 450001

摘要 整合在路由器上的数据包过滤功能是网络安全的第一道防线,对保护网络的安全有着至关重要的作用。当前的研究主要侧重于对数据包的检测和过滤,忽略了数据包过滤器的部署问题,如果对其部署得当,能极大改善网络的性能和效率。在大规模的网络中,数据包过滤器被部署在管理网络的边界上,即信任网络和非信任网络之间。一般情况下,数据包过滤器被部署在网络边界的每一个路由器上,由于需要遍历路径上的过滤器,大量的时间和空间被消耗,容易造成时延、拥塞、丢包等问题,降低网络的效率。本文依据最短路径优先的思想,提出了一种新的部署算法,通过对比边界上各种风险,移除有操作风险的边界,产生一条最短虚拟路径,更加有效地部署过滤器的位置,给内部节点之间的信任网络提供保护,并使用基于风险的方法对虚拟边界进行离线计算。实验结果表明,使用该算法后,过滤器的数量减少 20%~50%,网络的延迟时间变小,网络的连通性也得到改善。

关键词 网络安全;最优过滤器部署;数据包过滤器;最短路径

中图分类号 TP393 **文献标识码** A **文章编号** 1000-7857(2010)18-0080-04

An Efficient Disposition Algorithm for Packet Filters

WANG Jie, DONG Pengfei, LI Zhen

School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China

Abstract Packet filtering of the router is the first defense line for the network security and plays a crucial role. The current studies focus on the data packet detection and filtering, without due consideration of the packet filter device deployment. If the packet filter is deployed properly, the network performance and efficiency can be significantly improved. Packet filters are always placed on the border of the administrative network, the boundaries between trusted and non-trusted networks. Generally, the packet filters are placed on the router in the border of the network, and a tremendous amount of time and space would be consumed, that results in time-delay and congestion, and hampers the efficiency of the network. In this paper, a new disposition algorithm is proposed. To compare various kinds of risks in the border and to remove the edge with operation risks, the packet filters can be placed in an efficient way by generating a shortest border, to provide a perfect protection for trusted networks between the internal network nodes. Experiment results show that the number of packet filters is reduced by 20% to 50%. The network delay time is reduced and the network connectivity is improved.

Keywords network security; optimal filter placement; packet filter; the shortest path

0 引言

整合在路由器上的数据包过滤功能是网络安全的第一道防线,对保护网络的安全有着至关重要的作用。随着网络结构愈加复杂,网络规模日趋庞大,一般的方法只是对特定的网络环境有效,不能从根本上解决包过滤问题,已经不能适用于现在的大型网络[1]。目前,数据包过滤研究主要集中在检测方法上,忽视了过滤器部署方法的研究。事实上,数据包过滤器部署对整个网络性能有很重要的影响。如果部署得当,能极大提高网络的性能和效率,国外已有研究者开始重视包过滤器部署算法的优化问题[2-4]。本文依据最短路径优先(Shortest Path First,SPF)[5]思想,提出了一种最优过滤器部署算法(Optimal Filter Placement,OPF)过滤器部署的算法,在保证网络安全的前提下,能够减少数据包过滤器的数量,提高网络的运行效率,满足当前大型网络的应用要求。

收稿日期:2010-04-06;修回日期:2010-07-28

基金项目:教育部高等学校博士学科点专项科研基金项目(20094101120008);河南省杰出人才创新基金项目(074200510013)

作者简介:王杰,教授,研究方向为智能控制与智能计算、信息与计算机网络安全,电子信箱:wj@zzu.edu.cn

1 OPF 数学建模和部署算法设计

1.1 OPF 数学建模

假设网络被看作有向图 $G(N,A)$,其中 N 为网络节点的集合,$|N|\geqslant 3$,A 为弧(arc)的集合。定义通信集为所有用于数据包交换的节点的集合,则通信集 $C\subseteq N\times N-\{U_{i\in N}(i,i)\}$,例如 $(u,v)\in C$,当且仅当数据包在 $u\in N$ 条件下,能在给定的路径下到达 $v\in N$。

集合 C 由从 u 到 v 的所有可能路径组成,集合 $R(u,v)$ 由 C 确定。路由机制是基于目的地的逐跳路由机制。一条长度为 k 的路径 $p\in R(u,v)$,可看作是一系列点 $p_0,p_1,p_2,\cdots,p_k,p_0=u$,$p_k=v$。定义 $R^{all}=\bigcup_{(u,v)\in C}R(u,v)$,同时定义 $nodes(p)$ 为路径 p 中的节点,$arcs(p)$ 为路径 p 的边界集合。为了保证一般性,假设 A 是不可约分的,对于所有的 $(i,j)\in A$,得到 $(i,j)\in arcs(p)$。

然后定义网络数据包过滤系统的抽象模型,包括过滤器和网络中的数据包。过滤器允许放置在整个网络的所有节点上。每个数据包与其真正的源节点或是伪造的源节点,以及目标节点相关。因此,过滤器可以获取各个数据包的源节点和目的节点信息,以及数据包到达过滤器所经过的路径信息。一种基于路径的过滤器可以定义为

$$FM(o,a,d)=\begin{cases}0 & p\in R(o,d),a\in arcs(p)\\ 1 & \text{其他}\end{cases}\qquad(1)$$

如果一个从 o 到 d 的数据包在路径 R 下经过边 a,那么 $FM(o,a,d)$ 的返回值为 0,否则,返回值为 1。其中,o 为路由源点,d 为终点,a 为路由上的边。该模型的缺点是其复杂度 $O|N|^2$ 限制了它在更大型网络中的使用。因此,有了一个替代模型:

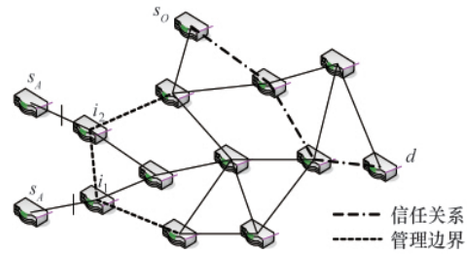
$$FS(o,a)=\begin{cases}0 & \exists d\in N,p\in R(o,d),a\in arcs(p)\\ 1 & \text{其他}\end{cases}\qquad(2)$$

在任何路径中,如果节点 o 使用边 a,则返回值为 0;否则,返回值为 1。当返回值为 1 时,可以断定数据包的源 IP 地址是伪造的,数据包被丢弃。注意到在使用遍历过滤器机制后,正常数据包不会被丢弃,因此安全是可以保证的。一般情况下,FM 模型比 FS 模型更有效,例如对于所有 $(o,d)\in C$,$a\in A$,有 $FS(o,a)\leqslant FM(o,a,d)$。当 $p\in R(o,b)$ 存在于 $(o,b)\in C$,$a\in arcs(p)$ 时,这种关系与 $(o,d)\in C,a\in A$ 紧密相关,但没有这样的路径存在于 $R(o,d)$。

无论使用什么类型的过滤器,过滤节点上的数据包是不可能被伪造源 IP 地址的。这种机制称为入口过滤,在技术上很容易实现。节点不会将伪造的数据包装转发出去。同时,如果数据包的源节点和目标节点对 $(o,d)\notin C$,那么将在节点 d 被自动丢弃,在那样的节点过滤器的存在与否无关重要。

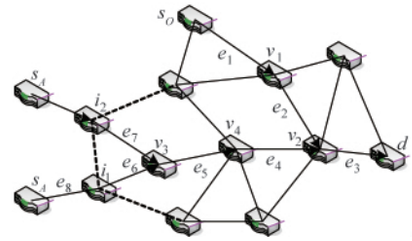
1.2 OPF 部署算法设计

传统的过滤器放置方法是尽可能地靠近攻击源,因此可将代表不信任网络的内部节点放在管理边界上,这对安全来说是简单易行的[6-7]。如图 1(a)所示,在内部节点 i_1 和 i_2 上的两个过滤器,可以保护 S_O 和 d 之间的可信任链接受到的可能攻击。为了得到在虚拟边界设置节点和接口的位置,首先必须找到与之对应的边界和顶点。在基础网络中,由路由算法决定的路由表的不同必须被考虑在内,如图 1(b)所示,对于任意节点,路由表都必须被访问,因为代表节点是没有能力执行过滤器的功能。当边缘不能提供过滤器时,通过合并其起点和终点,来自攻击路线的边界被移除,事件边缘的攻击表也被删除。如在图 1(c)中,由于边界 e_3 的高操作风险,该边界从攻击路径被移除;同时,由于在 v_2 和 v_4 点失去了过滤能力,边界 e_4 也被移除。结果,点 v_2 和 v_4 被合并成点 d。最后的



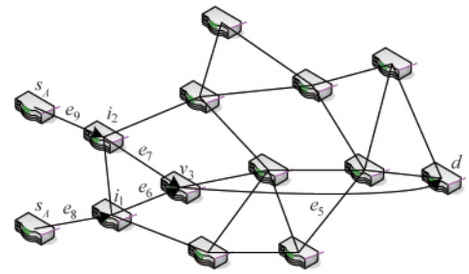
(a) 网络拓扑

(a) Network topology



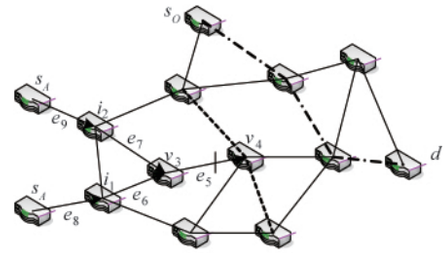
(b) 数据包路由图

(b) Packet routing map



(c) 边界去除图

(c) Border removal map



(d) 过滤器放置图

(d) Router deployment map

图 1 过滤器放置决策图

Fig. 1 Decision of router deployment

表被转化成一个 0–1 网络,可以使用最小化算法,即能用最少的过滤器保护网络免受欺骗性数据包的攻击。最后,图 1(d)表明可以通过改变过滤器的部署提高网络性能,原因是有潜在攻击点性的数据包的路径,可能会在网络内部的同一点合并。因此,通过网络内部一些点的合并,一些重要的基于目的地的逐跳路由的过滤配置被减少。

2 OPF 算法设计

2.1 风险计算

需要这样一个数据包部署算法,该算法不论从过滤器的数量上,还是操作风险和被攻击风险上,消耗都是最低的。因此,需定义以下内容:

1) 源节点 s_O(操作者)和 s_A(入侵者)和终点 d;

2) 路径 $p_O \in P_O$,P_O 为从 s_O 到 d 的所有路径的集合;

3) 路径 $p_A \in P_A$,P_A 为从 s_A 到 d 的所有路径的集合;

4) 假阳性[8]的概率 ω_{p_n},指过滤器错误的结束一个连接的概率;

5) 操作破坏性 D_O 为由于放置过滤器位置的错误产生的破坏性;

6) 操作风险 $R_O=\omega_{p_n} D_O$;

7) 假阴性[8]概率 φ_{p_A},指攻击者发出数据包和在路径 p_O 的任何处不被过滤出,这两种可能性混合就是假阳性概率;

8) 攻击破坏性 D_A 是由于控制节点的端口缺少过滤器引起的破坏;

9) 攻击风险 $R_A=\omega_{p_A} D_A$。

就以上风险的成本来说,与所有过滤器的配置是不匹配的,因为无论过滤器是否能被放置在某个节点上,总的成本是超过从 s_O 到 d 和从 s_A 到 d 的有效路径的集合:

$$R_{total}=\sum \omega_{p_n} D_O+\omega_{p_A} D_A \quad (p_O \in P_O, p_A \in P_A) \qquad (3)$$

现在的关键问题是,通过选择一个最优的分布式数据包过滤配置,最小化所有的风险。但是,无论使用操作系统内部的过滤系统,还是接口有过滤能力的硬件网络处理设备,都会对配置算法有所限制。

2.2 OPF 算法

所有的路径 p_O,p_A,对于计算过滤器的放置是一种合理需求,并同弹性需求和安全需求保持一致。如果每条路径上的元件不是很少,同时可用性相当高,那么缩小状态空间是可行的。因此,应将状态定向 Θ 限制在所有可能的故障中,这些错误的发生概率尽可能是预先的 t。首先确定同时失败的元素的数量,检验这种配置的阈值。然后,列举这个失败状态空间 Θ,并结合合法的使用终点和攻击终点,把路由算法用在网络的每一个状态上。因此,可以得到两个最大的路径集合,不仅最有可能性也是最有效的。最后,迭代所有的路径和所有的路径的边缘,给独立边界增加选择路径的确切的概率。

现在,通过对比每一个边界的 D_O 和 D_A,计算支配 D_O 的集合,支配 D_O 的边界集合如图 1(b)中$\{e_1,e_2,e_3\}$。如果希望删除一个点,那么所有代表节点的顶点必须在开始的时候分开,在开始点进入边界,把终点放在出口边界。通过合并它们的顶点,主要的有操作风险的边界被移除。然后,所有的边界上没有过滤功能的点,也通过合并终点被移除,例如图 1(c)的 e_4。攻击表中留下每一个边界,可能被放置过滤器阻挡诸如非法的数据包之类的入侵者。由于这些路径被移除后减少了图的连通性,所以被认为是最短路径集合。

具体算法如下。

步骤 1 提取状态空间 Θ。

步骤 2 提取路径集合。

对所有的 $s_O \in S_O$,$s_A \in S_A$ 和所有的状态 $\sigma \in \Theta$,$P_O \leftarrow R_\sigma(s_O, d)$,$P_A \leftarrow R_\sigma(s_A, d)$。

步骤 3 计算边界初始化。

对所有路径 $p \in P_O \cup P_A$ 和所有边界 $\varepsilon \in p$,如果 $\varepsilon \in P_O$,那么增加有效性 p 到 ω_ε;如果 $\varepsilon \in P_A$,那么增加有效性 p 到 φ_ε。

步骤 4 计算风险距离。

对所有路径 $p \in P_A$ 和所有边界 $\varepsilon \in p$,如果 $\omega_\varepsilon D_O > \varphi_\varepsilon D_A$,或者 $\varepsilon \in C$,然后合并 ε 的起点和终点,并用合并的点取代 ε 和它的终点。

步骤 5 计算最短路径。

对所有路径 $p \in P_A$ 和所有边界 $\varepsilon \in p$,增加 ε 和它的终点到 0–1 网络 N,$G \leftarrow N$ 中最小的不相交路径集合。对所有路径 $p \in G$ 和所有边界 $\varepsilon \in p$,如果在 ε 的删除部分减少流入 N 中,那么过滤边界集合 $F \leftarrow \varepsilon$ 从 G 中移除 p,循环步骤 4;如果 N 中流大于 0,那么,就是从 S_A 到 d 的没有过滤能力的边界。

3 OPF 算法仿真

3.1 仿真方法

在仿真试验中,由 OPNET 生成 10km×10km 的校园网络拓扑。节点选用 3C_SSII_110_3300_4s_ae52e_48_ge3 类型,链接方式是 Full mesh,链接模型选用 10Mb/s 以太网光纤线路(10BaseT),并从网络中随机的选出节点进行配置。源节点和目标节点随机选取。为了避免遍历所有网络元件的状态空间,需要在执行算法的时候,设置一个故障状态概率阈值。参数设置见表 1。

表 1 参数表

Table 1 Parameter list

参数	值
节点数量	n
故障状态概率阈值	t
内部连接点数量	i
虚拟边界上过滤器数量	g

3.2 仿真结果和分析

如图 2 所示,随着节点数目的增加,过滤器均值下降了 20%~50%。主要是因为从源节点到目标节点的平均路径长度

增加了。当攻击路径比较长的情况下,它与其他攻击路径合并的可能性就增大了。

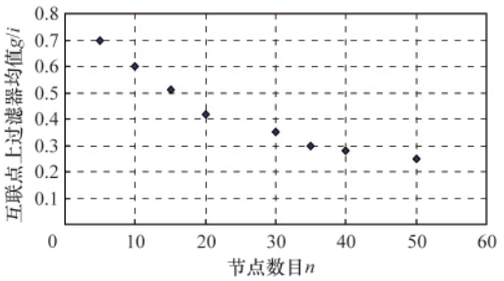


图2 节点数量与互联点上过滤器均值关系

Fig. 2 Number of nodes and filters per interconnection point

如图3所示,将文中的OPF算法与OSPF协议比较,OPF算法使用的过滤器数目只有OSPF协议的50%,就能保证内部网络安全,并且产生的错误更少。

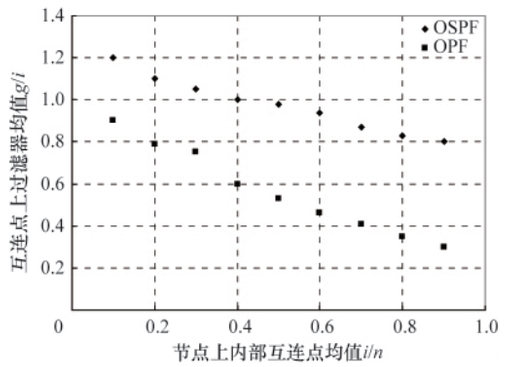


图3 过滤器使用情况对比

Fig. 3 Comparison of filter usages

图4给出了OPF算法与OSPF协议的延时仿真。可以看出,采用OPF算法时,文中选择的局域网的延时时间是0.002ms,而OSPF协议的延时时间是0.009ms,OPF算法比OSPF协议少用了0.007ms。

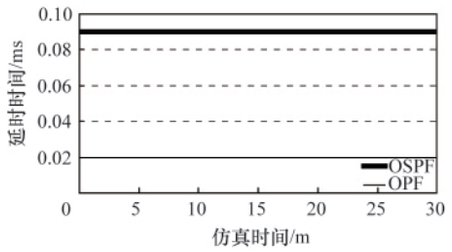


图4 OPF算法与OSPF协议延时仿真

Fig. 4 Time delay simulation of OPF and OSPF

4 结论

本文对大型网络管理服务提出了新的灵活的数据包过滤器配置机制。与一般把过滤器放在外部边界不同的是,本文算法通常在网络内部找到更高效率的虚拟边界,从而减少需要的过滤器的数量。同时,使用了基于风险的,离线计算办法,并综合了优良的过滤能力。最后,通过仿真过滤器的配置,证明在有大量的节点的网络环境中,能够极大地提高网络的性能和效率。

参考文献(References)

[1] 王杰, 石成辉. 基于节点共享计数型 Bloom Filter 高效数据包过滤算法[J]. 系统工程与电子技术, 2009, 31(9): 2227-2231.
Wang Jie, Shi Chenghui. *Systems Engineering and Electronics*, 2009, 31(9): 2227-2231.

[2] Armbruster B, Smith J C, Park K. A packet filter placement problem with application to defense against distributed denial of service attacks [J]. *European Journal of Operational Research*, 2007, 176(2): 1283-1292.

[3] Julkunen H, Chow C E. Enhance network security with dynamic packet filter [C]//Proceedings of the 7th IEEE International Conference on Computer Communications and Networks. Lafayette, IN: IEEE Press, 1998: 268-275.

[4] Todtmann B, Rathgeb E P. Requirements for managing distributed packet filter configurations in Carrier-grade Networks [C]. 10th IFIP/IEEE International Symposium on Integrated Network Management, Munich, 2007: 737-740.

[5] 朱彦波. 基于 IP 网络的 OSPF 路由协议实现机制的研究和解析[D]. 长春: 吉林大学通信工程学院, 2006.
Zhu Yanbo. Analysis and study on implement of OSPF protocol based on IP network [D]. Changchun: School of Communication Engineering, Jilin University, 2006.

[6] 李金平, 高东杰. 策略路由技术[J]. 计算机科学, 2002, 29(4): 84-85.
Li Jinping, Gao Dongjie. *Computer Science*, 2002, 29(4): 84-85.

[7] 胡滨, 夏欣, 任守刚. 数据包过滤模型的分析和研究[J]. 计算机工程与设计, 2007, 28(5): 1040-1042.
Hu Bin, Xia Xin, Ren Shougang. *Computer Engineering & Design*, 2007, 28(5): 1040-1042.

[8] Hooper E. An intelligent detection and response strategy to false positives and network attacks [C]//Proceedings of the 4th IEEE International Workshop on Information Assurance. Washington DC: IEEE Press, 2006: 16-21.

(责任编辑 朱宇)