

面向攻击图构建的网络连通性分析

黎 翰, 张少俊, 陈秀真, 陈晓桦

(上海交通大学信息安全工程学院, 上海 200240)

摘 要: 针对目前网络攻击图构建系统的需求, 设计网络连通性分析算法。通过对网络拓扑及防火墙规则进行离线分析, 可以判断网络中由若干台过滤设备分隔的任意 2 台主机间的连通性。引入关键实体集的概念, 结合经典的 Apriori 算法提出一种快速有效的获取关键实体集的方法。分析对比表明, 关键实体集可以在连通性分析过程中为网络中各节点的重要性评估提供有力依据。

关键词: 网络连通性; 攻击图; 关键实体集

Analysis of Network Connectivity for Attack Graph Construction

LI Han, ZHANG Shao-jun, CHEN Xiu-zhen, CHEN Xiao-hua

(Information Security Engineering Institute, Shanghai Jiaotong University, Shanghai 200240)

【Abstract】 This paper designs a network connectivity analysis algorithm according to the present techniques and the need of attack graph construction system. By using connectivity analysis, network topology and firewall rule analysis can be performed offline, which determines the connectivity between two hosts. It introduces a conception of Critical Entity Collection(CEC). An effective way of CEC detection is presented on the basis of classic Apriori algorithm. Deep analysis and comparison show that CEC provides effective information for the assessment of the importance of nodes in the network in the process of connectivity analysis.

【Key words】 network connectivity; attack graph; Critical Entity Collection(CEC)

1 概述

在计算机风险评估领域中, 网络信息系统的风险分析随着网络互联程度的日趋复杂变得愈来愈困难, 这种风险分析的成败关键在于能否准确而有效地检测出网络中的所有漏洞极其关联。近年来, 国内外的研究人员纷纷对网络漏洞的检测与关联展开研究。文献[1]提出一种网络漏洞的模型检测方法, 认为一次成功的攻击所必需的条件之一是保证各节点主机的基本连通, 并将连通性定义为一个主机与其他主机的通信能力, 在其模型中利用了静态矩阵来定义网络中各个节点主机的连通性, 并基于此连通性知识检测出网络漏洞。文献[2]提出的网络连通性模型描述了网络节点在 TCP/IP 结构模型中各协议层的通行状况, 在此工作的基础上, 文献[3]设计了一个拓扑漏洞分析工具, 并取得了较好的效果。由此可见, 充分理解和深入研究网络中各节点主机之间的连通性是实现网络漏洞检测及关联的重要环节。

2 网络连通性分析

在网络攻击图构建过程中, 连通性分析指根据网络拓扑、防火墙过滤规则列表等信息对网络中由若干台过滤设备分隔的任意 2 台主机间的协议连通性进行判断。对于攻击图的构建, 分析能否准确地反映出主机间的连通性将直接决定攻击图的构建结果。

2.1 术语定义

为了更好地描述, 先给出以下基本概念:

(1) 网络连通性分析器(Network Connectivity Analyzer, NCA)。

(2) 实体: 表示网络中的一台设备, 可以是一台普通的主机, 也可以是防火墙或者网关。

(3) 网关(Gateway): 指具有过滤规则的设备, 既可以是防火墙也可以是路由器(Router)。

(4) 全通实体: 指具有全通规则的网关。即其被攻陷后, 被攻击者加入了全通规则, 没有限制数据包的过滤规则, 允许全部数据包通过。

(5) 关键实体组合: 在用户给定的全通实体集合中, 能够使得连通性得到满足的全通实体的组合。其每个最小的组合都称为最小关键实体组合。

(6) 关键实体集: 所有最小关键实体组合的集合。

2.2 NCA 的初始化及数据的基本流向

在启动 NCA 之前, 需要初始化基本的网络拓扑信息以及过滤规则和 NAT(Network Address Translation)信息。这里, 拓扑信息、过滤及 NAT 规则通过数据库的数据表来体现。关于拓扑数据结构的描述详见文献[4]。

拓扑以及过滤规则初始化完成后, 即可以启用 NCA, 基本数据流向(Data Flow)如图 1 所示。

在 NCA 查询时, 主要通过调用 *isConnect()* 方法, 该方法输入为源目的实体、接口、IP 信息以及所需要查询的服务和当前的全通实体集合, 输出为服务的可用性以及关键实体集。该方法主要依据深度优先原则, 先进行 NAT 的判断, 然后开

基金项目: 国家自然科学基金资助项目(60605019); 国家“863”计划基金资助项目(2007AA01Z473); 教育部博士点基金资助项目(20070248002)

作者简介: 黎 翰(1984 -), 男, 硕士研究生, 主研方向: 计算机网络安全管理, 计算机网络安全态势感知; 张少俊, 博士研究生; 陈秀真, 讲师、博士; 陈晓桦, 博士

收稿日期: 2008-12-24 **E-mail:** vcfvct@163.com

始传播,在每个经过的节点上都进行规则检验,规则允许则取得与当前网关相连的网关,依此类推,对整个网关图进行遍历,直到找出所有可能的连接路径,图形界面(GUI)再根据结果通过反馈给用户查询的结果。在最坏的情况下,算法复杂度是指数级的;不过这种情况只发生在非常密集的网络环境下,而一般的网关图都是比较稀疏的,这是由于防火墙一般都被布置于网络中的战略关键点上,因此网关图一般是树状的。在这种情况下,算法复杂度是线性的。

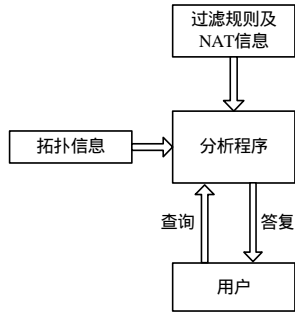


图1 NCA的基本数据流向

3 关键实体集的作用及求法

3.1 关键实体集的定义

关键实体集是在用户给定的全通实体集中分析网络攻击路径中所需最少的全通实体的集合。如用户给定的全通实体集中包含 B, C , 而 $A > B > C > D$ 和 $A > C > D$ 是 2 条连通路径, 对于前一条路径, 即使去掉 B 节点, 由于后一条路径的存在, A 和 D 之间仍能连通, 而 A 和 D 之间不能直接连通, 必须经过 C 节点, 因此, $\{C\}$ 就是一个最小关键实体组合。因为网络中连接 A, D 的路径可能不止 1 条, 所以关键实体集是指包含所有路径中的最小关键实体集的集合。

了解网络中各个节点在网络中的重要性(对于某服务)对于网络管理员是非常必要的。这样, 管理员可以根据这些节点的重要程度的不同在其上面布置不同等级的安全策略。而对于关键实体集的分析可以给网络管理员在这方面提供切实有效的信息以供参考, 因为每个关键实体集中的组合都代表该服务的一条路径上的关键节点。同时, 关键实体集分析还能提供关于特定实体上的规则变化对于整个端到端服务可用性的影响信息, 使得管理员能更深入地从宏观上认识各网络的结构以及为防火墙的部署位置提供关键信息。

3.2 关键实体集的求法

通过对关键实体集的描述, 可以看到它是由位于每条路径的最小关键实体组合组成的集合, 它们都来自于用户给定的全通实体集合。类似于 Agrawal 等人建立的用户事务数据库挖掘的项目集格理论^[5], 可以得出 2 点: (1) 关键实体组合的超集是关键实体组合; (2) 非关键实体组合的子集是非关键实体组合。

定理 1 如果全通实体组合 X 是关键实体组合, 那么它的所有超集都是关键实体组合。

证明: 设 X 是一个关键实体组合, 对于 X 的任一超集 Z , 都有 Z 对于任意规则的限制数小于 X 的限制数。这是因为根据全通实体的定义, X 包含于 Z , 所以在同样的拓扑下, Z 的限制将宽松于 X 。因此, Z 一定是一个关键实体组合。

定理 2 如果全通实体集合 X 是非关键实体组合, 那么它的所有子集都是非关键实体组合。

证明: 设 X 是一个非关键实体组合, 根据定义 X 应用于

网络时并不能使连通性得到满足。对于 X 的任一非空子集 Y , 在同样的拓扑下, 由于 Y 的限制条件严格于 X , 那么连通性条件也将不能满足, 因此 Y 一定是一个非关键实体组合。

根据定理 1 和定理 2, 可以得出类似于数据挖掘中 Apriori 算法的方法来求最小关键实体组合, 本文称之为升序方法。在此方法中, 从给定的全通实体集的每一个基本的元素开始验证, 若满足连通性, 则加入到关键实体集中。否则将其后面的元素连到其后组成新的评估组合。每个组合只要符合条件, 根据定理 1, 那么它就是最小关键实体组合, 它的超集就可以不再评估以提高程序运行效率。 $ascendMode(Q)$ 升序算法步骤如下:

输入 查询要素 Q

输出 关键实体集

```
(1) 取得所有的  $N$  个全通实体(AllPass[i])并加入待评估队列中;
(2) For each Allpass[i]
(3)   if isConnect(AllPass)为 true
        if 当前实体组合不包含已加入的关键实体组合, then
            将其加到关键实体中;
        else continue;
(4)   else 把当前元素后的  $N-1$  元素连到后面, 形成  $N-1$  个新组合 AllPass[j];
```

```
(5)       For each AllPass[j]
```

```
(6)         递归调用 ascendMode;
```

通过实验发现, 升序方法对于节点数较少, 规则限制不严格的查询有很高的效率。但是, 当遇到严格的网络环境时, 它的效率将急剧下降。因此, 在采用此方法的同时, 本文又引入了降序方法来处理规则限制较多的情况。 $descentMode(Q)$ 算法步骤如下:

输入 查询要素 Q

输出 关键实体集

```
(1) 取得所有的  $N$  个全通实体(AllPass[n])将其成体进行评估;
(2) if isConnect(AllPass[n])为 false
(3)   then 没有关键实体组合;
(4) else 取得  $N$  个由  $N-1$  个 AllPass 中的元素组成的新评估集合  $N-1$ AllPass[N-1]
(5)   if 所有的 isConnect( $N-1$ AllPass[i])都为 false
(6)     then AllPass[n]为最小关键实体组合, 加入;
(7)   else For each isConnect( $N-1$ AllPass[i])为 true 的组合
(8)     递归调用 descentMode;
```

在此方法中, 从给定的 N 个全通实体的全集入手开始评估, 如果连通性满足, 则取它的 N 个不同的 $N-1$ 的实体组合进行评估, 若都不满足连通性, 则说明这 N 个实体的组合是最小的关键实体组合。整个过程是个深度优先(DFS)的递归过程。若遇到不满足的实体组合, 根据定理 2, 则它的所有子集都不再评估以提高效率。

3.3 判断条件

只要采用的合适的判断条件, 2 种方法的结合后的效率将大大高于传统的 Apriori 算法。本文采用的判断条件由已通过的有限定规则全通实体的个数和源目的端的可用路径数的比值决定(即每条路径上有限定规则的全通实体的个数)。经过不同的拓扑及网络环境测试发现, 在一般情况下, 当此比值即平均数大于 2 时, 降序方法效率将高于升序方法, 反之升序方法更加快速。

4 实验结果与分析

为了测试 NCA 以及关键实体算法的有效性, 在攻击图构建系统中建立了如图 2 所示的实验环境。为了简便, 省去了

中间 Router 所连接的子网。网络由 3 个防火墙(FW)以及 5 个 Router 组成,简称 F_n 和 R_n 。它们都是可以包含过滤规则的。攻击者试图通过主机 IP0(61.4.3.134)访问 IP4(202.120.101.58),这里主要测试这 2 台主机之间服务的连通性以及此过程中的关键实体集。

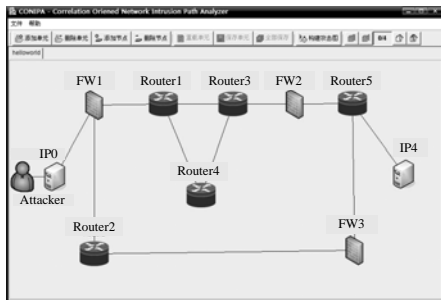


图 2 实验网络的拓扑环境

首先做一个典型的查询,通过 TCP 协议从 IP0 的 80 端口到 IP4 的 80 端口服务的可用性。当在所有实体上都加上限制规则的情况下,由于严格的规则限定,网络中的过滤设备将过滤掉数据包导致服务不可用。在 $F1, F3, R2, R5$ 没有限制规则的情况下,从拓扑图中可以看出,网络中将存在 1 条路径使得数据包能从源到达目的地。2 种情况下的测试结果分别如图 3、图 4 所示。



图 3 在所有网关都限制时的测试结果



图 4 在 $F1, F3, R2, R5$ 无限制规则时的测试结果

在对关键实体集的测试中,本文分别对拓扑中的网关设置各种不同的规则组合,以验证算法的有效性以及说明其优越性。

(1)环境 1:所有网关上都有对于此服务的限制规则。

(2)环境 2: $F1, F2, F3, R3, R4, R5$ 上有对于此服务的限制规则。

(3)环境 3: $F2, R1, R2, R3, R4$ 上有对于此服务的限制规则。

(4)环境 4: $F2, F3, R4$ 上有对于此服务的限制规则。

得到的 4 种环境中关键实体集如下:

(1)环境 1: $\{F1, F2, R1, R3, R5\}, \{F1, R2, F3, R5\}$;

(2)环境 2: $\{F1, F2, R3, R5\}, \{F1, R2, R5\}$;

(3)环境 3: $\{F2, R1, R3\}, \{R2\}$;

(4)环境 4: $\{F2\}, \{F3\}$ 。

可见,算法有效地发现了拓扑中的关键实体集,并且排除了干扰($R4$),达到了设计目的。

升序方法和降序方法在各个环境中的效率对比见图 5。

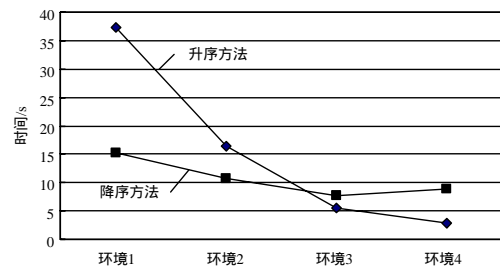


图 5 2 种方法的效率对比

升序方法和降序方法单独工作以及将它们结合的效率如图 6 所示。

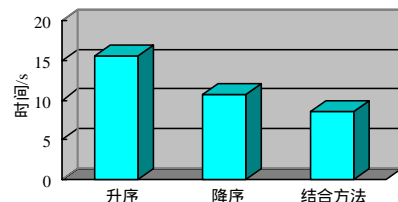


图 6 各种方法的效率对比

可以看出,本文的结合方法在升序和降序 2 种方法剪枝的基础上进一步有效地提高了查询的效率。

5 结束语

本文在实现攻击图构建系统的网络连通性分析的基础上,阐明了关键实体集在网络安全中的作用,并结合数据挖掘中的经典 Apriori 算法,提出一种有效的关键实体集求法。与传统算法相比,该算法能提供更快捷的查询,更高效地为网络管理员提供信息。

参考文献

- [1] Ritchey R, Ammann P. Using Model Checking to Analyze Network Vulnerabilities[C]//Proc. of IEEE Symposium on Security and Privacy. Oakland, California, USA: [s. n.], 2000.
- [2] Ritchey R, O'Berry B, Noel S. Representing TCP/IP Connectivity for Topological Analysis of Network Security[C]//Proc. of the 18th Annual Computer Security Applications Conference. Las Vegas, Nevada, USA: [s. n.], 2002.
- [3] Jajodia S, Noel S, O'Berry B. Topological Analysis of Network Attack Vulnerability[C]//Proc. of ASIACCS'07. Singapore: [s. n.], 2007.
- [4] Mayer A, Wool A, Elisha Z. Fang: A Firewall Analysis Engine[C]//Proc. of IEEE Symposium on Security and Privacy. Berkeley, California, USA: [s. n.], 2000.
- [5] Han Jiawei, Kamber M. 数据挖掘: 概念与技术[M]. 范明, 译. 北京: 机械工业出版社, 2001.

编辑 顾姣健