

基于内网行为分析的未知攻击检测模型

俞艺涵, 付钰, 吴晓平

(海军工程大学信息安全系, 湖北 武汉 430033)

摘要: 日益增多的未知攻击手段对内网造成安全威胁, 提出了一种基于内网行为分析的未知攻击手段检测模型。借助对内网信息资源充分可知的优势, 首先, 收集内网信息资源资料; 然后, 分析内网信息节点的行为异常风险要素; 最后, 以信息节点与信息资源获取路径为要素构建检测有向图模型。通过验证, 该模型可以达到预期的检测效果。

关键词: 内网行为; 未知攻击; 有向图; 检测模型

中图分类号: TP309.7

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2016.00058

Unknown attack detection model based on network behavior analysis

YU Yi-han, FU Yu, WU Xiao-ping

(Department of Information Security, Naval University of Engineering, Wuhan 430033, China)

Abstract: As for the intranet security threats of the increasing number of unknown attacks, an unknown attack detection model based on network behavior analysis was proposed. With the help of the information resources within the intranet, firstly, the information resources of the intranet were collected, then the risk factors of abnormal behavior of the internet information node were analyzed, finally, the information node and the information resources acquisition path as the key element were used to construct the detection model of the directed graph. By verifying, the model can achieve the desired detection results.

Key words: intranet behavior, unknown attacks, directed graph, detection model

1 引言

内网的安全性一直被人们广泛关注, 现有的内网安全防御技术在日益复杂多样的网络攻击手段面前, 显现出不能全方位以及长时间跨度内保证内网安全的情况。近年来, 发生的许多信息安全事件都是与内部网络相关的, 特别是日益猖獗

的APT (advanced persistent threat) 攻击^[1,2], 往往是针对一些蕴含信息价值高的内网实施的。

对已知网络攻击手段的检测技术相对成熟, 现阶段急需解决的是如何对那些未知的攻击手段进行入侵检测^[3,4]。通过分析发现, 任何对于内网的攻击, 在实施的过程中, 难免都会引起内网中网络行为特征的一些变化, 这些变

收稿日期: 2016-03-04; 修回日期: 2016-04-17。通信作者: 俞艺涵, cheniyike1992@163.com

基金项目: 国家自然科学基金资助项目 (No.61100042); 湖北省自然科学基金资助项目 (No.2015CFC867); 信息保障技术国防重点实验室基金资助项目 (No.KJ-13-111)

Foundation Items: The National Natural Science Foundation of China (No.61100042), The Natural Science Foundation of Hubei Province (No.2015CFC867), The National Defense Key Laboratory of Information Security Technology (No.KJ-13-111)

化可能是网络流量的增加，也可能是某个IP的频繁访问。

基于这一特点，本文构建了一种基于内网行为分析的有向图^[5]模型，通过利用对内网充分了解这一优势，收集内网行为资料，建立基于有向图的攻击检测模型，通过判断内网行为是否异常来对未知的网络攻击手段进行攻击检测。

2 相关定义

定义1 集合 $Tar = \{t_1, t_2, t_3, \dots\}$ 表示内网中的所有安全信息源，指的是内网中有价值的信息资源，这些信息源可能存在于内网中的任何一个设备中，也可能存在于内网操作者的思维中；集合 $IO = \{i_1\{b_1, b_2, b_3, \dots\}, \dots, i_n\{b_1, \dots, b_n, \dots\}, \dots\}$ 表示在内网中的所有信息节点，这些节点包括各类信息输入、输出设备、路由器、网络服务器以及内部人员等，任意一个 i_n 是唯一的， $i_n\{b_1, \dots, b_n, \dots\}$ 中的 b_n 表示在该信息节点可能出现的异常网络行为，包括频繁扫描、篡改权限、人员操作错误等，即信息节点的异常风险点；集合 $T\{\}$ 表示可能出现信息异常的节点；集合 $D\{\}$ 表示信息异常目的节点。

定义2 当某信息节点存在异常风险点时，所有可以利用此异常风险对信息异常节点进行信息获取的信息节点到信息异常节点存在一条有向边，用集合 $Z\{z_1, z_2, z_3, \dots\}$ 表示有向边。例如，信息节点可以通过异常风险点 b_1 对信息节点进行信息资源获取，则存在一条 i_2 到 i_1 的有向边 (i_2, i_1) 。

定义3 集合 $S\{s_1(i_n, i_m, i_l, \dots), \dots\}$ 表示信息资源获取路径， i_n, i_m, i_l 表示实施信息获取的节点，有向边 i_n 指向 i_m ， i_m 指向 i_l ； $Sre(s_n, i_l)$ ， $Sost(s_n, i_n)$ 分别表示信息资源获取路径 s_n 中的前驱信息节点序列和 i_n 的后续节点序列。对于路径 $s_n = (\dots, i_n, i_m, i_l, \dots)$ ， $Sre(s_n, i_l) = (\dots, i_n, i_m)$ ， $Sost(s_n, i_n) = (i_m, i_l, \dots)$ 。 $Otw(s_n, i_n, i_l)$ 表示信息获取路径 s_n 中 i_n 与 i_l 之间的节点，对于路径 $s_n = (\dots, i_n, i_m, i_l, \dots)$ ， $Otw(s_n, i_n, i_l) = (i_m)$ 。

3 未知攻击检测有向图模型建立

根据定义1~定义3，建立未知攻击检测有向

图，如图1所示，具体步骤如下。

1) 初始化。

$Tar = \{t_1, t_2, t_3, \dots\}$ ； $P = \{\}$ ； $S = \{\}$ ；

$IO = \{i_1\{b_1, b_2, b_3, \dots\}, \dots, i_n\{b_1, \dots, b_n, \dots\}, \dots\}$ 。

2) 安全信息源节点重构。

对 $Tar = \{t_1, t_2, t_3, \dots\}$ 进行重构，将 Tar 中的元素重构到不同的节点上，记为 $i_1\{b_1, b_2, b_3, \dots\}$ ， $i_2\{b_1, b_2, b_3, \dots\}, \dots, i_n\{b_1, \dots, b_n, \dots\}$ ，即异常网络行为节点。节点上所构建的安全信息源记为 $Tar_1, Tar_2, Tar_3, \dots, Tar_n$ ，且满足 $1 \leq m \neq l \leq n$ 时， $Tar_m \cap Tar_l = \emptyset$ ， $Tar_1 \cup Tar_2 \cup Tar_3 \cup \dots \cup Tar_n = Tar = \{t_1, t_2, t_3, \dots\}$ 。

3) 建立信息资源获取路径。

$T = \{i_1\{b_1, b_2, b_3, \dots\}, \dots, i_n\{b_1, \dots, b_n, \dots\}, \dots\}$

$D = \{i_1\{b_1, b_2, b_3, \dots\}, \dots, i_n\{b_1, \dots, b_n, \dots\}, \dots\}$

① 若 $IO \neq \emptyset$ ， $\forall i_m \in IO$ ， $Itmp = T$ 。

② 若 $\exists i_l \in Itmp$ 且 i_m 可以利用 i_l 的异常风险点对 i_l 进行信息资源获取，则 i_m 到 i_l 之间存在一条有向边 (i_m, i_l) ， $P = P \cup \{(i_m, i_l)\}$ 。

对于 S 中所有资源获取路径 s_n ，如果 $\exists i_{n_k} \in s_n, i_{n_k} = i_l$ ， $Sost(s_n, i_l) \neq \emptyset$ ，且 $(i_m, i_l, Sost(s_n, i_l)) \notin S$ 或者不存在资源获取路径 s_x 满足 $(i_m, i_l, Sost(s_n, i_l)) \in s_x$ ，那么 $S = S \cup \{(i_m, i_l, Sost(s_n, i_l))\}$ ；否则 $S = S \cup \{(i_m, i_l)\}$ 。

③ 若 $\exists i_l \in Itmp$ 且 i_l 到 i_m 存在一条有向边 (i_l, i_m) ，则 $P = P \cup \{(i_l, i_m)\}$ 。

对于 S 中所有资源获取路径 s_n ，如果 $\exists i_{n_k} \in s_n, i_{n_k} = i_m$ ， $Sre(s_n, i_l) \neq \emptyset$ ，且 $(Sre(s_n, i_l), i_l, i_m) \notin S$ 或者不存在资源获取路径 s_x 满足 $(Sre(s_n, i_l), i_l, i_m) \in s_x$ ，那么 $S = S \cup \{(Sre(s_n, i_l), i_l, i_m)\}$ ；否则 $S = S \cup \{(i_l, i_m)\}$ 。

④ $Itmp = Itmp - \{i_l\}$ ；若 $Itmp \neq \emptyset$ ，返回②继续计算。

⑤ $T = T \cup \{i_m\}$ ； $IO = IO - \{i_m\}$ ；若 $IO \neq \emptyset$ ，返回①继续计算。

4) 信息获取路径优化。

对于信息资源获取路径 S 中的某一条路径 s_x ，如果不存在信息资源节点 $i_m \in D$ ，则 $S = S - \{s_x\}$ ；如果存在 i_n 满足 $Otw(s_x, i_n, i_n) \neq \emptyset$ ，即信息资源获取路径 s_x 存在圈，则 $s_x = s_x - (i_n, Otww(s_x, i_n, i_n))$ 。

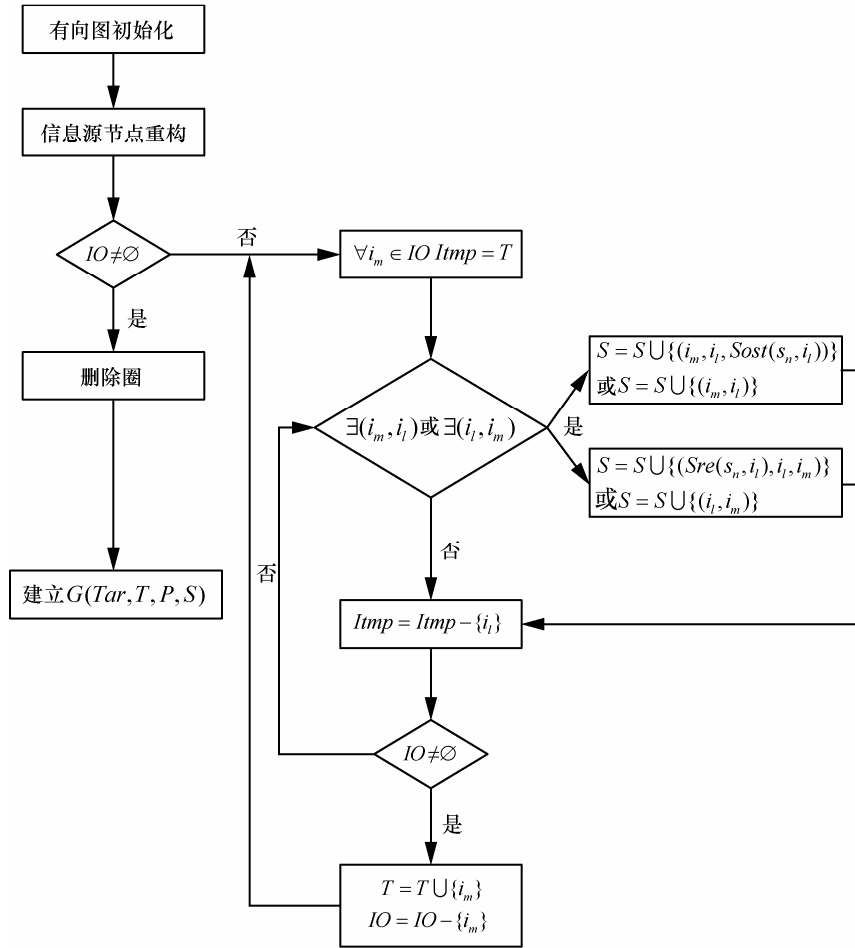


图 1 未知攻击有向图模型建立流程

5) 建立入侵检测有向图 $G(Tar, T, P, S)$ 。

4 未知攻击检测模型原理性验证

未知攻击检测有向图 $G(Tar, T, P, S)$ 是建立在对内网各个信息节点充分了解的前提下, 并通过对每个信息节点可能出现的异常风险的描述, 构建出信息资源获取路径, 用来反映当内网中出现行为异常时信息资源目标源被攻击所可能出现的情况。通过收集内网中的安全特征信息, 与未知攻击有向图模型进行路径匹配分析, 来判断内网是否遭受到网络攻击。

为了对模型进行原理性验证, 本文在实验室安全内网环境中模拟一种未知的网络攻击手段对内网进行攻击的过程。为了仿真体现出网络攻击手段的未知性, 本文着重设置了几个信息节点的异常风险因素, 来模拟未知攻击手段造成的内网网络中的异常。相关系统环境设置

如表 1 所示。

表 1 模拟环境设置		
节点名称	节点设备	利用的异常行为因素
i_1	操作者	权限扩大
i_2	Windows 计算机 1	频繁扫描
i_3	路由器 1	流量增加
i_4	Windows 计算机 2	源 IP 地址异常
i_l	交换机 1	目的 IP 地址异常

根据未知攻击检测有向图建立的过程可知, 该验证环境中最长的信息资源获取路径为 $s(i_1, i_2, i_3, i_4, i_l)$ 。所仿真的情况为内网操作者由于操作失误, 使操作权限扩大, 被 Windows 计算机 1 利用、进行频繁扫描, 使通过路由器 1 流量增加; Windows 计算机 2 通过修改源 IP 地址获取相关信息并由交换机 1 传出。

若在未知攻击检测的过程中, 设置异常参数

为 2, 则其信息资源获取路径分别为 $s_1(i_1, i_2)$ 、 $s_2(i_2, i_3)$ 、 $s_3(i_3, i_4)$ 和 $s_4(i_4, i_t)$, 4 条路径的相关性为 0, 与 $s(i_1, i_2, i_3, i_4, i_t)$ 的匹配程度也很低, 很难判断内网是否遭受了未知攻击; 如果设置异常参数为 5, 可以得到一条完整的信息资源获取路径 $s_x(i_1, i_2, i_3, i_4, i_t)$ 与 $s(i_1, i_2, i_3, i_4, i_t)$ 匹配, 可以判定内网遭受了网络攻击。

可以看出, 未知攻击检测有向图模型可以通过网络中异常行为来判断是否遭受网络攻击, 并不需要事先知道攻击的方式, 对于未知攻击检测是有效的。有向图模型对未知攻击检测的一个关键性因素为在获取内网异常特征时的异常参数设置, 异常参数设置越大, 其对入侵检测的准确性越高, 但是其检测过程中的匹配开销也越大^[6,7]; 另一方面, 内网信息节点中可利用进行信息资源获取的异常风险因素越多, 检测时的信息资源获取路径构建就越复杂, 检测的效果将降低。

由以上分析可知, 利用基于内网行为分析的未知攻击检测有向图模型进行内网未知攻击检测分析的有效性和检测开销与检测时的异常参数成正比, 与信息节点异常风险因素成反比。

5 结束语

针对日益增多的未知网络攻击手段, 本文通过分析当前内网安全防护需求, 以及整合内网在应对网络攻击时所拥有的资源, 提出了一种基于内网行为分析的未知攻击检测有向图模型。经过原理性验证, 该模型能够满足内网入侵检测需求, 达到了预期效果。然而, 在网络攻击手段越来越复杂, 越来越多样的今天, 内网的信息安全防护不可能仅仅依靠单一的技术, 本文所提的应对未知攻击的入侵检测模型可作为入侵检测的一个方面, 为后续研究起抛砖引玉的作用。

参考文献:

- [1] CHEN P, DESMET L, HUYGENS C. A study on advanced persistent threats[M]// Communications and Multimedia Security. Berlin Heidelberg: Springer, 2014:63-72.
- [2] 付钰, 李洪成, 吴晓平, 等. 基于大数据分析的 APT 攻击检测研

究综述[J]. 通信学报, 2015, 36(11): 1-14.

FU Y, LI H C, WU X P, et al. Detecting APT attacks: a survey from the perspective of big data analysis[J]. Journal on Communications, 2015, 36(11):1-14.

- [3] NIKOS V, DIMITRI G. The big four—what we did wrong in advanced persistent threat detection [C]// International Conference on Availability, Reliability and Security. c2013:248-254.
- [4] YANG G M Z, TIAN Z H, DUAN W L. The prevent of advanced persistent threat[J]. Journal of Chemical and Pharmaceutical Research, 2015, 6(1): 572-576.
- [5] 杨文君. 基于有向图的数据流模式发现算法研究[D]. 秦皇岛: 燕山大学, 2014.
YANG W J. Research on data stream pattern discovery algorithm based on directed graph[D]. Qinhuangdao: Yanshan University, 2014.
- [6] 杨蕊. 基于定量知识的分层有向图故障诊断方法及其应用[D]. 太原: 太原理工大学, 2014.
YANG R. Fault diagnosis method based on hierarchical directed graph with quantitative knowledge and its application[D]. Taiyuan: Taiyuan Institute of Technology, 2014.
- [7] 高东, 许欣, 马昕, 等. 基于符号有向图与趋势分析的故障诊断框架[J]. 计算机工程与应用, 2014(13):6-9.
GAO D, XU X, MA X, et al. Fault diagnosis framework based on sign directed graph and trend analysis[J]. Computer Engineering and Applications, 2014(13):6-9.

作者简介:



俞艺涵(1992-), 男, 浙江金华人, 海军工程大学硕士生, 主要研究方向为信息系统安全。



付钰(1982-), 女, 湖北武汉人, 博士, 海军工程大学副教授, 主要研究方向为信息安全风险评估。



吴晓平(1961-), 男, 山西新绛人, 博士, 海军工程大学教授, 主要研究方向为系统分析与决策。