网络资产测绘系统的实现

颜昭治

（中国移动通信集团广东有限公司汕头分公司，汕头 515043）

摘　要　网络资产测绘系统是网络安全信息收集的重要实现手段，能够收集网络上资产信息，管控安全风险，对安全治理起到举足轻重的作用。本文首先讲述了网络资产测绘涉及的端口扫描、指纹识别、情报收集和流量分析等关键技术，接着介绍了系统的实现方案及技术框架，最后重点详细阐述了系统在资产发现和指纹识别实现方面的具体做法。

关键词　网络安全；资产测绘；资产探测；指纹识别；端口扫描

中图分类号　TN918　　**文献标识码**　A　　**文章编号**　1008-5599（2022）07-0019-04

DOI:10.13992/j.cnki.tetas.2022.07.001

网络资产梳理是网络安全一项十分重要的基础性工作。网络环境错综复杂，存在着大量的无主资产、僵尸资产以及隐蔽资产，入侵攻击往往是以资产为载体或目标，时常威胁到网络的安全，如果缺乏技术手段自动发现识别网络资产，那么资产梳理难以全面、准确，网络安全将无法得到有效保障。网络资产测绘系统能够清晰收集网络上资产信息，及时发现未知资产和资产异动，有助于收敛资产暴露面，准确把控影响范围，进一步提高网络安全的目的。

1　关键技术

网络资产测绘是指追踪、掌握网络资产信息和资产属性的过程，在过程上可分为资产发现和指纹识别两个阶段，从形式上可分为主动和被动两种模式，涉及端口扫描、指纹识别、情报收集和流量分析等技术。

收稿日期：2021-09-10

1.1　端口扫描

端口扫描是网络资产主动探测的必要手段，通过扫描可以获取资产存活和端口开放等信息。在扫描技术类型上有 TCP 全连接、半连接、FIN、Null 和 UDP 扫描等。目前不同的扫描器所采用的扫描技术、扫描算法和扫描效果各不相同，在工作方式上可分为面向连接（同步模式）和无连接（异步模式）。面向连接扫描器可以识别丢弃的数据分组，具有较高的准确性，但缺点是其性能较慢。无连接的扫描器不依赖于当前被探测端口的完成来启动下一个端口，扫描速度较快，但由于无法检测丢弃的数据分组，扫描结果不太准确。

1.2　指纹识别

指纹识别是对暴露的资产做进一步探测，获取更多的资产属性信息。利用资产特有的指纹特征，识别出系统、设备、服务和应用的详细信息，可分为系统服务指纹识别和 Web 应用指纹识别两种类别。前者主要是识

别操作系统类型及版本、开放端口对应的服务名称及版本、设备类型及型号等；后者主要是识别网页标题、开发语言、开发框架、CMS 程序、第三方组件和中间件等。实现方法主要是判断返回分组差别、抓取标识语、解析 HTTP 头、比对特定文件 MD5 值以及匹配某些关键字等。

1.3 情报收集

情报收集主要是利用搜索引擎辅助完成网络资产信息获取，即依托 ZoomEye、FOFA 和 Shodan 等网络安全专用的搜索引擎对全网持续不断地扫描结果，通过调用 API 查询接口，间接地实现大规模网络资产测绘。不同于 Google、百度等传统搜索引擎是基于网络爬虫技术抓取网页内容，只能局限于 Web 相关资产的探测，专用搜索引擎更多的是对所有连接互联网的设备及其组件类型信息进行搜索。这种方式不仅高效、便捷、隐蔽，也避免了同目标资产直接交互，但受限于搜索引擎的数据获取能力，准确率相对较低，且无法应用到内网。

1.4 流量分析

流量分析是捕捉网络中流动的数据分组，基于流量分析技术，对数据分组的协议类型和数据内容进行综合分析，进而实现对网络资产信息的被动探测。在网络接口层，可从设备硬件差异识别网络资产类别；在网络层和传输层，可利用协议头部字段识别操作系统和服务端口信息；在应用层，可通过深度分析协议中的特征字段或负载数据，实现软件和组件等指纹检测。流量分析对目标网络运行的影响小，探测效果受限于所分析网络流量的全面性，并且由于需要获取目标网络的大量流量数据作为分析基础，该方法适用的网络规模有限。

2 实现方案

系统主要是利用 Python 调用 Nmap 端口扫描工具实现对资产存活、端口开放和服务信息的自动探测，并基于资产探测结果，通过抓取并分析标识语信息、URL 页面内容实现对设备类型、软件版本、开发语言、应用框架和网页标题等信息的识别。其技术框架如图 1 所示。

系统包含采集层、数据层和呈现层。在数据采集上，基于 Python 调用 Shell 命令和 PyPI 库自动下发采集任务；在数据处理上，对数据统一进行格式化处理并入库，通过数据库作为载体，提供上层数据应用；在数据呈现上，使用 Flask+Echarts 进行可视化呈现，将各种网络安全分析和统计数据以图表方式展示出来。

3 具体做法

3.1 资产发现的实现

Nmap、Masscan 和 Zmap 是目前常用的端口扫描工具，其各有优劣性。Masscan 和 Zmap 是基于异步无状态扫描工具，扫描速度超快，但准确性差且不支持操作系统和服务的探测。而 Nmap 是采用面向连接的基于响应协议栈指纹的扫描工具，虽然扫描速度较慢，但准确性高，不易造成漏报，并且能够获取更多资产属性信息，特别是针对端口服务类型的识别，能够通过对 HTTP/HTTPS 协议的发现，实现 IP 资产到 Web 资产的全量搜寻，因此系统使用 Nmap 作为资产发现的实现

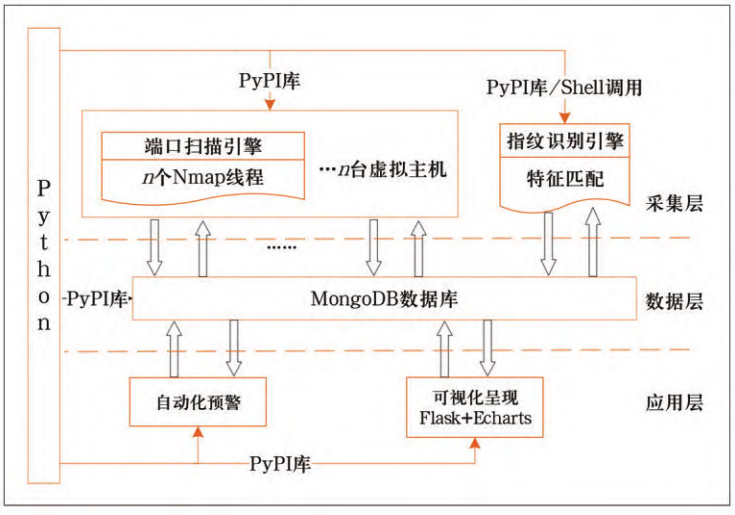

图1　系统技术框架图

手段。

系统引入Python-nmap和Python-libnmap库，利用Python自动化调用Nmap进行端口扫描，同时引入Pymongo库，将扫描结果保存在MongoDB数据库中，以便后续进行二次开发。为提高端口扫描效率，达到及时高效的目的，一方面采用多线程和分布式主从设计，单台主机实现多个Nmap线程调用，多台主机实现同时探测，并通过Docker/Vmware等虚拟化技术，提升服务器资源利用率，增强系统的并发能力；另一方面合理设置Nmap扫描参数，在确保不影响数据准确性的情况下，尽力缩短扫描探测时间，最终优化后的扫描参数是“-n -T4 -Pn -sSUV -O -open -host -timeout 10800s -pT:1-65535，U:53，U:69，…”。其中设置检测时间参数主要是考虑到网络性能较差或扫描目标限制等原因，导致某些IP需要很长的时间扫描，因此忽略这些IP扫描可减少时间消耗，提升整体扫描效率。UDP只针对常规端口进行扫描，主要是UDP全端口扫描时间长，扫描结果作用小。也可采用一种折中方法，即使用Masscan端口扫描工具作为辅助实现UDP全端口扫描。

3.2 指纹识别的实现

3.2.1 服务指纹识别

识别开放端口的服务名称和版本等信息可通过Nmap添加参数“V”实现，另外获取并分析标识语信息，识别结果会更加丰富完善。标识语通常包含了产品开发者、产品名称和版本、服务类型等敏感信息，与目标资产建立连接就可以获取。如图2所示，系统使用两种方法实现标识语信息识别，一种是利用Nmap脚本，只需要Nmap扫描时增加“-script=banner”参数；另一种是使用Python Socket编程，在端口扫描完成后，通过与开放端口建立Socket通信，接收目标资产发送过来的内容。目标资产可能对标识语信息进行了隐藏或修改，从而影响到识别效果。

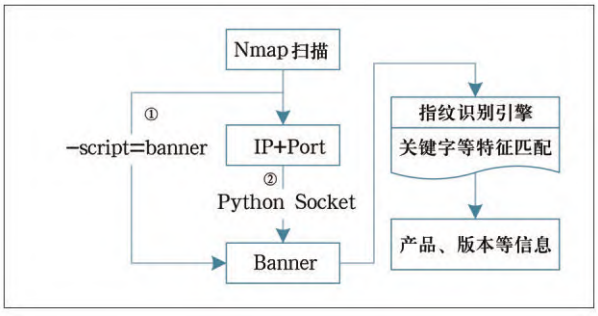

图2　标识语信息识别实现方法

3.2.2 网站指纹识别

系统实现网站指纹识别的方法是先从Nmap扫描结果中筛选出启用HTTP/HTTPS服务的资产，接着使用Python Requests GET请求爬取网页内容，再经指纹识别引擎对Header/Body内容进行分析和判断，进而识别出中间件、页面标题和CMS等信息，如图3所示。为使识别效果更加全面、准确，系统汲取了多个开源Web指纹库，同时也整合了Whatweb/Wappalyzer等工具识别结果，具体通过Python调用Shell命令Whatweb和引入Python-Wappalyzer库实现。由于指纹库越是丰富，识别效果越是理想，对于未识别出结果的目标资产，后续将利用手工分析不断追加指纹特征到自定义指纹库中。

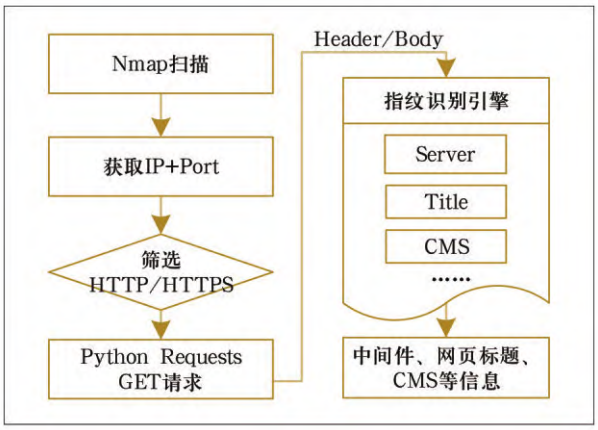

图3　网站指纹识别实现方法

4　结束语

通过对所辖IP地址进行网络资产测绘，可全面掌

握网络中资产分布、设备类型、网站类型、网站插件和开放端口等多维内容，并可结合漏洞扫描引擎和安全预警公告，甚至安全渗透测试，多方位评估资产漏洞情况和风险状况，进而对网络资产进行完整清晰的画像，令网络安全治理工作可以有的放矢，同时将安全风险的爆发扼杀在萌芽状态。

参考文献

[1] 王宸东，郭渊博，甄帅辉，等. 网络资产探测技术研究[J]. 计算机科学，2018(12).

Implementation of network asset probing system

YAN Zhao-zhi

(China Mobile Group Guangdong Co., Ltd. Shantou Branch, Shantou 515043, China)

Abstract Network asset probing system is an important means of network security information collection. It can help to find out the family background, control security risks, and play an important role in security management. This paper first describes the key technologies involved in network asset probing, such as port scanning, fingerprint identification, information collection and traffic analysis, then introduces the implementation scheme and technical framework of the system, and finally focuses on the specific methods of the system in asset discovery and fingerprint identification.

Keywords network security; asset probing; asset mapping; fingerprint identification; port scanning