

面向 VMware 的虚拟机远程检测技术

田雪梅, 蔡皖东

(西北工业大学 计算机学院, 陕西 西安 710129)

摘 要: 根据 MAC 地址的分配机制和构造原理, 自动识别网络中的 VMware 虚拟机, 同时, 根据 VMware 虚拟机的工作模式及通信机理, 对 VMware 进行分析和检测. 对比本技术与网络扫描嗅探工具 NMAP 的检测结果, 证明了该技术在识别和检测 VMware 虚拟机时有更高的准确率和完整性.

关键词: VMware; 虚拟机; 识别; 远程检测; MAC

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2014)01-0029-05

The Remote Detection Technique for VMware

TIAN Xue-mei, CAI Wan-dong

(College of Computer, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: A new remote detection method was proposed. By analyzing the allocation mechanism and construction of MAC address, this method can identify automatically whether there are any VMware in the network, at the same time, it can analysis and detect the VMware by research the working pattern and communication theory of VMware. By comparing the test results of this method and NMAP, the simulation results show that this method has a higher accuracy and completeness in the identification and detection of the VMware.

Key words: VMware; virtual machine; identify; remote detection; MAC

1 引言

随着云计算技术的发展, 作为云计算的核心技术——虚拟化技术应用日益广泛. 虚拟机成为大型企事业单位争相使用的资源, 导致了虚拟机的滥用问题. 这造成了管理虚拟机的困难, 如何检测到网络中虚拟机的存在并对其进行监测管理, 进而发现虚拟机中存在的安全隐患值得研究.

文献[1]中介绍了 VMware 虚拟机的三种网络工作模式, 但是并没有深入分析 VMware 虚拟机在桥接模式下的通信原理. 文献[2-3]中介绍了各种面向真实主机进行扫描检测的算法.

上述文献并没有提供虚拟机远程检测技术. 因此本文提出了一种虚拟机远程检测技术. 本文主要研究了以下四个方面: 分析了 VMware 虚拟机在桥

接模式下的通信原理; 自动获取目标主机群的 MAC 地址, 探测 VMware 虚拟机的存在; 完成了 VMware 虚拟机的远程检测; 通过实验对比了本方法与 NMAP 的检测结果, 证明了本方法的有效性.

2 相关概念

2.1 虚拟机

虚拟机(Virtual Machine)是指可以通过软件模拟出完整硬件系统功能, 建立一个完整的计算机系统的应用软件. 通过虚拟机, 我们可以在一台物理计算机上模拟出一台或者多台虚拟的计算机, 这些虚拟机可以像真正的计算机那样安装操作系统、安装应用程序、访问网络资源等. 如今主流的虚拟机产品主要有 VMware、Virtual Box^[4]、Virtual PC^[5]等, 其中应用最广的是 VMware, 因此本文主要研究面

收稿日期: 2013-03-20; **修回日期:** 2013-05-22

基金项目: 西北工业大学基础研究基金(JC201149); 2013 年陕西省科学发展计划项目(2013K06-19); 西北工业大学研究生创业种子基金资助(Z2013124)

向 VMware 的虚拟机远程检测技术.

2.2 MAC 地址

MAC(Media Access Control)地址是用来定义网络设备的位置,由 48 比特长,12 位的十六进制数字组成,0 到 23 位是厂商向 IETF 等机构申请用来标识厂商的代码,是唯一的标识符,24 位到 47 位是由厂商自行分派的,是各个厂商中唯一的编号.

3 VMware 虚拟机的工作模式

虚拟机网卡的工作模式主要有三种^[1]:桥接模式、NAT 模式、Host-only 模式.本文涉及到的虚拟机工作模式是桥接模式.

桥接模式是三种工作模式中最简单的模式,宿主机和虚拟机之间是通过 VMnet0 虚拟交换机进行桥接,如图 1 所示.在这种模式下,宿主机和虚拟机同处于一个网段,具有相同的子网掩码,虚拟机在局域网中就像是一台真正的主机一样,享受宿主机的各种服务,比如打印、文件共享服务等,它能够访问局域网中的任何客户端,也可以访问广域网.

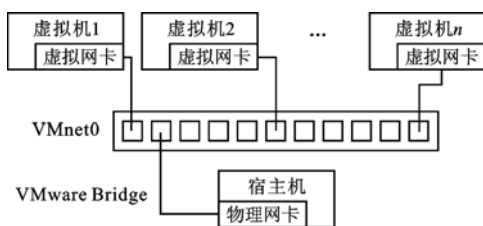


图 1 桥接模式的网络拓扑图

当虚拟机的网卡工作模式为桥接模式时,实际上就是将虚拟机直接接入到物理网络中,将虚拟机和宿主机放到了同等的位置.接入到物理网卡上的网线可以理解为直接接到 VMnet0 虚拟交换机上.此时物理网卡处于混杂模式,接收网络中的所有数据包,物理网卡同时监听包含宿主机和所有安装在该宿主机上的虚拟机的多个 IP 地址,遵循 Ethernet 的 CSMA/CD 协议进行通信的争用共享.

虚拟机在桥接模式下的通信涉及两个方面:一是接收数据,二是发送数据.

(1)接收数据

当数据到达物理网卡时,如果目的地址在监听的 IP 地址范围内,则物理网卡接收该数据.当目的地址与宿主机地址匹配时,物理网卡直接将数据传递给上层协议;当目的地址为某个虚拟机的地址时,物理网卡通过 VMware Bridge Protocol(桥接协议)将数据转发到虚拟交换机,再由虚拟交换机将数据转发给真正的目标虚拟机.

(2)发送数据

当虚拟机需要发送数据到网络中的另一台目标主机时,数据经由自己的虚拟网卡首先到达虚拟交换机,再经由宿主机的物理网卡发送.

4 VMware 虚拟机远程检测技术

本文的主要目的是提出一种识别某一网段内是否存在 VMware 虚拟机,进而实现 VMware 虚拟机远程检测的技术.本技术的核心步骤如下:

- (1)获取所有开启主机的 MAC 地址;
- (2)自动识别 VMware 虚拟机;
- (3)远程检测第二步探测到的所有 VMware 虚拟机.

为了方便下文中算法的描述首先将本技术中用到的主要变量和文件做以下定义,如表 1 所示.

表 1 变量及文件说明

变量名称	类型	变量含义
ipBegin	ip 地址	待扫描区域的起始 IP
ipEnd	ip 地址	待扫描区域的终止 IP
ip	ip 地址	目标主机 IP 地址
mac	字符串	响应信息中的原始 MAC 字符串
lineList	动态数组	读取 mac.txt 中的信息
port	整数	目标主机的端口
mac.txt	文档	存储自动获取的 MAC 地址
vir.txt	文档	存储探测到的虚拟机信息
port-service.txt	文档	存储常用端口及其对应的服务

4.1 MAC 地址自动获取

为了检测 VMware 虚拟机的存在,必须获得网段内所有主机的 MAC 地址.本文是在已知扫描区域的前提下,由测试主机向所有主机发送 ARP 广播,根据响应信息判断网段内所有主机的开启情况,进而提取各个主机的 MAC 地址. ARP 数据包格式如图 2 所示.



图 2 ARP 数据包格式

算法首先等待用户输入待扫描区域,即 ipBegin 和 ipEnd;初始化变量 ip 为 ipBegin.具体算法伪代码如下:

```
While (ip<=ipEnd)
```

```

{
    向所有主机发送 ARP 广播,查询 ip 对应主机的 MAC
    地址;
    如果返回的数据不为空
    {
        提取 ARP 应答中 ip 对应的 MAC 地址;
        存储 ip 对应主机的 IP 和 MAC 地址于 mac.txt 文
        档中;
    }
    ip++;
}

```

例如,对应主机 ip 为 10.13.32.198 时,发送的 ARP 广播如图 3(a)所示,返回结果如图 3(b)所示。

0000	ff	ff	ff	ff	ff	ff	00	13	d3	2b	63	3a	0a	0d	01		..+C1....
0010	08	00	06	04	00	01	00	13	d3	2b	63	3a	0a	0d	20	ba	
0020	00	00	00	00	00	00	0a	0d	20	c6	00	13	d3	2b	63	3a	0a

(a) ARP 请求包

0000	00	13	d3	2b	63	3a	0a	e8	2e	14	ed	0a	0d	20	c6		..+C1....
0010	08	00	06	04	00	02	00	0a	e8	2e	14	ed	0a	0d	20	c6	
0020	00	13	d3	2b	63	3a	0a	0d	20	ba	00	00	00	00	00	00	
0030	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	

(b) ARP 应答包

图3 ARP 数据包

从图 3(b)可以看出,28 字节的 ARP 应答为十六进制字符串:00 01 08 00 06 04 00 02 00 0A EB 2E 14 ED 0A 0D 20 C6 00 13 D3 2B 63 3A 0A 0D 20 BA。

对比上述字符串与图 2 所示的 ARP 数据包格式,0X0001 代表硬件类型为以太网,0X0800 表示所用协议为 IP 协议,0X06、0X04 分别代表硬件地址(MAC 地址)长度为 6,协议地址长度为 4,OP 为 0X0002 代表 ARP 应答,紧接着的字符串 0X000AEB2E14ED 即为 ip 为 10.13.32.198 对应的 MAC 地址,由此可知只要提取 ARP 应答中的第 9~14 个字节即可取得 ip 对应主机的 MAC 地址。

通过上述算法即可获取待扫描区域内所有开启主机的 MAC 地址。在提取目标主机的 MAC 地址时,需注意两个问题:

(1) 逆序转换

从目标主机返回的响应信息中提取的 mac 字符串是逆序的,例如获取的某一个虚拟机的 MAC 地址字符串“A67503290C00”,但是真实的 MAC 地址字符串为“000C290375A6”,因此需要将 MAC 地址进行特殊的逆序转换。

(2) 缺位补零

所谓缺位补零是指响应信息中的 MAC 地址信息在转换为字符串时会省略信息前端的“0”字符,致使转换后的 MAC 地址与实际不符。因此在进行转换之前首先要检测 mac 字符串的长度是否为 12,不足时需在 mac 字符串之前添“0”补足。

例如,目标主机的 MAC 地址为“00:0C:29:E2:C5:0D”,从响应信息中提取的 mac 字符串为“DC5E2290C00”,经过逆序转换之后为“00:0C:29:E2:C5:D0”,与实际不符。究其原因是由于 mac 字符串长度不足 12,在进行逆序转换之前首先要检查 MAC 地址字符个数是否为 12 位,不足时要在字符串起始位置补“0”凑足 12 位,再进行逆序转换。

4.2 VMware 虚拟机识别

根据 MAC 地址的分配机制和构造原理,本文在分析大量 VMware 虚拟机特征信息后发现,不管是哪种版本的 VMware 虚拟机,在创建新的虚拟机时,分配给 VMware 虚拟机的 MAC 地址是有一定规律的,即分配给 VMware 虚拟机的 MAC 地址的前三个字节是固定不变的,绝大多数 VMware 虚拟机的 MAC 地址前三个字节为“00:0C:29”。

结合各种资料和实践经验,统计发现分配给 VMware 生产厂商的标识符还有“00:05:69”、“00:1C:14”、“00:50:56”三种。因此,一般情况下如果获得的 MAC 地址中前三个字节若为以上四种字符串中的一种,即可断定该 MAC 地址对应的主机为 VMware 公司生产制造的。

本文根据上述理论分析和实践验证,证明确实可以根据 MAC 地址来判断目标主机是否为 VMware 虚拟机。识别 VMware 虚拟机的具体步骤如下:

第一步,将 mac.txt 文档中的信息读取到动态数组 lineList 中;

第二步,从 lineList 中读取一条记录,提取记录中 MAC 地址的前三个字节,对比所提取的字节与前文中提到的 VMware 厂商标识,如果相同即为 VMware 虚拟机,将该记录存储到文档 vir.txt 文件中,方便进行虚拟机远程检测。

第三步,重复第二步直至将 lineList 中的记录遍历完。

通过以上步骤即可识别出网段内的所有 VMware 虚拟机,各 VMware 虚拟机的信息存储在 vit.txt 文档中,为后续的虚拟机远程检测提供原始数据。

4.3 虚拟机远程检测

本文提出的虚拟机检测方法主要是通过主动探测^[6]的方式,向被确认为是 VMware 虚拟机的目标主机发送自定义的数据包,提取响应信息中虚拟机的特征信息。

经过第二步,已得到了待扫描网段内的所有

VMware 虚拟机的 IP-MAC 地址信息,并存储在 vir.txt 文档中,随后即可对各 VMware 虚拟机进行检测,获取特征信息.

本文重点从两个方面实现目标虚拟机检测:一是远程探测目标虚拟机的主机名;二是检测目标虚拟机开启的端口及服务.除却检测到的 IP、MAC、虚拟机类型、主机名和端口状态及服务,还包括操作系统等特征信息.分析检测到的特征信息,可以初步断定目标虚拟机的安全状况:探测主机名可以基本判断目标虚拟机有无开启防火墙,扫描端口可以探测目标虚拟机是否开启了未知服务,两者综合起来可以简单判定目标虚拟机是否存在安全隐患.

(1) 获取主机名

获取目标主机名是向目标主机的固定端口发送自定义的数据包,提取响应信息中的特征信息.如果可以直接从响应信息中提取目标主机的主机名,则证明目标主机未开启系统防火墙,存在安全隐患;反之,目标主机相对安全.具体步骤如下:

Step1:向目标主机的 137 端口发送如图 4 中标注的 UDP 数据;

Step2:抓取目标主机的响应信息;

Step3:若响应信息为空,转至第四步,否则,提取响应信息中的主机名,提示用户目标主机存在安全隐患.

Step4:提示用户目标主机已开启系统防火墙,相对安全.

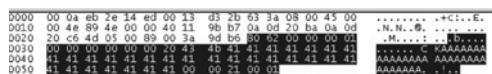


图 4 UDP 请求(标注部分为 UDP 数据内容)

(2) 端口扫描

端口扫描^[2-3]包括 TCP 端口扫描和 UDP 端口扫描,均是向目标主机的服务端口发送自定义的探测数据包,根据响应信息判断端口的状态.如果端口的状态是开启,需遍历 port-service.txt 文档,若查找到直接输出端口对应的服务;否则需提示用户查看端口是否存在安全隐患,如图 5 所示.

算法首先初始化变量 port 为 1,具体步骤如下:

①判断变量 port 是否小于 1024,若小于跳至第二步;否则,算法结束.

②发送相应的协议数据包,接收响应信息;

③根据响应信息判断目标主机端口的状态,若端口开启转至第四步;否则转至第七步;

④遍历 port-service.txt 文档,若查找到 port,转至第五步;否则,转至第六步;

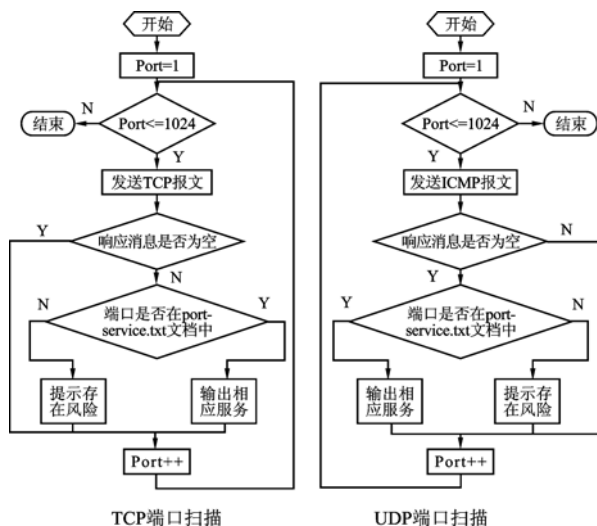


图 5 端口扫描流程图

⑤输出端口对应服务,跳至第七步;

⑥提示存在风险,跳至第七步;

⑦执行 port++,跳至第一步循环;

在上述算法中需要注意以下几点:

第一,在第二步中,当进行 TCP 端口扫描时,直接发送 TCP 数据包.由于向目标主机端口发送 UDP 数据包时,不管端口状态如何,均无响应信息返回,因此,在进行 UDP 端口扫描时,需发送 ICMP 数据包.如图 5 中标注的部分.

第二,在第三步中,判断端口的开启状态时, TCP 和 UDP 是不同的,接收到 TCP 数据包的响应信息说明目标 TCP 端口是开启的,而 UDP 端口只有在未接收到 ICMP 报文的响应信息时才是开启状态.

第三,进行 UDP 端口扫描时,要将测试主机 IP 于某一端口进行绑定,便于收发数据.本文将测试主机 IP 绑定在测试主机的 80 端口.

5 实验结果及分析

5.1 实验环境

本实验主要在局域网中进行测试,具体实验环境如表 2 所示.

表 2 实验环境

资源名称/类型	配置
测试 PC	Windows XP 系统、硬盘 80 GB、内存 960 MB、主频 3.0 GHz
测试范围	10.13.32.0/24
应用软件	VS2008、NMAP、WireShark

5.2 实验结果及分析

实验用现有的技术 NAMP 与本技术进行对

比.表3显示了NMAP和本技术在同种环境下的一次检测结果及性能.其中准确率指的是准确识别的VMware虚拟机的比率;检测到的信息包括VMware虚拟机的IP、MAC、主机名、虚拟机类型、操作系统和开启端口及服务;完整性是指检测到信息的完整度.分析表3中的数据,说明在同等条件下,本技术可以准确识别到更多的VMware虚拟机,检测到更完整的VMware虚拟机信息.

表3 性能对比

方法	面向对象	识别虚拟机数	准确率	检测到的信息	完整性
实际	/	20	/	328	/
NMAP	真实主机	18	0.9	238	0.727
本技术	VMware	20	1	322	0.982

为了尽量减少由于客观或主观方面原因造成的影响,证明本技术的优越性,两种技术均在不同时段下进行了多次测试,结果如图6所示.

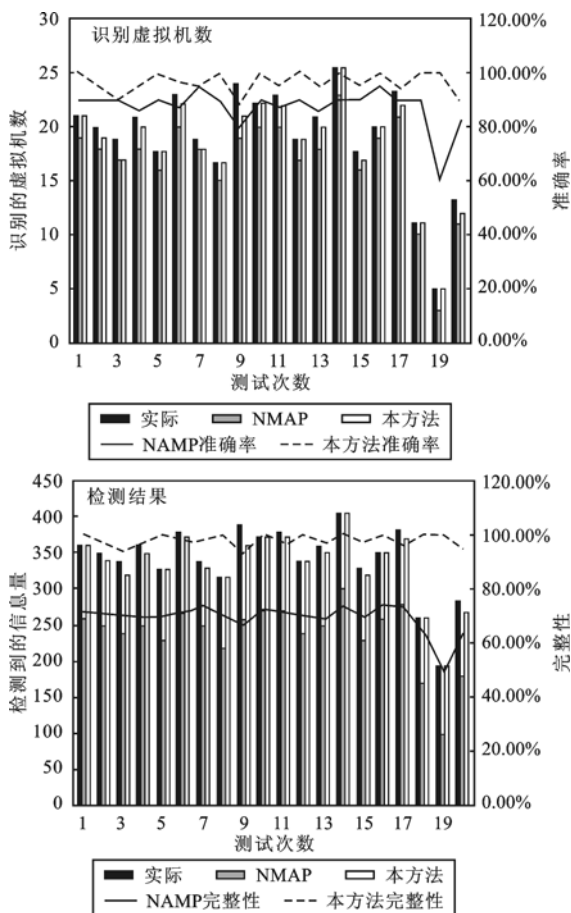


图6 性能比较

在图6中,横坐标表示测试的次数,主纵坐标分别代表识别的虚拟机数和检测到的虚拟机特征信息量,次纵坐标分别代表完整识别虚拟机的准确度和检测到特征信息的完整性,柱形表示数量,折线代表概率.

分析图6所示的结果,可以看出本技术在识别虚拟机时的准确度可达到95%以上,检测虚拟机信息的完整性高达90%以上.实验结果证明本技术在识别和检测VMware虚拟机时有更高的准确率和完整性.

6 结束语

提出了识别VMware虚拟机的方法,同时,实现了VMware虚拟机的远程检测,为实现虚拟机的监测管理和漏洞扫描奠定了基础.对比分析本技术与NMAP检测结果,证明了本技术有更高的准确度和完整性.

参考文献:

- [1] 庄城山. VMware虚拟机网络工作模式的研究[J]. Computer Knowledge and Technology, 2011(9):2008-2009.
- [2] Bhuyan M H, Bhattacharyya D K, Kalita J K, et al. Surveying port scans and their detection methodologies [J]. The Computer journal, 2011, 54(10):1565-1581.
- [3] Wang Liang. Design and application of network scanning system [J]. Technology and Application, 2011 (12):44-46.
- [4] Cameron Sturdevant. xVM virtual box poses new challenges to VMware [J]. Network world, 2009, 26(9): 41-42.
- [5] Microsoft's virtual PC 7.0 [J]. Network world, 2009, 26 (23):38.
- [6] Genevieve Bartlett, John Hedemann, Christos Papadopoulos. Understanding passive and active service discovery [C]// 7th ACM SIGCOMM Internet Measurement Conference. Korea, Seoul, 2007.

作者简介:

田雪梅 女, (1988—), 硕士研究生. 研究方向为网络信息安全.

蔡晓东 男, (1955—), 教授, 博士生导师. 研究方向为网络信息安全.