

基于 NetFlow 的蠕虫病毒监控系统设计与实现

汪精明, 赵晓峰, 王平水

(安徽财经大学 网络中心, 安徽 蚌埠 233041)

摘 要: NetFlow 是 Cisco 公司在其交换、路由体系中采用的一种三层交换技术。NetFlow 服务能够提供包括地址、协议、端口和服务类型等详细的数据流统计信息。由于蠕虫病毒传播过程中会发起大量的扫描连接, 使用 Flow-tools 等工具统计分析 NetFlow 数据流, 可以很容易地找出染毒计算机 IP 地址, 再利用 SNMP 准确地定位该 IP 的位置并关闭其所连交换机端口, 就可以将染毒计算机与网络隔离。同时对染毒计算机相关信息的详细记录, 为网络管理员对染毒计算机的处理提供了准确的信息。

关键词: NetFlow; SNMP; Flow-tools; 蠕虫病毒

中图分类号: TP393.07

文献标识码: A

文章编号: 1673-629X(2007)05-0117-04

Design and Implementation of System for Supervising Worm Virus Based on NetFlow

WANG Jing-ming, ZHAO Xiao-feng, WANG Ping-shui

(Network Center, Anhui University of Finance & Economics, Bengbu 233041, China)

Abstract: NetFlow is one kind of layer 3 switch technologies which is used in switch and route system of Cisco. The NetFlow service can provide the detailed data stream statistics information about the address, the protocol, the port and the service type and so on. Using the tools such as Flow-tools to analyze the data stream of NetFlow, it is easy to discover the IP addresses of virus-infected computers, since the worm virus can initiate massive scanning connection during their spreading process. Using SNMP, the switch ports linked with the computers which use these IP addresses can be located and the virus-infected computers can be isolated from the network by closing the corresponding switch ports. At the same time the recorded detailed information of the contamination computers may help administrator to deeply analyze and process.

Key words: NetFlow; SNMP; Flow-tools; worm virus

0 引言

1988 年莫里斯从实验室放出第一个蠕虫病毒以来, 计算机蠕虫病毒以其快速、多样化的传播方式不断给网络世界带来灾害, 成为网络中严重的安全隐患。1999 年以来, 蠕虫病毒进一步演化, 结合漏洞攻击与黑客技术病毒大量出现, 对网络威胁进一步加大。在企事业单位、学校这类园区网中, 因网络速度快、计算机集中、用户防范力差等原因, 蠕虫非常容易传播、爆发, 网络在大范围病毒攻击时, 短时间内就会瘫痪! 针对这一巨大网络安全威胁, 网络管理者应建立一套有效的蠕虫探测与隔离机制, 通过自动探测、自动隔离, 将蠕虫的攻击扼杀在萌芽状态。

1 蠕虫病毒扫描探测行为特征分析

蠕虫病毒通过攻击有漏洞的主机才能实现自身传播, 网络上有漏洞的主机蠕虫必须查找才能发现, 通常蠕虫会在很短时间内, 发送成千上万的探测包到某网段。如果蠕虫是通过 TCP 进行扫描传播(绝大部分蠕虫都是通过这种方式), 在扫描传播时, 会发送大量 TCPSYN 包以查找目标网段存在漏洞主机。一般来讲, 蠕虫查找目标系统漏洞是通过随机扫描、顺序扫描实现的。随机扫描是指蠕虫随机选择一段 IP 地址进行主动探测, 以试图发现存在漏洞的可攻击主机系统; 顺序扫描则是蠕虫在扫描一段 IP 地址时, IP 地址以递增或递减的方式进行。很明显, 当一台感染蠕虫病毒的主机跨网段进行漏洞扫描时, 在路由器上显示出最明显特征是: ①在一个短的时间段, 向不同 IP 地址发起首次连接数很多; ②收到首次连接请求响应很少; ③针对一个或几个目的端口产生首次连接请求; ④发

收稿日期: 2006-09-11

基金项目: 安徽省 2006 年教育厅自然科学基金项目(2006kj017C)

作者简介: 汪精明(1957-), 男, 江苏江阴人, 实验师, 研究方向为网络管理、网络安全。

出连接请求数据包数量和大小都是有固定值。因此,蠕虫病毒的这种非理性扫描探测行为是其网络行为中最明显的特征。

2 监控模式设计

蠕虫病毒跨网段大规模扫描探测,经路由器或核心交换机三层路由模块与大量目标 IP 产生连接,定期采集流经路由器或三层路由模块数据流,对数据流的连接源 IP 地址、目的 IP 地址、源端口号、目的端口号、发送包数量、发送字节数等信息,进行统计、分析,很容易就能发现蠕虫病毒产生的大量异常连接。

2.1 NetFlow 介绍

NetFlow 是 Cisco 公司首创的为提高路由设备路由转发能力而在交换、路由体系中采用的一种三层交换技术,目前已成为事实工业标准,被包括 Juniper, Extreme, Foundry 等大多数主流厂商路由器和三层交换机支持^[1]。其工作过程如图 1 所示。

从图中可以看到经路由器建立的每一个连接,其每一个数据分组仍然采用一般的三层路由/交换方式,

NetFlow 高速缓存中,再进行通常路由转发。启动 Cisco 路由器或三层交换机路由模块提供的 NetFlow 服务功能,可定期将 NetFlow 条目及相关统计信息以 UDP 报文发送给一台接收主机。NetFlow 流记录格式目前主要有 V1, V5, V6, V7, V8, V9 六个版本^[2],其中 V5 是路由器上使用最广泛的格式,对于 V5 版本,每个从路由器上发送到接收主机 UDP 报文中包含 1 个 Flow 包头和 30 条 Flow 记录,每个 Flow 记录包含的主要字段在表 1 列出。

表 1 版本 5 流记录主要内容

字节	内容	描述
0~3	srcaddr	源 IP 地址
4~7	dstaddr	目的 IP 地址
8~11	nextthop	下一跳的路由器 IP 地址
12~13	input	输入接口的 SNMP 索引
14~15	output	输出接口的 SNMP 索引
16~19	dPkts	流中的报文数
20~23	dOctets	在流的报文中第 3 层字节总数
32~33	srcport	TCP/UDP 源端口号或等价值
34~35	dstport	TCP/UDP 目的端口号或等价值
37	Tcp-flags	TCP 标记的累积 OR
38	prot	IP 协议(例如 1=ICMP,6=TCP,17=UDP)

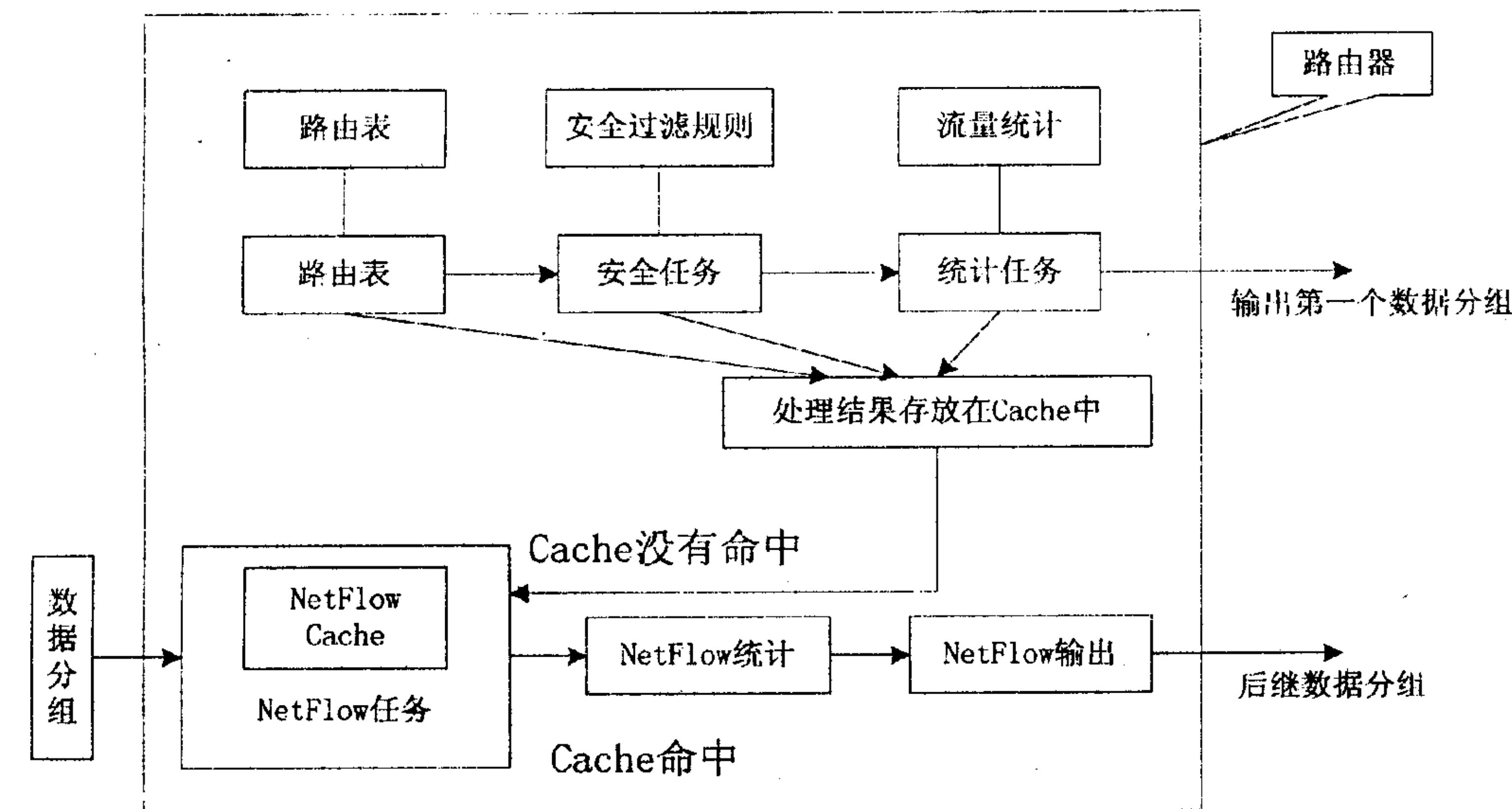


图 1 NetFlow 工作原理图

当路由器接收到某一连接的第一个数据分组时,会将该分组信息记录在 NetFlow 高速缓存中,形成一条 Flow 条目,Cisco 公司在文档中描述各 Flow 是通过序列<源 IP 地址,目的 IP 地址,源端口号,目的端口号,协议类型,TOS,输入逻辑接口号>来区分的,因此所谓流(Flow)就是由上面 7 个关键元素唯一确定的两个通信终端之间的单向网络连接。路由器判断某一分组是否是第一个分组,是通过扫描缓存中上面 7 个字段组成的序列来判断的,如果缓存条目中有此分组匹配条目,就使用缓存条目路由信息及相关访问控制策略将此分组直接交换转发,否则该分组信息会记录到

由于路由器存放 NetFlow 条目缓存有限,所以它会通过一些策略清除过期 Flow,以下三种情况的 Flow 会被作为过期 Flow 清除:

- 1)传输已经完毕(TCP FIN 或 RST);
- 2)Flow 空闲计时器已超过 15 秒;
- 3)Flow 活动计时器超过 30 分钟。

NetFlow 服务会在清除过期 Flow 之前,将 Flow 条目信息向接收主机发送。目前,Windows 下处理 NetFlow 信息的软件不多,但在 Linux 系统有不少,Fullmer 的 Flow-tools 目前使用者较多。

2.2 Flow-tools 介绍

Flow-tools 是 Linux 下一套处理 Flow 信息的软件集合,flow-capture 命令具有接收 NetFlow 信息功能,flow-print、flow-stat 等命令用来查看、统计接收 Flow 信息,通过 Flow-tools 提供的工具结合脚本程序,已经可以简单实现对网络流信息分析和统计,但要进行复杂分析、统计,光依靠这些工具和脚本很难达到。好在 Flow-tools 提供了 Flow 信息导出功能,利

用 Flow - export 命令可将接收的 Flow 信息导出到 MySQL 或 PostgreSQL 数据库中,利用数据库快速、强大数据处理功能,可以迅速方便统计、分析海量接收的 Flow 信息^[3]。

2.3 基于 NetFlow 对蠕虫病毒的监测

针对蠕虫病毒进行大规模探测这一鲜明特征,对收集的 NetFlow 信息进行分析,可过滤出染毒计算机 IP 地址,以下介绍三种实现方法。假设定期收集的 NetFlow 信息导入 MySQL 的 test 库 test 表中, test 表中各属性名与表 1 内容列名称相同。

2.3.1 TOPN 模式

蠕虫病毒不断扫描探测,会对大量目标 IP 一个或多个端口发出超出正常数量的连接请求,并在很短时间里,对大量目标 IP 发出大量数据包。因此只要统计出源 IP 建立连接数 TOPN 位,或源 IP 采样时间内发送数据分组 TOPN 位,就可找出染毒计算机 IP 地址^[4]。

使用以下 SQL 查询,可从大到小列出建立连接请求排前十位计算机 IP 地址及其连接数。

```
select srcaddr, count(*) from test.test group by srcaddr order by 2 desc limit 10;
```

使用以下 SQL 查询,可以从大到小列出单位时间内发送数据分组数排前十位计算机 IP 及其发送分组数。

```
select srcaddr, sum(dpkts) from test.test group by srcaddr order by 2 desc limit 10;
```

使用以下 SQL 查询,可以从大到小列出针对某一端口的连接排前十位计算机 IP 及其与该端口连接数,以下以很多蠕虫喜欢连接的端口 445 为例。

```
select srcaddr, count(*) from test.test where dst-port = 445 group by srcaddr order by 2 desc limit 10;
```

2.3.2 模式匹配

每种蠕虫病毒,在扫描或攻击时,都有固定模式,根据其典型特征,分析 NetFlow 数据流,可找出其 IP。如:中冲击波病毒计算机建立的 NetFlow 数据流,具有以下典型特征:目的端口:135;协议类型:6;字节数:48。

使用下面 SQL 查询,可将满足上面典型特征主机 IP 及其连接数统计出来。

```
select srcaddr, count(*) from test.test where dst-port = 135 and prot = 6 and doctets = 48 group by srcaddr order by 2 desc;
```

2.3.3 TCP 标志位分析

TCP 连接在建立前,要经过三次握手。第一次握手:主机 A 发送一个 SYN 标志位 TCP 报文给主机 B;

第二次握手:主机 B 向主机 A 反馈回 SYN/ACK 报文,确认收到;第三次握手:主机 A 确认(ACK)主机 B 的 SYN/ACK 报文。如果主机 A 同主机 B 非开放端口连接,主机 B 会回送一个 RST/ACK 报文。对于正常建立的 TCP 连接来说,其 TCP 标志位是不单一的,也就是说在整个连接过程中,ACK/SYN/FIN 都会出现,不会出现大量单一 TCP 标志位连接。

蠕虫病毒如果是基于 TCP 进行目标主机漏洞探测、攻击,其使用 TCP 探测时,必然会产生大量 SYN 标志位 TCP 连接。伴随大量 SYN 标志位连接, DOS 攻击和恶意扫描也会产生,这些和蠕虫都是对一个健康网络的危害,因此,都属于被严格控制的行为,在后面实现中,将这些行为视同蠕虫一样处理。

针对存入 MySQL 中的 NetFlow 数据,通过以下 SQL 查询,可以从大到小列出 TCPSYN 连接排前十名 IP 及其 TCPSYN 连接数。

```
select srcaddr, count(*) from test.test where tcp-flags = 2 group by srcaddr order by 2 desc limit 10;
```

2.4 基于 SNMP 对染毒主机定位、隔离

要实现染毒主机定位与隔离,首先必须在园区网核心交换机或路由器上配只读团体名,各接入层交换机要支持 SNMP 和端口 MAC 地址自学习功能(现在除了傻瓜交换机外,智能交换机都支持这两种功能),各接入交换机要配读写权限团体名。

2.4.1 获取染毒机 MAC 地址

RFC1213 定义的 MIB-2 IP 组下有一个 ipNetToMediaTable 表,该表下的 ipNetToMediaPhysAddress 项提供了 IP 地址和 MAC 地址的对应,该项 OID 值为 1.3.6.1.2.1.4.22.1.2,管理站针对路由器 SNMP 代理服务该 OID 项,连续发送 getNextrequest 请求报文,就可得到 ARP 缓存中所有 IP-MAC 对应值。

得到全部 IP-MAC 对应值后,再从中过滤出染毒机 IP 地址的 IP-MAC 值,以获取染毒机 MAC 地址。

2.4.2 定位染毒机连接的接入层交换机端口

因交换机具有端口 MAC 地址自学习功能,只要通过 SNMP 轮询所有接入层交换机,就可以准确定位染毒机 MAC 地址所在交换机端口。

2.4.3 关闭染毒机与接入交换机连接端口

确定染毒机所连接接入交换机端口后,通过 SNMP 关闭该端口, RFC1213 定义的 MIB-2 接口组 ifTable 中 ifAdminStatus 表示端口的管理状态,其 OID 为 1.3.6.1.2.1.2.2.1.7,当其值为 1 时,表示端口打开,值为 2 时表示端口关闭,值为 3 时表示测试,该值可以使用读写团体名修改。管理站向被管设备 SNMP 代理服

务某一需要关闭的 OID, 发送值为 2 的 setrequest 请求报文, 就可将该 OID 对应端口关闭^[5]。

2.5 监控流程设计

结合上面论述, 给出具体蠕虫病毒监控流程(如图 2 所示)。

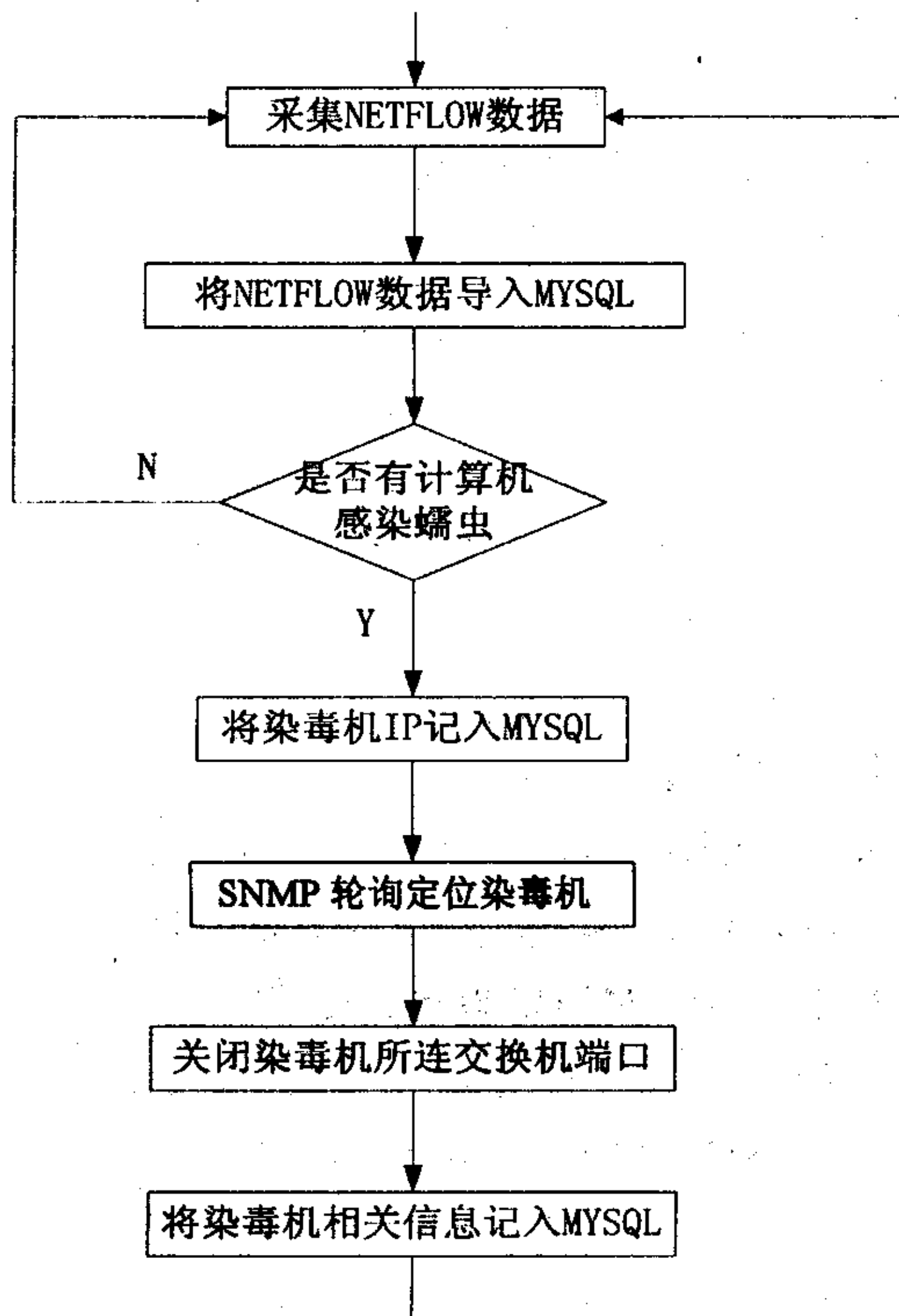


图 2 蠕虫病毒监控流程图

3 监控系统具体实现

参照图 2 蠕虫病毒监控流程, 通过以下 4 个模块实现监控系统:

(1) 监测引擎: 本模块为监控系统主模块, 其担负的主要功能是启动监控系统, 协调各模块之间工作, 并定期将 Flow - tools 采集的 NetFlow 信息导入 MySQL。

(2) 染毒机分析与过滤: 本模块主要功能是综合利用监控模式设计所述 4 种检测方法, 通过 MySQL 强大的数据分析与统计功能, 过滤出感染已知蠕虫病毒或疑似感染病毒主机的 IP。首先通过 TOPN 模式、TCP 标志位分析的方法, 分别统计出排前十位的 IP, 然后对各排前十位的 IP 通过模式匹配方法, 与模式库中定义的病毒模式进行匹配, 检查其是否感染已知蠕虫病毒, 如果确定该 IP 已感染蠕虫病毒, 将该 IP 记入 MySQL 感染蠕虫病毒 IP 表(以下简称染毒表)。对于模式匹配不能判定是否感染病毒的 IP, 当这些可疑 IP 所连目的端口是一些蠕虫病毒喜欢攻击的端口, 如 445、139、135 等, 且其连接数大于 500 时(此数字可以根据需要调整, 笔者在实验中发现单一 IP 正常情况下

与某一目的端口连接, 决不会瞬间达到成百上千), 就把其当成感染蠕虫看待, 将其 IP 记入染毒表, 同时, 把该 IP 产生的相关 NetFlow 信息存入 MySQL 未知病毒数据表(以下简称可疑表)。当可疑 IP 大量连接的某一目的端口, 虽非蠕虫喜欢攻击的高危端口, 但连接数过大, 超过某一阈值时, 虽不将其 IP 记入染毒表, 但要将该 IP 产生的相关 NetFlow 信息存入可疑表。

(3) 染毒机定位与隔离: 本模块主要功能是在监测引擎协调下, 根据染毒表中 IP, 对感染机定位并将其与网络隔离。监测引擎在染毒机分析与过滤模块运行结束后, 引导本模块读取染毒表中各 IP, 通过 SNMP 查询路由器 ARP 缓存, 获取染毒机 IP 对应 MAC 地址, 轮询各接入层交换机, 定位染毒机 MAC 地址所在交换机端口, 通过 SNMP 关闭该端口。

(4) 日志管理与显示: 本模块主要功能是将染毒机 IP、MAC、所连交换机端口、病毒扫描特征等信息记入 MySQL 病毒记录日志表中, 连同可疑表中的相关信息, 通过图形化界面显示给管理员。管理员通过该模块可知每日有哪些计算机感染蠕虫, 并根据可疑表中的相关信息, 通过实地检查分析等手段, 判断是否有新蠕虫出现, 如果判定是新蠕虫, 将其特征加入蠕虫病毒模式库。

4 结束语

随着园区网不断发展, 蠕虫病毒威胁成为亟待解决的问题。对蠕虫病毒攻击的检测与控制, 保障计算机网络安全成为刻不容缓的重要课题。文中针对蠕虫病毒传染特征, 介绍了利用 NetFlow 协议查找感染蠕虫计算机的一些实用方法, 及确定染毒机 IP 后, 利用 SNMP 定位并隔离染毒计算机方法, 综合利用这些方法建立起的监控系统, 不但可以有效地扼制蠕虫病毒泛滥, 对于 DOS 攻击、恶意扫描等非正常网络行为也有很好的扼制效果。

参考文献:

- [1] Udupa D K. Network Management Systems Essentials[M]. San Francisco: McGraw - Hill, 1996.
- [2] Cisco. NetFlow Services and Applications. Cisco White Paper [EB/OL]. 1999. <http://www.cisco.com>.
- [3] 何海涛, 罗笑南, 郭清顺. Netflow 在边界网流量测量中的应用研究[J]. 计算机工程与应用, 2004(11): 11 - 14.
- [4] 杨 嵘, 张国清, 韦 卫, 等. 基于 NetFlow 流量分析的网络攻击行为发现[J]. 计算机工程, 2005(13): 137 - 139.
- [5] Stallings W. SNMP 网络管理[M]. 北京: 中国电力出版社, 2001.