doi: 10.3969/j.issn.1001 - 893x.2015.02.011
引用格式: 管涛,王科人,徐正国. 基于 Hough 变换的 NAT 规模被动估计新方法[J]. 电讯技术,2015,55(2):175 - 181. [GUAN Tao, WANG Keren, XU Zhengguo. A New Passive NATted Hosts Counting Method Based on Hough Transform[J]. Telecommunication Engineering, 2015, 55(2): 175 - 181.]

# 基于 Hough 变换的 NAT 规模被动估计新方法*

管 涛**,王科人,徐正国
(盲信号处理重点实验室,成都 610041)

**摘 要:** 针对网络安全领域中网络地址转换器(NAT)规模被动估计这一问题,提出了一种新的 NAT 规模被动估计方法。在被动接收条件下,通过利用 TCP 数据包时间戳值与接收时间之间表现出的线性关系,将 NAT 规模估计转换为坐标图上的直线数目检测问题。然后基于 Hough 变换递归放大坐标图像来检测直线数目,从而估计出 NAT 规模大小。与已有算法相比,该方法不仅可以解决初始类别的选择问题,还能提高时间戳序列发生交叉或距离较近时的判断精度。实验测试结果表明,该方法可以准确地检测出线性关系的数目,进而估计出 NAT 规模大小,并且性能优于已有算法。

**关键词:** 网络安全; NAT 规模估计; TCP 时间戳; 时间序列分析; 线性关系估计; Hough 变换
**中图分类号:** TP393.08 **文献标志码:** A **文章编号:** 1001 - 893X(2015)02 - 0175 - 07

# A New Passive NATted Hosts Counting Method Based on Hough Transform

GUAN Tao, WANG Keren, XU Zhengguo
(Key Laboratory of Science and Technology on Blind Signal Processing, Chengdu 610041, China)

**Abstract:** To solve the problem of NATted (Network Address Translator, NAT) hosts counting in network security, a new method based on Hough transform is proposed. By exploiting the linear relation between TCP timestamp and received time in the packets, the problem is transformed to counting lines in the coordinate graph. Hough transform is introduced to detect lines in the graph. To get more precise result, zooming in the graph iteratively is used in the algorithm. Compared with existing algorithms, the proposed method can solve the problem of choosing initial clusters, and improve the detection accuracy with intersected or short distance timestamp sequence. Experiment results show that the algorithm based on Hough transform can detect the number of NATted hosts precisely which outperforms existing methods.

**Key words:** network security; NATted hosts counting; TCP timestamp; time series analysis; linearity extraction; Hough transforms

## 1 引 言

网络地址转换器(Network Address Translator, NAT)允许多个内网主机使用相同公网地址连接互联网,它可以缓解 IPv4 地址数量紧张的问题[1]。目前,NAT 已经在互联网中大量部署,范围涵盖了小型的家庭网络以及大型的企业内网。NAT 提高了网络使用的隐私性,隐藏了内部网络的大小和拓扑结构。但是,从网络管理和网络安全的角度来说,NAT 的大量使用严重影响了网络的正常管理,并且造成了潜在的安全隐患。在被动接收条件下实现 NAT 检测对网络管理和网络安全威胁检测具有重要作用,例如对僵尸网络的规模检测和未授权的设

* 收稿日期: 2014 - 12 - 16; 修回日期: 2015 - 01 - 26 Received date: 2014 - 12 - 16; Revised date: 2015 - 01 - 26
** 通讯作者: guantau@163.com Corresponding author: guantau@163.com

备接入等。因此,被动接收条件下的 NAT 流量检测及其规模估计成为工业界和学术界关注的问题。

NAT 流量检测需要判断某个 IP 是否使用了 NAT,而 NAT 规模估计还要对 NAT 后面主机数量进行估计。目前已有大量工作对 NAT 流量检测及规模估计进行了研究。

在 NAT 流量检测方面,文献[2-3]综合了初始 TTL(Time To Live)、IPID(IP Identification)、TCP SYN、TCP 源端口、TCP 时间戳等特征进行 NAT 流量检测。文献[4-10]则采用的是流量统计的方法,通过训练学习流量特征对 NAT 流量进行检测。实验结果表明,这些方法均达到了较高的检测率。

在 NAT 规模估计方面,文献[11]最早提出利用主机发送 IPID 序列的连续性来估计 NAT 主机数目。当 IPID 随机产生或者为固定值(如 0)时,文献[11]方法会失效,文献[12]在此基础上进一步关联 IPID、TCP 序列号和源端口序列进行综合判断。文献[13]给出了一种基于 TCP/IP 协议栈指纹的 NAT 检测方法,其适用于 NAT 后主机具有不同操作系统的情形。文献[14-15]利用了 HTTP 协议中的 User-Agent 和 Cookie 信息出现的种类数进行估计。文献[16]提取了 ICMP 和 TCP 中的时间戳值,采用时钟漂移特征区分不同主机。文献[17-18]指出了通过主机启动时间和 TCP 时钟频率可以唯一标识一台主机,因此采用 TCP 时间戳序列可以估计 NAT 规模大小。

本文主要关注的是 NAT 规模估计问题,即在给定一段时间内的网络数据的条件下,研究如何准确估计 NAT 后面主机数量。总的来说,现有 NAT 规模估计的两类方法均存在一定问题:统计 IPv4、TCP 或 HTTP 头部中字段种类,如 TTL 值、User-Agent 等,其使用条件比较有限,且分辨率较差;通过提取序列特征可以更准确地估计 NAT 规模大小,如对 IPID 序列、TCP 序号序列和 TCP 时间戳序列的连续线段进行检测,但目前采用的方法是密度聚类或线性递归,当线段发生交叉或距离较近时误判率较高,且无法解决初始类别的选择问题。

因此,为了更好地估计出 NAT 规模,我们将采用 Hough 变换对基于 TCP 时间戳序列的估计方法进行改进。我们的解决思路为:对于 TCP 时间戳和数据包接收时间,由于不同主机的开机时间和 TCP 时钟频率存在一定的差异,它们体现出不同的线性关系;该线性关系在坐标图中表现为直线,基于计算机视觉中的 Hough 变换方法递归放大该坐标图像,从而检测出现的直线数目即可获得 NAT 规模大小。区别于以往估计方法的地方在于,我们改变了序列连续性检测方式,利用计算机视觉的方法解决了初始类别的选择问题,并改进了序列发生交叉或距离较近时的判断精度。

本文组织结构如下:第 2 节对 NAT 条件下 TCP 时间戳与数据包接收时间之间的关系进行建模分析;第 3 节介绍基于 Hough 变换的 NAT 主机数目估计算法;第 4 节给出了算法在仿真和实际数据条件下的实验结果;第 5 节对全文进行总结。

## 2 模型建立

TCP 时间戳在 RFC1323 中引入,它作为 TCP 头部的选项字段,长度为 10 Byte[19]。其主要有两个作用,一是精确测量往返传输时延,二是防止高速传输网络下 TCP 序号的冲突。TCP 时间戳选项中包含长度为 32 bit 的 TSval 子字段,它代表当前数据包发送主机设置的 TCP 虚拟时钟值。下文无特殊说明时,TCP 时间戳均指的是 TSval 子字段。

在被动接收条件下,假设我们获得从某个 IP 发出的一系列包含 TCP 时间戳的 IP 数据包,其中第 $i$ 个数据包的发送时间为 $s_i$,TCP 时间戳值为 $t_i$,它们构成一个 TCP 时间戳序列。根据 RFC1323 规定,发送方数据包中的 TCP 时间戳 $t_i$ 与发送时间 $s_i$ 之间应保持线性关系 $t_i = ks_i + t_0$。其中:$k$ 为 TCP 虚拟时钟频率,它与操作系统内核直接相关,RFC 建议取值范围为 1 ~ 1000 Hz,常见的时钟频率有{2 Hz,10 Hz,100 Hz,250 Hz,500 Hz,1000 Hz};$t_0$ 为初始计数值,它与开机时间直接相关,一般从开机起始为 0 或某个随机值。

对于被动捕获方,数据包发送时间 $s_i$ 是未知的,我们只能获取捕获时间。假设第 $i$ 个数据包的捕获时间为 $r_i$,数据包经历的路径时延为 $\tau_i$,那么有 $r_i = s_i + \tau_i$,捕获时间 $r_i$ 与时间戳 $t_i$ 满足 $t_i = kr_i - k\tau_i + t_0$。定义相对捕获时间为 $x_i$,TCP 时间戳值为 $y_i$ 如式(1)所示:

$$\begin{aligned} x_i &= r_i - r_1, \\ y_i &= t_i。 \end{aligned} \tag{1}$$

那么有 $y_i = k \cdot x_i + b$,其中 $b = t_0 + kr_1 - k\tau_i$。假设在时间跨度 $T$ 中,收到的含时间戳的数据包个数为 $N$,

即数据集合为$\{(x_i,y_i)\mid i=1,2,\cdots,N\}$。路径时延$\tau_i$为随机变量，当接收时间跨度$T$较小时，$\tau_i$可近似为常数，此时，$y_i$与$x_i$能够较好地满足线性关系。

在 TCP 时间戳$y$与相对捕获时间$x$存在的线性关系$y=kx+b$中，斜率$k$是 TCP 虚拟时钟频率，截距$b$由主机的开机时间和数据包传输时延决定。因此，采用 TCP 时间戳序列进行 NAT 规模估计基于三个假设条件：NAT 不修改 TCP 时间戳值；具有相同 TCP 时钟频率的主机不在同一时刻开机；数据包传输时延不发生剧烈变化。也就是说，用 TCP 时钟频率和时间戳起始值可以标识一台主机，这在实际条件下通常是满足的[17-18]。理论上讲，由于时钟频率最大值为1000 Hz，最小可分辨的开机时间差异为1 ms。

对某包含 6 台主机的 NAT 设备采集 TCP 时间戳序列，其中主机 1 和主机 2 为 Windows XP 操作系统，主机 3 ~ 6 为 Ubuntu 12.04 操作系统，特别地，主机 4 ~ 6 为同一型号设备。画出相对捕获时间和 TCP 时间戳的散点图如图 1 所示，线性关系在坐标图中表现为直线，可以看出不同主机表现出不同的线性关系。

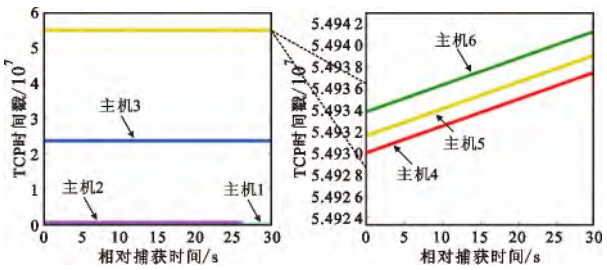



图 1　NAT 中不同主机表现出不同的线性关系

Fig. 1 Different host behind NAT shows different linearity

假设 NAT 包含的主机数量为$M$，对数据包相对捕获时间$x$和 TCP 时间戳$y$建立如下的线性混合模型：

$$y=f(x)=\begin{cases}k_1x+b_1 & \text{主机 1}\\ k_2x+b_2 & \text{主机 2}\\ \vdots & \\ k_Mx+b_M & \text{主机 }M\end{cases} \tag{2}$$

式中，主机数量$M$未知，模型参数$\{(k_1,b_1),(k_2,b_2),\cdots,(k_M,b_M)\}$未知。这是一个典型的无监督聚类问题，本文将其转化为检测坐标图中的直线数目，通过基于 Hough 变换的方法估计直线数量。

## 3　算法描述

为了估计出 NAT 规模大小，只需要检测 TCP 时间戳与相对捕获时间坐标图上直线的数目。从计算机视觉角度出发，图像上直线检测可以通过 Hough 变换完成。由于图像分辨率的原因，利用 Hough 变换进行检测时，需要进行多级迭代放大，从而更加准确地检测出直线数目。

### 3.1　数据预处理

由于线性关系的稳定性与路径时延抖动相关，为尽量减小时延抖动对检测的影响，数据的时间跨度需要限制在比较小的长度内。在数据预处理时，我们将数据按照一定的时间长度$T$分段进行处理。通常时间长度$T$根据目标的平均流量来确定，从而保证用于识别的数据量足够且具有一定的反应速度。此外，相同 TCP 时间戳为重复信息，我们只保留第一次出现的数据点。

进行 Hough 变换前，还需要将分段后的数据二值化映射为图像。假设映射后图像的大小为$W\times H$，即宽度为$W$，高度为$H$，那么图像分辨率为

$$\begin{aligned}x_{\text{res}}&=(\max(x)-\min(x))/W,\\ y_{\text{res}}&=(\max(y)-\min(y))/H。\end{aligned} \tag{3}$$

归一化映射后得到原数据点在图像上的坐标点为

$$\begin{aligned}x'&=\text{round}((x-\min(x))/x_{\text{res}}),\\ y'&=\text{round}((y-\min(y))/y_{\text{res}})。\end{aligned} \tag{4}$$

将图像上对应坐标置为 1 即可生成二值图像。生成图像后，再对图像做一次边缘检测处理，完成数据预处理。

### 3.2　Hough 变换

标准 Hough 变换采用如下参数形式表示一条直线[20]：

$$\rho=x\cdot\cos(\theta)+y\cdot\sin(\theta)。 \tag{5}$$

式中，变量$\rho$表示从原点到直线的垂直距离，变量$\theta$表示原点到直线的垂向量与$x$轴的夹角。

参数空间$(\rho,\theta)$需要离散化，Hough 变换后可以获得离散化参数空间上的分布矩阵，矩阵的每个元素代表落在相应参数位置的图像点数，其峰值点则代表图像上可能存在对应参数的直线。假设输入数据为$(x,y)$，Hough 变换后得到

$$\boldsymbol{H}=\text{Hough}(x,y)。 \tag{6}$$

对 Hough 变换的矩阵$\boldsymbol{H}$检测参数空间中出现

的峰值点，并得出当前图像中的直线数目。

参数空间($\rho$,$\theta$)的离散精度决定着检测结果的精度，距离 $\rho$ 的离散化精度记为 *Rho*，夹角 $\theta$ 离散化区间记为 *Theta*。此外，还需要设置 $\boldsymbol{H}$ 矩阵的峰值判决门限 *V*。

### 3.3 递归放大

受图像分辨率的限制，需要对检测得到的直线进行递归放大进而获得更精确的结果。首先按照检测结果对图像进行分割，选择已检测到的直线邻域的数据作为新的输入数据重新检测。对于不归属于任何已检测直线的数据，同样作为新的输入数据重新检测。

在选择已检测到直线的邻域数据时，当数据点与直线的距离小于邻域半径 $\varepsilon$ 时，认为数据点归属该直线的领域。已知 TCP 虚拟时钟频率大于 0，当图像上检测到的直线斜率显著大于 0 时，我们认为不需要再放大。为了减小奇异点的影响，对于获取的数据量小于门限值 *Th* 的不做检测处理。

### 3.4 具体步骤

基于 Hough 变换的 NAT 规模被动估计算法的具体步骤如图 2 所示。

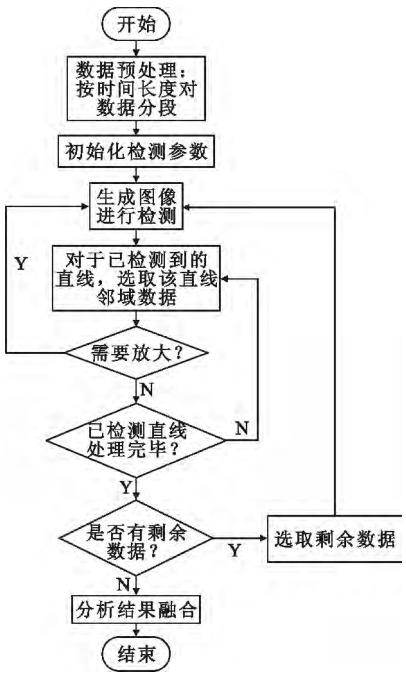



图 2 算法具体步骤

Fig. 2 Detailed procedure of the proposed algorithm

算法中涉及的关键检测参数及实验中采用的典型值如表 1 所示。

表 1 算法关键参数列表

Table 1 Key parameters of the proposed algorithm

| 算法步骤 | 参数名 | 参数值 |
|---|---|---|
| 数据预处理 | 时间长度 *T*/s | 10 |
| | 图像宽度 *W* | 256 |
| | 图像高度 *H* | 1024 |
| Hough 变换 | 离散化精度 *Rho* | 0.5 |
| | 离散化区间 *Theta* | [-90: 0.5: 0] |
| | 峰值判断门限 *V* | 0.95 · max($\boldsymbol{H}$) |
| 分级放大 | 邻域半径 $\varepsilon$ | 10 |
| | 数据量门限 *Th* | 10 |

## 4 实验结果

本节首先在真实数据环境下，对本文算法和已有算法进行验证。为了进一步测试算法性能，我们搭建实验网络，对比本文算法和已有算法的性能。

下面实验将对两种常用的针对 TCP 时间戳序列实现 NAT 规模估计的算法进行对比。

(1) Bursztein 算法[17]

通过 TCP 时间戳在特定 TCP 虚拟时钟频率下增长的误差值判断两个数据包是否属于同一主机，设定 TCP 虚拟时钟频率集合为{2 Hz，10 Hz，100 Hz，250 Hz，500 Hz，1000 Hz}，时间间隔为 10 ms，误差范围为 0.1%。

(2) Wicherski 算法[18]

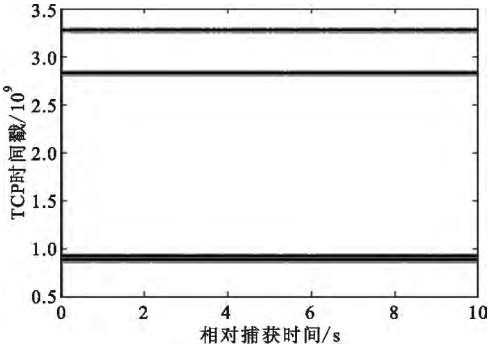
以同一 TCP 连接(即 TCP 五元组相同)的数据包作为初始类，通过最小均方误差线性回归方法求取其对应的 TCP 虚拟时钟频率及估计的开机时间，并将相同时钟频率下开机时间小于 $\delta_{\text{boot}}$ 的归为同一主机，设置 $\delta_{\text{boot}}=2$ ms。

注意到，Bursztein 算法中将 TCP 虚拟时钟频率作为先验信息，Wicherski 算法则假设同一连接的所有数据包属于同一主机，而本文算法并没有添加这些限制条件。

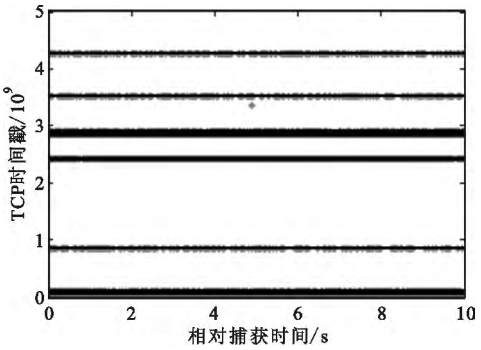
### 4.1 真实数据验证

采集真实环境下某网络出口的数据，取其中两个 IP 地址的 TCP 时间戳序列进行检测。第 1 个 IP 地址的流量和流数分别为4.1 Mbit/s和 7192，图 3(a)所示的为第 1 个 IP 地址数据的检测结果，三种算法检测得到主机数目均为 4，与人工分析结果一致。第 2 个 IP 地址的流量和流数分别为 35.6 Mbit/s和43 734，图3(b)所示的为第2个IP地址数据的检测结果，本文算法检测得到主机数目为 21，与人工分析结果一致，而 Bursztein 算法和 Wich-

erski 算法检测结果分别为 14 和 18。




(a) IP 地址 1




(b) IP 地址 2

**图 3　真实数据检测结果**

Fig. 3 Experiment results on real data

### 4.2　性能对比测试

采用实验网络产生数据进行对比测试，实验采用的网络连接如图 4 所示。在 NAT 设备后面连接若干台主机，这些主机持续地随机访问服务器，数据采集点位于 NAT 设备与服务器之间。

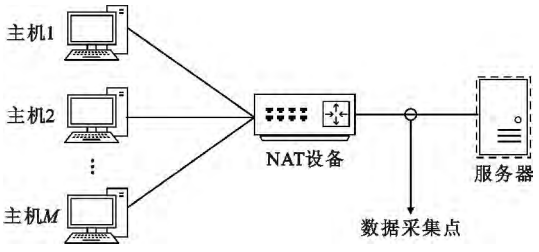


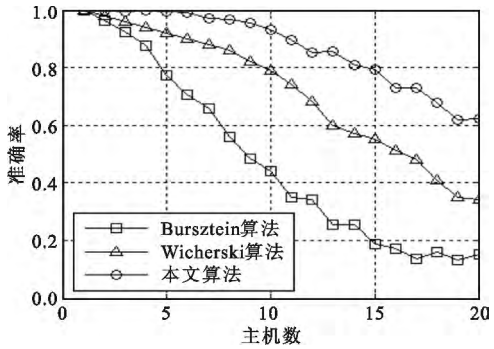

**图 4　实验网络连接图**

Fig. 4 Experimental network setup

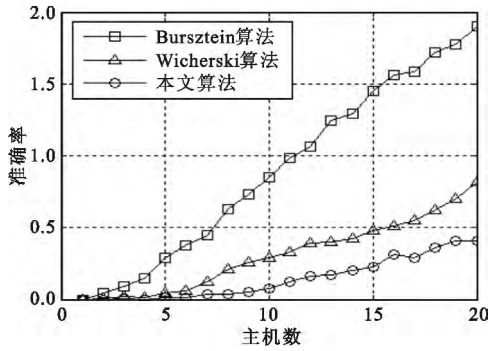
我们采用检测准确率 $\alpha$ 和偏差值 $\delta$ 评价算法性能。对于特定主机数目 $M$，当检测得到 $M$ 个主机时，检测正确，否则检测错误。假设检测次数为 $N$，检测正确的次数为 $P$，第 $n$ 次检测得到的主机数为 $X_n$，准确率 $\alpha$ 和偏差值 $\delta$ 定义如式(7)所示:

$$\alpha = P/N,\qquad \delta = \frac{1}{N}\sum_{n=1}^{N}|X_n - M|\ 。 \tag{7}$$

选取主机数目 $M$ 范围为 $[1,20]$，检测次数 $N=500$，统计算法的检测准确率和偏差值如图 5 所示。从图中可以看出，本文算法检测的准确率要高于 Bursztein 算法和 Wicherski 算法，检测的偏差值要小于 Bursztein 算法和 Wicherski 算法，这说明本文算法性能要优于这两种传统算法。




(a) 准确率统计结果




(b) 偏差值统计结果

**图 5　实验结果对比图**

Fig. 5 Comparison results on experimental network

选定主机数目为 $M=10$，查看检测结果的分布如图 6 所示。从检测的分布图可以看出本文算法的分布较为集中，而其他两种算法得到分布较为分散，这进一步表明本文算法性能更好。

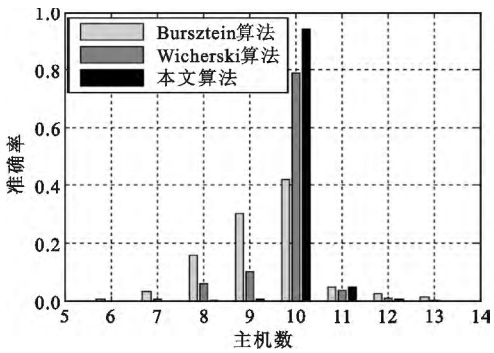
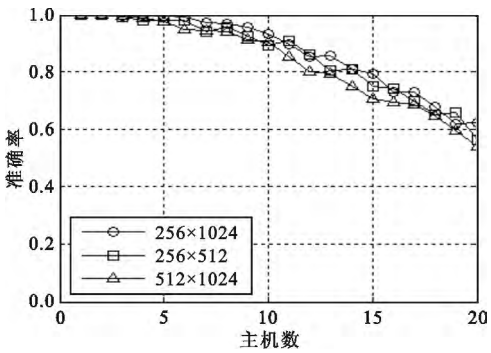



**图 6　主机数目为 10 时检测结果分布对比图**

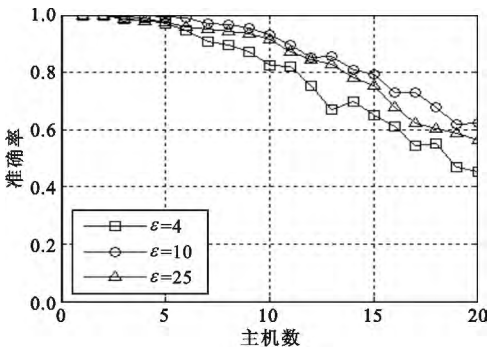
Fig. 6 Distributions comparison with 10 hosts

由于 Bursztein 算法是根据数据点之间的时间戳差和接收时间差来判断是否其是否属于同一主机，当多个主机的时间戳相差较小时，Bursztein 算法会将这些数据都归于同一主机，从而发生误判。Wicherski 算法是基于同一连接属于同一主机这一假设，在短连接条件下数据点较少使得拟合误差较大，进而引起误判。而本文算法是基于 Hough 变换进行直线检测，即使存在主机时间戳相差较小或存在大量短连接的情况，从图像上依然可以较为准确地区别出不同的直线。

检测图像大小 $W \times H$ 和邻域半径 $\varepsilon$ 是本文算法非常关键的参数，选择不同的参数组合对其灵敏度进行分析。图 7( a) 为单独改变图像大小的实验结果，图 7( b) 为单独改变邻域半径的实验结果。从理论上讲，基于 Hough 变换对直线进行检测要求映射后的图像分辨率适中，当图像上的点过于稀疏时，检测会发生一定偏差。实验结果显示图像大小对算法准确率影响不大。检测邻域半径 $\varepsilon$ 决定着放大的区域，当 $\varepsilon$ 过小时，放大区域包含数据可能不完整，而过大时，放大区域可能包含额外的数据，这些都会对检测结果产生影响。实验结果表明，适中的邻域半径能够达到最好的检测效果。




( a) 不同图像大小




( b) 不同邻域半径

图 7 不同参数条件下检测结果对比

Fig. 7 Comparison results with various parameter settings

## 5 结束语

针对 NAT 主机数目检测问题，本文利用 TCP 时间戳与数据包接收时间之间存在的线性关系，提出了一种基于 Hough 变换的 NAT 主机数目自动识别算法。与以往工作相比，该算法解决了多个主机时间戳相距较近以及短连接导致误识别的问题。实验测试结果表明了算法的有效性，且性能优于已有算法。

本文算法并不局限于对 TCP 时间戳序列进行分析，它可以很容易地扩展至 IPID 序列和 TCP 初始序号序列的线性关系自动识别中。下一步工作将考虑利用更多的可用先验信息，从而进一步提高算法的性能。

## 参考文献:

[1] RFC 1631, The IP Network Address Translator( NAT) [S].

[2] 焦程波，郑辉，黄宇. 被动式远程网络地址翻译器识别系统 [J]. 电子科技大学学报，2012( 6) : 899 - 904.
JIAO cheng - bo, ZHENG Hui, HUANG Yu. Novel passive remote network address translation detecting system [J]. Journal of University of Electronic Science and Technology of China, 2012( 6) : 899 - 904. ( in Chinese)

[3] Detection of NAT devices [EB/OL]. [2014 - 06 - 07]. http: //www. muni. cz/ics/research/projects/4622/web/natdet.

[4] Li R, Zhu H L, Xin Y, et al. Remote NAT detect algorithm based on support vector machine [C] //Proceedings of International Conference on Information Engineering and Computer Science( ICIECS) . Wuhan: IEEE, 2009: 1 - 4.

[5] 高骥翔. 基于网络流量特征的 NAT 识别方法 [D]. 成都: 电子科技大学，2012.
GAO Jixiang. NAT Detection Based on Network Traffic Feature [D]. Chengdu: University of Electronic Science and Technology of China, 2012. ( in Chinese)

[6] Detecting NAT Devices using sFlow [EB/OL]. [2014 - 06 - 07]. http: //www. sflow. org/detectNAT/.

[7] Abt S, Dietz C, Baier H, et al. Passive remote source NAT detection using behavior statistics derived from NetFlow [M] //Emerging Management Mechanisms for the Future Internet. Berlin: Springer, 2013: 148 - 159.

[8] Krmicek V, Vykopal J, Krejci R. Netflow based system for NAT detection [C] //Proceedings of the 5th International Student Workshop on Emerging Networking Experiments and Technologies. New York: ACM, 2009: 23 - 24.

[9] Li R, Zhu H L, Xin Y, et al. Passive NATted hosts detect algorithm based on directed acyclic graph support vector machine [C] // Proceedings of 2009 International Conference on Multimedia Information Networking and Security. Wuhan: IEEE, 2009: 474 - 477.

[10] Gokcen Y, Foroushani V A, Heywood. Can we identify NAT behavior by analyzing Traffic Flows [C] // Proceedings of 2014 IEEE Security and Privacy Workshops. San Jose: IEEE, 2014: 132 – 139.

[11] Bellovin S. A technique for counting NATted hostes [C] //Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurement. New York: ACM, 2002: 267 – 272.

[12] Mongkolluksamee S, Fukuda K, Pongpaibool P. Counting NATted hosts by observing TCP/IP field behaviors [C] // Proceedings of 2012 IEEE International Conference on Communications (ICC). Ottawa: IEEE, 2012: 1265 – 1270.

[13] Beverly R. A robust classifier for passive TCP/IP fingerprinting [M] //Passive and Active Measurement. Berlin: Springer, 2004: 158 – 167.

[14] Maier G, Schneider F, Feldmann A. NAT Usage in Residential Broadband Networks [M] //Passive and Active Measurement. Berlin: Springer, 2011: 32 – 41.

[15] 白雪,钱步仁,梁华庆. 一种检测 NAT 后主机数目的方案 [J]. 计算机安全, 2009(4): 46 – 48.
BAI Xue, QIAN Buren, LIANG Huaqing. A scheme for counting NATted hosts [J]. Computer Security, 2009(4): 46 – 48. (in Chinese)

[16] Kohno T, Broido A, Claffy K. Remote physical device fingerprinting [J]. IEEE Transactions on Dependable and Secure Computing, 2005, 2(2): 93 – 108.

[17] Bursztein E. Time has something to tell us about network address translation [EB/OL]. (2007 – 07 – 08) [2014 – 06 – 07]. http://cdn.1y.tl/publications/time – has – something – to – tell – us – about – network – address – translation.pdf.

[18] Wicherski G, Weingarten F, Meyer U. IP agnostic real – time traffic filtering and host identification using TCP timestamps [C] //Proceedings of 2013 IEEE 38th Conference on Local Computer Networks. Sydney: IEEE, 2013: 647 – 654.

[19] RFC 1323, TCP extensions for high performance [S].

[20] Duda R O, Hart P E. Use of the hough transformation to detect lines and curves in pictures [J]. Communications of the ACM, 1972(15): 11 – 15.

## 作者简介:

**管　涛**(1986—),男,江西抚州人,2010 年获硕士学位,现为博士研究生,主要研究方向为网络数据挖掘;

GUAN Tao was born in Fuzhou, Jiangxi Province, in 1986. He received the M. S. degree in 2010. He is currently working toward the Ph. D. degree. His research concerns network data minig.

Email: guantau@163.com

**王科人**(1986—),男,四川成都人,博士,工程师,主要研究方向为网络挖掘、视频隐写与隐写分析、图像视频分类;

WANG Keren was born in Chengdu, Sichuan Province, in 1986. He is now an engineer with the Ph. D. degree. His research interests include network mining, video steganography and steganalysis, image and video classification.

**徐正国**(1985—),男,湖北荆州人,2011 年获硕士学位,现为博士研究生,主要研究方向为网络数据挖掘。

XU Zhengguo was born in Jingzhou, Hubei Province, in 1985. He received the M. S. degree in 2011. He is currently working toward the Ph. D. degree. His research concerns network data minig.