SSI：一种 IPv6/IPv4 多址同源识别模型

王 轩 王振兴 王 禹 张连成
（数学工程与先进计算国家重点实验室 郑州 450001）

摘 要 IPv6/IPv4 共存环境下多址同源识别是共存网络管理与拓扑发现的一个关键问题。现有研究主要集中于子网内部的双栈发现及单一 IP 协议栈中的别名解析，难以识别远程 IPv6/IPv4 共存网络中的多址同源。通过分析同源地址间的本质联系，提出一种 IPv6/IPv4 多址同源识别模型（SSI），该模型综合利用特殊地址格式匹配、TCP 时钟指纹比对和上层协议短时致瘫等多种方式来提高同源地址的识别能力。实验结果表明，上述方法均可有效识别 IPv6/IPv4 多址同源；SSI 模型具有较理想的识别率和正确率。

关键词 IPv6，共存环境，多址同源，格式匹配，时钟指纹，短时致瘫

中图法分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.08.031

SSI：A Same Source Identification Model for Multiple IPv6/IPv4 Addresses

WANG Xuan WANG Zhen-xing WANG Yu ZHANG Lian-cheng
（State Key Laboratory of Mathematic Engineering and Advanced Computing，Zhengzhou 450001，China）

Abstract Same source identification for multiple addresses is a key problem of network management and topology discovery under the IPv6/IPv4 coexistence environment. Existing research has focused on the discovery of dual-stack in the subnet and alias resolution of a single IP protocol stack，so it is difficult to identify same-source addresses in the remote network with IPv6/IPv4 coexistence. First，the essential connections between same-source addresses were analyzed，then a same source identification (SSI) model for IPv6/IPv4 addresses was proposed，which improves the identification ability by combing methods including special address pattern matching，TCP clock fingerprint matching and upper-protocol short-time paralyzation. The experimental results indicate that the methods above can effectively identify multiple IPv6/IPv4 addresses with same source. SSI model has an ideal recognition rate and accuracy.

Keywords IPv6，Coexistence environment，Multiple addresses with same source，Pattern matching，Clock fingerprint，Short-time paralyzation

1 引言

互联网已进入 IPv4 与 IPv6 长期共存时代，为了实现 IPv4 网络向 IPv6 网络的平滑演进，出现了越来越多的适用于不同场景的过渡机制，其基本思想可归为双栈、翻译及隧道 3 类。而双栈作为另两种过渡机制的实现基础，更是被广泛应用。互联网中部署了大量同时配置并运行 IPv4 与 IPv6 协议的双栈节点，且 IPv4、IPv6 协议都允许在同一节点配置多个 IP 地址。本文将上述多个 IPv6/IPv4 地址属于同一物理节点的现象称为 IPv6/IPv4 多址同源，此类节点称为 IPv6/IPv4 多址同源节点，同源节点上的 IPv6/IPv4 地址称为同源地址。

随着 IPv4 向 IPv6 网络的过渡，IPv6、IPv4 两类网络在多址同源节点处交叉重合，多址同源现象愈发普遍[1]。运用现有的网络拓扑发现技术，只能获得 IPv6/IPv4 共存网络逻辑上分离的 IPv6 与 IPv4 网络拓扑，难以呈现其真实结构，致使网络监管难度进一步增大。采用 IPv6/IPv4 多址同源识别技术能够发现 IPv6/IPv4 多址同源节点，从而将逻辑上分离的 IPv6 与 IPv4 网络拓扑进行有效关联，呈现其交叉重叠区域，从而更加全面地描述 IPv6/IPv4 共存环境下的网络结构。同时，IPv6/IPv4 多址同源识别也是共存环境下网络关键节点发现、过渡机制探测等技术的重要支撑。因此，IPv6/IPv4 多址同源识别是网络管理领域的一个关键问题。

目前，在国内外公开研究中尚未发现不依赖特殊权限且能够远程识别 IPv6/IPv4 多址同源的有效方法。已有相关研究主要集中在远程识别纯 IPv6、IPv4 网络中的多址同源（即别名归并），以及在本地链路中的 IPv6/IPv4 多址同源识别。前者中比较有代表性的是基于源地址的识别方法：向地址 x 发送 UDP 高端口探测报文，若回复的不可达信息的源地址是 y，则判断 x、y 属于同一节点。Iffinder[2]、Mercator[3] 等工具中使用的就是这种方法。显然，该方法不适用于 x 与 y 分别是 IPv4、IPv6 地址的情况。Ally[4] 提出了基于 IP 标识的识别

到稿日期：2013-10-14 返修日期：2013-12-15 本文受国家 863 计划项目（2012AA012902）资助。

王 轩（1988－），男，硕士生，主要研究方向为 IPv6 网络安全，E-mail：xuan_w @126.com；王振兴（1959－），男，教授，博士生导师，主要研究方向为 IPv6、网络安全；王 禹（1984－），男，博士生，主要研究方向为网络安全；张连成（1982－），男，博士，讲师，主要研究方向为流量分析、网络安全。

方法，这种方法的实现原先是针对指定的源地址、目的地址对及协议，同一会话中的分片报文被按顺序分配 IP 标识码。首先分别发送高端口号的 UDP 探测包给两个待识别地址，且回应的不可达报文中分别包括了各自的 IP 识别码 x 和 y。然后发送第三个探测包到前一阶段首先回复的地址。假设这次的响应包的 IP 识别码是 z，首先回应的 IP 识别码是 x。若 $x<y<z$，且 $(z-x)$ 很小，则判断这对地址是同源。在此基础上，提出 MIDAR[5]，用于实现对超大规模 IPv4 地址的别名归并。而 Robert 等人通过诱使响应报文在待测节点处分片，使得该方法能够用于 IPv6 的多址同源识别[6]。然而由于 IPv4 与 IPv6 不共用 IP 标识，该方法无法判断 IPv4 与 IPv6 地址是否同源。此外，还可以通过 DNS 反向查询，根据多个 IP 地址是否映射为同一域名来判断其是否属于同一网络设备[7]。但该方法并非总是有效，因为互联网中的很多设备没有注册名字，且域名相同也不总是意味着对应同一台网络设备（例如，集群服务器）。对于本地链路中的双协议栈多址同源识别，较有代表性的是通过邻居发现协议获取〈IPv6 地址，MAC 地址〉，通过 ARP 协议获取〈IPv4 地址，MAC 地址〉，把对应相同 MAC 的 IPv6/IPv4 地址判断为同源地址。然而这种方法不适用于远程识别。

针对上述问题，本文提出了 SSI 模型（Same-Source Identification Model）用以识别 IPv6/IPv4 多址同源，该模型包含 3 个子模块：基于特殊地址格式匹配的识别模块 SAPI（Specific Address Patterns-based Identification）、基于 TCP 时钟指纹比对的识别模块 TCFI（TCP Clock Fingerprint-based Identification）和基于上层协议短时致瘫的识别模块 USPI（Upper-protocol Short-time Paralyzation-based Identification），分别从 IPv6/IPv4 地址是否满足地址嵌入格式、是否对应相同的设备指纹以及是否共用传输层协议栈等 3 个层面判断 IPv4/IPv6 地址的同源性。为评价各模块及 SSI 模型的性能，本文提出识别率和正确率两种评价指标。通过实验验证，3 个子模块均可独立识别多址同源，且根据优势互补原则组成 SSI 模型后，能有效提高识别率和正确率。

2 IPv6/IPv4 多址同源分析

基于已有研究，IPv6/IPv4 多址同源识别的难点可以归纳为：

（1）同源的 IP 地址可能包含 3 种情况：纯 IPv4 地址、纯 IPv6 地址，以及 IPv4 与 IPv6 协议地址。

（2）同源节点拥有多个 IP 地址，属于多个网络。

（3）从网络层的不同协议 IP 地址出发，难以判断其在物理层的同一性。

针对上述难点，通过分析本文发现：

（1）IPv6/IPv4 同源地址最本质的联系是对应同一物理设备；

（2）IPv6/IPv4 同源节点上的 IPv6 与 IPv4 协议在网络层分离，而共用网络层以上协议栈；

（3）为实现 IPv4 向 IPv6 的平滑演进，很多 IPv6/IPv4 多址同源节点上使用了过渡隧道和协议翻译，它们中的一些在 IPv6 地址中嵌入 IPv4 地址以实现地址映射。

IPv6/IPv4 同源地址在以上 3 个层面的本质联系为识别 IPv6/IPv4 多址同源提供了依据。

3 SSI 模型

本文根据同源 IPv6/IPv4 地址的本质联系，提出了 SSI 模型（见图 1）。

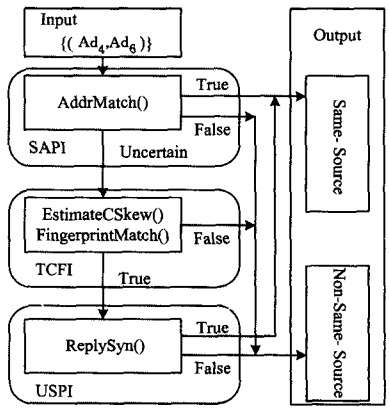

图 1 SSI 模型

该模型分为 3 个子模块：

第一个模块是基于特殊地址格式匹配的识别模块 SAPI，根据一些过渡机制规定在 IPv6 地址的特定位置嵌入 IPv4 地址这一特征来识别 IPv6/IPv4 同源地址。

第二个模块是基于设备指纹比对的识别模块 TCFI，使用 TCP 时钟漂移标识设备指纹，根据 IPv6 与 IPv4 地址对应设备指纹的匹配程度来识别 IPv6/IPv4 多址同源。

第三个模块是基于上层协议短时致瘫的识别模块 USPI，根据对传输层协议短时致瘫过程中待测 IPv6 与 IPv4 地址对 TCP 请求是否同步响应来判断 IPv6 与 IPv4 地址是否同源。

在 SSI 模型中，各子模块只处理其擅长识别的样本，将不擅长的交给其他模块处理。在对（IPv6、IPv4）地址对组成的样本集进行识别的过程中，各子模块需识别的样本数量递减。通过这种方式的优势互补，能够有效提高 IPv6/IPv4 多址同源识别性能。

3.1 基于特殊地址格式匹配的识别模块（SAPI）

为了实现 IPv4 向 IPv6 的平滑演进，一些过渡机制规定了特殊的地址结构，将节点的 IPv4 与 IPv6 地址嵌套。SAPI 模块通过这种特殊格式的地址识别多个 IPv6/IPv4 地址是否同源。例如，兼容地址、IVI、6to4、ISATAP 等过渡机制通过将 IPv4 地址嵌入 IPv6 地址（见表 1）实现跨协议栈通信。

表 1 特殊地址格式及特征

过渡机制	地址格式	特征提取	IPv4 嵌入位置
兼容地址	::a.b.c.d	0－95 位为 0	96－127 位
IVI	Prefix:FF: a.b.c.d::	32－39 位 1； 72－127 位 0	40－71 位
6to4	2002:a.b. c.d::/48	2002 前缀	16－47 位
ISATAP	Prefix:0:5EFE: a.b.c.d	64－95 位 是 0:5EFE	96－127 位

我们根据这些过渡机制的地址格式，提取其特征并确定 IPv4 地址嵌入的位置，从而识别出过渡机制类型，找到内嵌的 IPv4 地址，以识别同源的 IPv6、IPv4 地址。具体过程见算法 1。

算法 1 AddrMatch()

输入：(Ad_4，Ad_6)

输出：true/false/uncertain

Begin

1. Ad_6 与特征库匹配
2. if 匹配成功
3. 提取嵌入的 IPv4 地址 $addr_4$
4. if Ad_4 == $addr_4$
5. return true /* 同源 */
6. else
7. return false/* 非同源 */
8. endif
9. else
10. return uncertain /* 不确定 */
11. endif

End

本模块识别准确度高，资源开销小，但能够识别的地址类型有限，对大量样本无法判断，因此作为 SSI 模型的第一阶段。

3.2 基于 TCP 时钟指纹比对的识别模块(TCFI)

网络设备指纹是一种根据网络设备实现与配置等因素造成的个体差异来标识设备的技术[8,9]。本模块使用 TCP 时钟漂移作为设备指纹，根据 IPv6 与 IPv4 地址对应设备指纹的匹配程度来识别 IPv6/IPv4 多址同源。

定义 1(时钟偏移，clock offset) 时钟 C 在 t 时刻的时钟偏移 $offset[C](t)$，是指时钟 C 在 t 时刻报告的时间 $R[C](t)$ 与真实时间 t 之差。即：$offset[C](t)=R[C](t)-t$，$t\geqslant i[C]$，单位：μs，$i[C]$表示时钟 C 的初始时间。

定义 2(时钟漂移，clock skew)[10] 假设 $R[C](t)$是 t 的简单可微函数，则时钟偏斜 $s[C]$等于时钟偏移 $offset[C](t)$在时间 t 上的一阶导数，即 $s[C]=offset[C](t)/dt$，单位：μs/s (或 ppm)。

根据 Kohno 等人的实验结论，网络设备 TCP 时钟漂移恒定，且不受接入拓扑、跳数及网络延时等因素的影响，可用作网络设备的指纹[8]。RFC1323[11] 规定，如果 TCP 流的两端都实现 TCP 时间戳选项，且在流的初始 SYN 包中包含了这个选项，则这个 TCP 流将使用 TCP 时间戳选项。指纹提取过程见算法 2。

算法 2 EstimateCSkew()

/* 估算时钟漂移作为设备指纹 */

输入：地址 addr

输出：时钟漂移 α

Begin

1. 以 Δt 为周期向 addr 发送带有时间戳的 TCP 探测报文
2. 根据第 i 对往返报文的时间戳计算 t_i 时刻的时钟偏移 $offset(t_i)$
3. 在 t-offset(t)坐标系中绘出点集

$T=\{(t_i,offset(t_i))|i\in(0,N]\}$

4. 使用线性规划估算方法描绘设备特征线 l：$offset(t)=\alpha t+\beta$
5. return α/* α 用作设备指纹 */

End

算法 2 中，步骤 3 执行后，可见坐标系中点集呈带状分布。根据两点连线的斜率估算时钟偏斜是不准确的。在估算 TCP 时钟偏斜过程中，使用基于线性规划的方法与基于最小二乘法的方法得到的偏斜值相同，而在网络延时变化的环境中前者具有更好的稳定性。因此，步骤 4 使用基于线性规划的估算方法绘制点集上边界线作为设备特征线[12]，得出时钟漂移 α 作为节点的设备指纹。再根据 IPv6 与 IPv4 设备指纹的比对情况判断其是否同源。指纹比对过程见算法 3。

算法 3 FingerprintMatch()

/* 设备指纹比对 */

输入：(Ad_4，Ad_6)

输出：true/false

Begin

1. α_4←EstimateCSkew(Ad_4)/* 获取 Ad_4 指纹 */
2. α_6←EstimateCSkew(Ad_6)/* 获取 Ad_6 指纹 */
3. $\theta(\alpha_4,\alpha_6)=\tan^{-1}\left|\frac{\alpha_4-\alpha_6}{1+\alpha_4\alpha_6}\right|$
4. if $\theta\in(\tau,+\infty)$
5. return false /* 非同源 */
6. else
7. return ture /* 同源 */
8. endif

End

算法 3 中，步骤 2 用设备指纹特征线的夹角 θ 表征指纹差异。理论上 $\theta=0$ 说明指纹无差异，即 Ad_4、Ad_6 地址属于同一节点。然而在现实环境中由于 TCP 时钟频率、探测误差等因素的影响，使得同源地址的 θ 在小范围内浮动。因此根据精度需求取 τ，当 $\theta\in(0,\tau)$时，判定 Ad_4 与 Ad_6 同源；当 $\theta\in(\tau,\pi)$时，判定 Ad_4 与 Ad_6 非同源。在实验结果图中表现为，将特征线基本重合的 IPv6/IPv4 地址判定为同源，将特征线夹角明显的 IPv6/IPv4 地址判定为非同源。在实际环境中经测试得知，当 τ 取 1×10^{-3}时有较理想的区分度。

3.3 基于上层协议短时致瘫的识别模块(USPI)

IPv4 与 IPv6 是网络层协议，共用网络层以上的协议栈。对于同源 IPv6 与 IPv4 地址，从网络层任一侧(IPv4 侧/IPv6 侧)对传输层协议栈产生的异常将间接影响另一侧。根据此原理，本模块从 IPv6 侧针对 TCP 协议栈发起短时致瘫，通过测试 IPv4 与 IPv6 地址响应是否同步来判断其同源性。假设短时致瘫有效，即从开始实施的时刻起目标节点无法响应 TCP 报文，且停止致瘫后目标节点立即恢复响应能力，则该模块识别过程见算法 4。

算法 4 ReplySyn()

输入：(Ad_4，Ad_6)

输出：true/false

Begin

1. Attack to Ad_6
2. sent TCP SYN to Ad_4
3. if received TCP ACK from Ad_4
4. return false /* 非同源 */
5. else
6. Stop Attack
7. sent TCP SYN to Ad_4
8. if received TCP ACK from Ad_4
9. return ture /* 同源 */
10. else
11. return false /* 非同源 */

12. endif

13. endif

End

影响该模块识别能力的主要因素为短时致瘫是否有效。考虑其资源开销及实施难度，将本模块放在 SSI 识别模型的最后阶段，只对前阶段不擅长识别的少量样本进行判断。

4 实验与分析

4.1 性能评估指标

使用某种方法从样本集中识别特定类型样本时，应关注样本识别的完备程度，以及识别结果与真实情况的相符程度。因此，本文定义识别率、正确率两个指标，用于评估识别方法的性能。

定义 3(识别率，ρ) 用于表示某种方法识别多址同源的完备性，即识别出的数量(λ)与实际样本总数(ε)的比例：$\rho=\frac{\lambda}{\varepsilon}\times 100\%$。

由此，同源识别率为 $\rho_{hom}=\frac{\lambda_{hom}}{\varepsilon_{hom}}$，非同源识别率为 $\rho_{non\_h}=\frac{\lambda_{non\_h}}{\varepsilon_{non\_h}}$，总识别率 $\rho=\frac{\lambda_{hom}+\lambda_{non\_h}}{\varepsilon}=1-\frac{\lambda_{un}}{\varepsilon}$，$\lambda_{hom}+\lambda_{non\_h}+\lambda_{un}=\varepsilon_{hom}+\varepsilon_{non\_h}=\varepsilon$。

定义 4(正确率，η) 用于表示使用某种方法识别多址同源地址的可靠程度，即正确识别的地址组数量(ω)占识别出地址组总数(λ)的比例：$\eta=\frac{\omega}{\lambda}\times 100\%$。

由此，同源识别正确率为 $\eta_{hom}=\frac{\omega_{hom}}{\lambda_{hom}}$，非同源识别正确率为 $\eta_{non\_h}=\frac{\omega_{non\_h}}{\lambda_{non\_h}}$，总正确率为 $\eta=\frac{\omega_{hom}+\omega_{non\_h}}{\lambda_{hom}+\lambda_{non\_h}}=\frac{\omega}{\lambda-\lambda_{un}}$。

正确识别总数＝样本总数×总识别率×总正确率，即 $\omega=\varepsilon\rho\eta$。

4.2 子模块有效性验证

SAPI 模块根据过渡时期特殊地址格式识别多址同源，当 IPv6 地址 Ad_6 内部嵌入 IPv4 地址 Ad_4 时，判断 Ad_6 与 Ad_4 同源。此方法显然有效，此处不再赘述。

为验证 TCFI 与 USPI 模块的有效性，搭建图 2 所示实验环境：主机 A 为探测源，S 为辅助探测节点，用于实施短时致瘫。C2 为双栈节点，与 C1 属于同一 IPv6 网络，与 C3 属于同一 IPv4 网络。Ad_{6-2}，Ad_{4-1} 是待识别地址，Ad_{6-1} 与 Ad_{4-2} 用作对照。

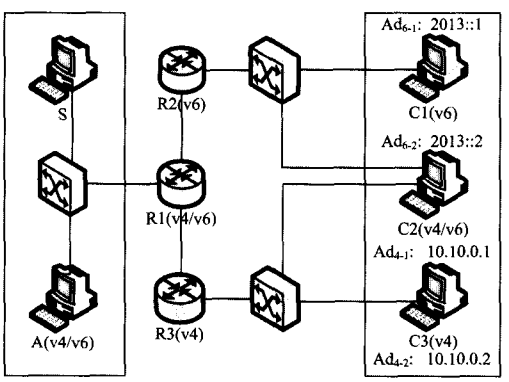

图 2 实验环境拓扑

4.2.1 TCFI 模块有效性验证

探测源 A 每间隔 100s 向 Ad_{4-1}、Ad_{4-2}、Ad_{6-1}、Ad_{6-2} 各发送一组带有 TCP 时间戳的探测包(4 个/组)，记录每个探测包的发送时间及对应的时间偏移，并在坐标系中描出这些点，使用线性规划方法绘制点集的下边界直线，求得直线的斜率作为对应地址所属节点的 TCP 时钟指纹。

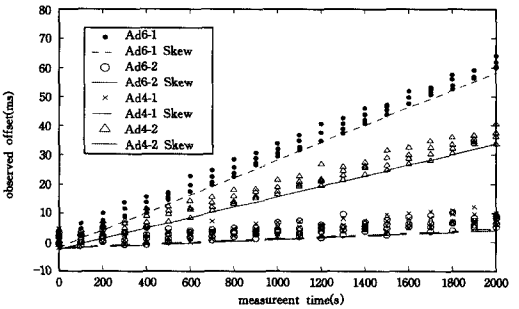

图 3 TCFI 模块识别结果

从图 3 所示实验结果可以看出，Ad_{6-2}、Ad_{4-1} 的时钟特征线基本重合($\alpha_{6-2}=34.63$ppm，$\alpha_{4-1}=34.21$ppm，$\theta=0.42\times 10^{-3}$)，与 Ad_{4-2}、Ad_{6-1} 时钟特征线的夹角明显($\theta\gg 1\times 10^{-3}$)，据此判定 Ad_{6-2}、Ad_{4-1} 是组同源地址。

实验结果与真实情况相符，证明基于 TCP 时钟指纹的识别模块能够有效识别 IPv6/IPv4 多址同源。

4.2.2 USPI 模块有效性验证

使用 Nmap 向待测地址发送 TCP 请求包探测其存活性，使用 THC-IPv6 工具集实施短时致瘫。实验过程如下：

1)将待测地址、探测模式及输出格式写入 Nmap 脚本；

2)运行探测脚本，记录输出结果；

3)辅助探测源 S 向地址 Ad_{6-2} 发起针对 TCP 协议栈的短时致瘫；

4)运行探测脚本，记录输出结果；

5)停止致瘫；

6)运行探测脚本，记录输出结果。

该实验输出结果如表 2 所列。

表 2 USPI 模块输出结果

探测阶段	Ad_{4-1}	Ad_{4-2}	Ad_{6-1}	Ad_{6-2}
初始化探测	√	√	√	√
致瘫中探测	×	√	√	×
恢复后探测	√	√	√	√

注：√：有响应；×：无响应

测试过程中，短时致瘫前后 Ad_{4-2} 与 Ad_{6-1} 正常响应；而致瘫中均无响应，即在致瘫过程中上层传输层协议响应同步，由此判定为同源地址组。

实验结果与真实环境相符。证明基于上层协议短时致瘫的识别模块能够有效识别 IPv6/IPv4 多址同源。

4.3 SSI 模型识别性能测试

为进一步验证 SSI 模型与各子模块的有效性及识别能力。我们从教育网中抽取 100 对 IPv6/IPv4 地址作为测试样本进行实验。实验中分别使用 3 个子模块以及 SSI 模型对样本集地址对进行识别，并将实验结果与网管中心提供的真实数据(该样本集中同源地址 17 对，非同源地址 83 对)比对，测试其识别率和正确率。

在使用SSI模型识别的过程中，各子模块识别样本数量及性能见表3。第一阶段从100对样本地址中识别出4对同源，阶段识别率为15%，正确率为100%。第二阶段从未确定的85对样本中识别出54对非同源(1对判断错误)，阶段非同源识别率为63.53%，正确率为98.14%。该阶段出现判断错误是因为判别临界值是根据大量实验数据获得的统计值，在实际样本中存在不符合该特征的个例。在第三阶段中，对最后31对样本进行判断，其中23对短时致瘫成功，10组被识别为同源(1组判断错误)，13组被识别为非同源(1组判断错误)，致瘫失败的8组标记为不确定，阶段识别率为74.19%，正确率为91.30%，2组判断错误。出现判断错误的原因可能是被识别的IPv6/IPv4地址处于同一链路且位置相近，而短时致瘫过程导致目标地址同一链路上的相近节点出现短时失效情况；也可能是因为探测过程中存在丢包情况。综上，使用模型识别出14组同源(1组识别错误)，识别出78组非同源(2组识别错误)，8组不确定，总识别率为92%，正确率为96.74%。各子模块单独识别性能与SSI模型识别性能对比如表4所列。

表3 SSI识别过程中子模块工作情况对比

模块	样本	同源组	识别率	正确率	非同源组	识别率	正确率	不确定	阶段识别率	正确率
SAPI	100	4	23.52%	100%	11	13.25%	100.00%	96	15.00%	100%
TCFI	85	—	—	—	54	75.00%	98.14%	31	63.53%	98.14%
USPI	31	10	75.00%	90%	13	68.42%	92.31%	8	74.19%	91.30%
SSI	100	14	82.34%	92.85%	78	78.00%	97.40%	8	92.00%	96.74%

表4 子模块与SSI模型识别性能对比

识别方法	同源组	识别率	正确率	非同源组	识别率	正确率	不确定	总识别率	正确率
实际数量	17	—	—	83	—	—	—	—	—
SAPI	4	23.50%	100.00%	11	13.25%	100.0%	85	15%	100.0%
TCFI	35	205.88%	45.71%	65	78.31%	98.46%	0	100%	81.00%
USPI	11	64.70%	90.91%	51	61.40%	98.04%	38	62%	96.77%
SSI	14	82.34%	92.85%	78	78.00%	97.40%	8	92%	96.74%

从实验结果看出SAPI模块识别正确率最高，而识别率较低；TCFI模块识别率最高，但正确率一般；USPI模块识别率略低，正确率尚可；SSI模型综合识别能力更胜一筹。

本节实验进一步证明了各子模块及SSI模型识别IPv6/IPv4多址同源的有效性。3个子模块能够协同工作，且相对于各子模块的独立识别过程，SSI模型在保证正确率的同时，总识别率得到有效提高。

结束语 本文通过分析同源IPv6/IPv4地址之间的本质联系，提出了一种IPv6/IPv4多址同源识别模型(SSI)，该模型包含3个子模块：SAPI×TCFI及USPI。分别根据IPv6/IPv4地址是否满足特殊的嵌套格式，是否对应相同的TCP时钟指纹，是否共用传输层3个层面判断IPv6/IPv4地址是否同源，提出了两种评价指标用以评估各模块及SSI模型的识别性能。最后，通过实验证明：SSI模型及其子模块均能有效识别IPv6/IPv4多址同源；SSI模型的识别率和正确率优于任一独立的子模块；在真实网络环境中SSI模型具有较理想的识别率和正确率。

下一步将研究如何在保证识别率和正确率的前提下，降低识别过程对网络流量的影响。另外，本文是从多IPv6/IPv4地址出发，验证其是否同源。如何从相逆的角度，识别某节点是否存在多个IPv6/IPv4地址，也是下一步工作的重点之一。

参考文献

[1] Dhamdhere A, Luckie M, Huffake B R, et al. Measuring the deployment of ipv6: topology, routing and performance[C]// Proceedings of the 2012 ACM Conference on Internet Measurement Conference. ACM, 2012: 537-550

[2] Keys K. "iffinder," a tool for mapping interfaces to routers [OL]. http://www.caida.org/tools/measurement/iffinder, 2000

[3] Govindan R, Tangmunarunkit H. Heuristics for Internet map discovery[C]// Proceedings Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000) IEEE. 2000, 3: 1371-1380

[4] Augustin B, Friedman T, Teixeira R. Measuring load-balanced paths in the Internet[C]// Proceedings of the 7th ACM SIGCOMM conference on Internet measurement. ACM, 2007: 149-160

[5] Keys K, Hyun Y, Luckie M, et al. Internet-Scale IPv4 Alias Resolution with MIDAR[J]. IEEE/ACM Transactions on Networking, 2013, 21(2): 383-399

[6] Beverly R, Brinkmeyer W, Luckie M, et al. IPv 6 Alias Resolution via Induced Fragmentation[C]// Passive and Active Measurement. Springer Berlin Heidelberg, 2013: 155-165

[7] Cho K, Luckie M, Huffaker B. Identifying IPv6 network problems in the dual-stack world[C]// Proceedings of the ACM SIGCOMM workshop on Network troubleshooting: research, theory and operations practice meet malfunctioning reality. ACM, 2004: 283-288

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

[9] de Donato W, Marchetta P, PescapéA. A hands-on look at active probing using the ip prespecified timestamp option[C]// Passive and Active Measurement. Springer Berlin Heidelberg, 2012: 189-199

[10] 徐朝农，徐勇军，邓志东. 线性传感器网络时间同步协议[J]. 软件学报，2009，20：266-277

[11] Arnold G, Nelson R. (RCF 1312) Message Send Protocol 2[S]. 1992

[12] Moon S B, Skelly P, Twosley D. Estimation and removal of clock skew from network delay measurements[C]// Proceedings Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM' 99) IEEE. 1999, 1: 227-234