Back to the Future: Revisiting IPv6 Privacy Extensions

DAVID BARRERA, GLENN WURSTER, AND P.C. VAN OORSCHOT

David Barrera is a PhD student in computer science at Carleton University under the direction of Paul Van Oorschot. His research interests include network security, data visualization, and smartphone security.
dbarrera@ccsl.carleton.ca

Glenn Wurster completed his PhD in computer science (2010) at Carleton University under the direction of Paul Van Oorschot. His interests include software security, system administration, operating systems, and Web security.
gwurster@scs.carleton.ca

Paul Van Oorschot is a professor of computer science at Carleton University in Ottawa, where he is Canada Research Chair in Authentication and Computer Security. He was Program Chair of USENIX Security '08, Program Co-Chair of NDSS 2001 and 2002, and co-author of the *Handbook of Applied Cryptography* (1996).
paulv@scs.carleton.ca

Network stacks on most operating systems are configured by default to use the interface MAC address as part of the IPv6 address. This allows adversaries to track systems as they roam between networks. The proposed solution to this problem—IPv6 privacy extensions—suffers from design and implementation issues that limit its potential benefits. Our solution creates a more usable and configurable approach to IPv6 privacy extensions that helps protect users from being tracked.

With more people adopting IPv6, some features of the protocol are slowly being explored by a small user-base. Security issues related to IP packet fragmentation and malicious route headers [4] have been identified, and new RFCs addressing those issues have been published (e.g., RFC 5095 and RFC 5722). Over many years, the iterative process of identifying flaws and creating fixes led to IPv4 becoming a stable and mature protocol. Since IPv6 is much newer and only now being broadly deployed, many of its features have not enjoyed broad testing or security analysis. In this article we concentrate on one such feature: IPv6 privacy extensions.

IPv6 provides the option for clients to assign themselves an IPv6 address based on a 64-bit prefix periodically broadcast by a local server, and a 64-bit value derived from the network interface identifier—usually the MAC address of the network card. Having a globally routable IP address which includes (and therefore reveals to remote servers) the MAC address of a client was regarded as a potential privacy issue, leading to the development of IPv6 privacy extensions (RFC 4941). Through the use of these extensions, a host can generate and assign itself a partially randomized (but still valid) IP address at fixed intervals, allowing connectivity without revealing its MAC address. The existing IPv6 privacy extensions are not only important for personal privacy, but also for hiding information which can otherwise allow wide-scale targeted malware attacks such as Internet-scale hit lists.

Paradise Lost

The initial IPv6 address design choices have had a detrimental impact on privacy. The proposed privacy extensions can also fail to provide the benefits they were designed for. Having recorded multiple IPv6 addresses, an adversary can trivially map two or more of these addresses to the same client, sometimes even when privacy extensions are in use. In practice, the adversary could be a corporation wishing to provide targeted services only to users that fit a specific profile (e.g., users who have visited more than three coffee shops in the past week or have been at five airports in the past month). Other adversaries may include governments or eaves-

droppers who wish to follow users as they roam through multiple locations. While the tracking of users in IPv6 is partially addressed by the IPv6 privacy extensions, the specification has a number of design issues which can cause implementations to fall short of the goal of keeping the client-to-IP-address mapping private across locations.

Without privacy extensions, tracking is possible because the last 64 bits of a client's IPv6 address are constant: if a client has IPv6 address 1::2345:6789 while on network 1, the client may have IPv6 address 2::2345:6789 when using network 2. This provides the means to track users as they move between IPv6-capable networks. Existing privacy extensions essentially randomize the last 64 bits every *x* seconds.

Paradise Regained

We propose a new technique and prototype implementation for generating private IPv6 addresses. Our proposal differs from current IPv6 privacy extensions in that it can be configured for different privacy requirements and is capable of providing private addresses even if an administrator has configured the network to deter their use. Our proposal provides as much privacy as IPv4 and has minimal overhead. We also describe the implementation of a Linux kernel prototype of our proposal.

In this article we identify issues with the current state of IPv6 privacy extensions that could lead to a downgrade attack, enabling eavesdroppers to track IPv6 users as they move through networks. We also identify issues with currently deployed implementations of IPv6 privacy extensions in modern operating systems, and we propose a more flexible and robust algorithm for generating private IPv6 addresses.

IPv6 Background

Before we explain the details of our proposal, we will review relevant terminology and background on how clients are assigned IPv6 addresses and on the originally proposed privacy extensions. We will use the generally accepted terminology.

- **Prefix:** the first (most significant) 64 bits of an IPv6 address. A prefix can be learned through periodic router advertisements, assigned by DHCPv6, or self-assigned (e.g., for loopback and link local addresses).
- **Interface identifier:** the least significant 64 bits of an IPv6 address. The prefix and interface identifier together fully specify an IPv6 address.
- **Preferred lifetime:** a lifetime associated with a particular IPv6 address during which the address should be used to initiate connections. Once the lifetime expires, the IPv6 address is deprecated, but still active for the remaining open connections. The address remains deprecated until the valid lifetime expires.
- **Valid lifetime:** a lifetime associated with a particular IPv6 address. When it expires, the IPv6 address is removed from the network interface by the kernel and no longer used.

Obtaining IPv6 Addresses

Clients can obtain IPv6 addresses through one of three methods: (1) the user manually assigns a valid IPv6 address to an interface; (2) a periodically advertised

prefix is prepended to the self-generated interface identifier; or (3) a DHCP server is queried and the received response used. We review methods (2) and (3).

STATELESS ADDRESS AUTO-CONFIGURATION

IPv6 provides a method for clients to automatically assign themselves a valid IPv6 address based on periodically broadcast router advertisements. In a typical setup, a router broadcasts the IPv6 prefix that all clients should prepend to their auto-configured interface identifier. In 1998 the authors of RFC 2462 suggested that clients use their network MAC address in the generation of the interface identifier. The rationale was that this provided sufficient uniqueness and would require no persistent storage. Nine years later, RFC 4862 removed the MAC address suggestion, allowing hosts to choose their own method for generating interface identifiers. The interface identifier is always appended to the network prefix, after which the Duplicate Address Detection (DAD) algorithm is run by the client to ensure that the address is unique to the network segment, and therefore globally unique (as prefixes are also unique).

In cases where a MAC address is used as the interface identifier (still currently the default behavior of Linux and Mac OS), the IPv6 address reveals information which can be used to identify the client hardware. This ability to determine the hardware configuration of a machine may lead to additional information about the client being revealed on the network. For example, Mac OS runs on a specific underlying hardware platform, allowing the identification of Apple users based only on MAC addresses. This ability to determine the characteristics of a client through the MAC address can be used in a targeted attack on a user or organization (e.g., sending a malicious PDF that only exploits Mac OS). Bellovin et al. [3] argue that the MAC address could also be used by IPv6 worms to target specific hosts.

INADVERTENT IPV6 USERS

We define *inadvertent IPv6 users* as users who unknowingly use IPv6 to connect to remote servers. While the vast majority of Internet users currently use IPv4, modern OSes attempt to use IPv6 by default when resolving hosts. In a typical network, connecting (and therefore revealing the source IP address of the connection) to a remote server over IPv4 will not typically allow the server to track the individual or identify the network hardware. Connecting to the same server over IPv6 may reveal sufficient information to track the individual and identify hardware. Because stateless address auto-configuration does not depend on additional client software (other than an updated kernel), it is likely to cause inadvertent IPv6 use. This increases the importance of IPv6 privacy extensions that truly provide protection and information hiding.

DHCPV6

With the addition of stateless address auto-configuration for IPv6, hosts can obtain network information and learn how to route packets without installing additional software. While this may seem ideal from a network management standpoint (e.g., set up a route prefix advertisement daemon and IPv6 just works), there may be other configuration parameters needed by hosts in order to actually communicate with external hosts. These parameters will vary from network to network, but some include WINS, NTP, NETBIOS, and DNS.

There are cases where administrators may choose to replace stateless auto-configuration with DHCPv6, or use both simultaneously. When using DHCPv6 for address assignment, the server keeps track of assigned addresses and the hosts using them, as DHCP did in IPv4. When using both, a host obtains its IPv6 address through stateless auto-configuration and other information through a server on the local network. The issue of tracking clients using IPv6 is specific to those who obtain an address through stateless address auto-configuration.

Original Privacy Extensions

IPv6 privacy extensions for stateless address auto-configuration were proposed specifically to address privacy concerns with having a static and globally unique interface identifier. Concerns that a well-placed sniffer (or prolific ad network) might track users as they roam through different networks are partially mitigated by privacy extensions through using periodically changing random interface identifiers. RFC 4941 specifies the algorithm used to generate a random identifier, as well as when to update it. As shown in Figure 1, a hash function (MD5 is suggested in the RFC) is used to generate the interface identifier. The first 64 bits of output are used as the interface identifier, while the last 64 are stored for the next iteration of the algorithm, which takes place every *x* seconds (or when a duplicate address is detected by the client). The first iteration of the function uses a random value as the history value.

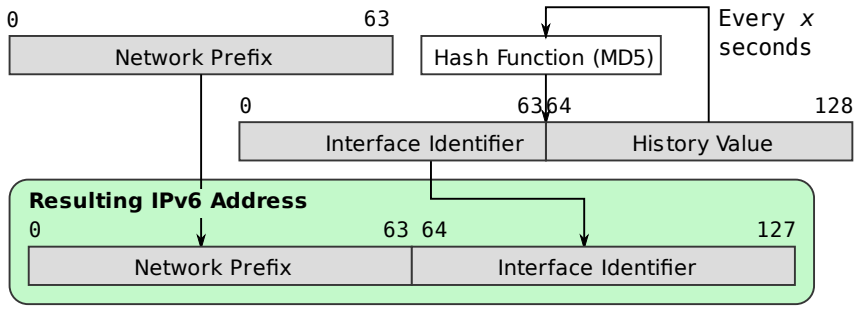

Figure 1: Original privacy extension address generation

The current specification has two important limitations. The only configurable parameter is the interval at which new random interface identifiers are generated. The default interval is to generate a new identifier every 24 hours. This still allows a user moving between two or more IPv6 networks in a 24-hour window to be tracked by an adversary (since the client's interface identifier will not change during this time, even if the network prefix does). The expert user can configure the regeneration interval to be smaller, at the expense of no longer maintaining long-lasting connections (e.g., SSH or movie downloads).

The intervals are dependent on the configuration of the network. If a user has configured the interval for regeneration of addresses to be small, but the network advertises smaller intervals, the smallest takes precedence. This means that if the network is configured to advertise prefixes with valid lifetimes of 60 seconds, a user with privacy extensions enabled will generate a new and different IPv6 address roughly every 60 seconds. This will severely impact user experience: no connection made will last more than 60 seconds.

The latest RFC for privacy extensions also specifies that system implementers should add an option for the user to enable or disable random interface identifiers on a per-prefix basis. This is similar to our proposal in that a new full IPv6 address is generated when the prefix changes (the user changes networks), but differs in that they rely on the client to maintain a list of networks for which privacy extensions should be enabled (or disabled) and do not use the prefix directly in the generation of random interface identifiers.

New Privacy Extensions Proposal

In this proposal we focus on protecting clients who configure their IPv6 address through router advertisements from being tracked as they move between IPv6-enabled networks. We do not address clients configured through DHCPv6 or clients with static IPv6 addresses. We assume that each IPv6 network visited by a client is associated with a distinct prefix (routing problems result if two networks share the same IPv6 prefix).

We assume that the attacker does not have access to the LAN segment, and hence cannot associate IPv6 addresses with MAC addresses, but we do not assume that the network administrator is totally benign. We assume the network administrator is capable of modifying router advertisements, forcing users to renew their IPv6 address often. We assume an attacker attempting to track the client can see traffic generated with each IPv6 address the client uses. We do not attempt to protect against tracking clients using higher-level protocols such as HTTP [5].

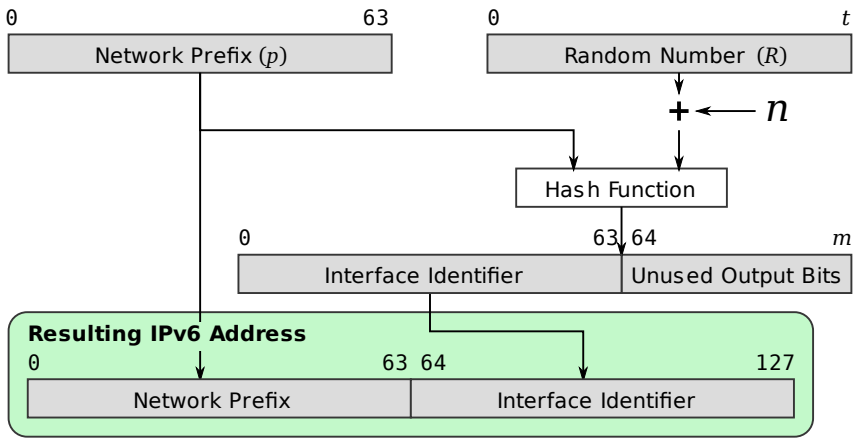

Figure 2: Generation of new IPv6 addresses

The new proposal is for clients to generate IPv6 address interface identifiers *(I)* through hashing the IPv6 prefix advertised by the router advertisement daemon *(p)* with a t-bit random number *(R)* incremented by *n* (in order to resolve duplicate addresses). *R* is generated locally and does not leave the client. The generated interface address is composed of the first 64 bits of the resulting m-bit hash value, as illustrated in Figure 2.

$$I = H(p \mid (R + n))$$

We require a pre-image resistant hash function H [8] so that given the prefix and interface identifier, an attacker cannot determine R. Keeping R hidden prevents an attacker from determining that two distinct IPv6 addresses correspond to the same client. There are no compatibility or interoperability concerns should two clients choose to use different hash functions in generating the interface identifier.

To ensure that a new I is generated for every new network prefix (which is not possible in the current privacy extensions), p is included in the hash. There are several options regarding when to change R, allowing the client control over when to start using a new interface identifier. To prevent known attacks [11, 8] against guessing R, its length (t) should be sufficient (e.g., t = 128 or 256 bits should certainly be enough), and it should be set from a cryptographically secure random number generator.

Should a client discover (through duplicate address detection) that it is attempting to use the same generated IPv6 address as another client on the local network (an unlikely scenario), the client should generate a new I (and hence IPv6 address) by incrementing n and recomputing the hash. The client should reset n to 0 on reboot and whenever p or R changes.

Generation of New Random Numbers

Our proposal includes several options on when to (re)generate R, resulting in a changed IPv6 address. The different options provide different levels of privacy protection, which we now discuss.

1. **Generate *R* on OS install**. If R is generated during system install and then never changed, I will change when the network prefix advertised by the router changes. As long as p remains constant, so will I. This option is useful for laptops in enterprise environments. As long as the laptop is on the corporate network, the IPv6 address will be fixed. When the laptop is removed from the network (e.g., the employee goes to a USENIX workshop), the interface identifier I will change, preventing the employee from being tracked as they roam between networks. When they rejoin the enterprise environment, they will re-obtain the original interface identifier.
2. **Generate *R* on OS reboot**. This option results in a new IPv6 address every time the computer is rebooted, even if the client receives the same IPv6 prefix from the broadcast daemon.
3. **Generate *R* on network interface change**. This option results in a new interface identifier I being generated whenever the client computer brings up the network interface. Since interfaces are brought up on boot and when connecting to a wireless network, a client will use a different I each time it joins a network broadcasting the same IPv6 prefix.
4. **Generate *R* when the user chooses**. This option results in a new interface identifier I being generated based on user involvement (e.g., the user regenerates R when transitioning between tasks). While we include this option for completeness, we do not suggest defaulting to this option.
5. **Generate *R* every *x* seconds**. In this option, the client generates a new IPv6 address every x seconds. This approach closely parallels the current IPv6 privacy scheme. Unlike the approaches discussed above, a new IPv6 address may be generated while network connections are open, causing these connections to be dropped. To reduce the number of dropped connections, the kernel can avoid deleting old IPv6 temporary addresses associated with active network connec-

tions. As long as the active network prefix is the same as that contained in the old temporary address, the temporary address can continue to be used. While we note that x does not need to stay fixed (i.e., a new x can also be chosen when the random identifier is updated), we currently see no additional benefit in randomly changing x. A default x of one day mirrors the current default with IPv6 privacy extensions.

We suggest the generation of a new random number (and hence interface identifier) whenever the network interface is brought up (option 3). This method generates IPv6 addresses as frequently as possible without interrupting open connections (since connections are terminated when the interface goes down).

Our proposal is designed so that given two distinct IPv6 addresses, it should be hard for an attacker to answer the question, "Did the same client use both IPv6 addresses?" To answer this question, the attacker must be able to determine that the same R value was used in the generation of both addresses (since two clients sharing the same R value is extremely unlikely).

In answering this question, we assume that the attacker has access to the generated interface identifier as well as to the prefix. The security, therefore, depends on the difficulty of determining the random number provided to the pre-image resistant hash function—which is assumed to be hard for a sufficiently large R. An attacker attempting to track a client would need to keep trying random values for R until finding one which generates multiple distinct and observed interface identifiers; therefore a birthday attack [2, 8] does not seem to help.

Because the interface identifier changes whenever the prefix changes, a client connecting through two networks with different prefixes will also connect with different interface identifiers, leaking no information in the IPv6 address.

One potential attack against our proposal involves a network administrator (as attacker) broadcasting target prefixes in an effort to detect what interface identifier would be used by the client on that network (e.g., the administrator broadcasts prefix 1:2:3::/64 to determine what IPv6 address the client would attempt to use on that network). The attack would be successful if R was not updated by the client before visiting the target network. One way to defend against this attack is to configure the client to generate a new R whenever it enables or makes a change to the network interface.

The proposal for generating interface identifiers relies on an appropriate random number generator. If R can be guessed, an attacker can determine whether a client using the same R value generated multiple distinct IPv6 addresses using distinct prefixes. The proposal also depends on a pre-image resistant hash function (SHA-2 should suffice).

Our approach does not protect against an attacker identifying two addresses used by a client through correlating the time at which one IPv6 address stops being used and another starts. We expect that as IPv6 privacy extensions are deployed, the volume of IPv6 address churn will make correlations more difficult.

As an extension to the core approach, it may be possible to use multiple IPv6 temporary addresses concurrently on a host. As an example, a new IPv6 address could be used for every application running on the client (e.g., Web browsing would use one IPv6 address, DNS queries would use another, and an active SSH connection may use a third). While not directly privacy related, a server may also choose to use multiple IPv6 addresses (e.g., as a method for distributing firewalls [12]).

Implementation

For our prototype implementation, we used version 2.6.34 of the Linux kernel. We modified the currently implemented IPv6 privacy extensions. The modified kernel provides several sys-controls which can be read and written to by user-space programs, controlling the operation of IPv6 privacy extensions. These sys-controls are as follows:

use_tempaddr: controls whether or not to enable privacy extensions. Possible values are listed in Table 1.

temp_valid_lft: the maximum amount of time a temporary address is valid. In the original approach, a new distinct temporary address would be created. In this proposal, the lifetime of an already existing temporary address will be extended when router advertisements are received if both p and R are unchanged.

temp_prefered_lft: the maximum amount of time a temporary address will be the preferred address for the interface. As with temp_valid_lft, the lifetime will be extended if both p and R are unchanged when a router advertisement is received.

temp_random (new): a 32 byte (256 bit) random value R used as input to the hash function. This sys-control is specific to our proposal.

We tested our prototype by switching between several IPv6 networks and verifying that the generated IPv6 addresses were different at each network. We did not notice any impact on activities such as Web browsing and SSH. During the course of implementing the proposed IPv6 privacy extensions in Linux, we found several bugs which cause IPv6 privacy extensions to be disabled and/or have all temporary addresses deleted. We have a patch in version 2.6.37 of the Linux kernel which addresses these implementation deficiencies. We will be submitting our proposed revisions to privacy extensions to the Linux kernel in hopes that this will help improve adoption.

Value	*Meaning*
0	Privacy extension disabled
1	Original privacy extension enabled but not used by default for new outgoing connections
2	Original privacy extension enabled and used by default for new outgoing connections
5	Proposed privacy extension enabled but not used by default for new outgoing connections
6	Proposed privacy extension enabled and used by default for new outgoing connections

Table 1: Possible values and their meanings for the *use_tempaddr* sys-control in the modified kernel.

Other Related Work

Cryptographically Generated Addresses [1] (CGA) were proposed to prevent stealing and/or spoofing IPv6 addresses. CGAs define a method for securely associating a public key to an IPv6 address. The interface identifier of the IPv6 address is a cryptographic hash of the public key, which can later be verified by the recipient of the packet or message. Because CGAs tie a public key to an IPv6 address, even as hosts switch networks, they are uniquely identifiable through use of the public key. CGAs, like our proposal, use a cryptographic hash to generate the interface identifier, but the purpose of CGAs is contrary to ours and our proposal does not involve public keys.

Mobility extensions (RFC 3775) define ways in which a mobile host can move between networks while maintaining open connections, even if the networks use different link layer technologies (WiMAX, LTE, WiFi). This is accomplished by establishing a tunnel (usually with IPSec) to the *home network*. The mobile host is then reachable through the proxy home network. Route optimization (RFC 4866) allows the correspondent node (server) to communicate directly with the mobile host, even though the mobile host's IPv6 address may continue to change. IPv6 Mobile extensions with route optimization are illustrated in Figure 3.

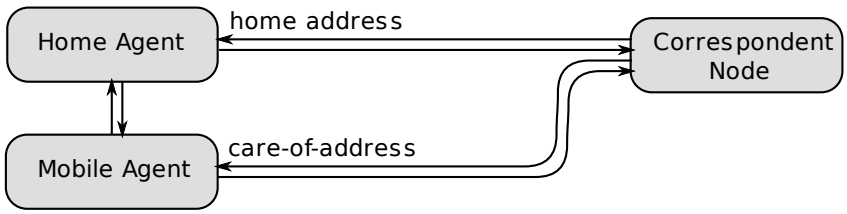

Figure 3: Communication with Mobile IPv6 agents

For mobile agents, implementing route optimization is mandatory. The correspondent node receives both the home address and care-of address and can track the location of the mobile agent as it moves between IPv6 networks. Because the mobile agent is also tied to the fixed home address, enabling privacy extensions at the mobile agent does not prevent the correspondent node from tracking the mobile agent. To prevent tracking the mobile agent, the home address and care-of address must be changed at the same time. Otherwise, the correspondent node can tie the old address to the new address and continue to track the mobile node.

Mobility extensions aim to address connection persistence problems rather than privacy concerns. The former have also been solved independently in the cell phone industry at the link layer, where cell towers hand off connections to prevent active phone calls from being dropped when a device switches towers.

Tracking at Other Layers

Guha and Francis [6] demonstrate how to track users through DNS. Their analysis shows that dynamic DNS updates combined with geolocation [9] can provide a passive attacker with all the approximate locations visited by a victim. Geolocation in IPv6 may also reveal more accurate results compared to IPv4 due to address allocation recommendations (http://www.apnic.net/policy/ipv6-address-policy).

Users with private IPv6 addresses may be tracked at the application layer through cookies [7] or browser characteristics [5]. Protecting privacy at all layers is clearly a difficult problem and beyond our scope, but we argue that in order to have privacy at higher-level protocols, underlying protocols must also be private.

While we do not address the problem of tracking users on the LAN specifically in this article, we note that tracking users at the Ethernet level is possible due to clients broadcasting their MAC addresses [10]. Because MAC addresses in the Ethernet header are overwritten on a hop-by-hop basis, attackers outside the LAN do not obtain the MAC address of a client. This article focuses on the network layer, where tracking can be performed across the Internet.

Conclusion

We have proposed a new way of generating the interface identifier used in temporary IPv6 addresses. The use of temporary addresses prevents tracking clients as they move between IPv6 networks. Our approach does not use MAC addresses, which can be used to identify client hardware.

Our proposal has several benefits over the current IPv6 privacy extension scheme, including: (1) the ability to maintain a consistent IPv6 address over an extended period regardless of the lifetime specified by a router advertisement (as long as the prefix being advertised does not change); (2) the ability to always use the same interface identifier while connected to a network broadcasting an unchanging prefix; (3) the ability to configure when a new interface identifier should be created (e.g., whenever the network interface is brought up); and (4) not being able to track a client through the use of a common interface identifier across networks broadcasting different IPv6 prefixes. We have implemented and tested the approach in Linux and found that it generates new interface identifiers as designed while not impacting Internet activities.

A version of this article is also available as Technical Report TR-10-17 (September 9, 2010), Carleton University, School of Computer Science.

References

[1] T. Aura, "Cryptographically Generated Addresses (CGA)," in *Proceedings of the 6th International Information Security Conference (ISC '03)*, pp. 29–43.

[2] M. Bellare, O. Goldreich, and H. Krawczyk, "Stateless Evaluation of Pseudorandom Functions: Security beyond the Birthday Barrier," 19th International Conference on Cryptology (Crypto '99).

[3] S. Bellovin, B. Cheswick, and A. Keromytis, "Worm Propagation Strategies in an IPv6 Internet," *;login:*, vol. 31, no. 1 (February 2006), pp. 70–76.

[4] P. Biondi, A. Ebalard, M. Balmer, and V. Manral, "Ipv6 Protocol Type 0 Route Header Denial of Service Vulnerability," April 23, 2007: http://www.securityfocus.com/bid/23615.

[5] P. Eckersley, "A Primer on Information Theory and Privacy": https://www.eff.org/deeplinks/2010/01/primer-information-theory-and-privacy.

[6] S. Guha and P. Francis, "Identity Trail: Covert Surveillance Using DNS," in *Proceedings of the Privacy Enhancing Technologies Symposium*, 2007.

[7] D. Kristol and L. Montulli, "HTTP State Management Mechanism," RFC 2965 (Proposed Standard), October 2000.

[8] A.J. Menezes, P.C. Van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography* (CRC Press, 1996).

[9] J.A. Muir and P.C. Van Oorschot, "Internet Geolocation: Evasion and Counter-evasion," *ACM Computing Surveys* (CSUR), vol. 42, no. 1 (2009), pp. 1–23.

[10] L. Peterson and B. Davie, *Computer Networks: A Systems Approach* (Morgan Kaufmann, 2007).

[11] B. Preneel and P.C. van Oorschot, "On the Security of Iterated Message Authentication Codes," *IEEE-IT*, vol. 45, no. 1 (January 1999), pp. 188–99.

[12] H. Zhao, C.-K. Chau, and S.M. Bellovin, "ROFL: Routing as the Firewall Layer," in New Security Paradigms Workshop (NSPW), 2008.