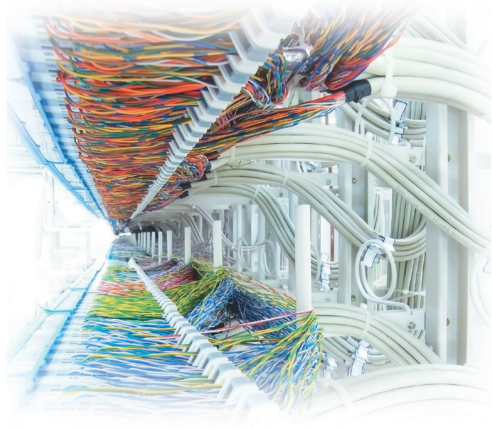
Anonymization Technology Takes a High Profile



Neal Leavitt

Increasingly, governments and various types of organizations are trying to either block or track Internet access and online communications by dissidents, employees, or others. To sidestep these activities, users are turning to anonymization technology.

Censorship and the increased tracking of users online have become important topics. Numerous governments censor computer- and network-based communications to keep their citizens from freely getting news from or transmitting information to the outside world.

Dissidents and everyday Internet users—as well as criminals and others who want their online identities to be secret—have turned to anonymization technology to keep from being identified. This has made the technology more important and widely used in recent years, sparking the start up of anonymization companies and the development of new techniques.

"The evolving threats, the introduction of new technologies and applications, and the emergence of Internet censorship are really driving [the approach] right now," said Lance Cottrell, the founder and chief scientist of anonymization vendor Anonymizer. "And events like the recent elections in Iran have really drawn attention to it." In Iran, protesters against the results of the recent presidential elections fought government censorship to communicate with the outside world.

Anonymization technology faces numerous challenges to increased adoption and commercial success. Moreover, the technology has generated controversy among those concerned that terrorists, pedophiles, criminals, and others could take advantage of it.

THE BASICS

Anonymity systems prevent observers from discovering the source of online communications. Typically, the system keeps a recipient or observer of a transmission from seeing the IP address of a source or tracking a message back to its originator.

"The name of the game is to keep the servers you visit from knowing your IP address, which means not connecting to them directly. This means going through one or more other computers [called *proxies*] to arrive at the desired destination," said James Marshall, an independent consultant and software developer who created CGIProxy, a free Web proxy.

Some anonymity systems also encrypt data.

History

The first popular anonymization tool was the Penet remailer developed by Johan Helsingius of Finland in the early 1990s. Penet was not totally safe for users because it kept a potentially accessible record of their names.

Some members of the Cypherpunk privacy and cryptography developers' group released their eponymous remailer in 1992.

Cottrell wrote the Mixmaster remailer in 1993. In 1995, he launched Anonymizer—the first Web-based anonymity system, initially a free service but now a commercial product.

Driving forces

Concerns about communications privacy are driving anonymization technology's increased adoption.

And as Internet use has grown, criminals have increasingly gone online to break the law, noted Rob Enderle, principal analyst with the

0018-9162/09/$26.00 © 2009 IEEE Published by the IEEE Computer Society

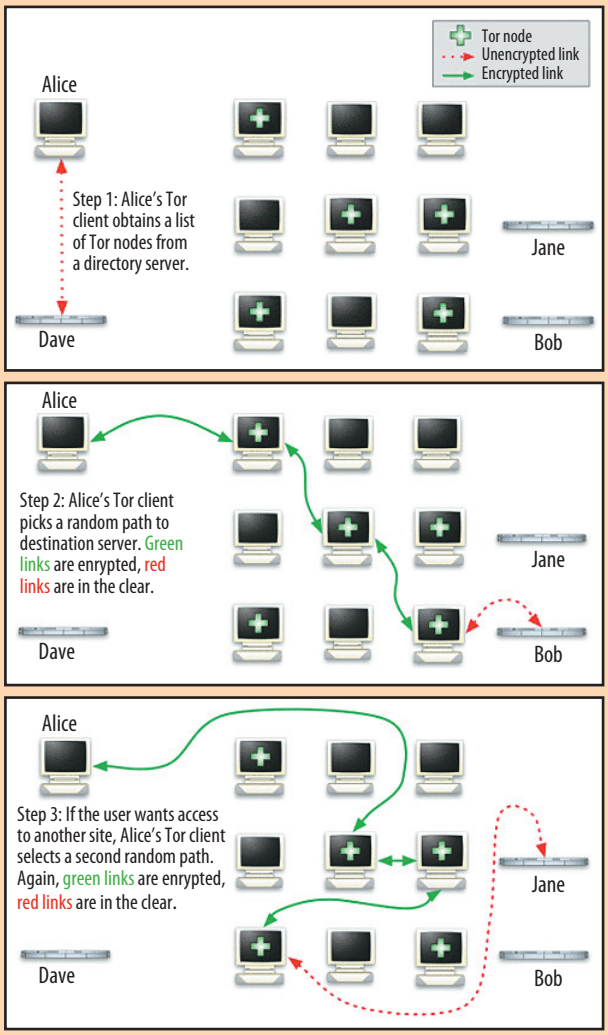


Figure 1. The Tor anonymization system hides a user's identity by sending traffic through a series of participating nodes.

Enderle Group, a market research firm. In the process, they have looked for ways to keep law-enforcement agencies from identifying them.

In some countries, libraries and employers block content, and some ISPs and websites record people's Web habits for marketing purposes.

Dissidents and other protesters also want ways to communicate without governments being able to identify or trace transmissions back to them.

The 2006 OpenNet Initiative (http://opennet.net)—a research project by Harvard University and the Universities of Cambridge, Oxford, and Toronto—studied Internet censorship and surveillance in 46 countries. The study found that 25 of the nations filtered various types of communications—including political content, religious sites, and pornography—by blocking transmission to and, in some cases, from specific IP addresses.

Officials identify sources they want to block by tracking back communications and identifying their senders via tools such as traceroute and services such as whois. They also find sites to block by using search engines and monitoring discussion groups and chat rooms.

Some individuals and organizations identify various sources' IP addresses and sell the information to governments or other interested parties.

UNDER THE HOOD

Over time, demand for anonymization has grown and the types of uses, applications, and business models have evolved.

Commercial anonymization systems, such as SwissVPN (www.swissvpn.net), charge subscription fees for their services.

Noncommercial anonymization systems don't charge fees but instead generate revenue by selling advertising that appears on their webpages.

Home-brewed anonymization systems are based on anonymizing proxy packages, such as CGIProxy and Freenet, available online for free. They are popular with college students who use them to circumvent school networks' URL filtering systems.

Types of anonymizers

Users can install software to implement *simple virtual private network (VPN) systems*, such as Anonymizer Total Net Shield and Perfect Privacy. VPNs create encrypted tunnels through which traffic passes. Recipients or observers cannot read the encrypted traffic and thus cannot track it back to the sender.

Users can also install software for *simple proxies*, also known as *open proxies* and *anonymous proxies*. Users enter the proxy's IP address or hostname in their browser's network settings, and when they point their browser at a website, the browser tells the proxy which site to visit. The proxy visits the site on the users' behalf and sends the content back to them.

The systems remove the users' IP information from packets and replace it with their own IP information, said Rolf Wendolsky, a director of anonymization vendor JonDos.

The typical proxy provider sets up a server on the Internet through

which users can relay traffic, which some anonymization applications encrypt. This *single-hop architecture* is easy to implement and maintain. However, users all enter and leave through the same server, thereby creating a single point of failure.

Daisy-chaining anonymization, which uses a *multihop* approach, sends a user's traffic through a series of participating nodes, as Figure 1 shows. The traffic travels a path which either the user or the anonymization software selects, depending on the application. The goal is to route traffic through nodes owned by different individuals or organizations. That way, no one organization can see enough packet information to identify the user.

With *form-based proxies*—such as Anonymizer Anonymous Surfing and Anonymouse—users enter the URL of websites they want to visit into a form field on an anonymization provider's page. The provider then takes the user to the desired site. The anonymization software rewrites links on the delivered page so that they connect to the provider, preventing anyone from tracing the transaction back to the original user.

Form-based proxies are written via either common-gateway-interface scripts, designed to transfer information from forms and other online sources between a Web server and a browser; or PHP Hypertext Preprocessor scripts, which run on a Web server and enable dynamic Web content such as forms.

Form-based proxies are popular because users don't have to configure or install any software.

However, they are the most insecure of all anonymization systems, said Wendolsky. For example, attacks could use form-based proxies to replace links on websites with URLs that send users to malicious sites.

"The disadvantage of such systems is [slower] performance, both because of the multiple hops and because of the poor performance of many nodes," explained Cottrell. Also, he added, users can't always judge the trustworthiness of the participating nodes' owners.

Protocol support

Protocol-specific systems like Anonymouse and PHProxy anonymize online activities—for example, e-mail or Web access—based on only one or several application-layer protocols, such as HTTP or the Simple Mail Transfer Protocol, thus they are not versatile.

Anonymization technology faces numerous challenges to increased adoption and commercial success.

But because they are designed to work in detail with only certain types of applications, they can effectively recognize and strip out all user-specific data from the traffic they send.

Protocol-independent systems such as JonDonym use approaches such as SOCKS—designed to send TCP traffic via a proxy server—which supports many communications protocols. They can also take advantage of VPNs, which also work with many protocols.

Although these systems obscure the path that traffic takes, they don't generally "understand" traffic well enough to actually change data in packets, which could reduce their effectiveness.

APPLICATIONS

The leading anonymizing applications include the following:

- Anonymizer (www.anonymizer.com): VPN- and form-based systems, supported by user payments, open to anyone on the Internet, protocol-independent, encrypts communications
- Anonymouse (http://anonymouse.org): form-based system, supported by on-site advertisements and user payments, open, protocol-specific
- I2P (Invisible Internet Project, www.i2p2.de): VPN system, free, closed to all but those on subscribing networks, open source, protocol-independent, encrypts communications
- JonDonym (https://www.jondos.de/en): multihop proxy system; free and commercial versions; open source; open; encrypts communications; originally developed by the Technical University of Dresden, the University of Regensburg, and JonDos
- Megaproxy (www.megaproxy.com): VPN system, supported by user payments, open, protocol-independent, encrypts communications
- Proxify (http://proxify.com): form-based proxy system, supported by advertisements or user payments, protocol-independent, encrypts communications
- Tor (www.torproject.org): multihop proxy system; free; open; open source; protocol-independent; encrypts communications; started in 2003 with 30 proxies on two continents, now has 2,000 on five continents and up to 500,000 users at any one time
- XeroBank (https://xerobank.com): multihop-proxy and VPN systems, supported by user payments, partially open source, open, protocol-independent, encrypts communications

CHALLENGES AND CONTROVERSY

Criminals could take advantage of improved anonymization technology to hide their identities, said analyst Enderle.

Also, anonymizers aren't foolproof. For example, if the first and last proxies in a system are malicious or compromised, the first proxy would know the client's identity and the last proxy would know the server's identity, explained Indiana University assistant professor Apu Kapadia. If the same person owns both proxies or if their separate owners communicate, this could break anonymity, he said.

Most open source projects publish enough information about their workings, including node addresses, to let governments or other organizations block traffic from at least some of those nodes, noted Cottrell.

According to Seth Schoen, staff technologist for the Electronic Frontier Foundation, a privacy and Internet-user-rights organization, there is a risk that some single-proxy anonymizer services may log users' IP addresses. If governments order them to turn over information or hackers break into their servers, users could lose their anonymity, he explained.

However, he noted, providing greater security would hurt performance because additional proxies and encryption increase overhead.

In fact, performance overhead sometimes causes anonymization to slow users' Internet access.

Expanding the number of nodes in anonymization systems could be difficult because users serving as nodes will experience a lot of traffic flowing through their computers.

Some ISPs block nodes to control spam. If, in the process, they block those used by anonymizers, Marshall said, this would hurt anonymization.

Browser complexity and the need to maintain browsing functionality could help proficient hackers sidestep anonymization, noted Wendolsky. Hackers could accomplish this in some cases, he explained, by exploiting browser plug-ins, JavaScript, cookies, caches, or HTML parsing engines.

Analyst Enderle stated, "Anonymizers are wrongheaded." The technology conceals identities, he said, which makes it attractive to criminals.

The technology's two biggest marketplace challenges are cultural and legal, according to Cottrell. "The legal challenge is that some countries are outlawing or could decide to prohibit the use of privacy tools and require all Internet providers to keep detailed access records. The cultural issue is the trend toward [openness on the Internet]."

However, proponents say that privacy and the desire to communicate online without fear of identification or government retribution are among the good reasons to use anonymization and that this will drive the technology's continued development and adoption.

Anonymizer, for example, has reported a 20 percent annual growth in its business over the past few years.

Marshall predicted that anonymization will have a bright future, with more organizations developing systems as people become aware of its importance. He said, "The demand is there."

Neal Leavitt *is president of Leavitt Communications (www.leavcom.com), a Fallbrook, California-based international marketing communications company with affiliate offices in Brazil, France, Germany, Hong Kong, India, and the UK. He writes frequently on technology topics and can be reached at neal@leavcom.com.*

Editor: Lee Garber, *Computer*, l.garber@computer.org