First Monday, Volume 1, Number 2 - 5 August 1996

first monday

PEER-REVIEWED JOURNAL ON THE INTERNET

PROSPECTS FOR REMAILERS
Where is Anonymity Heading on the Internet?

By SAMEER PAREKH

Abstract

Remailers have permitted Internet users to take advantage of the medium as a means to communicate with others globally on sensitive issues while maintaining a high degree of privacy. Recent events have clearly indicated that privacy is increasingly at risk on the global networks. Individual efforts have, so far, worked well in maintaining for most Internet users a modicum of anonymity. With the growth of increasingly sophisticated techniques to defeat anonymity, there will be a need for both standards and policies to continue to make privacy on the Internet a priority.

Contents

Introduction
Early tools for anonymous postings
A Cypherpunk solution
Watching the traffic go by
Timers
Spamming networks for identities
Mixmaster
Remailers for everyone
Conclusion
Notes

In 1979, Tom Truscott and James Elliott of Duke University experimented with a way to share files with UNIX, using computers at Duke and the University of North Carolina. From these humble and unpretentious beginnings, where it was expected that traffic would amount to just one or two messages a day, Usenet grew into a number of groups dedicated to specific topics.

The growth of Usenet led to a re-organization of the newsgroups a decade ago, into seven areas dedicated to computer science (comp.), social issues (soc.), science (sci.), recreation (rec.), news (news.), discussions (misc.), and debate (talk.). This growth was also stimulated by the development of specific protocols to allow the distribution of Usenet information over the Internet, and the rapid interest in the Internet and electronic communication itself. Usenet has become one of the most successful vehicles for global discourse, with nearly 100,000 Usenet sites and 3 million users scattered around the world [1].

There was some concern early in the history of Usenet for newsgroups outside the Usenet universe. The alternative hierarchy in fact began over objections by Brian Reid, the moderator of a recipe newsgroup; his alt.gourmand group was the first of now many alternative groups dedicated to many diverse topics. The enthusiasm for these alternative newsgroups, and other more traditional groups, led for a need for anonymity on the Internet.

Early tools for anonymous postings

Readers of Usenet groups which dealt with sensitive topics - the alt.support.* hierarchy in particular - developed a need for anonymous mail. To solve this problem, Karl Kleinpaste built a pseudonym server, which created pseudonymous identities. Those posting to sensitive newsgroups could protect their identities from their coworkers, employers, and even their families.

The system was set up so that it would work transparently and very easily. To make an anonymous posting to a newsgroup, a Usenet reader would send an electronic mail message to the pseudonym server, directing the server to send the contents of a post to the Usenet group. The actual identity of the sender would be replaced with a pseudonymous ID, allocated by the server. This identification would allow responses from the newsgroup to the posting. The server in turn maintained a database that correlated each pseudonymous ID to the real e-mail address of every user. When a mail message was sent to the pseudonymous ID on the server, the message itself would also receive the same anonymous treatment by the server. The author of a given message to the server would also be allocated a pseudonymous ID and then forwarded to the recipient.

This pseudo-anonymous service was used heavily, and a number of systems were set up to handle the traffic. Some of these servers lasted only weeks or months. Discussions raged in newsgroups over the ethics of anonymous postings. There was abuse of the initial server itself, resulting in its migration from site to site and from owner to owner. Eventually, Johan 'Julf' Helsingius in 1992 established one of the most heavily used pseudo-anonymous servers in Finland, the anon.penet.fi server.

Julf's server is the most prominent pseudo-anonymous server on the Internet with over 500,000 users sending about 8,000 messages every day [2]. Julf continues to run the server as a free service to the Internet community in the face of threats and abuse, at the cost of at least $1,000/month [3].

Anon.penet.fi provides only very minimal security. Since it maintains a database that maps anonymous identities to real e-mail addresses, it is susceptible to compromises. The best example of this sort of invasive manipulation involved the Church of Scientology last year. On February 8, 1995, Helsingius turned over to Finnish officials the identity of an anonymous user of his server, based on a complaint initially filed by the Church of Scientology with the Los Angeles, Calif. police and the U. S. Federal Bureau of Investigations. Helsingius, as a Finnish citizen operating anon.penet.fi on Finnish soil, was required to surrender the specific information or lose his server completely [4]. Thus, anon.penet.fi - and other pseudonym servers this type - was shown to be not secure, not from a technical angle but from a legal one.

Helsingius has instituted a number of mechanisms to reduce the susceptibility of his server to compromise, but the anon.penet.fi server still provides only a low level of security. For instance, it doesn't have a distributed trust model, depending on just a single server [5]. Disable the server and no one can use it. Infiltrate the server and compromise hundreds of thousands of users.

A Cypherpunk solution

In late 1992, a group of cryptography enthusiasts and professionals was founded in a meeting in the San Francisco area. Calling themselves "cypherpunks" (a term invented by Judith Milhon as a play on the 'cyberpunk' science fiction genre and the word cipher), a mailing list was started. It now generates at least 50 messages a day and includes almost 2,000 members worldwide [6]. At the same time, two of the group's founders, Eric Hughes and Hal Finney, built a secure anonymous mail system for the Internet. Their system was a very weak version of the so-called "digital mix" envisioned in the early 1980s by David Chaum [7].

Hughes and Finney's 'cypherpunk remailer' system fixes the problems of the anon.penet.fi-style anonymous server with distributed trust and public-key cryptography [8]. While the security provided by this system was much stronger than anon.penet.fi, it did not provide users with sufficient safeguards to defeat a determined and technically sophisticated attacker.

The cypherpunk remailer system is more complicated in that it is a network, not just a single server. In order for the security of the entire network to be compromised, every node on the network must be compromised. How does it work? A user desiring anonymity sends a message to one of the remailers in the network. The message includes instructions to the remailer, requesting it to send the payload to another remailer in the same network. The payload is then shipped to the next remailer, removing in the process all identifying information on the originating user. The next remailer in the network follows the same instructions from the previous hop and forwards the payload along. When the last remailer in the network is reached, the message is sent to the intended recipient, either to a specific e-mail account or to a Usenet newsgroup.

Because the system uses public-key cryptography at each 'hop' along the network in the process, a remailer can only read its own commands. The message itself cannot be read. Here's an example. Teresa sends a message to Victor, chaining her message through remailers W, X, Y, and Z. In the sequence of events, remailer W would only know to send some arbitrary encrypted message to remailer X. Remailer X can decrypt the instruction to send the message to its next hop, remailer Y. Remailer X has no idea what the final destination of the message will be. Eventually, when remailer Z follows the instructions to deliver the message to Victor's mail box, it has no idea that Teresa originally sent it. In order for the source of a message through the network to be discovered, the entire chain must be compromised. Sophisticated traffic analysis might reveal some details however and I will describe those techniques shortly.

Could this system provide security against non-technical but legal demands? First, if all remailers in a given chain were located in different jurisdictions (countries or states), it would be a significant hurdle for any one agency to obtain cooperation in all jurisdictions and force some information to be revealed. More important, the remailer network provides distributed trust, that is the system is not dependent on just one remailer. Even if all of the operators of all the remailers in the network were forced to cooperate by legal injunctions, as long as they did not cooperate with each other, messages would remain anonymous.

Eventually, the ability to reply through a remailer chain was implemented. A sender wishing to anonymously receive mail from a recipient could create what is now called a reply block [9]. The reply block is merely a multi-encrypted remailer message itself. The final destination of the message is the actual identity of the anonymous entity. Someone wishing to preserve their identity in this way can create such a reply block and publish it, or, include it in their e-mail. People wishing to send messages to this person use the reply block to send messages through the remailer network. The messages reach their destination anonymously.

Functionally, the reply block is very cumbersome. It requires that those corresponding with anonymous users take special steps to include the reply blocks in their electronic replies. In order to make such reply blocks easier to use, Matt Ghio wrote the alpha.c2.org nymserver, running at Community ConneXion [10].

The alpha.c2.org nymserver provides the full functionality of the anon.penet.fi server, while protecting privacy to a much greater degree. By not storing the actual identities of its users, the alpha.c2.org server is safe against even legal recourses.

In order to create an identity, our hypothetical user, Teresa, merely creates a reply block that points right back to herself. The reply block is then sent in encrypted form through a remailer chain, to the administrative address for the alpha.c2.org nymserver. The server software takes the anonymous mail message, and allocates an anonymous identity according to the instructions in the actual message. The reply block is then stored on the server. When mail for the anonymous identity arrives, it is forwarded through the remailer network, according to the instructions in the reply block.

The system is secure against legal compromises because the nymserver itself has no knowledge of the identity of the anonymous user. The nymserver is only aware of the existence of the encrypted reply block. It is encrypted with the key of a foreign remailer, which is not under the control of the nymserver. Given a court order or other legal directive to reveal the identity of a given pseudonym, the server can only reveal information regarding the first remailer in the chain and the block encrypted for the next remailer. In order to reveal the identity of the user, properly authorized officials would have to track every remailer in the chain. If the remailers are multi-jurisdictional, the process becomes very difficult.

Watching the traffic go by

Sophisticated traffic analysis could trivially defeat the protection provided by the remailer network in its current configuration. An agent could monitor the network itself, examining and recording message sizes, message timing, and message contents. Sophisticated traffic analysis would require an analysis of all of the network traffic in and out of all the remailers. For some agencies and organizations, with sufficient manpower and computing resources, this sort of approach is highly possible.

How would this sort of traffic monitoring work? What kind of information would be secured? First, an agent would monitor message times. Simply, the time at which a message enters and leaves a network would be recorded. Second, an agent would analyze message sizes. In a standard remailer chain, message sizes decrease by a predictable amount at each hop along the chain. Finally, an agent could determine the actual identity of a reply block by using a spamming the network. Spamming - sending a large number of unwanted messages - in this case would use a given reply block. In combination with analyses of message times and sizes, this approach could provide much useful information to an agent.

Timers

In order to correlate messages to time, an agent will study traffic going into and out of a specific group of remailers. This agent will monitor which messages enter and leave the remailers at approximately the same time. For example, the agent would note that a message from teresa@someisp.n et was sent to remailerone@abcdefg.net at 04:50. It would record that another message was sent from remailerone@abcdefg.net to finalhop@remailer.net at 04:52. It would also register that a third message was seen sent from finalhop@remailer.net to victor@xyz.gov at 04:53. Based on this cursory evidence, the agent could reasonably assume that teresa@someisp.net sent a message to victor@xyz.gov

An additional tactic for an agent would be to examine message sizes in the course of routine traffic analysis. Suppose the message between teresa@someisp.net and remailerone@abcdefg.net was one kilobyte long. The message was reduced between remailerone@abcdefg.net to finalhop@remailer.net to 800 bytes. On its final digital diet, the message between finalhop@remailer.net and victor@xyz.gov fell to 600 bytes. The agent could conclude that teresa@someisp.net sent a message to victor@xyz.gov Combining both of these methodologies would provide a high degree of confidence to the agent that the initial analysis of traffic was correct.

Spamming networks for identities

In addition to using the previously described strategies, an agent could also reveal the identity of a alpha.c2.org-style pseudonym through a simple spam attack. By sending a large quantity of messages through the network to a given pseudonym, the agent would once again monitor the network, this time tracing the spam "lump" through the system. The information gleaned by the agent in the course of the processing of this "lump" by the network could be used to trace a given recipient.

Mixmaster

As these scenarios demonstrate, Hughes and Finney's remailer network is not sufficient robust to prevent a sophisticated and dedicated attack. In response to these possibilities, Lance Cottrell developed the Mixmaster remailer system [11]. Mixmaster uses reordering and message padding to protect against traffic analysis [12].

Mixmaster was created to solve traffic analysis problems in older remailer approaches to anonymity. It protects users against higher level threats from organizations with strong eavesdropping capabilities. Mixmaster's philosophy assumes that all network links are compromised and that even some specific nodes in the chain are compromised.

While both Mixmaster and the Hughes/Finney approach assume that some nodes in the network may be compromised, Mixmaster goes one step further by assuming that every network connection is being monitored. In order to protect against those with the computing resources to monitor all network traffic, Mixmaster creates specific mechanisms to overcome agents studying traffic patterns. These mechanisms include reordering and message padding.

One means by which an agent, with access to all network traffic, could violate the remailer network is by correlating message times. For example, suppose Teresa sends a message at 15:45 to remailer A. This remailer in turn sends another message at time 15:46 to remailer B, which then sends a message to Victor at 15:47. The agent does not need to spend any time looking at individual nodes or encryption schemes, thanks to logs of the message times. The agent knows that Teresa and Victor are in communication.

Mixmaster protects against this approach by reordering messages flowing through the remailer. It simply does not deliver messages immediately. Mixmaster waits until a large number of messages are stored in the remailer. It then randomly mixes the order of all the messages, sending all of the messages out in one large and random batch. This processing defeats a basic message-timing analysis of traffic.

A second compromising approach would analyze message sizes. In the Hughes/Finney remailer, message sizes decrease by a certain amount at every hop. Suppose Teresa sends a message measuring 513 bytes to Remailer A. This remailer then sends a message with 490 bytes to remailer B. Remailer B then packages a message of 467 bytes to Victor. Again, an agent would be able to conclude that Teresa and Victor are sending messages to each other, violating their privacy.

Mixmaster protects against this approach by making every message exactly the same size. A Mixmaster message has a message format which includes 20 header fields, each the same size, and a body field, which is a predetermined size. Any messages sent through the network are padded to the size of the body field. Every message is the exact same size.

Another approach involves tracking the actual message bodies. This method is possible in servers that preserve a message body as it passes from one remailer to another. It is possible because some servers alter just the reply block and do not disguise the message in any way. An agent would simply trace message bodies through the remailer network in order to detect users. Mixmaster re-encrypts the body of the message with a new key at every hop, confusing an agent by changing the contents of the message body.

Remailers for everyone

It may seem, from these descriptions, that the actual, everyday use of remailers is difficult. This is hardly the case. Remailers are actually very easy to use thanks to a number of client front-ends available for several computing platforms. For example, with UNIX premail is just an example of one client program. For Microsoft Windows, there are several options, including Private Idaho and John Doe [13].

Premail is an excellent program written by Raph Levien for use with remailers [14]. It fully supports Finney/Hughes-style remailers, alpha.c2.org, and Mixmaster. In addition to being a remailer client, premail supports PGP and S/MIME for standard secured Internet e-mail without any identity and traffic analysis protections. Premail works with mh, elm, and newer versions of Netscape Navigator for UNIX.

Conclusion

As the remailer network grows, a number of new developments will be needed in the near future if anonymity will survive. First, it requires real standards, with the actual support of the Internet Engineering Task Force (IETF). Only by the actual labor of an IETF Working Group establishing an official standard will the remailer network gain the respectability it needs for ubiquitous deployment.

Second, a secure anonymous payment system must be incorporated into the network. In order for privacy to survive further commercialization of the Internet, it must be profitable. In order to make remailing profitable it must be offered as a fee-based service.

Some work is being done towards making privacy on the Internet economically possible. For example, there is one Web-based remailer interface that now accepts electronic cash in payment for its services [15]. The very nature of communications on the Internet will depend on the successful deployment of a number of schemes to preserve anonymity, in messages, purchases, and file transfers. Research by a number of groups on revised remailer designs, and the availability of free and low-cost options, clearly indicate the importance of anonymity to the Internet community as a whole. Only organized and standardized efforts will ensure continued privacy in the future.

The Author

Sameer Parekh
sameer@c2.net,
snail mail: Community ConneXion
3038A Mabel Street, Berkeley, Calif. 94702 U. S. A.
A recent interview with the author can be found at http://www-e1c.gnn.com/gnn/wr/dec22/reviews/bio/

Notes

[1] For general information on the development of Usenet, see, for example, Jenny A. Fristrup, USENET: Netnews for everyone. Englewood Cliffs, N. J.: Prentice Hall PTR, 1994, pp. 10-21, and Noel Estabrook, Kate Gregory, Jim Mann, and Tim Parker, Using UseNet newsgroups. Indianapolis, Ind.: Que, 1995, pp. 12-31.

[2] For an overview of remailers, see Andre Bacard, "Anonymous Remailer FAQ," at http://www.well.com/user/abacard/remail.html

Statistics on the server according to Arnoud "Galactus" Engelfriet at http://www.stack.urc.tue.nl/~galactus/remailers/index-penet.html

[3] As reported by Daniel Akst, "The Helsinki incident and the right to anonymity," Los Angeles Times (Feb. 22, 1995).

[4] Details on this incident can be found at in a posting dated Feb. 19, 1995 to the alt.privacy group at http://www.tezcat.com/~wednsday/penet.michigass gass and also http://www.tezcat.com/~wednsday/penet.pr

This incident led to several reports in the traditional media, such as in Time magazine (March 6, 1995) and at http://pathfinder.com/@@hTX*lQUAveSNGIPN/time/magazine/domestic/1995/950306/950306.technology.html and to several anti-anonymity editorials and stories in American newspapers. A summary of these reports can be found at Avi Baumstein's summary at http://www.clas.ufl.edu:80/~avi/NII/anonymity.html

[5] A brief description of distributed trust can be found at Douglas Armati, "Tools and standards for protection, control and presentation of data," at http://www.lmcp.jussieu.fr/icsu/Information/Proc_0296/armati.html

[6] For a history of cypherpunks, see Timothy C. May, "Crypto Anarchy and Virtual Communities," at http://www.c2.org/~arkuat/consent/Anarchy.html and posted at numerous other locations.

[7] See "Anonymity, digital Mixes, and remailers: Dining cryptographers," at http://www.oberlin.edu/%7Ebrchkind/cyphernomicon/chapter8/8.11.html

[8] For a general description of public-key cryptography, see Bruce Schneier, Applied cryptography: Protocols, algorithms, and source code in C. 2nd ed. New York: Wiley, 1996, pp. 31-34.

[9] See Arnoud "Galactus" Engelfriet, "Creating encrypted reply blocks," http://www.stack.urc.tue.nl/~galactus/remailers/reply-blocks.html

[10] See Andre Bacard, "ALPHA.C2.ORG Remailer FAQ," http://www.well.com/user/abacard/alpha.html

[11] Lance Cottrell's home page can be found at http://www.obscura.com/~loki/

[12] For general information on Mixmaster, see Lance Cottrell, "Frequently asked questions about Mixmaster remailers," http://www.obscura.com/~loki/remailer/mixmaster-faq.html

[13] Private Idaho is available at www.eskimo.com/~joelm/ and John Doe can be found at http://www.compulink.co.uk/~net-services/jd.htm

[14] On premail, see http://www.c2.org/~raph/premail.html

[15] This interface can be found at http://www.c2.net/remail/by-www.html

Prospects for remailers by Sameer Parekh.
First Monday, volume 1, number 2 (August 1996),
URL: http://www.firstmonday.org/?journal=fm&page=article&op=view&path[]=476