UNTRACEABLE ELECTRONIC MAIL,

RETURN ADDRESSES, AND DIGITAL PSEUDONYMS

by

D. L. Chaum

Memorandum No. UCB/ERL M79/9

22 February 1979

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Untraceable Electronic Mail,

Return Addresses, and Digital Pseudonyms*

D. L. Chaum

ABSTRACT

A cryptographic technique is presented that allows an electronic mail system to hide who a participant communicates with, as well as the content of communications--in spite of an un-secured underlying telecommunication system. The technique does not require a universally trusted authority. One correspondent can remain anonymous to a second, while allowing the second to respond via an untraceable return address.

The technique can also be used to form rosters of untraceable digital pseudonyms from selected applications. Applicants retain the exclusive ability to make digital signatures corresponding to their pseudonyms. Elections in which any interested party can verify that the ballots are properly counted are possible if anonymously-mailed ballots are signed with pseudonyms from a roster of registered voters. Another use allows an individual to correspond with a record keeping organization under a unique pseudonym that appears in a roster of acceptable clients.

Key Words and Phrases: electronic mail, public key cryptosystems, digital signatures, traffic analysisss, security

CR Categories: 2.12, 3.81

Introduction

Cryptology is the science of secret communication. Cryptographic techniques have been providing secrecy of message content for thousands of years [Kahn 67]. Recently, some new solutions to the "key distribution problem" (the problem of providing each communicant with a secret key) have been suggested [Diffie and Hellman 76, and Merkle 78], under the name of public key cryptography. Another cryptologic problem, "the traffic analysis problem" (the problem of keeping confidential who converses with whom, and when they converse), will become

* This work was partially supported by the National Science Foundation under NSF Grant MCS75-23739.
Author's address: Computer Science Division, Electrical engineering and Computer sciences Department, University of California, Berkeley, CA 94720. (415) 642-1024.

increasingly important with the growth of electronic mail. This paper presents a solution to the traffic analysis problem that is based on public key cryptography. The military has solved the traffic analysis problem in their electronic communications networks [Baran 64], but their approach requires each participant to trust a common authority. In contrast, systems based on the solution advanced here can be compromised only by subversion or conspiracy of all of a set of authorities. Each participant is an authority in the limiting case.

Notation

Someone becomes a participant in a public key cryptosystem (like that of [Rivest, Shamir, and Adleman 78]) by creating a pair of keys, K and K^{-1}, from a randomly generated seed. The public key K is made known to the other participants, or anyone else who cares to know it; the private key K^{-1} is never divulged. The encryption of X with key K will be denoted $K(X)$, and is just the image of X under the mapping implemented by the cryptogrphic algorithm using key K. The increased utility of these algorithms over conventional algorithms results from the two keys being inverses of each other, in the sense that

$$K^{-1}(K(X)) = K(K^{-1}(X)) = X. \qquad (1)$$

A message X is *sealed* with a public key K so that only the holder of the private key K^{-1} can discover its content. If X is simply encrypted with K, then anyone can verify a guess that $Y = X$ by checking whether $K(Y) = K(X)$. This threat can be eliminated by attaching a key-sized string of random bits R to X before encrypting. The sealing of X with K, then, is denoted $K(R,X)$. A participant *signs* some material X by prepending a key-sized constant C (all zeros, e.g.) and then encrypting with its private key, denoted $K^{-1}(C,X) = Y$. Anyone can verify that Y has been signed by checking that $K(Y) = C,X$.

Mail System

The members of the cryptosystem will include not only those who wish to correspond, but computers called *mixes* which will perform the actual shuffling of correspondences en route. A participant prepares a message M for delivery to a participant at address A by sealing it with the addressee's public key K_a, appending the address, and then sealing the result with the mix's public key K_1. The left hand side of the following expression denotes this input:

$$K_1(R_1,K_a(R_a,M),A) \rightarrow K_a(R_a,M),A. \tag{2}$$

The → denotes the transformation of the input by the mix into the output shown on the right hand side. The mix decrypts its input with its private key, throws away the random information R_1, and outputs the remainder. The purpose of a mix is to hide the correspondence between the items it receives as input, and the items it outputs. The correspondence is not revealed by the values of the uniformly-sized items, because the input items are sealed, whereas the output items are not. The order of arrival of items is hidden by outputing items in lexicographically-ordered batches. One might imagine a mechanism that forwards the sealed messages $K_a(R_a,M)$ of a batch to the addressees, who can then decrypt them with their private keys.

If a participant gets signed receipts for messages it submits to a mix, then it can provide substantial evidence that the mix failed to output an item properly. Only a wronged participant can supply the receipt Y $(=K_1^{-1}(C,K_1(R_1,K_a(R_a,M),A)))$, the missing output X $(=K_a(R_a,M),A)$, and the retained string R_1, such that $K_1(Y) = C,K_1(R_1,X)$. A mix signs each batch of output as a whole, so the presence or absence of an item in the output should not be the subject of dispute.

The use of multiple mixes offers the advantage that any single constituent mix is able to provide the secrecy of the correspondence between the inputs and the outputs of the entire cascade of mixes. Incrimination of a particular mix that failed to properly process an item also remains possible. An item is prepared for a cascade of n mixes the same as for a single mix; then it is successively sealed for each succeeding mix:

$$K_n(R_n,K_{n-1}(R_{n-1}, \ldots ,K_2(R_2,K_1(R_1,K_a(R_a,M),A))\ldots)) \rightarrow. \tag{3}$$

The first mix yields a lexicographically ordered batch of items, each of the form

$$K_{n-1}(R_{n-1}, \ldots ,K_2(R_2,K_1(R_1,K_a(R_a,M),A))\ldots) \rightarrow. \tag{4}$$

The final outputs of a cascade are of the form $K_a(R_a,M),A$, the same as those of a single mix.

Return Addresses

We have seen how participant **x** can send messages to participant **y**--without **y** knowing who **x** is. Now we will provide a way for **y** to respond to **x**, while still keeping the identity of **x** secret from **y**. First, **x** will form an untraceable return address $R_a, K_1(R_1, A)$, where A is its own real address, and R_a and R_1 are random strings. Then, **x** will send this return address to **y** as part of a message sent by the techniques already described. (Ultimately, two participants can exchange return addresses through a chain of intermediaries, where at least one member of each adjacent pair knows the identity of the other member of the pair.) The following indicates how **y** uses this untraceable return address to form a response to send to **x**, via a new kind of mix:

$$K_1(R_1,A), R_a(M) \rightarrow A, R_1(R_a(M)). \qquad (5)$$

This mix uses the string of bits R_1, that it finds after decrypting the address part, as a key to re-encrypt the message part. (Conventional as opposed to public key cryptography could be used for both encryptions of *M*.) Only the addressee can decrypt the resulting output, because it created both R_1 and R_a.

With multiple mixes, the message part is prepared as for a single mix, and the input is of the form

$$K_1(R_1, K_2(R_2, \ldots, K_{n-1}(R_{n-1}, K_n(R_n, A))...)), K_1(R_a(M)) \rightarrow. \qquad (6)$$

The result of the first mix is

$$K_2(R_2, \ldots, K_{n-1}(R_{n-1}, K_n(R_n, A))...), R_1(R_a(M)) \rightarrow, \qquad (7)$$

and the final result of the remaining $n-1$ mixes is

$$A, R_n(...R_2(R_1(R_a(M)))...). \qquad (8)$$

Any single mix can ensure the anonymity of the issuer of a return address, provided the address is used only once. The problem with schemes that reuse return addresses is that a repeated use will be evident to every mix, because they will see the same address part repeated. One way to stop multiple uses is for mixes to record address parts, and refuse to pass any repeats. Alternatively, mixes could change their public keys, signing a new one with the previous private key, before each batch; so that a return address would only be valid for a particular batch. A completely different approach is to insist that each address be used once in each batch. In this approach, mixes seal message parts (instead of simply encrypting them), and create dummy messages with address parts that are absent from the input. Filling unused channel capacities with null messages, however, is also necessary to eliminate some potentials for analysis.

Other Applications

Mixes can be used to form *rosters* of untraceable digital *pseudonyms.* A pseudonym is a public key used to verify signatures made by the anonymous holder of the corresponding private key. A roster of pseudonyms is created by an authority who can decide which applications to accept, but who will be unable to trace the pseudonyms in the completed roster.

Each application contains both information the authority requires for the acceptance decision, and an un-mailed digital letter. The letter is addressed to the authority and contains the applicant's proposed pseudonym as the message. Letters from accepted applications are mailed by the authority. The mail system will process these letters as a single batch. It will be possible to use the output of the final mix as the roster because the pseudonyms it contains will not have been sealed with K_a by the applicants.

Applicants could be protected, or at least supplied with incriminating evidence, if the application and the un-mailed letter it contains are both *registered.* The address (A) that is incorporated in a registered letter contains the usual address of the recipient along with an untraceable return address for the sender. The final mix uses this return address to mail the sender a receipt which includes the message, the address to which it was delivered, and an indication of the output batch in which it appeared.

If only registered voters are accepted for a particular roster, then it can be used to carry out an election. Voters sign their ballots with the private inverses of their pseudonyms, before preparing the ballots for the same sort of special handling received by the letters originally containing the pseudonyms. Anyone can count the votes in the output batch of the final mix, and make sure that no one has voted twice, by using the public keys in the roster to check the ballots' signatures.

An individual might be known to an organization by a pseudonym that appears in a roster of acceptable clients. The organization would be assured that the same client could not come to it under different names. Even though clients are only known to an organization by their pseudonyms, the organization can correspond with the clients via untraceable return addresses.

Conclusion

We have seen a decentralized solution to the traffic analysis problem; it is well suited to electronic mail and other applications in which universally trusted authorities do not exist, and in which anonymity is important.

Acknowledgement

Considerable thanks are due Bob Fabry for support, and for reading drafts of a paper from which this paper has been extracted.

References

Baran, P., On distributed communications: IX. security secrecy and tamper-free considerations. Memo RRM-3765-PR, Rand Corp., Santa Monica, Ca., August 1964.

Diffie, W., and Hellman, M.E., New directions in cryptography. *IEEE Trans. Inform. Theory IT-22,* 6 (Nov. 1976), 644-654.

Kahn, D., *The Code Breakers, The Story of Secret Writing.* Macmillan Co., New York, 1967.

Merkle, R.C., Secure communications over insecure channels. *Comm. ACM 21,* 4 (Apr. 1978), 294-299.

Rivest, R.L., Shamir, A., and Adleman, L., A method for obtaining digital signatures and public key cryptosystems. *Comm. ACM 21,* 2 (Feb. 1977), 120-126.