Computers & Security, 7 (1988) 483–488

Refereed Article

Solutions to the Multidestination Secure Electronic Mail Problem

Chris Mitchell[1] and Michael Walker[2]

[1] Hewlett–Packard Ltd., Bristol, U.K.
[2] Racal Research Ltd., Reading, U.K.

Providing security for electronic mail messages sent to more than one destination can be a difficult problem, particularly when authentication is required. Previous attempts to solve this problem have been shown to be flawed. In this paper we describe two approaches which can be used to solve the problem in an efficient and secure way.

Keywords: Cryptography, Data security, Electronic mail, Key management.

1. Security and Multidestination Mail

In most if not all of today's electronic mail systems there exists the capability for sending a message to a list of users simultaneously. When such a message is sent, it will often only be replicated when it really needs to be, so that a single message sent from the U.K. to two recipients in the U.S.A. will only be made into two copies after it has crossed the Atlantic. Such a process is obviously desirable, not least because of the savings involved in transferring information using potentially heavily loaded communications links. However, a problem arises when a message containing sensitive information needs to be protected against disclosure or alteration while in transit.

In this note we focus on the problem of authenticating a multidestination mail item so as to protect it against alteration, even by one of the legitimate recipients. To authenticate a message using conventional symmetric cryptography, and we do not consider the use of asymmetric (public key) cryptography here, requires the use of a key known only to the message originator and recipient.

Thus one approach to the multidestination authentication problem is to append to a message separate authentication data computed using the key of each of the intended recipients. This idea is the basis of two distinct types of solution to the problem described in this paper. The aim of these solutions is to avoid some of the subtle weaknesses which can exist in apparently secure authentication schemes of this type, a topic we explore further in the next section, where we briefly discuss a previously proposed solution.

2. A Proposed Scheme and its Weakness

A recent *Request For Comments* for the U.S. Defense Advanced Research Projects Agency (DARPA) "Internet" electronic mail system [9], contains an apparently simple and elegant solution to the problem of authenticating multidestination mail. However, as we shall see, this solution has a flaw which, in certain circumstances, allows the malicious construction of apparently authentic messages. A more detailed discussion of the DARPA scheme can be found in a recent paper [10].

In the DARPA scheme, every pair of users, e.g. A and B, wishing to exchange secure mail are equipped with a secret *Interchange Key* (IK), which we denote by IK{A, B}. It is not important how this key is distributed, but we assume that each key is known only to the appropriate pair of users (and perhaps to a Key Distribution Centre, if one exists). All messages are to be authenticated using a block cipher in *Cipher Block Chaining* (CBC) mode (see, for example, the standard modes of use for the DES algorithm [3, 8]). The block cipher algorithm to be used is immaterial, but it could, for example, be the DES algorithm [7, 2]. A message from one user A to a pair of users B and C is protected in such a way that both recipients can

0167-4048/88/$3.50 © 1988, Elsevier Science Publishers Ltd.

authenticate the message origin and validate the message contents. To afford this protection A proceeds as follows.

(1) A random *Data Authenticating Key* (DAK) is obtained by A. This key is used to secure one and only one message.

(2) Two encrypted versions of the DAK are produced using the block cipher in Electronic Codebook Mode [3, 7], once under IK{A, B} and once under IK{A, C}.

(3) The message is encrypted using the block cipher in CBC mode under the DAK with *Initialization Vector* (IV) set to all zeros. All the ciphertext blocks except the last are discarded, with the remaining block forming the MAC which is used to authenticate the message to the receiver; this is a standard method for computing MACs [4].

(4) Two encrypted versions of the MAC are produced using the block cipher in ECB mode, once under IK{A, B} and once under IK{A, C}.

(5) The message is sent, preceded by the two encrypted versions of the DAK and the two encrypted versions of the MAC, to both B and C.

With this scheme, each of B and C can use their own IK to recover the DAK used to authenticate the message, compute the MAC for the message, encrypt this under their own IK and confirm that the resulting value corresponds to the received value. So far so good. However, we now show how user C can subvert the scheme and send a new message to B which B will believe to have come from A.

First C creates the message which is to be sent to B as if from A. Then the following procedures are executed.

(1) C recovers the MAC for the original message (this can be done since it was sent encrypted under IK{A, C} which is known to C).

(2) In a similar way C recovers the DAK used to authenticate the original message.

(3) C decrypts the MAC using the DAK to obtain a block we call x.

(4) C encrypts the new message using CBC under the control of the DAK to obtain a new MAC we call y.

(5) C joins the block $x + y$ onto the end of the new message as an additional "garbage" block, where the + denotes exclusive—or of blocks.

(6) C sends to B the new message (augmented by $x + y$) preceded by the DAK encrypted under IK{A, B} and the MAC encrypted under IK{A, B}, both of which are taken from the original message.

A message prepared using steps (1)–(6) above will pass B's authentication check and will therefore be accepted as coming from A. The reason for this is straightforward: the CBC encryption of the new message using the DAK will produce final block y, which, when added to the "garbage" block $x + y$, will give block x. This is the MAC for the old message and so it encrypts to the value expected by B. The one thing suspicious in the message is the "garbage" block, although even this could be moved to the middle or even the beginning of a message using an extension of the procedure described above.

We have therefore discovered a major flaw in the suggested scheme for authenticating mail to a multiplicity of users. Although it would be desirable to try and "repair" the above scheme to preclude the type of fraud described, it is by no means obvious how to do this. The basic problem is that user C knows both the MAC and the key used to generate the MAC for the original message and it is this, in combination with the fact that the CBC function can be inverted, which makes the fraud possible. In the following two sections we examine two possible approaches to resolving the problem. The first method involves attempting to stop the inversion procedure by using a different function to compute the MAC. In the second method the original function is retained, but a different MAC is computed for each possible recipient. Interestingly, the basic idea for solving the multidestination secure mail problem has been independently invented at least twice, although it does not appear in any literature known to the authors.

Before proceeding, we should note that the general type of attack described in the last section is well known, albeit in a slightly different context. For example, Akl [1], has described weakenesses in a whole range of methods for computing what he calls "compressed encodings", (of which MACs are an example), mostly within the context of digital signatures. Some of the attacks which Akl [1] describes are based on the possibility of replacing both the message and the encoding. This is ruled out in the DARPA scheme by the encryption of the MAC under the IK.

3. Single Encoding Based Solutions

3.1 A General Scheme

To describe our first solution requires looking at the problem from a more general viewpoint. The basic idea behind the flawed scheme we have just described and the method we describe here is as follows. The sender of a message computes a *digest* of the message using a predetermined *hash function*. The digest is then encrypted using each of the recipient keys in turn (suppose there are *r* of them) and all of the *r* encrypted digests are sent with the message. These recipient keys are session keys, which may already be known by both parties or may be randomly generated for each message. In the latter case the session key would need to be encrypted under an existing key encrypting key shared by the two parties; the message would then be augmented by the addition of *r* encrypted digests and *r* encrypted session keys.

The effectiveness of the scheme depends upon the hash function, which must satisfy a number of properties, and the encryption of the digests. We consider first the hash function. Suppose this is h, i.e. h maps a message M onto h(M), where the digest h(M) is of a fixed (small) length, e.g. 128 bits. The first property we require of h is that h(M) depends upon all of M. This is clearly essential, for otherwise an adversary could merely change those parts of a message which do not affect the digest. The second requirement is that h should exhibit the following "one way" property.

(H1) Given any possible candidate for a digest, e.g. C, then it should be computationally infeasible to find a message M such that h(M) = C.

The reason that we require this property is to prevent attacks of the type which flawed the DARPA scheme. Suppose user A sends an authenticated message M to users B and C. Then user C can decrypt his copy of the encrypted digest to obtain h(M). If H1 does not hold then C may be able to construct a new message M′ such that h(M′) = h(M). User C then sends this message together with the encrypted MAC from the old message to B, and B will accept this as a genuine message from A.

In the DARPA scheme, the hash function is made up from DES used in CBC mode with a key that is secret to all but the intended recipients of the message. Unfortunately, this means H1 does not hold for any of the message recipients and hence the scheme is weak. In fact using DES in CBC mode has an additional feature we do not require here, namely the use of a secret key. Given that h is well chosen then it can be made public and the same function used universally.

In general, H1 is not sufficient to guarantee a secure scheme. In certain circumstances, for example if h(M) is not long enough (i.e. there are not sufficiently many different digests), a "birthday paradox" attack may make it feasible to find two messages, M and M′, such that h(M) = h(M′). As has been discussed by many authors, e.g. Akl [1] and Davies and Price [5], to obtian a high probability of obtaining such a pair of messages requires $O(n)^{\frac{1}{2}}$ steps, where h(M) contains *n* bits, regardless of the choice of h. Moreover, it is possible to construct the pair in such a way that the contents of the pair of messages can be pre-selected. Now suppose that user C has constructed two such messages, M and M′, with the property that A will be happy to send M to user B. If A authenticates M by computing h(M) and encrypting it with a key known to B, then C can intercept the message and replace M by M′ so that B will accept M′ as an authentic message from A (since h(M) = h(M′)).

There are two possible approaches to the prevention of this type of attack; one involves making $(n)^{\frac{1}{2}}$ sufficiently large and the other involves always using the hash function with great care. Akl [1] recommends the former approach and suggests that digests should always contain at least 128 bits (64 bits being insufficient). Davies and Price [5] favour the second approach, whereby no users ever authenticate a message just as it is presented to them. One possible method to achieve this is to insist that every message is always prefixed with a random value before computing the digest, where this value is always selected by the authenticating party. In either case, h must be chosen and used so as to make the replacement of one message by another impossible without changing the digest.

Having discussed the hash function, we now consider the role of the encryption of the digest. This encryption is present solely to prevent substitution of one digest for another. Given that the key used to perform the encryption is known only to the sender and the receiver of a message (and given a "good" encryption function), then the malicious interceptor will not be able to substitute one digest for another except in a "random" way.

3.2 Realizations of the Scheme

Having discussed the construction of a secure scheme in a general way, we now consider techniques which might be used to implement it, using available technology.

Considerable research has already been done in this area, mainly with reference to hashing for *digital signatures* and this is the emphasis in both Akl's paper [1] and Davies and Price's book [5]. It turns out that the requirements for a hash function for digital signature and for this multidestination authentication problem are identical (which is not surprising considering the similarity of application).

Much work has gone into devising "good" hash functions using DES (or any other block cipher), since the technology is readily available. However, the history of work in this area is littered with failures (see Akl [1] and Winternitz [11]). One scheme, however, appears to be a good candidate.

This scheme, described in a 1983 paper of Winternitz [11] and attributed to Davies, is also described in a 1985 paper of Davies and Price [6] and attributed there to Meyer; a brief discussion of this type of hash function can also be found in Section 9.3 of Davies and Price's book [5]. For convenience we refer to the method here as the Davies– Meyer–Winternitz (DMW) scheme. In this scheme the message M is first divided into a sequence of 56-bit blocks: $M_1, M_2, \ldots, M_n$. Next, using a fixed value for C(0), iteratively compute

$$C(i) = C(i-1) + E_{M(i)}\{C(i-1)\}$$

where $E_{M(i)}\{C(i-1)\}$ means the result of encrypting block $C(i-1)$ using key $M(i)$ with the DES encryption algorithm, and the + sign denotes the exclusive-or of 64-bit blocks. The digest of the message is then simply $C(n)$. This scheme is particularly valuable, since, under some reasonable assumptions about the nature of DES, Winternitz [12] has proved that it satisfies H1.

However, the DMW system produces digests of only 64 bits and it is by no means obvious how to obtain digests containing 128 bits, except by replacing DES with a 128-bit block cipher. For example using $C(n-1)$ concatenated with $C(n)$ as the digest is not an improvement, since any message having final block $M(n)$ and penultimate digest $C(n-1)$ will also always have final digest value $C(n)$.

Following on from our earlier theoretical discussion, there are two possible solutions. First, we could use another hash function which produces 128-bit digests, such as one of the schemes suggested by Akl [1]. However, these schemes lack the desirable feature of proof of security which the DMW system has. Secondly, we could use the DMW scheme with the added proviso that messages must always be prefixed with a truly random value before hashing. This latter scheme, although perhaps a little complex to implement, seems a very good candidate for the hashing function.

We have yet to consider the type of encryption operation to use to encrypt the digest of a message. This is a relatively simple problem and can easily be solved using DES (or other good block cipher). If the digest is only 64 bits then a single code book encryption will suffice. If the digest contains 128 or more bits, then the encryption can be performed using DES in CBC mode.

To summarize, we have shown how a secure multidestination authentication scheme can be produced using a secure hash function and a block cipher. The DMW scheme offers a reliable technique for message hashing using the DES algorithm, which means that the whole scheme can be implemented using a single encryption function together with a random noise source for message prefixing.

4. Multiple Encoding Solutions

An alternative approach derives from observing that, even for the scheme described in the last section, it is necessary to include within a secure message an encrypted version of the digest for each of the proposed recipients. It would therefore not alter the length of the transmitted message if a keyed function were to be used to compute the authentication code, with a different session key for each message recipient. In this case we assume that the authentication data added to the message will then contain *r* authentication codes (computed using the *r* different session keys) and in addition will possibly contain *r* encrypted session keys. In the sequel we shall assume that the encrypted session keys are to be added to the message.

We now describe this type of scheme in more detail. It should be noted that we discuss the scheme in a very specific way, referring to DES and CBC throughout. However, it would be possible to replace DES with any other "good" block cipher and CBC is by no means the only feasible mode of use. The scheme is therefore more general then it might appear.

Suppose that user A wishes to send message M to a list of *r* other users using the authentication facility. Under this second scheme, A generates *r* different random session keys (one for each recipient) and using each of the *r* keys in turn generates a MAC for the message using DES in CBC mode. Each key is then encrypted under the corresponding recipient's key using DES in codebook (ECB) mode. When the message is sent, appended to it are the *r* pairs of MACs and encrypted keys, one pair for each recipient.

Given that DES is secure, it is now impossible to attack this system, since only the correct recipient of a message will know the key used to generate the MAC from the message. It will therefore be impossible for anyone to generate a message with the same MAC, even other authorised recipients of the message.

This system has the virtue of simplicity when compared with the previously described method. However, to send a message to *r* recipients requires the sender to process the entire message *r* times and hence for a long message sent to a large number of recipients it could be very time consuming to compute all the relevant MACs. This possible handicap needs to be weighed against the complexities of the previous solution.

5. Conclusion

We have described two types of solution to the problem of authenticating electronic mail messages sent to a multiplicity of recipients. Both solutions have advantages; the first solution involves less processing time, while the second solution is perhaps simpler. We have also described ways in which both schemes can be implemented using established cryptographic technology. The choice of which type of scheme to use is with the implementor, although the first solution would probably be favoured if hashing functions are required elsewhere in the system.

Acknowledgments

This paper has been prepared under the auspices of the LOCATOR project, a part of the Alvey Mobile Information Systems Large Demonstrator Project. The authors would also like to acknowledge the helpful comments of many of their colleagues in this project.

References

[1] S.G. Akl, On the security of compressed encodings, *Advances in Cryptology: Proc. Crypto 83,* Plenum, New York, 1984, pp. 209–230.

[2] *ANSI X3.92–1981,* Data encryption algorithm, American National Standards Institute, New York, 1981.

[3] *ANSI X3.106–1983,* Modes of operation of the DEA, American National Standards Institute, New York, 1983.

[4] *ANSI X9.9–1982,* Financial institution message authentication, American Bankers Association, Washington, DC, April 1982.

[5] D.W. Davies and W.L. Price, *Security for Computer Networks,* Wiley, Chichester, 1984.

[6] D.W. Davies and W.L. Price, Digital signatures—an update. In J.M. Bennett and T. Pearcey (eds.), *The New World of the Information Society,* Elsevier, Amsterdam, 1985, pp. 843–847.

[7] *FIPS 46,* Data encryption standard, Federal Information Processing Standard, National Bureau of Standards, Washington, DC, January 1977.

[8] *FIPS 81,* DES modes of operation, Federal Information Processing Standard, National Bureau of Standards, Washington, DC, December 1980.

[9] J. Linn, Privacy enhancement for Internet Electronic Mail: part I: message encipherment and authentication procedures, *Request for Comments 989 (RFC 989), IAB Internet Privacy Task Force,* February 1987.

[10] C. Mitchell, Multi-destination secure electronic mail, *Comput. J.,* to appear.

[11] R.S. Winternitz, Producing a one-way hash function from DES, *Advances in Cryptology: Proceedings of Crypto 83,* Plenum, New York, 1984, pp. 203–207.

[12] R.S. Winternitz, A secure one-way hash function built from DES, *Proc. 1984 IEEE Symp. on Security and Privacy, Oakland, April–May 1984,* IEEE, 1984, pp. 88–90.

Chris Mitchell has been a member of the technical staff at the Bristol Research Centre of Hewlett-Packard Laboratories since 1985 (Filton Road, Stoke Gifford, Bristol BS12 6QZ, U.K.), having previously been Chief Mathematician at Racal Comsec Ltd. He received his B.Sc. (1975) and Ph.D. (1979) degrees in mathematics from Westfield College, University of London. His research interests are in cryptology, data security and combinatorial mathematics. He is a member of the British Computer Society, the Institution of Electrical Engineers and the London Mathematical Society and a fellow of the Institute of Mathematics and its Applications.

Michael Walker is Head of Mathematics at Racal Research Ltd. (Worton Drive, Worton Grange Industrial Estate, Reading RG2 0SB, U.K.), where he is responsible for much of the company's work in cryptography, data security and coding for error control. Before joining Racal in 1983, he was a lecturer in mathematics at the University of Tübingen, where he undertook research in finite geometries, groups and combinatorics. He was educated at Royal Holloway College, University of London, where he received a B.Sc. in 1969 and a Ph.D. in 1973, both degrees in mathematics. He is a member of the London Mathematical Society and a fellow of the Institute of Mathematics and its Applications.