Cryptanalysis of the Gemmell and Naor Multiround Authentication Protocol*

Christian Gehrmann

Dept. of Information Theory, Lund University,
Box 118, S-221 00, Lund, Sweden

Abstract. Gemmell and Naor proposed a new protocol for the authentication of long messages which was based on block codes and which used a transmission channel k times. This multiround authentication makes it possible to limit the key size independently of the message length. We propose a new attack and show that the probability analysis made by Gemmell and Naor, which was only based on the minimum distance property of the codes, does not hold for our attack. Considering also the impersonation attack we conclude that the number of rounds have to be odd.

1 Introduction

The first treatment of codes that detect deception was given by Gilbert, MacWilliams and Sloane [1]. The use of universal hashing for authentication codes (A-codes) without secrecy, so called Cartesian A-codes, was first described in [2]. The general authentication problem was formulated in information theoretic terms by Simmons [3]. Many constructions and bounds have been derived for Cartesian A-codes [4] [5], [6], [7] and it is possible to construct such codes, which are close to the theoretical bounds. However all these constructions only deal with single transmission authentication. Gemmell and Naor [9] proposed a multiple round authentication protocol. Let n denote the message length, $H(K)$ the key entropy and P_s the probability for a successful substitution attack. For single round Cartesian authentication codes it was shown that [4]

$$H(K) \approx \log(n) + 2\log(\frac{1}{P_s}) - \log\log(\frac{1}{P_s}). \tag{1}$$

The Gemmell and Naor k-round protocol obtains:

$$H(K) \approx \log^{(k-1)}(n) + 5\log(\frac{1}{P_s}) \tag{2}$$

and Gemmell and Naor proved the existence of a k-round protocol such that

$$H(K) \approx \log^{(k-1)}(n) + 2\log(\frac{1}{P_s}). \tag{3}$$

* This work was supported by the TFR grant 222 92-662.

Y.G. Desmedt (Ed.): Advances in Cryptology - CRYPTO '94, LNCS 839, pp. 121-128, 1994.

Hence it would be possible to limit the key size independently of the message length by using multiround authentication. We will start with describing the essential properties of the protocol suggested by Gemmell and Naor. Section 3 and 4 consist of an analysis of the described scheme. We first bring to attention an impersonation attack when the number of rounds is even. Next we describe a substitution attack. Finally we give an example of the proposed substitution attack for a construction based on RS-codes.

2 The Gemmell and Naor protocol

We assume two participants A and B, who want to communicate over an insecure channel where an opponent O may introduce a new message m'_j (impersonation attack) or substitute message m_j sent by A or B for m'_j (substitution attack). Here m_j denotes the message sent by A or B in the j-th round. The first message m_0 is the information message and the rest of the messages only "check messages" in the protocol. Hence the general goal for the opponent is to send an own message m_0 or to substitute for a transmitted one. However in his attempts to succeed with this purpose he may manipulate the "check messages" as well. A and B share a secret information, i.e., the key, unknown to the opponent.

In the Gemmell and Naor protocol:

C^j:	the error correcting code used in the j-th round.
$C^j(m)$:	the codeword corresponding to message m when using the code C^j.
$C^j_i(m)$:	the i-th code symbol of the codeword $C^j(m)$.
C^A:	a Cartesian A-code.
$C^A(m)$:	the authentication tag corresponding to message m when using the code C^A.
$x \circ y$:	concatenation of string x and y.
p:	as defined in [9].

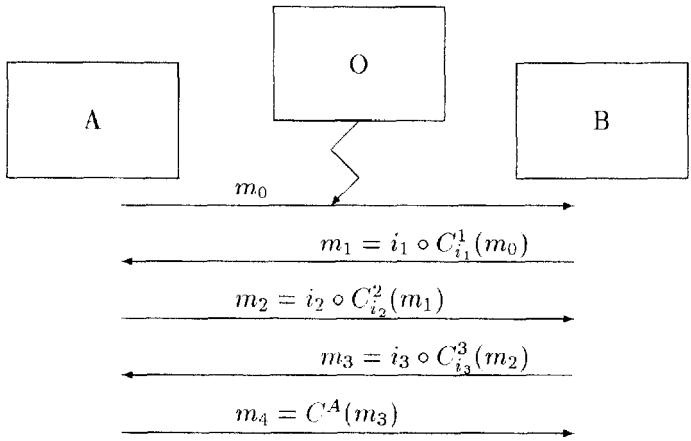


Figure 1: Multiround authentication for $k = 5$.

Assume we will use a k-round protocol and that: $C^j : \{0,1\}^n \rightarrow GF(Q_j)^{n_j}$ is a code with properties:

(i) $Q_j \geq \frac{2^{k-1-j}}{p}$

(ii) The minimum distance d_j of C^j satisfies: $d_j \geq n_j - n_j p/2^{k-2-j}$.

The slightly modified(see below) Gemmell-Naor protocol may be described as follows:

(i) A sends message $m_0 = m, j = 0$.

(ii) $j = j + 1$, B receives message m'_{j-1} and chooses a random number $i_j, 1 \leq i_j \leq n_j$. B sends message $m_j = i_j \circ C^j_{i_j}(m'_{j-1})$.

(iii) If $j = k - 2$ then step vi).

(iv) $j = j + 1$, A receives message m'_{j-1} and chooses a random number $i_j, 1 \leq i_j \leq n_j$. B sends message $m_j = i_j \circ C^j_{i_j}(m'_{j-1})$.

(v) if $j = k - 2$ then step (vi), else back to step ii).

(vi) If k is even(odd) A(B) receives message $m'_j = m'_{k-2}$ and use a Cartesian A-code with $P_s \leq p$, to transmit $m_{k-1} = C^A(m'_{k-2})$.

We have changed the last step (vi) from the protocol in [9], by letting A(B) just send the authentication tag for the last message m'_j. This is possible because B(A) already knows the message to be authenticated. It is also important to notice that it is not possible for the opponent to freely choose the substitution message in the last step and that that this decreases the restrictions on the A-code.

For the protocol described above Gemmell and Naor stated the following:

Proposition 1 (Gemmell-Naor). *Let p be the parameter as it appears in [9]. Then*

$$\forall k \geq 2, \quad P_s \leq 2(1 - \frac{1}{2^{k-1}})p, \tag{4}$$

where P_s is the probability of a successful substitution attack.

3 An impersonation attack

First we will bring to attention an impersonation attack for the case k is even. Consider first the two round protocol($k = 2$):

(i) O sends another message $m_0 = m$.

(ii) B receives message m_0 and uses a Cartesian A-code to transmit $m_1 = C^A(m_0)$.

(iii) O absorbs the message sent by B.

This case may easy be generalised to higher order even round protocol just letting O act like A.

Proposition 2. *When the number of rounds is even, for the impersonation attack above we have*

$$P_I = 1, \tag{5}$$

where P_I is the probability of a successful impersonation attack.

Proof. A never receives the last message, i.e., the authentication tag. Hence O never will be caught.

Remark: For k odd, O succeeds in an impersonation attack if and only if he successfully authenticates the last message sent by B. But, by the definition of the A-code the probability for this event $\leq P_s$.

In the sequel we will assume k to be odd and we deal only with the substitution attacks.

4 A substitution attack

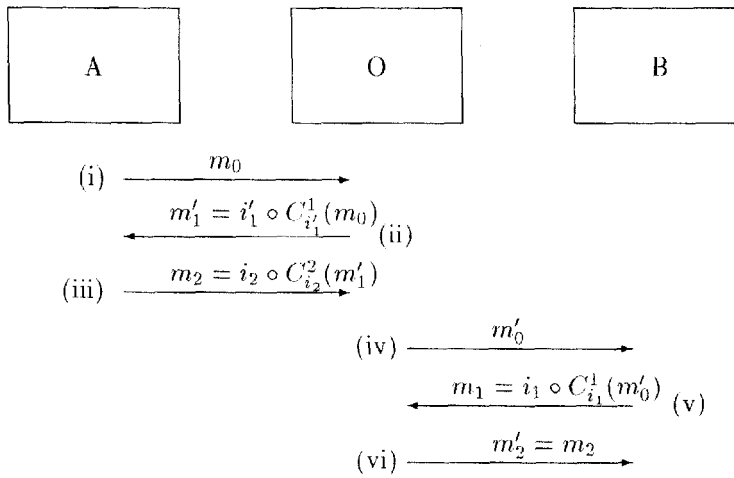


Figure 2: The attack scheme.

We will consider a specific substitution attack on the system above in the case $k \geq 5$. The attack is described by the scheme above.

(i) A sends an arbitrary message m_0 over the channel.

(ii) The opponent receives the message m_0 and chooses a random number $i_1', 1 \leq i_1' \leq n_1$ and sends $m_1' = i_1' \circ C_{i_1'}^1(m_0)$ to A.

(iii) A receives the message m_1' and chooses a random number $i_2, 1 \leq i_2 \leq n_2$ and sends $m_2 = i_2 \circ C_{i_2}^2(m_1')$ over the channel.

(iv) The opponent receives message m_2 from A and now substitutes the message m_0 for m_0' and sends this to B.

(v) B receives message m_0' and chooses a random number $i_1, 1 \leq i_1 \leq n_1$ and sends $m_1 = i_1 \circ C_{i_1}^1(m_0')$ over the channel.

(vi) The opponent receives the message m_1 from B and just absorbs it and sends message $m_2' = m_2$ to B.

Proposition 3: For the attack scheme above

$$P_s \geq \max_{m_0' \neq m_0} \frac{|\{i_1 : C_{i_2}^2(m_1) = C_{i_2}^2(m_1')\}|}{n_1}, \tag{6}$$

where

$$m_1 = i_1 \circ C_{i_1}^1(m_0'), \quad 0 \leq i_1 \leq n_1,$$
$$m_1' = i_1' \circ C_{i_1'}^1(m_0'), \quad 0 \leq i_1' \leq n_1,$$

(See Figure 2).

Proof. Clearly if for one particular i_1 that B might choose we have

$$C_{i_2}^2(m_1) = C_{i_2}^2(m_1'),$$

then

$$m_2' = i_2 \circ C_{i_2}^2(m_1) = i_2 \circ C_{i_2}^2(m_1') = m_2$$

and adding the fact that the opponent knows i_2 before sending m_0' and that B chooses i_1 at random the result follows.

This result differs from (4) given by Gemmell and Naor and hence when $k \geq 5$ this must be taken into account when constructing the codes in the protocol. We will now construct an example that will illustrate the consequences of this result.

5 Example

Before describing an example of the G-N protocol we recall some simple facts on Reed-Solomon codes (RS-codes) [8]. We use the polynomial description of RS-codes as it appeared in the original paper by Reed and Solomon. Denote by $GF(Q)$ the Galois field with Q element. Consider the polynomial $P(x)$ of degree at most $k-1$ over $GF(Q)$, i.e.,

$$\mathcal{P} = \{P(x); m_0 + m_1 x + \cdots + m_{k-1} x^{k-1}, m_i \in GF(Q)\} \tag{7}$$

Let $\alpha \in GF(Q)$ be a primitive root. The RS-code C over $GF(Q)$ is now obtained as the set of Q-tuples

$$C = \{(P(0), P(\alpha), P(\alpha^2), ..., P(1)); P(x) \in \mathcal{P}\}$$

It is a code with k information symbols (over $GF(Q)$) and blocklength Q. We see that each codeword can be regarded as the k-tuple $m = (m_0, m_1, \cdots, m_{k-1})$ onto the Q-tuple $(P(0), P(\alpha), P(\alpha^2), \cdots, P(1))$. Let $C(m)$ denote the image of

m, i.e. $C(m) = (P(0), P(\alpha), P(\alpha^2), ..., P(1))$. As we already have seen we need for the protocol not the whole codeword but only one of it's Q coordinates. We will use the mapping

$$C_\gamma(m) = P(\gamma), \gamma \in GF(Q). \tag{8}$$

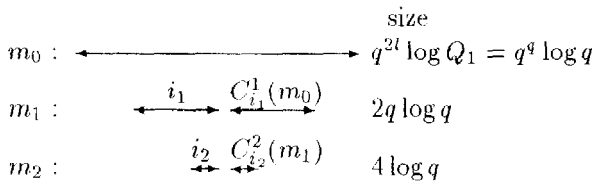
Finally we recall that the minimum distance d of the code C satisfies

$$d = Q - k.$$

We now describe an example of the G-N protocol for $k = 5$. Thus, see Figure 2, we need two codes C^1 and C^2. We will show that we can give a construction that gives a protocol for which $P_s = 1$ using the substitution attack described in Section 3.2.

We begin with setting up the protocol by chosing the codes[2] C^1 and C^2. Let q be a power of 2 and $l = q/2$. The code C^2 is chosen to be an RS-code over $GF(Q_2), Q_2 = q^2$, with $k_2 = 2l$. Hence it has block length $n_2 = q^2$ and distance $d_2 = q^2 - 2l = q^2 - q$. The code C^1 is chosen to be an RS-code over $GF(Q_1), Q_1 = Q_2^l$, with $k_1 = q^{2l-1}$. Hence it has block length $n_1 = q^{2l}$ and distance $d_1 = q^{2l} - q^{2l-1} = q^q(1 - 1/q)$. The coordinates of the codewords of C^1 are obtained by evaluating a polynomial, associated with the q^{2l-1} tuple m over $GF(Q_1)$ as specified by (8). Similar we obtain the codewords of C^2 by evaluating a polynomial associated with the $2l$ tuple over $GF(Q_2)$.

Recall that m_0 constitutes the actual information message that A wants to send to B. The second message m_1 consists of the index i_1 chosen by B (of size $\log Q_1$) and the coordinate of the corresponding codeword from C^1 selected by this index. Thus m_1 has size $2 \log Q_1 = 4l \log q$. Similary, the third message m_2 consists of the index i_2 chosen by A and the corresponding coordinate. It has size $2 \log Q_2 = 4 \log q$. Thus we see that the original message is "reduced" as illustrated below



Now recall from the definition of the G-N protocol property (ii) that $n_i - n_i p_i = d_i$, hence we have

$$p_1 = 1 - \frac{d_1}{n_1} = 1 - \frac{q^{2l} - q^{2l-1}}{q^{2l}} = \frac{1}{q},$$

$$p_2 = 1 - \frac{d_2}{n_2} = 1 - \frac{q^2 - q}{q^2} = \frac{1}{q},$$

[2] Actually we need also C^3 and C^A but these are irrelevant for our attack.

and hence according to proposition 1. (4) (Claim!) we would have for the probability of a successful attack in the two first rounds

$$P_s = p_1 + p_2 - p_1 p_2 < 2/q.$$

However let the received message in the first round be $m_2 = [i_2 \circ C^2_{i_2}(m'_1)]$. For simplicity define

$$\alpha = i_1, \beta = i_2, \text{ and } c = C^2_{i_2}(m'_1)$$

i.e. $m_2 = [\beta \circ c]$. The attack given in Section 4 succeeds if $C^2_\beta(m_1) = C^2_\beta(m'_1)$.

Now write

$$m_0 = m_{0,0}, m_{0,1}, ..., m_{0,q^{2l-1}-1}, \quad m_{0,j} \in GF(Q_1),$$
$$m_1 = m_{1,0}, m_{1,1}, ..., m_{1,2l-1}, \quad m_{1,j} \in GF(Q_2).$$

Thus $C^1_\alpha(m_0) = m_{0,0} + m_{0,1}\alpha + m_{0,2}\alpha^2 + \cdots + m_{0,q^{2l-1}-1}\alpha^{q^{2l-1}-1}$.

We have $\alpha \in GF(Q_1)$ and $C^1_\alpha(m_0) \in GF(Q_1)$ and since $GF(Q_1)$ is an extension field of $GF(Q_2), \alpha$ and $C^1_\alpha(m_0)$ may be represented as a l-tuple over $GF(Q_2)$.

Now recall that α consists of the first l coordinates of m_1 and $C^1_\alpha(m_0)$ of the second batch of the l coordinates, i.e.,

$$\alpha = \alpha_0, \alpha_1, ..., \alpha_{l-1} = m_{1,0}, m_{1,1}, ..., m_{1,l-1},$$

$$C^1_\alpha(m_0) = C^1_\alpha(m_0)_0, C^1_\alpha(m_0)_1, ..., C^1_\alpha(m_0)_{l-1} = m_{1,l}, m_{1,l+1}, ..., m_{1,2l-1}, m_{1,j}.$$

Thus for the whole m_1 given by
$m_1 = [\alpha \circ C^1_\alpha(m_0)] = m_{1,0}, m_{1,1}, ..., m_{1,l-1}, m_{1,l}, m_{1,l+1}, ..., m_{1,2l-1}$ we get that
$C^2_\beta(m_1) = m_{1,0} + m_{1,1}\beta + m_{1,2}\beta^2 + \cdots + m_{1,2l-1}\beta^{2l-1}, \beta \in GF(Q_2)(\subseteq GF(Q_1)!)$.

Suppose now that we choose to replace message m_0 by $m'_0 = c(\beta^l)^{-1}, -(\beta^l)^{-1}, 0, 0, ..., 0$. Note that $c(\beta^l)^{-1}, -(\beta^l)^{-1}$ are elements of the small field $GF(Q_2)$.Then

$$C^1_\alpha(m'_0) = c(\beta^l)^{-1} - (\beta^l)^{-1}\alpha.$$

Thus for m_1 we get

$$m_1 = [\alpha, C^1_\alpha(m'_0)] = \alpha_0, \alpha_1, ..., \alpha_{l-1}, (C^1_\alpha(m'_0))_0, (C^1_\alpha(m'_0))_1, ..., (C^1_\alpha(m'_0))_{l-1} =$$
$$= \alpha_0, \alpha_1, ..., \alpha_{l-1}, (c(\beta^l)^{-1} - (\beta^l)^{-1}\alpha)_0, -(\beta^l)^{-1}\alpha_1, ..., -(\beta^l)^{-1}\alpha_{l-1},$$

where we used the fact that $(c(\beta^l)^{-1} - (\beta^l)^{-1}\alpha)_i = -(\beta^l)^{-1}\alpha_i$ for $i = 1, 2, ..., l-1$. (recall that both $c(\beta^l)^{-1}$ and $-(\beta^l)^{-1}$ are elements of $GF(Q_2)$). Hence

$$C^2_\beta(m_1) = \alpha_0 + \alpha_1\beta + \alpha_2\beta^2 + \cdots + \alpha_{l-1}\beta^{l-1} +$$

$$
\begin{aligned}
&+ (c(\beta^l)^{-1} - (\beta^l)^{-1}\alpha)_0\beta^l + -(\beta^l)^{-1}\alpha_1\beta^{l+1} + \cdots + -(\beta^l)^{-1}\alpha_{l-1}\beta^{2l-1} = \\
&= \alpha_0 + \alpha_1\beta + \alpha_2\beta^2 + \cdots + \alpha_{l-1}\beta^{l-1} + \\
&+ c - \alpha_0 - \alpha_1\beta - \alpha_2\beta^2 - \cdots - \alpha_{l-1}\beta^{l-1} = \\
&= c = C'^2_\beta(m'_1) \quad !
\end{aligned}
$$

and independently of the index α chosen by B, the substitution attack will succeed, i.e., $P_s = 1$.

6 Conclusion

We have analysed the Gemmell and Naor multiround protocol. We have shown that the number of rounds have to be an odd number. Furthermore, we have given a counter example to the Claim by Gemmell and Naor for the probability of a successful substitution attack.

Acknowledgements

I want to thank T. Johansson for his helpful comments and suggestions regarding this work.

References

1. E. Gilbert, F.J. MacWilliams, N. Sloane, "Codes Which Detect Deception". Bell System Technical Journal. Vol. 53. No. 3. March 1974, pp. 405-424.
2. J.L. Carter, M.N. Wegman, "New hash functions and their use in authentication and set equality", *J. Computer and System Sci.*, Vol 22, 1981, pp. 265-279.
3. G.J. Simmons, "A survey of Information Authentication", in *Contemporary Cryptology, The science for information integrity*, ed. G.J. Simmons, IEEE Press, New York, 1992.
4. T. Johansson, G. Kabatanskii, B. Smeets, "On the relation between A-codes and codes correcting independent errors", *Proceedings of Eurocrypt '93*, 1993, pp. 1-11.
5. J. Bierbrauer, T. Johansson, G. Kabatanskii, B. Smeets, "On Families of Hash Functions via Geometric Codes and Concatenation", *Proceedings of CRYPTO '93*, 1993, pp. 331-342.
6. D.R. Stinson, "Universal hashing and authentication codes", to appear in IEEE Transaction on Information Theory.
7. C. Gehrmann, "Long Message Authentication by using Pseudo-Random Functions", *Proceedings of IEEE ISIT 94*, to appear (preprint).
8. I.S. Reed, G. Solomon, "Polynomial Codes over certain Finite Fields", J. Soc. Ind. Appl. Math., vol. 8, June 1960, pp. 300-304.
9. P. Gemmell, M. Naor,"Codes for interactive authentication", *Proceedings of CRYPTO '93*, 1993, pp. 355-367.