On the construction of authentication codes with secrecy and codes withstanding spoofing attacks of order $L \geq 2$

Ben Smeets, Peter Vanroose and Zhe-xian Wan

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

Abstract

We present an analysis of some known cartesian authentication codes and their modification into authentication codes with secrecy, with transmission rate $R = r/n$, where $n = 2, 3, \ldots$, and $1 \leq r \leq n-1$ using $(n-r)(r+1)$ q-ary key digits. For this purpose we use a grouping technique.

Essentially the same key grouping technique is used for the construction of codes that withstand spoofing attacks of order $L \geq 2$. The information rate of this scheme is also r/n, and it requires $(L+r)(n-r)$ q-ary key digits. Moreover these codes allow that previously transmitted source states can be reused.

I.B. Damgard (Ed.): Advances in Cryptology - EUROCRYPT '90, LNCS 473, pp. 306-312, 1991.

1 Introduction

Authentication codes are codes which allow two trusting parties to communicate information to each other in the presence of an opponent who might submit false messages and/or substitute legally transmitted messages by false ones that will misinform the receiver. One of the first constructions of such codes were the codes given by Gilbert, MacWilliams and Sloane [1]. Their constructions stem from the projective spaces over finite fields.

In this paper we address two different problems. The first problem is the construction of authentication codes which also achieve perfect secrecy. That is, given an observed transmitted message, the opponent obtains no information about the source state that was transmitted by the legal sender. Recently various authors have presented some constructions most of them stemming from combinatorial designs [2,3]. The other problem is the construction of codes for multiple authentication. The latter type of codes have to withstand the so-called spoofing attacks of order $L \geq 2$. Most researchers investigated this problem under the ad-hoc assumption that previously transmitted source states were excluded from subsequent transmissions. For a discussion of the origin of this assumption and its effects we refer to [4]. In this paper we extend the concept of spoofing attack [5] so that it allows the transmitter to repeat previously used source states.

In [1] a modification of the code is presented allowing a simpler logic circuitry than for the orginal code. In Section 2, we analyse this construction in more detail and introduce a grouping technique of the encoding rules. This grouping is then used in the code constructions given in the subsequent sections.

In Section 3 we use the grouping technique to obtain codes that achieve perfect secrecy and in Section 4 we discuss how to use it to obtain codes for multiple authentication.

Finally, Section 5 contains a description of a generalization of the construction allowing a tradeoff between the security requirements and the desire to achieve high information rates.

2 The Modified GMS Scheme and a Grouping Technique

Let q be a prime power. Gilbert, MacWilliams and Sloane [1] constructed an authentication code from the projective plane $\mathbf{PG}(2, \mathbf{F}_q)$ over the field $\mathbf{F}_q$ with q elements as follows. Fix a line ℓ in $\mathbf{PG}(2, \mathbf{F}_q)$. The points on ℓ are regarded as the source states, the points not lying on ℓ are regarded as the encoding rules and the lines different from ℓ are regarded as the messages. Given a source state s and an encoding rule e, there is a unique line passing through s and e, which will be called m, i.e., $m = \Updownarrow(s, e)$. The source state s will now be encoded into the message m by the encoding rule e. This is an authentication code with $q + 1$ source states, q^2 encoding rules and $q^2 + q$ messages.

The probabilities of successfull impersonation and substitution attack of this code are

$$P_I = P_S = \frac{1}{q}.$$

This scheme will be referred to as the GMS scheme and the corresponding code is a so-called cartesian code because each message conveys the transmitted source state. For $q = 2$, the GMS scheme can be represented by the following encoding table:

	a	b	c	d	e	f
e_0	0			1	2	
e_1		0	1		2	
e_2	0		1			2
e_3		0		1		2

Table 1: The GMS Scheme derived from $\mathbf{PG}(2, \mathbf{F}_2)$

where 0, 1, 2 are the 3 source states, being all points of a given line ℓ, e_0, e_1, e_2, e_3 are the encoding rules, being the 4 other points, and a, b, c, d, e, f are the messages, being the lines different from ℓ.

In Section V of reference [1] the authors also gave a modification of their scheme. For this modified scheme one chooses in addition to the line ℓ also a fixed point P on ℓ. The source states (they are q in number) are then all the points on ℓ different from P, the encoding rules are all the points not on ℓ and the messages are all the lines not passing through P. We obtain a code with q source states, q^2 encoding rules and q^2 messages. This scheme will be referred to as the Modified GMS scheme. For this code we also have $P_I = P_S = 1/q$. In Table 2 we illustrate the modification with the case $q = 2$.

	a	b	c	d		
e_0	0		1			P
e_1	0			1	P	
e_2		0	1		P	
e_3		0		1		P

Table 2: The Modified GMS scheme derived from $\mathbf{PG}(2, \mathbf{F}_2)$.

For this example, we observe that the four encoding rules can be grouped into two groups, $\{e_0, e_3\}$ and $\{e_1, e_2\}$, such that each message contains exactly one encoding rule from each group [4]. This kind of *grouping* holds in the general case. The groups are the lines through P.

Theorem 1: In the Modified GMS scheme, if we group the q^2 encoding rules into q groups, such that each group consists of the q rules lying on a line through P, then each message contains exactly one encoding rule from each group.

Proof: This follows from the property of the projective plane that any two lines meet in exactly one point. □

3 Authentication codes with perfect secrecy

We continue with the example of the foregoing section:

Cartesian

	a	b	c	d	
e_0	0		1		P
e_1	0			1	P
e_2		0	1		P
e_3		0		1	P

Perfect Secrecy

	a	b	c	d	
e_0	0		1		P
e_1	1			0	P
e_2		1	0		P
e_3		0		1	P

Table 3:The Construction of a code with perfect secrecy from the Modified GMS scheme derived from $\mathbf{PG}(2, \mathbf{F}_2)$.

Above we interchanged 0 and 1 in the rows of the second group of encoding rules $\{e_1, e_2\}$. In the new code, each message contains two source states, one 0 and one 1. Thus perfect secrecy is achieved. This is the well-known "simplest" example of an authentication code with secrecy, see [7].

For the general case, we label the source states of the Modified GMS scheme with $0, 1, \ldots, q-1$ and the groups of encoding rules also with $0, 1, \ldots, q-1$. We change the encoding table of the Modified GMS scheme according to the following procedure. Replace source state s lying in the row of the encoding rule e and the column of the message m in the encoding table, where e belongs to the group i, by $s - i (\mathrm{mod}\ q)$. That is, m will be used to transmit the source state $s - i (\mathrm{mod}\ q)$ under the encoding rule e. We can prove the following:

Theorem 2: Each message in the new authentication code contains each of the q source states $0, 1, \ldots, q-1$ exactly once, thus we have perfect secrecy for this code.

Proof: In the encoding table of the original Modified GMS construction, each column, i.e., each message, contains one source state q times. Theorem 1 guarantees that these are replaced by q different source states. □

Finally we note that the information rate of this code is 1/2, i.e., the same as the rate for the Modified GMS scheme. The key information required to specify the encoding rule used is two q-ary digits (considering the case that the encoding rules are selected according to a uniform probability distribution).

4 Authentication codes withstanding spoofing attacks of order $L \geq 2$

The previous geometrical scheme (with reduced source state set and with grouping of keys) can also be turned into a code that can be used L consecutive times, having probability of deception $1/q$ at each use. The information rate of this scheme is also 1/2, and it requires $2 + (L-1)$ q-ary key digits. But we can no longer guarantee the source state to be secret from the second use on.

At the first time slot, the scheme of the previous paragraph (with or without secrecy) is used, consuming the first 2 q-ary key digits. At any of the next $L-1$ time slots, the group label of the key used at the previous slot is incremented by one (modulo q), and 1 extra q-ary key digit is used to determine the key inside this new group to be used for this time slot, using the same geometrical scheme as at the first time slot. The number of key digits required per transmission is asymptotically optimal ($L \to \infty$), i.e., one q-ary digit per transmission. These codes have a probability $1/q$ of deception at every transmission.

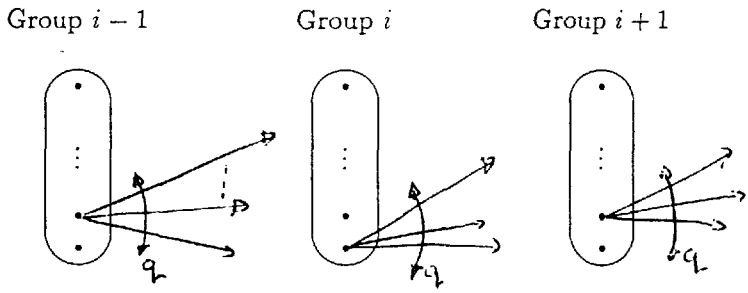

Figure 1: Illustration of the paths of possible encoding rules

5 Generalizations

The modification technique of Gilbert, MacWilliams and Sloane and our grouping technique can be applied to several generalizations of the GMS scheme. For instance, we

can use some authentication codes given by Beutelspacher [6]. We shall present a new generalization below, to which both the modification technique and the grouping technique can be applied. It will give us a generalization of the two dimensional schemes of the previous sections to the n-dimensional projective space.

Fix an r-flat $\mathcal{L}$ of the n-dimensional projective space $\mathbf{PG}(n, \mathbf{F}_q)$, where $1 \leq r \leq n-1$. The source states are all the t-flats, $0 \leq t < r$, contained in $\mathcal{L}$. Encoding rules are all $(n-r-1)$-flats which have empty intersection with $\mathcal{L}$. The messages are the $(n-r+t)$-flats intersecting $\mathcal{L}$ in a t-flat. The message corresponding to source state s and encoding rule e is the unique $(n-1)$-flat containing s and e. For this code we have:

The number of source states:	$\begin{bmatrix} r+1 \\ t+1 \end{bmatrix}_q$ [1)]
The number of encoding rules:	$q^{(n-r)(r+1)}$
The number of messages:	$q^{(n-r)(r-t)} \begin{bmatrix} r+1 \\ t+1 \end{bmatrix}_q$
The probability of impersonation:	$P_I = q^{-(n-r)(r-t)}$
The probability of substitution:	$P_S = q^{-(n-r)}$
The information rate:	$\geq \frac{t+1}{n-r+t+1}$ (for large q).

Substitution is done with the highest probability of success by replacing the legal message by a message having an as large as possible intersection with the legal one, i.e., an $(n-r+t-1)$-flat. For $r = n-1$ and $t = n-2$, this is the generalization given in [1].

Now take $t = r-1$. The modification of the code is carried out as follows. Fix a point P on $\mathcal{L}$. Source states are all $(r-1)$-flats of $\mathcal{L}$ *not containing* P. Encoding rules are the same as above and the messages are the $(n-1)$-flats which do not contain P. For the modified code we have:

The number of source states:	q^r
The number of encoding rules:	$q^{(n-r)(r+1)}$
The number of messages:	q^n
The probability of impersonation:	$P_I = \frac{1}{q^{n-r}}$
The probability of substitution:	$P_S = P_I$
The information rate:	$\frac{r}{n}$.

The grouping of the encoding rules is done as follows. Consider all possible $(n-r)$-flats through P, intersecting $\mathcal{L}$ only in P. The encoding rules are grouped such that encoding rules lying in a common $(n-r)$-flat of the above type belong to the same group. This results in the partitioning of the $q^{(n-r)(r+1)}$ encoding rules into $q^{(n-r)r}$ groups of q^{n-r} encoding rules such that each message contains exactly one encoding rule from each group. Thus from this modified code we can construct an authentication code with

[1)] A Gaussian binomial coefficient $\begin{bmatrix} r+1 \\ t+1 \end{bmatrix}_q \stackrel{\text{def}}{=} \frac{(q^{r+1}-1)\ldots(q^{r-t+1}-1)}{(q^{t+1}-1)\ldots(q-1)}$

perfect secrecy as well as an authentication code withstanding spoofing attacks of order $L \geq 2$ as before. But in the case of codes for multiple authentication we have to spend additional $(n-r)$ q-ary digits per additional use of the code. That is $(L+r)(n-r)$ q-ary key digits in total. Hence, the asymptotic number of key digits per transmission is $n-r$.

References

[1] E.N. Gilbert, F.J. MacWilliams and N.J.A. Sloane, "Codes which detect deception", *Bell Syst. Techn. J.* **53**, pp. 405–424, 1974.

[2] M. De Soete, "Some constructions for authentication - secrecy codes", Eurocrypt'88, Davos, Switzerland, May 25-27, 1988, in *Advances in Cryptology — Eurocrypt '88*, Ed. C.G. Günther, Springer-Verlag, Berlin, pp. 469-472, 1988.

[3] D.R. Stinson, "A construction for authentication codes/secrecy codes from certain combinatorial designs", J. Cryptology, **1**, pp.119-127, 1988.

[4] G. Simmons, B. Smeets, "A paradoxical result in unconditionally secure authentication codes – and an explanation", *IMA Conference on Cryptography and Coding*, Dec. 18-20, 1989, Cirencester, England. to appear.

[5] J.L. Massey, "Cryptography – A selective survey", in *Digital Communications*, pp.3-21. 1986.

[6] A. Beutelspacher, "Perfect and essentially perfect authentication schemes", *Advances in Cryptology — EUROCRYPT '87*, **LNCS 304**, Springer-Verlag, Berlin, pp. 167–170, 1987.

[7] J.L. Massey, "An introduction to contemporary cryptology", *Proc. IEEE* **76**, pp. 533–549, 1988.