Authentication Codes in Plaintext and Chosen-content Attacks

R. Safavi-Naini *
L. Tombak **

Department of Computer Science University of Wollongong
Northfields Ave., Wollongong 2522, AUSTRALIA

Abstract. We study authentication codes (A-codes) where it is assumed that the enemy has access to the content of the intercepted cryptogram. This is similar to plaintext attack in secrecy systems. Enemy's success is defined in two ways. The first is as in Simmons' model. We will also consider *chosen-content* attacks in which the success is by constructing a fraudulent cryptogram with a *given content*. We will obtain information theoretic bounds, define perfect protection and obtain lower bounds on the number of encoding rules for codes with perfect protection against chosen-content impersonation and chosen-content plaintext substitution. We characterize these A-codes when the number of encoding rules is minimum. We give methods for making an A-code resistant against plaintext and chosen-context plaintext attack.

1 Introduction

A basic assumption in secrecy and authenticity systems is that the encoding procedure and probability distribution of the source and key space is known to the enemy. In secrecy systems the extra information available to the enemy is used to classify the severity of attacks, so that the ciphertext only attack is the easiest, followed by the plaintext and chosen plaintext attack which are the more stringent ones. In authenticity systems the same approach can be used. However a second crucial factor in this classification is the way *success* is defined. In Simmons' model of authentication [1, 2], the enemy succeeds if he/she can construct a cryptogram acceptable to the receiver. A more demanding type of success is when the receiver is deceived by a cryptogram with a *given content*. This is called *chosen-content attack*. This model of attack is also mentioned in [13] but no analysis of the model is given. An example of such attack is when the value figure of a financial transaction is substituted by a value chosen by the opponent. Table 1 gives a classification of attacks in A-systems for various degrees of the enemy's power and the two types of success mentioned above.

* Support for this project was partly provided by Australian Research Council grant A49030136.

** Support for this project was provided by Australian Research Council grant A49030136.

Table 1. Attacks in A-systems

Type of success	Information available				
	No information	ctext	Chosen ctext	ptext	Chosen ptext
Simmons' model	*Impersonation*	*Substitution*	*Johansson et al.*	√	
Chosen-content	√			√	

where ctext = ciphertext and ptext = plaintext.

In this table columns, from left to right, correspond to the increasing degree of the enemy's power while rows, from top to bottom, correspond to the increased difficulty of success in the attack. The following is a glossary of the labels used for the table's columns:

- *No information*: enemy has no extra information.
- *Ciphertext*: enemy has access to a cryptogram (ciphertext only attack);
- *Chosen ciphertext* : enemy can choose the valid cryptogram sent in the channel;
- *Plaintext* : enemy has intercepted a cryptogram and knows its content;
- *Chosen plaintext*: enemy can choose a source state and will be given the corresponding cryptogram.

It is noted that the enemy's power and the type of success are independent attributes of an attack and hence each cell of the table represents a possible type of attack. Simmons' impersonation and substitution correspond to cells $(1,1)$ and $(1,2)$ of the table respectively and the way probability of deception is calculated by Johansson et al [12] corresponds to cell $(1,3)$ of the table. We use P_0 and P_1 to denote probability of success in impersonation and substitution attack in Simmons' model of attack.

In this paper we will consider attacks corresponding to the cells $(2,1)$, $(1,4)$, $(2,4)$ of the table. We call them chosen-content impersonation, plaintext, and chosen-content plaintext respectively and use P_0^c, P_1^p, P_1^{cp} to denote probability of success in each case. We obtain information theoretic bounds on probability of deception and define A-codes that provide perfect protection. We will show that perfect protection for plaintext attack is closely related to perfect one-fold secrecy and use a transformation, on an arbitrary A-code, to increase its resistance against plaintext attack. It is known, [3], [6], that optimal perfect t-fold secrecy codes are equivalent to perpendicular arrays (PAs). Our transformation uses $PA_1(1,k,k)$s to increase P_1^p of an A-code without affecting P_1 and P_0 of the code. Codes that provide protection against chosen-content attacks are related to ordered designs (ODs). We give a second transformation, using ordered designs, that increases resistance of A-codes against chosen-content attacks. The two transformations, mentioned above, produce A-codes with the best P_1^p and P_1^{cp}, respectively, for the given P_1. We will show that A-codes with perfect protection against chosen-content impersonation and chosen-content plaintext and minimum number of encoding rules are equivalent to ODs. Stinson [7] has proved similar types of results for cartesian A-codes and A-codes with secrecy for Simmons' model of attack.

It is interesting to note that the lower bound on P_1^{cp} is always greater than the lower bound on P_0^{cp} and so for codes that satisfy these bounds substitution is always the better game to play.

Finally we use a composition method used by Bierbrauer and Edel [10] for authentication PAs, to combine ODs and obtain more efficient A-codes.

Let A be an (M, k, E) A-code with probability of substitution equal to P_1. The main results of this paper include:

- construction of an (M, k, kE) A-code, using PAs, with $P_1^p = P_1$; that is, an A-code for which the knowledge of the cryptogram content is not useful to the enemy and for the given P_1, P_1^p has its lowest value;
- construction of an $(M, k, k(k-1)E)$ A-code, using ODs, with the same P_1 and having $P_1^{cp} = \frac{P_1}{k-1}$. Again this is the best possible value of P_1^{cp} (for the given P_1) and corresponds to the case when the knowledge of cryptogram content is not of assistance to the enemy. The construction is possible only if k is a prime power;
- a bound on the number of encoding rules for A-codes that provide perfect protection against impersonation and plaintext substitution and a similar bound when chosen-content attack is used;
- a characterisation of A-codes that provide perfect protection against chosen-content impersonation and chosen-content plaintext substitution with minimum number of encoding rule;
- information theoretic and combinatorial bounds on P_1^p and P_1^{cp}.

2 Preliminaries

We consider an authentication scenario with three participants: a transmitter and receiver (*communicants*) who want to communicate over a publicly exposed channel and an *enemy* who tries to deceive the receiver into accepting a fraudulent message as genuine. We are only concerned with honest communicants. An (M, k, E) *authentication code* (A-code) is a collection $\mathcal{E}$, $|\mathcal{E}| = E$, of mappings called *encoding rules*, from the set $\mathcal{S}$, $|\mathcal{S}| = k$, of *source states* into the set $\mathcal{M}$, $|\mathcal{M}| = M$, of *codewords*. The code provides protection only if $k < M$. The *encoding matrix* of the code is an $E \times M$ matrix, denoted by B in this paper, whose rows and columns are labeled by the elements of $\mathcal{E}$ and $\mathcal{M}$ and $B(e, m)$ is the source state s with $e(s) = m$ and zero otherwise. We denote by $\mathcal{E}(m)$ the subset of keys that are incident with the cryptogram m, by $\mathcal{E}(m, s)$ the subset of keys that map the source state s into the cryptogram m, i.e, $\mathcal{E}(m, s) = \{e_j \in \mathcal{E} | a_{msj} = 1\}$, and by $\mathcal{E}((m, s), m')$ the subset of keys that map the source state s to the cryptogram m and are incident with cryptogram m'. Also $\mathcal{M}(e)$ is the subset of cryptograms that are valid for the encoding rule e.

For a set $\mathcal{X}$ we use X to denote its cardinality. For example $E(m, s)$ is the cardinality of the set $\mathcal{E}(m, s)$. The *incidence matrix* of an A-code is a binary matrix $A = [a_{(ms),e}]$ whose rows are labeled by the elements of the set $\mathcal{M} \times \mathcal{S}$

and whose columns are labeled by the elements of $\mathcal{E}$ and $a_{mse} = 1$ if $e(s) = m$ and zero otherwise. We note that $E(m,s)$ might be zero for some pairs (m,s) and so $\sum_{j=1}^{E} a_{msj} = 0$. We consider A-codes without splitting. In such A-codes, an encoding rule is a one to one mapping from the set $\mathcal{S}$ of source states to the subset $\mathcal{M}(e)$ of $\mathcal{M}$ and so $\sum_{m\in\mathcal{M}} \sum_{s\in\mathcal{S}} a_{msj} = k$ and $\sum_{s\in\mathcal{S}} a_{msj} = 1$ or 0.

The communicants choose an encoding rule (key) according to the probability vector $\pi = (\pi_1, ..., \pi_E)$. The enemy uses a plaintext attack or a chosen-content plaintext attack. Both of these attacks are variations of the substitution attack. We have also considered the chosen-content impersonation attack.

3 Plaintext Attack

Plaintext attack is a substitution attack in which the content of the cryptogram is known to the enemy. For cartesian A-codes this is the same as traditional substitution. However for A-codes with secrecy knowledge of the content of the cryptogram is extra information available to the enemy and hence:

$$P_1^p \geq P_1. \tag{1}$$

Let $P(m,s)$ denote the probability of a source state s being mapped into a cryptogram m. We have $P(m,s) = P_S(s) \times \sum_{j=1}^{E} \pi_j a_{msj}$, where $P_S(s)$ is the source probability distribution. Probability of the enemy's success if he/she intercepts cryptogram m, knows the corresponding source state s, and introduces a fraudulent message m' into the channel, is given by $payoff((m,s),m')$,

$$\begin{aligned} payoff((m,s),m') &= P(m' \text{ valid } |(m,s) \text{ received }), \\ &= \frac{\sum_{s'\in\mathcal{S}\setminus s} \sum_{j=1}^{E} \pi_j a_{msj} a_{m's'j}}{\sum_{j=1}^{E} \pi_j a_{msj}}. \end{aligned} \tag{2}$$

The enemy's strategy q can be represented as a collection of probability vectors $q = \{q^{m,s}\}$, where,

$$q^{ms} = (q^{ms}_{m_1}, \cdots, q^{ms}_{m_{M-1}}),\ m_i \neq m,\ q^{ms}_{m_i} \geq 0, \sum_i q^{ms}_{m_i} = 1.$$

The probability of the enemy's success when the enemy has intercepted a pair (m,s) and uses the best strategy is

$$P_1^p = \sum_{m\in\mathcal{M}} \sum_{s\in\mathcal{S}} \sum_{m'\in\mathcal{M}\setminus m} \sum_{s'\in\mathcal{S}\setminus s} \pi_j a_{msj} a_{m's'j} q^{ms}_{m'} P_s(s) \tag{3}$$

$$= \sum_{m,s} P(m,s) Max_{m'}(payoff(m',(m,s))) \tag{4}$$

3.1 Bounds and perfect protection

An A-code provides *perfect protection* against plaintext attack if the enemy's best strategy is uniform, i.e, $q_{m'}^{ms} = 1/(M-1)$ for every pair (m,s) for which $E(m,s) > 0$. Note that we do not require $E(m,s) \neq 0$ for all pairs (m,s); rather it is required that the strategy be uniform for all (m,s) with $P(m,s) > 0$.

Similar to the usual authentication scenario we have:

Proposition 1. *An A-code provides perfect protection against plaintext attack if and only if,*

$$payoff((m,s),m') = \frac{k-1}{M-1},$$

for every pair (m,s) with $P(m,s) > 0$. In this case $P_1^p = \dfrac{k-1}{M-1}$.

Let C denote the set of pairs (m,s) with $P(m,s) > 0$. C is the number of non-zero (with at least one non-zero element) rows of matrix A.

Proposition 2. *Let $P_1^p = \dfrac{k-1}{M-1}$. Then*

$$2M \leq C \leq kM.$$

Equality in the right hand side is if $P(m,s) > 0$ for all $m \in \mathcal{M}$ and $s \in \mathcal{S}$. In this case $C = kM$.

Using a simple counting argument it is easy to show that if an A-code provides perfect protection for plaintext attack then the number of encoding rules is lower bounded by the following expression,

$$E \geq \frac{C(M-1)}{k(k-1)}. \tag{5}$$

Expression 5 shows that the minimum number of encoding rules for an A-code that provides perfect protection for plaintext attack is at least *twice* the corresponding number for Simmons' model of attack.

Theorem 3. *Let $P_1^p = (k-1)/(M-1)$, $C = kM$ and let E satisfy bound 5 with equality. Then $P_0 = k/M$ and the code provides perfect one-fold secrecy. The communicants' best strategy is uniform.*

Theorem 3 leads to a construction of A-codes, with perfect protection against plaintext attack. The construction was originally used by Stinson (theorem 4.2,[6]) to construct codes that provide 1-fold secrecy and are 1-fold secure against spoofing. In section 3.2 we generalize this construction to a transformation on an arbitrary A-code to make it resistant against plaintext attack.

The main information-theoretic bound on P_1 is due to Pei and Rosenbaum [9, 11]. Similar result can be proved for P_1^p:

Theorem 4.

$$P_1^p \geq 2^{-(H(\mathcal{E}|\mathcal{MS})-H(\mathcal{E}|\mathcal{M}^2\mathcal{S}))} \tag{6}$$

and equality holds if and only if

(i) $P_1^p = payoff((m,s),m') = const$ *for all* (m,s) *and* m' *such that* $E((m,s),m') > 0$;

(ii) the conditional probability $P(m'|e,(m,s))$ *that* m' *is the next cryptogram sent by the transmitter, given that* e *is the actual encoding rule and pair* (m,s) *has already been sent, is constant for all* $e \in \mathcal{E}((m,s),m')$.

Equality in bound 6 implies that

$$P_1^p = \frac{k-1}{V},$$

where for any (m,s), V is the number of $m' \in \mathcal{M}$ with $\mathcal{E}(m') \bigcap \mathcal{E}(m,s) \neq \emptyset\}$, and is independent of (m,s).

3.2 Construction of *A*-codes resistant against plaintext attack

A perpendicular array $PA_\lambda(t,k,v)$ is a $b \times k$ array of elements of a v-set V such that each row of the array consists of k distinct elements of V and a set of t columns contains every t-subset of V, λ times.

Theorem 5. *Consider an* (M,k,E) *A-code with uniform source and* $P_1 = \epsilon$. *Then we can construct an* (M,k,kE) *A-code, with* kE *encoding rules, and* $P_1^p = P_1 = \epsilon$.

If the enemy knows the content of the cryptogram he/she is in a more powerful position compared to the Simmons' model of attack. However theorem 5 shows that it is always possible to transform an arbitrary A-code to one with $P_1^p = P_1$, that is, one for which the knowledge of the content of the cryptogram is not useful to the enemy and his/her chance of success is not affected by this extra knowledge. The transformation substitutes a row e_i of the encoding matrix by k rows, each with k non-zero elements, such that restriction of these rows to $M(e_i)$ is a $PA_1(1,k,k)$. Using bound 1 it can be seen that the code constructed in theorem 5 has the lowest possible P_1^p for the given P_1.

Corollary 6. *In theorem 5 if* $P_0 = k/M$ *the resulting code will have* $P_0 = k/M$, $P_1 = P_1^p = \epsilon$. *The code will also provide perfect one-fold secrecy.*

It is shown [7] that for an A-code with $P_0 = k/M$ and $P_1 = (k-1)/(M-1)$ the number of encoding rules E is lower bounded by

$$E \geq E_0 = \frac{M(M-1)}{k(k-1)}.$$

If the code has $E = E_0$ the transformation of theorem 5 results in an A-code for which E satisfies bound 5 with equality and hence has the minimum number of encoding rules for A-codes with $P_0 = k/M$ and $P_1^p = (k-1)/(M-1)$.

We note that for codes with $P_0 = k/M$ the increase in the number of encoding rules, due to the construction given in theorem 5, may contribute to providing secrecy, protection against plaintext attack, or in general both. We consider two extreme cases. If the code originally provides perfect secrecy but does not provide any protection against plaintext attack, increasing the number of encoding rules will only increase protection against plaintext attack (see example 1). On the other hand codes without secrecy always provide protection against plaintext attack. Thus increasing the number of encoding rules only results in perfect secrecy (see example 2).

Example 1. Consider the following A-code with $P_0 = k/M$, $P_1 = (k-1)/(M-1)$ that provides perfect one-fold secrecy, but has $P_1^p = 1$

E/M	0	1	2	3	4	5	6
0	s_1	0	s_2	s_3	0	0	0
1	0	s_1	0	s_2	s_3	0	0
2	0	0	s_1	0	s_2	s_3	0
3	0	0	0	s_1	0	s_2	s_3
4	s_3	0	0	0	s_1	0	s_2
5	s_2	s_3	0	0	0	s_1	0
6	0	s_2	s_3	0	0	0	s_1

Applying the transformation of theorem 5 results in an A-code with the same values of P_0 and P_1, which provides perfect one-fold secrecy and has $P_1^p = P_1$.

E/M	0	1	2	3	4	5	6
0	s_1	0	s_2	s_3	0	0	0
1	s_2	0	s_3	s_1	0	0	0
2	s_3	0	s_1	s_2	0	0	0
3	0	s_1	0	s_2	s_3	0	0
4	0	s_2	0	s_3	s_1	0	0
5	0	s_3	0	s_1	s_2	0	0
6	0	0	s_1	0	s_2	s_3	0
7	0	0	s_2	0	s_3	s_1	0
8	0	0	s_3	0	s_1	s_2	0
9	0	0	0	s_1	0	s_2	s_3

E/M	0	1	2	3	4	5	6
10	0	0	0	s_2	0	s_3	s_1
11	0	0	0	s_3	0	s_1	s_2
12	s_3	0	0	0	s_1	0	s_2
13	s_1	0	0	0	s_2	0	s_3
14	s_2	0	0	0	s_3	0	s_1
15	s_2	s_3	0	0	0	s_1	0
16	s_3	s_1	0	0	0	s_2	0
17	s_1	s_2	0	0	0	s_3	0
18	0	s_2	s_3	0	0	0	s_1
19	0	s_3	s_1	0	0	0	s_2
20	0	s_1	s_2	0	0	0	s_3

□

Example 2. Consider an A-code without secrecy with $P_0 = P_1 = P_1^p = k/M$.

E/M	0	1	2	3
0	s_1	s_2	0	0
1	s_1	0	s_2	0
2	0	s_2	0	s_1
3	0	0	s_2	s_1

After transformation the A-code will have $P_0 = P_1 = P_1^p = k/M$ and the code also provides perfect one-fold secrecy.

E/M	0	1	2	3
0	s_1	s_2	0	0
1	s_2	s_1	0	0
2	s_1	0	s_2	0
3	s_2	0	s_1	0
4	0	s_2	0	s_1
5	0	s_1	0	s_2
6	0	0	s_2	s_1
7	0	0	s_1	s_2

□

4 Chosen-content attacks

In these types of attacks the enemy succeeds if he/she can construct a valid cryptogram for a particular source states. In impersonation attack the enemy sends the cryptogram m into the channel and expects that it will be decoded to the source state s. For substitution we only consider a plaintext attack, that is, the enemy intercepts a cryptogram m, knows its corresponding source state s, and wants to substitute it with a valid cryptogram $m' \neq m$ which will be decoded to a particular source state $s' \neq s$.

The enemy's impersonation strategy is represented by a $k \times M$ dimensional probability vector $q = (q_{m_1 s_1}, \cdots, q_{m_M s_k})$ and probability of deception is given by

$$P_0^c = \sum_{(m,s)} \sum_j \pi_j a_{msj} q_{ms}.$$

4.1 Bounds and perfect protection

An A-code provides *perfect protection against chosen-content impersonation* attack if the enemy's best strategy is random selection among all pairs $(m, s) \in \mathcal{M} \times \mathcal{S}$ and $q_{(ms)} = 1/(kM)$. It is easy to see that if an A-code provides perfect protection against chosen-content impersonation it provides one-fold secrecy.

Proposition 7.

$$P_0^{cp} \geq 1/M,$$

and equality holds if and only if the code provides perfect protection against chosen-content impersonation attack. In this case the code provides perfect one-fold secrecy.

In substitution the enemy intercepts a pair (m, s) and introduces a cryptogram m' for a chosen source state s'. In this case $payoff((m, s), (m', s'))$ is the probability of enemy's success.

$$payoff((m, s), (m', s')) = P((m', s')valid)|(m, s)received),$$
$$= \frac{\sum_{j=1}^{E} \pi_j a_{msj} a_{m's'j}}{\sum_{j=1}^{E} \pi_j a_{msj}}. \tag{7}$$

Summing over all possible $m' \in \mathcal{M} \backslash m$ and $s' \in \mathcal{S} \backslash s$ it is easy to show that,

$$payoff((m, s), (m', s')) \geq 1/(M-1).$$

The enemy's strategy q can be represented as collection of probability vectors $\{q^{ms}\}$, where q^{ms} is a $(k-1) \times (M-1)$ dimensional vector and $q^{ms}_{m's'}$ is the probability of choosing (m', s') when (m, s) is received. The probability of the enemy's success in this case is,

$$P_1^{cp} = \sum_{m \in \mathcal{M}} \sum_{s \in \mathcal{S}} \sum_{m' \in \mathcal{M} \backslash m} \sum_{s' \in \mathcal{S} \backslash s} \pi_j a_{msj} a_{m's'j} q^{ms}_{m's'} P_s(s), \tag{8}$$
$$= \sum_m \sum_s P(m, s) Max_{m', s'} payoff((m', s'), (m, s)), \tag{9}$$

An A-code provides *perfect protection against chosen-content plaintext attack* if and only if for any intercepted pair (m, s) the enemy's best strategy is random selection from all pairs $(m', s'), m' \neq m, s' \neq s$, i.e., $q^{ms}_{m's'} = \dfrac{1}{(k-1)(M-1)}$.

Proposition 8.

$$P_1^{cp} \geq \frac{1}{(M-1)}$$

and equality holds if and only if

$$payoff((m, s), (m', s')) = \frac{1}{(M-1)}, \tag{10}$$

for all (m, s) *with* $P(m, s) > 0$ *and* (m', s'), $m' \neq m, s' \neq s$. *Also if* $C = kM$, *we have* $P_0^c = 1/M$.

It is easy to see that $P_0 \geq P_0^c$. However the relation between P_1 and P_1^{cp} is not so obvious. In example 1 we have $P_1 = 1/3$ but if the content of a cryptogram is known the encoding rule is uniquely determined and we have $P_1^{cp} = 1 > P_1$. On the other hand for cartesian A-codes P_1^{cp} is a more restricted attack than traditional substitution and $P_1^{cp} \leq P_1$.

Proposition 9. *For an A-code with probability of deception in impersonation and substitution equal to P_0 and P_1 respectively, we have*

$$\frac{P_0}{k} \leq P_0^c \quad \leq \quad P_0 \tag{11}$$

$$\frac{P_1}{k-1} \leq \frac{P_1^p}{k-1} \leq P_1^{cp} \leq P_1^p \tag{12}$$

It is interesting to note that for an A-code that provides perfect protection against chosen-content impersonation and chosen-content plaintext, substitution is always the better game as $P_1^{cp} = \frac{1}{M-1} > P_0^{cp} = \frac{1}{M}$. Theorem 10 shows that the number of encoding rules in this case is quite large. We need the following definitions.

An *ordered design* $OD_\lambda(t,k,v)$ is a $b \times k$ array of ordered k-subsets of a v-set V, such that every set of t columns contains every ordered t-subset of V exactly λ times.

Encoding rules of an A-code can be written as a $E \times k$ matrix, denoted by D, whose rows are indexed by encoding rules and columns indexed by source states. Entries of matrix D are cryptograms and we have $D(e,s) = m$ if $e(s) = m$.

Theorem 10. *Let $P_1^{cp} = 1/(M-1)$ and $C = kM$. Then $E \geq M \times (M-1)$ and equality holds if and only if matrix D of the code is an ordered design $OD_1(2,k,M)$. In this case $P_0^{cp} = 1/M$ and the code provides perfect one-fold secrecy.*

Theorem 10 is an interesting characterisation of A-codes with perfect protection and the minimum number of encoding rules in terms of known combinatorial structures. Stinson [7] has proved similar results for cartesian A-codes and codes with secrecy for Simmons' model of attack.

Theorem 11 gives the main information theoretic bound on P_1^{cp}.

Theorem 11.

$$P_1^{cp} \geq 2^{-(H(\mathcal{E}|\mathcal{MS}) - H(\mathcal{E}|\mathcal{M}^2\mathcal{S}^2))}, \tag{13}$$

and equality holds if and only if

(i) $P_1^{cp} = payoff((m,s),(m',s')) = const$ for all (m,s) and (m',s') for which $P((m,s),(m',s')) > 0$;

(i) conditional source probability $P((m',s')|e,(m,s)$ is independent of e_i, where $e_i \in \mathcal{E}((m,s),(m',s'))$ for two arbitrary pairs $((m,s),(m',s'))$ for which $P((m,s),(m',s')) > 0$.

In the case of equality we have

$$P_1^{cp} = \frac{k-1}{U},$$

where for any (m,s), with $P(m,s) > 0$, U is the number of $(m',s'), m' \in \mathcal{M}, s' \in \mathcal{S}$ with $\mathcal{E}(m,s) \bigcap \mathcal{E}(m',s') \neq \emptyset\}$ and is independent of (m,s).

4.2 Construction of *A*-codes that provide resistance against chosen-content plaintext attack

In chosen-content plaintext enemy has more information (knows the content of the cryptogram) but a more difficult goal to achieve. In proposition 9 we noted that $P_1^{cp} \geq \frac{P_1}{k-1}$. Theorem 12 shows that an arbitrary A-code can be transformed into one for which P_1^{cp} is at its minimum and the enemy cannot benefit from the knowledge of the cryptogram content. The transformation replaces each row e_i of the encoding matrix with $k(k-1)$ rows, in a way similar to the one briefly described for theorem 5, but this time the restriction of the $k(k-1)$ rows to $M(e_i)$ is an $OD_1(2,k,k)$. Such designs exist when k is prime power [8].

Theorem 12. *Consider an (M,k,E) A-code with $P_1 = \epsilon$ with a uniform source and let k be a prime power. Then there exists an $(M,k,k(k-1)E)$ A-code with $P_1^{cp} = P_1/(k-1) = \epsilon/(k-1)$. If $C = kM$ then the code will have $P_0^c = 1/M$ and will provide perfect one-fold secrecy.*

Corollary 13. *In theorem 12 if $P_0 = k/M$ then the new code will have $P_0^c = 1/M$ and will provide perfect one-fold secrecy.*

An example of this construction is given below.

Example 3. The original code is a code without secrecy with $P_0 = P_1 = k/M$.

E/M	0	1	2	3	4	5
0	s_0	s_1	s_2	0	0	0
1	s_0	0	0	s_2	s_1	0
2	0	s_1	0	s_2	0	s_0
3	0	0	s_2	0	s_1	s_0

The new code preserves the properties of the original code but also provides perfect secrecy and $P_0^c = 1/M$, $P_1^{cp} = k/((k-1)M)$.

E/M	0	1	2	3	4	5
0	s_0	s_1	s_2	0	0	0
1	s_1	s_2	s_0	0	0	0
2	s_2	s_0	s_1	0	0	0
3	s_1	s_0	s_2	0	0	0
4	s_0	s_2	s_1	0	0	0
5	s_2	s_1	s_0	0	0	0
6	s_0	0	0	s_2	s_1	0
7	s_2	0	0	s_1	s_0	0
8	s_1	0	0	s_0	s_2	0
9	s_2	0	0	s_0	s_1	0
10	s_0	0	0	s_1	s_2	0

E/M	0	1	2	3	4	5
11	s_1	0	0	s_2	s_0	0
12	0	s_1	0	s_2	0	s_0
13	0	s_2	0	s_0	0	s_1
14	0	s_0	0	s_1	0	s_2
15	0	s_2	0	s_1	0	s_0
16	0	s_1	0	s_0	0	s_2
17	0	s_0	0	s_2	0	s_1
18	0	0	s_2	0	s_1	s_0
19	0	0	s_1	0	s_0	s_2
20	0	0	s_0	0	s_2	s_1
21	0	0	s_1	0	s_2	s_0
22	0	0	s_2	0	s_0	s_1
23	0	0	s_0	0	s_1	s_2

□

A-codes with $P_0^c = 1/M$ and $P_1^{cp} = 1/(M-1)$ and suitable parameters can be combined. The result is an A-code with the same values of P_0^c and P_1^{cp} for a larger source and hence increased efficiency. The composition is based on a method used by Bierbrauer et al [10] for composition of perpendicular arrays. Proposition 14 gives the details of this construction.

Proposition 14 [10]. *If an $OD_\lambda(2,k,M)$ and an $OD_\mu(2,\ell,M-k)$ exist, where $\ell > 1$ then there is an $OD_{\lambda\times\mu\times(M-k)(M-k-1)}(2,k+\ell,M)$.*

Hence having an (M,k,E_1) A-code and an $(M-k,\ell,E_2)$ A-code that provide perfect protection for chosen-content impersonation and chosen-content plaintext substitution implies existence of a $(M,k+\ell,E)$ A-code with $P_0^c = 1/M$ and $P_1^{cp} = 1/(M-1)$ and for which $S = k+l$. We note that P_0 and P_1 will increase for the new code.

References

1. G. Simmons, *A game theory model of digital message authentication*, Congressus Numerantium 34 (1982), 413-424.
2. G.J. Simmons, *Authentication theory/coding theory*, Lecture Notes in Comput. Sci. **196**, Proceedings of Crypto 84, Springer-Verlag , 1985, pp. 411-431.
3. C.E. Shannon, *Communication theory of secrecy systems*, Bell System Technical Journal, 28 (1949), 656-715.
4. B. Smeets, P. Vanrose, Zhe-Xian Wan, *On the construction of Authentication codes with secrecy and codes withstanding spoofing attack of order $L \geq 2$*, Lecture Notes in Comput. Sci. **473**, Proceedings of Eurocrypt '90, Springer-Verlag, (1990), 307-312.
5. D.R. Stinson, *Some constructions and bounds for authentication codes*, Journal of Cryptology 1, (1988), 37-51.
6. D.R. Stinson, *The combinatorics of authentication and secrecy codes*, Journal of Cryptology 2, (1990), 23-49.
7. D.R. Stinson, *Combinatorial characterization of authentication codes*, Lecture Notes in Comput. Sci. **576** , Proceedings of Crypto 91, Springer-Verlag , 1992, 62-72.
8. J.H. Dinitz, D. Stinson, *Contemporary Design Theory. A Collection of Surveys*, A Wiley Interscience Publications, JOHN WILEY & SONS, INC, 1992.
9. D. Pei *Information-Theoretic bounds for authentication codes and PBIB* , Proceedings Asiacrypt, (1991), Rump Session.
10. J. Bierbrauer, Y. Edel *Theory of Perpendicular Arrays*, submitted to Journal of Combinatorial Designs.
11. U. Rosenbaum, *A lower bound on authentication after having observed a sequence of messages*, Journal of Cryptology, No 3, Vol 6, (1993), 135-156.
12. T. Johansson, B. Smeets, G. Kabatianskii, *On the relation between A-codes and codes correcting independent errors*, Preproceedings of Eurocrypt '93, Norway, (1993), M1-M10.
13. Y. Desmedt, M. Yung, *Unconditional subliminal-freeness in unconditional authentication systems*, In preparation, Abstract appeared in Proceedings 1991 IEEE International Symposium on Information Theory, p. 176.