J. Cryptology (1988) 1: 37–51

Journal of Cryptology

Some Constructions and Bounds for Authentication Codes

D. R. Stinson
Department of Computer Science, University of Manitoba, Winnipeg, Manitoba, Canada

Abstract. We investigate authentication codes, using the model described by Simmons. We review and generalize bounds on the probability that an opponent can deceive the transmitter/receiver by means of impersonation or substitution. Also, we give several constructions for authentication codes that meet one or more of these bounds with equality. These constructions use combinatorial designs, such as transversal designs, group-divisible designs, and BIBDs (balanced incomplete block designs).

Key words. Authentication code, Combinatorial design.

1. Introduction

We shall use the model of authentication theory as described by Simmons [4]–[6]. In this model there are three participants: a transmitter, a receiver, and an opponent. The *transmitter* wants to communicate some information to the *receiver*, whereas the *opponent* wants to deceive the receiver. The opponent can either impersonate the receiver, making him accept a fradulent message as authentic, or modify a message which has been sent by the transmitter.

More formally, we have a set of source states **S**, a set of messages **M**, and a set of encoding rules **E**. A *source state* $s \in \mathbf{S}$ is the information that the transmitter wishes to communicate to the receiver. The transmitter and receiver will have secretly chosen an *encoding rule* $e \in \mathbf{E}$ beforehand. An encoding rule e will be used to determine the *message* $e(s)$ to be sent to communicate any source state s. It is possible that more than one message can be used to determine a particular source state (this is called *splitting*). However, in order for the receiver to be able to uniquely determine the source state from the message sent, there can be at most one source state which is encoded by any given message $m \in \mathbf{M}$.

We assume that the opponent will play either *impersonation* or *substitution*. When the opponent plays impersonation, he sends a message to the receiver, attempting to have the receiver accept the message as authentic. When the opponent plays substitution, he waits until a message m has been sent, and then replaces m with another message m' so that the receiver is misled as to the state of the source.

There will be a probability distribution on the set of source states **S**. Given the probability distribution on **S**, the receiver and transmitter will determine a probability distribution on **E**, called an *encoding strategy*. If splitting occurs, then they will also determine a *splitting strategy* to determine $m \in \mathbf{M}$, given $s \in \mathbf{S}$, and $e \in \mathbf{E}$. The

transmitter/receiver will choose the encoding and splitting strategies to minimize the chance that the opponent can deceive them.

This defines two possible games, which we refer to as the impersonation game and the substitution game. Each game has a *value*, which is the possibility that the opponent can deceive the transmitter/receiver, given that they are using the optimal encoding and spltting strategies. We denote the values of these games by v_I (for impersonation) and v_S (for substitution).

Many of these bounds depend on entropies of the various probability distributions. For a probability distribution on a set $\mathbf{X}$, we define the *entropy* of $\mathbf{X}$, $H(\mathbf{X})$, as follows:

$$H(\mathbf{X}) = -\sum_{x \in \mathbf{X}} p(x) \cdot \log p(x).$$

As well, the conditional entropy $H(\mathbf{X}|\mathbf{Y})$ is defined to be

$$H(\mathbf{X}|\mathbf{Y}) = -\sum_{y \in \mathbf{Y}} \sum_{x \in \mathbf{X}} p(y) \cdot p(x|y) \cdot \log p(x|y).$$

An authentication code is said to be *Cartesian* if every message uniquely determines the source state, independent of the particular encoding rule being used. In terms of entropy, this is expressed by the equation $H(\mathbf{S}|\mathbf{M}) = 0$. Note that in a Cartesian authentication code, there can be no secrecy.

In this paper we primarily consider authentication systems without splitting. We use the following notation. Denote the number of source states by k, and let $\mathbf{S} = \{s_i: 1 \leq i \leq k\}$. Denote the number of messages by v, and let $\mathbf{M} = \{m_j: 1 \leq j \leq v\}$. Denote by b the number of encoding rules, and write any encoding rule $e \in \mathbf{E}$ as $e = (e_i: 1 \leq i \leq k)$, where e_i is the message used to communicate source state s_i, for $1 \leq i \leq k$. Then, the authentication system can be represented by the $b \times k$ matrix A, where row e of A consists of the entries $e_1, \ldots, e_k$. Given an encoding rule $e \in \mathbf{E}$, we define $M(e) = \{e_i: 1 \leq i \leq k\}$, where $e = (e_i: 1 \leq i \leq k)$. Also, for each encoding rule e, define $f_e(m) = s$ if and only if $e_s = m$ (if message m does not occur in encoding rule e, then $f_e(m)$ is undefined).

2. Bounds on the Values of the Impersonation and Substitution Games

Several bounds have been proven on the values of the games v_I and v_S. In this section we review the known bounds, and prove some new results. We also determine some necessary conditions for the various bounds to be met with equality.

Theorem 2.1 [5, Theorem 1]. *In an authentication system without splitting,* $v_I \geq k/v$.

Proof. Suppose the opponent sends message m. We denote the probability that the message m is accepted by the receiver by payoff(m). Then we have that

$$\text{payoff}(m) = \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e).$$

It follows that

$$\sum_{m \in \mathbf{M}} \text{payoff}(m) = k.$$

Hence, there must be some m such that payoff$(m) \geq k/v$. That is, we have an impersonation strategy in which the transmitter/receiver can be deceived with probability at least k/v. □

It is clear from the above proof that $v_I = k/v$ if and only if $\sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) = k/v$ for all $m \in \mathbf{M}$.

Theorem 2.2 [5, Theorem 0]. *In any authentication system,* $v_I \geq 2^{H(\mathbf{MES})-H(\mathbf{E})-H(\mathbf{M})} = 2^{H(\mathbf{M}|\mathbf{ES})+H(\mathbf{S})-H(\mathbf{M})}$. *In an authentication system without splitting,* $H(\mathbf{M}|\mathbf{ES}) = 0$, *so* $v_I \geq 2^{H(\mathbf{S})-H(\mathbf{M})}$.

An authentication system which satisfies the bound of Theorem 2.2 with equality is said to be *perfect*. It is also possible to determine some properties of perfect authentication codes (see Theorem 1 of [1]). In the case of authentication codes without splitting, we must have the following:

Lemma 2.3. *In a perfect authentication code without splitting, the following properties hold*:

(i) *For all messages* m, $v_I = \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) = k/v$.
(ii) *For any message* m, $p(s)$ *is a constant for all* s *such that there is an* e *such that* $e_s = m$.

Let us next turn our attention to bounds on v_S. The following bound is for substitution with secrecy.

Theorem 2.4 [1, Theorem 3]. $v_S \geq 2^{-H(\mathbf{E}|\mathbf{M})} = 2^{H(\mathbf{M})-H(\mathbf{E})-H(\mathbf{S})+H(\mathbf{M}|\mathbf{ES})}$. *In an authentication system without splitting,* $H(\mathbf{M}|\mathbf{ES}) = 0$, *so* $v_S \geq 2^{H(\mathbf{M})-H(\mathbf{E})-H(\mathbf{S})}$.

Brickell has also given the conditions under which equality is attained in Theorem 2.4. We state these conditions in the case of codes without splitting.

Lemma 2.5. *If equality is attained in Theorem* 2.4 *for an authentication code without splitting, then the following properties are satisfied*:

(i) *For all* e, *and for all* m *such that* $m \in M(e)$, $p(m) \cdot v_S = p(e) \cdot p(S = f_e(m))$.
(ii) *For any* m *and* m', $m \neq m'$ *there is at most one* e *such that* $m, m' \in M(e)$.

The first thing we do is give a generalization of this bound, which will include cases where condition (ii) does not hold. We prove the bound in the case of authentication systems without splitting; the same bound holds for systems with splitting.

We require some notation. Given any encoding rule e', and given any m, $m' \in M(e')$, define

$$\delta(e', m, m') = \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m))/(p(e') \cdot p(S = f_{e'}(m))).$$

Then, let $\delta = \min\{\delta(e', m, m'): m, m' \in M(e'), m \neq m'\}$. Observe that $\delta \geq 1$, and $\delta = 1$ if and only if condition (ii) of Lemma 2.5 is satisfied.

Theorem 2.6. *In an authentication system without splitting,* $v_S \geq \delta \cdot 2^{-H(\mathbf{E}|\mathbf{M})}$, *where* δ *is defined as above.*

Proof. The proof is essentially the same as that of Theorem 3 of [1]. □

Suppose the opponent substitutes message m with message m' $(m \neq m')$. We denote the probability that the message m' is then accepted by the receiver by payoff(m, m'). Then we have that

$$\begin{aligned}\text{payoff}(m, m') &= \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m)) \Big/ \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \\ &= \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m))/p(m).\end{aligned}$$

If we define $v_S(m) = \max\{\text{payoff}(m, m'): m' \neq m\}$, then

$$v_S = \sum_{m \in \mathbf{M}} p(m) \cdot v_S(m).$$

For any m, m', and e', such that $m \neq m'$ and $m, m' \in M(e')$, observe that we have

$$\begin{aligned}v_S(m) &\geq \text{payoff}(m, m') = \delta(e', m, m') \cdot p(e') \cdot p(S = f_{e'}(m))/p(m) \\ &\geq \delta \cdot p(e') \cdot p(S = f_{e'}(m))/p(m)\end{aligned}$$

and, hence,

$$p(e') \cdot p(S = f_{e'}(m))/p(m) \leq v_S(m)/\delta.$$

Let us calculate $H(\mathbf{E}|\mathbf{M})$. By definition we have

$$\begin{aligned}H(\mathbf{E}|\mathbf{M}) &= -\sum_{m \in \mathbf{M}} \sum_{e \in \mathbf{E}} p(m) \cdot p(e|m) \cdot \log p(e|m) \\ &= -\sum_{m \in \mathbf{M}} \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(m|e) \cdot \log p(e|m) \\ &= -\sum_{m \in \mathbf{M}} \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \cdot \log(p(e) \cdot p(S = f_e(m))/p(m)) \\ &\geq -\sum_{m \in \mathbf{M}} \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \cdot \log(v_S(m)/\delta) \\ &= -\sum_{m \in \mathbf{M}} \log(v_S(m)/\delta) \cdot \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \\ &= -\sum_{m \in \mathbf{M}} \log(v_S(m)/\delta) \cdot p(m) \\ &\geq -\log\left(\sum_{m \in \mathbf{M}} v_S(m) \cdot p(m)/\delta\right) \\ &= -\log(v_S/\delta).\end{aligned}$$

Lemma 2.7. *In an authentication system without splitting, where* $v_S = \delta \cdot 2^{-H(\mathbf{E}|\mathbf{M})}$, *the following properties must be satisfied*:

(i) $\delta = \delta(e, m, m')$ *for any* m, m', *and* e *such that* $m \neq m'$ *and* $m, m' \in M(e)$.
(ii) *For any* m *and* e *with* $m \in M(e)$, $v_S = v_S(m) = \delta \cdot p(e) \cdot p(S = f_e(m))/p(m)$.

Using this information we can calculate δ and v_S in the case of equality. We do this as follows. From the definition of δ we have

$$\delta = \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m))/(p(e') \cdot p(S = f_{e'}(m)))$$

for any $m, m' \in M(e')$. It follows that $p(e) \cdot p(S = f_e(m))$ is a constant ($\neq 0$) for all e such that $m, m' \in M(e)$. Denote this common value by X. Then, we have

$$\delta \cdot X = \lambda(m, m') \cdot X, \qquad \text{where} \quad \lambda(m, m') = |\{e \in \mathbf{E}: m, m' \in M(e)\}|.$$

Hence,

$$\delta = \lambda(m, m'),$$

where m and m' are any two messages which occur in at least one encoding rule together.

Now, we calculate v_S. Recall that we have the relation

$$v_S = v_S(m) = \delta \cdot p(e) \cdot p(S = f_e(m))/p(m)$$

for any m and any e such that $m \in M(e)$. Define $r_m = |\{e \in \mathbf{E}: m \in M(e)\}|$ for any m. Now, fix any m, and sum this equation over all e such that $m \in M(e)$. We have

$$\sum_{\{e \in \mathbf{E}: m \in M(e)\}} v_S = \sum_{\{e \in \mathbf{E}: m \in M(e)\}} \delta \cdot p(e) \cdot p(S = f_e(m))/p(m),$$

$$r_m \cdot v_S = \delta.$$

Hence, r_m is a constant for any m. Hence, $r_m = r = b \cdot k/v$ for all m. Then, $v_S = \delta/r$.

Summarizing, we have

Theorem 2.8. *In an authentication code without splitting, where* $v_S = \delta \cdot 2^{-H(\mathbf{E}|\mathbf{M})}$, *the following properties are satisfied*:

(i) *For every message* m, $r_m = |\{e \in \mathbf{E}: m \in e\}| = r = b \cdot k/v$.
(ii) *For every pair of messages* m *and* m', *either* $\lambda(m, m') = |\{e \in \mathbf{E}: m, m' \in e\}| = 0$ *or* λ, *where* λ *is a constant.*
(iii) *For every* m, m', *and* e *such that* $m, m' \in e$, *we have* $\delta(e', m, m') = \lambda$.
(iv) $v_S = \delta/r$.

The value of v_S proved in (iv) is always a lower bound on v_S. We have the following:

Theorem 2.9. *In an authentication system without splitting,* $v_S \geq \delta/r$, *where* $r = \max\{r_m: m \in \mathbf{M}\}$.

Proof. Recall that, for any $m \in M(e')$, we have

$$p(m) \cdot v_S(m) \geq \delta \cdot p(e') \cdot p(S = f_{e'}(m)).$$

It then follows that

$$\begin{aligned} v_S &= \sum_{m \in \mathbf{M}} p(m) \cdot v_S(m) \\ &\geq \delta \cdot \sum_{m \in \mathbf{M}} (1/r_m) \cdot \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \\ &\geq \delta \cdot \sum_{m \in \mathbf{M}} (1/r) \cdot \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \\ &= (\delta/r) \cdot \sum_{e \in \mathbf{E}} p(e) \cdot \sum_{m \in M(e)} p(S = f_e(m)) \\ &= (\delta/r) \cdot \sum_{e \in \mathbf{E}} p(e) \\ &= \delta/r. \end{aligned}$$

□

Now, we prove a variation on the bound of Theorem 2.8. Given any encoding rule e', and given any $m, m' \in M(e')$, define

$$\gamma(e', m, m') = \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m))/p(e').$$

Then, let $\gamma = \min\{\gamma(e', m, m'): m, m' \in M(e'), m \neq m'\}$.

Theorem 2.10. *In an authentication system without splitting,* $v_S \geq \gamma \cdot 2^{H(\mathbf{M}) - H(\mathbf{E})}$, *where* γ *is defined as above.*

Proof. We use the same notation as before. For any m, m', and e' such that $m \neq m'$ and $m, m' \in M(e')$, observe that we have

$$v_S(m) \geq \text{payoff}(m, m') = \gamma(e', m, m') \cdot p(e')/p(m) \geq \gamma \cdot p(e')/p(m)$$

and, hence,

$$p(e')/p(m) \leq v_S(m)/\gamma.$$

Calculating $H(\mathbf{E}|\mathbf{M})$, we have

$$\begin{aligned} H(\mathbf{E}|\mathbf{M}) &= -\sum_{m \in \mathbf{M}} \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \cdot \log(p(e) \cdot p(S = f_e(m))/p(m)) \\ &\geq -\sum_{m \in \mathbf{M}} \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \cdot \log(v_S(m) \cdot p(S = f_e(m))/\gamma) \\ &= -\sum_{m \in \mathbf{M}} \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \cdot (\log(v_S(m)/\gamma) + \log(p(S = f_e(m)))) \\ &= -\sum_{m \in \mathbf{M}} \log(v_S(m)/\gamma) \cdot \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \\ &\quad -\sum_{m \in \mathbf{M}} \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \cdot \log(p(S = f_e(m))) \\ &= -\sum_{m \in \mathbf{M}} \log(v_S(m)/\gamma) \cdot p(m) \\ &\quad -\sum_{e \in \mathbf{E}} p(e) \cdot \sum_{\{m \in M(e)\}} p(S = f_e(m)) \cdot \log(p(S = f_e(m))) \\ &\geq -\log\left(\sum_{m \in \mathbf{M}} v_S(m) \cdot p(m)/\gamma\right) + \sum_{e \in \mathbf{E}} p(e) \cdot H(\mathbf{S}) \\ &= -\log(v_S/\gamma) + H(\mathbf{S}). \end{aligned}$$

Using the entropy identity $H(\mathbf{E}|\mathbf{M}) = H(\mathbf{S}) + H(\mathbf{E}) - H(\mathbf{M}) + H(\mathbf{M}|\mathbf{E}, \mathbf{S})$ (see Theorem 2 of [1]), we obtain

$$\log(v_S/\gamma) \geq -H(\mathbf{E}|\mathbf{M}) + H(\mathbf{S}) + H(\mathbf{M}|\mathbf{E}, \mathbf{S}) = H(\mathbf{M}) - H(\mathbf{E}) + H(\mathbf{M}|\mathbf{E}, \mathbf{S}).$$

In an authentication system without splitting, $H(\mathbf{M}|\mathbf{E}, \mathbf{S}) = 0$, so we get

$$\log(v_S/\gamma) \geq H(\mathbf{M}) - H(\mathbf{E}).$$

Hence, $v_S \geq \gamma \cdot 2^{H(\mathbf{M})-H(\mathbf{E})}$. □

We have the following consequences of equality in this bound.

Lemma 2.11. *In an authentication system without splitting, where* $v_S = \gamma \cdot 2^{H(\mathbf{M})-H(\mathbf{E})}$, *the following properties are satisfied*:

(i) *For all* e', m, *and* m', *where* $m, m' \in M(e')$, *we have* $\gamma = \gamma(e', m, m')$.
(ii) *For any* e *and any* $m \in M(e)$, $v_S = v_S(m) = \gamma \cdot p(e)/p(m)$.

Property (ii) says that given an encoding rule e, $p(m)$ is constant for all messages $m \in M(e)$, and given a message m, then $p(e)$ is constant for all encoding rules e with $m \in M(e)$. Suppose we construct a graph with vertex set $\mathbf{E}$, and join two vertices e and e' by an edge if and only if there is an $m \in M(e) \cap M(e')$. If this graph has more than one connected component, then the authentication system can be considered to be the "union" of the authentication codes corresponding to each component. We will only consider authentication codes where this graph has one connected component; such codes will be called *connected*. Hence, we have the following.

Theorem 2.12. *In a connected authentication system without splitting, where* $v_S = \gamma \cdot 2^{H(\mathbf{M})-H(\mathbf{E})}$, *we must have* $H(\mathbf{M}) = \log v$, $H(\mathbf{E}) = \log b$, and $v_S = \gamma \cdot v/b$.

Proof. All encoding rules must have the same probability, so $p(e) = 1/b$ for every $e \in \mathbf{E}$. It then follows that $p(m) = 1/v$ for every $m \in \mathbf{M}$. Hence, $v_S = \gamma \cdot v/b$. □

We can now determine some relations between γ and probabilities of source states.

Lemma 2.13. *In a connected authentication system without splitting, where* $v_S = \gamma \cdot 2^{H(\mathbf{M})-H(\mathbf{E})}$, *the following properties are satisfied*:

(1) *For every* e', m, *and* m' *with* $m, m' \in M(e')$, $\gamma = \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(S = f_e(m))$.
(2) *For every* m, $\sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(S = f_e(m)) = b/v$.
(3) *For every* m, *there are precisely* X *messages* m' *with which* m *occurs in at least one encoding rule, where* $X = b \cdot (k - 1)/(v \cdot \gamma)$.

Proof. By property (i) of Lemma 2.11, for every e', m, and m' with $m, m' \in M(e')$, we must have

$$\begin{aligned}\gamma &= \gamma(e', m, m') \\ &= \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m))/p(e') \\ &= \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(S = f_e(m)),\end{aligned}$$

proving (1). Also, we can calculate

$$\begin{aligned}1/v &= p(m) \\ &= \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \\ &= (1/b) \cdot \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(S = f_e(m)),\end{aligned}$$

so, for every m, we have

$$\sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(S = f_e(m)) = b/v.$$

This proves (2).

Finally, pick any m, and let $\mathbf{M}' = \{m' \in \mathbf{M}$: there exists an e with $m, m' \in M(e)\}$. Then, we calculate

$$\begin{aligned}|\mathbf{M}'| \cdot \gamma &= \sum_{m' \in \mathbf{M}'} \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(S = f_e(m)) \\ &= (k-1) \cdot \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(S = f_e(m)) \\ &= b \cdot (k-1)/v.\end{aligned}$$

The result (3) follows. □

Next, we prove another new bound on the value of the substitution game in an authentication code without splitting.

Theorem 2.14. *In an authentication system without splitting,* $v_S \geq (k-1)/(v-1)$.

Proof. Suppose the opponent substitutes message m with message m' $(m \neq m')$. We denote the probability that the message m' is then accepted by the receiver by payoff(m, m'). As before, we have that

$$\text{payoff}(m, m') = \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m)) \Big/ \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)).$$

It follows that

$$\sum_{m' \neq m} \text{payoff}(m, m') = k - 1.$$

Hence, there must be some m_0 such that payoff$(m, m_0) \geq (k-1)/(v-1)$. For every m, determine such an m_0. This defines a substitution strategy in which the transmitter/receiver can be deceived with probability at least $(k-1)/(v-1)$. □

3. Constructions for Authentication Systems

Our interest in this section is in constructing authentication systems which meet one or more of the bounds of the previous section with equality. We are interested in the existence of authentication codes with a specified number of source states, and specified upper bounds on the number of encoding rules, messages, v_I, and v_S. Therefore, we define an AC(k, v, b, α, β) to be an authentication code with k source states, at most v messages, and at most b encoding rules, and where $v_I \leq \alpha$ and $v_S \leq \beta$.

Then, we define

$$\varepsilon(k, \alpha, \beta) = \min\{b: \text{there exists an AC}(k, v, b, \alpha, \beta)\}$$

and

$$v(k, \alpha, \beta) = \min\{v: \text{there exists an AC}(k, v, b, \alpha, \beta)\}.$$

That is, we are attempting to minimize the number of encoding rules (or messages) required in an authentication code for k source states, with upper bounds α and β on the impersonation and substitution games, respectively.

First, observe that we have an easy lower bound on $v(k, \alpha, \beta)$.

Theorem 3.1. $v(k, \alpha, \beta) \geq \max\{k/\alpha, 1 + (k-1)/\beta\}$.

Proof. This is an immediate corollary of Theorems 2.1 and 2.14. □

Next, we mention a lower bound on $\varepsilon(k, \alpha, \beta)$ due to Brickell [1, Theorem 4].

Theorem 3.2. $\varepsilon(k, \alpha, \beta) \geq 1/(\alpha \cdot \beta)$.

This bound can be strengthened, using the quantity δ defined in Section 2.

Theorem 3.3. *If an* AC(k, v, b, α, β) *exists, then* $b \geq \delta/(\alpha \cdot \beta)$.

Proof. We have $\alpha \geq v_I \geq 2^{H(\mathbf{S})-H(\mathbf{M})}$ and $v_S \geq \delta \cdot 2^{-H(\mathbf{E}|\mathbf{M})} = \delta \cdot 2^{H(\mathbf{M})-H(\mathbf{E})-H(\mathbf{S})}$. Hence, we have $\alpha \cdot \beta \geq \delta \cdot 2^{-H(\mathbf{E})}$. Since $H(\mathbf{E}) \leq \log b$, the result follows. □

In the remainder of this paper we describe constructions for authentication codes, which will enable us to put upper bounds on ε and v. For our first construction we require the following definition. A *transversal design* TD$(k, \lambda; n)$ is a triple $(X, \mathbf{G}, \mathbf{A})$, which satisfies the following properties:

(1) X is a set of $k \cdot n$ elements called *points*.
(2) $\mathbf{G}$ is a partition of X into k subsets of n points, called *groups*.
(3) $\mathbf{A}$ is a set of $\lambda \cdot n^2$ subsets of X (called *blocks*) such that a group and a block contain at most one common point.
(4) Every pair of points from distinct groups occurs in exactly λ blocks.

We usually denote a TD$(k, 1; n)$ by TD(k, n). It is well known that a TD(k, n) is equivalent to $k - 2$ mutually orthogonal Latin squares of order n.

Theorem 3.4 [1, Theorems 5 and 6]. *If there is a transversal design* TD(k, n) *then there is a Cartesian authentication system with* $v_S = 2^{-H(\mathbf{E}|\mathbf{M})} = 1/n$, $v_I = 2^{H(\mathbf{S})-H(\mathbf{M})} = 1/n$, $|\mathbf{S}| = k$, $|\mathbf{M}| = k \cdot n$, *and* $|\mathbf{E}| = n^2$, *with no splitting. (Hence, if there exists a* TD(k, n), *then there is an* AC(k, $k \cdot n$, n^2, $1/n$, $1/n$), *and we have* $\varepsilon(k, 1/n, 1/n) \leq n^2$ *and* $v(k, 1/n, 1/n) \leq k \cdot n$.) *Conversely, if there is a Cartesian authentication system with no splitting with* $v_S = 2^{-H(\mathbf{E}|\mathbf{M})} = \alpha$, $v_I = 2^{H(\mathbf{S})-H(\mathbf{M})} = \alpha$, *and* $|\mathbf{S}| = k$, *then* $n = 1/\alpha$ *is an integer and there exists a transversal design* TD(k, n).

We can prove a generalization of this result, using transversal designs with $\lambda \geq 1$.

Theorem 3.5. *If there is a transversal design* TD(k, λ; n), *then there is a Cartesian authentication system with* $v_S = \lambda \cdot 2^{-H(\mathbf{E}|\mathbf{M})} = 1/n$, $v_I = 2^{H(\mathbf{S})-H(\mathbf{M})} = 1/n$, $|\mathbf{S}| = k$, $|\mathbf{M}| = k \cdot n$, *and* $|\mathbf{E}| = \lambda \cdot n^2$, *with no splitting. (Hence, if there exists a* TD(k, λ; n), *then there exists an authentication code* AC(k, $k \cdot n$, $\lambda \cdot n^2$, $1/n$, $1/n$), $\varepsilon(k, 1/n, 1/n) \leq \lambda \cdot n^2$, *and* $v(k, 1/n, 1/n) \leq k \cdot n$.) *Conversely, if there is a Cartesian authentication system with no splitting with* $v_S = \delta \cdot 2^{-H(\mathbf{E}|\mathbf{M})} = \alpha$, $v_I = 2^{H(\mathbf{S})-H(\mathbf{M})} = \alpha$, *and* $|\mathbf{S}| = k$, *then* $n = 1/\alpha$ *is an integer and there exists a transversal design* TD(k, δ; n).

Proof. The proof is essentially the same as the proof of Theorems 5 and 6 of [1]. First, it is not difficult to see that the TD(k, λ; n) gives rise to the desired authentication code, by associating each group of the transversal design with a particular source state, and using each encoding rule with probability $1/(\lambda \cdot n^2)$, as in [1]. Let us prove the converse assertion.

Since we have assumed that we have a Cartesian authentication code, each message m determines a unique source state s_m. Thus, the probability that message m is sent is

$$\begin{aligned} p(M = m) &= p(S = s_m) \cdot \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \\ &= p(S = s_m) \cdot v_I \qquad \text{(from Lemma 2.3(i))} \\ &= p(S = s_m) \cdot \alpha. \end{aligned}$$

On the other hand, given any m and e with $m \in M(e)$, we have, from Lemma 2.7(ii), that

$$\begin{aligned} \alpha = v_S &= \delta \cdot p(e) \cdot p(S = s_m)/p(m) \\ &= \delta \cdot p(e)/\alpha, \end{aligned}$$

from above. So, $p(e)$ is independent of e, and $p(e) = 1/b = \alpha^2/\delta$ for every e. Hence, $b = \delta/\alpha^2$.

Since the system is Cartesian, and every message occurs in $r = b \cdot k/v$ encoding rules (Theorem 2.8(i)), we can partition the message space into k subsets M_i ($1 \leq i \leq k$), each of size v/k. However,

$$\begin{aligned} v/k &= b/r \\ &= (\delta/\alpha^2)/(\delta/\alpha) \qquad \text{(since } r = \delta/\alpha, \text{ by Theorem 2.8(iv))} \\ &= 1/\alpha \end{aligned}$$

is an integer, so $\alpha = 1/n$, where $n = v/k$. Then, $b = \delta \cdot n^2$ and $r = \delta \cdot n$.

Now, using Theorem 2.8(ii) and (iii), it is easy to see that any two messages from different M_i's occur together in precisely δ encoding rules (see Theorem 2.6 of [1]). Hence, we can construct the desired transversal design. □

Suppose our desire is to construct an authentication code AC($k, k \cdot n, b, 1/n, 1/n$). We can construct such a code if a TD($k, \lambda; n$) exists, for $b = \lambda \cdot n^2$. (Note that this satisfies the bound $b \geq \delta/(\alpha \cdot \beta)$ with equality, where $\alpha = \beta = 1/n$ and $\delta = \lambda$.) Thus, given k and n, we are interested in the smallest λ such that a TD($k, \lambda; n$) exists. First, we observe that there is a simple numerical bound on k in terms of λ and n.

Theorem 3.6 [3]. *If a* TD($k, \lambda; n$) *exists, then* $k \leq (\lambda \cdot n^2 - 1)/(n - 1)$.

Consequently, if we use a TD($k, \lambda; n$), then we have a lower bound on b, namely

$$b = \lambda \cdot n^2 \geq kn - k + 1.$$

We present an infinite example of transversal designs which meet this bound with equality.

Theorem 3.7. *For all prime powers* $n \geq 2$, *and for any* $d \geq 1$, *there is an* AC($k, k \cdot n, n^d, 1/n, 1/n$), *where* $k = (n^d - 1)/(n - 1)$; *hence* $\varepsilon((n^d - 1)/(n - 1), 1/n, 1/n) \leq n^d$ *and* $v((n^d - 1)/(n - 1), 1/n, 1/n) \leq k \cdot n$.

Proof. In [3] Hanani shows that for any prime power n, and for any $d \geq 1$, there exists a TD($(n^d - 1)/(n - 1), n^{d-2}; n$). □

Corollary 3.8. *For any* $\alpha > 0$, $\varepsilon(k, \alpha, \alpha)$ *is* $O(k/\alpha^2)$ *and* $v(k, \alpha, \alpha)$ *is* $O(k/\alpha)$.

Proof. Let $n = 2^j$, where $2^j \geq 1/\alpha \geq 2^{j-1}$. Then n is $O(1/\alpha)$. Now, choose d so that $n^d \geq k(n - 1) + 1 > n^{d-1}$. Since $k \leq (n^d - 1)/(n - 1)$, we have $\varepsilon(k, \alpha, \alpha) \leq n^d$. But, $n^d \leq k(n^2 - n) + n = O(k \cdot n^2)$. Since n is $O(1/\alpha)$, therefore $\varepsilon(k, \alpha, \alpha)$ is $O(k/\alpha^2)$. Also, $k \cdot n$ is $O(k/\alpha)$. □

As another example of the use of transversal designs with $\lambda > 1$, let us consider codes with parameters AC($k, v, b, \frac{1}{6}, \frac{1}{6}$). For $k = 4$, we cannot construct such a code from a TD(4, 6), since this TD does not exist (this is the famous 36 officers problem of Euler, i.e., a (nonexistent) pair of orthogonal Latin squares of order 6). In [1] Brickell constructs an example of an AC(4, 30, 36, $\frac{1}{6}, \frac{1}{6}$) with splitting. However, we can employ a TD(7, 2, 6), which is constructed in [3, p. 49], to obtain an AC(7, 42, 72, $\frac{1}{6}, \frac{1}{6}$).

More generally, we have the following class of authentication codes with seven source states.

Theorem 3.9. *For all* $n \geq 2$, *there is an* AC($7, 7 \cdot n, 2n^2, 1/n, 1/n$); *hence* $\varepsilon(7, 1/n, 1/n) \leq 2n^2$.

Proof. For these n, there is a TD(7, 2; n) (see [3]). □

The authentication codes obtained from Theorem 3.5 are Cartesian. Hence, the opponent, on seeing a message being sent, knows the source state. Therefore, no secrecy is possible in such an authentication system. We also want to be able to construct good authentication codes with secrecy. Ideally, we would like to have $H(\mathbf{S}|\mathbf{M}) = H(\mathbf{S})$; i.e., the message gives absolutely no clue as to the state of the source. If this happens, then we say that the authentication code is *perfectly* non-Cartesian.

Our main construction for perfectly non-Cartesian authentication codes uses group-divisible designs, which are a generalization of transversal designs. A *group-divisible design* GD(k, λ, n; v) is a triple (X, $\mathbf{G}$, $\mathbf{A}$) which satisfies the following four properties:

(1) X is a set of v elements called *points*.
(2) $\mathbf{G}$ is a partition of X into v/n subsets of n points, called *groups*.
(3) $\mathbf{A}$ is a set of subsets of X (called *blocks*), each of size k, such that a group and a block contain at most one common point.
(4) Every pair of points from distinct groups occurs in exactly λ blocks.

Note that a TD(k, λ; n) is equivalent to a GD(k, λ, n; $k \cdot n$). Also, a (v, b, r, k, λ)-BIBD (balanced incomplete block design) is equivalent to a GD(k, λ, 1; v).

We have the following construction:

Theorem 3.10. *Suppose there exists a* GD(k, λ; n; v). *Then there is a perfectly non-Cartesian* AC(k, v, $\lambda \cdot v \cdot (v-n)/(k-1)$, k/v, $(k-1)/(v-n)$).

Proof. Let (X, $\mathbf{G}$, $\mathbf{A}$) be a GD(k, λ, n; v). By simple counting, each point occurs in $r = \lambda \cdot (v-n)/(k-1)$ blocks, and the total number of blocks is $\lambda \cdot v \cdot (v-n)/(k \cdot (k-1))$. What we do is construct k encoding rules from every block of the group-divisible design: for each block $A = \{x_1, \ldots, x_k\}$ of the group-divisible design, and for each i, $0 \le i \le k-1$, we define an encoding rule $e(A, i) = (e_j: 1 \le j \le k)$, where $e_j = x_{(j+i) \text{ modulo } k}$.

There are $b = \lambda \cdot v \cdot (v-n)/(k-1)$ encoding rules in the resulting authentication code. We shall use each encoding rule with probability $(k-1)/(\lambda \cdot v \cdot (v-n))$.

Let us first verify that $v_1 = k/v$. Let m be any message. There are $r = \lambda \cdot k \cdot (v-n)/(k-1)$ blocks A containing m. For each such A, $m \in M(e(A, i))$ for every i, $1 \le i \le k$. We calculate

$$\text{payoff}(m) = r/b = k/v,$$

as desired.

Next, we verify that $v_S = (k-1)/(v-n)$. Let m and m' be two distinct messages. If m and m' are in different groups, then there are no encoding rules that contain m and m', so payoff(m, m') = 0. If m and m' are in the same group, then there are λ blocks A for which m, $m' \in A$. For each such block A, and for each source state j, there is exactly one encoding rule $e(A, i)$ where m, $m' \in M(e(A, i))$ and $f_{e(A,i)}(m) = j$.

Then,

$$\begin{aligned}\text{payoff}(m, m') &= \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m)) \Big/ \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) \\ &= \sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(S = f_e(m)) \Big/ \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(S = f_e(m)) \\ &= \lambda / r \\ &= \lambda / (\lambda \cdot (v - n)/(k - 1)) \\ &= (k - 1)/(v - n),\end{aligned}$$

as desired.

Finally, the authentication code is perfectly non-Cartesian since $p(s|m) = p(s)$ for every $s \in \mathbf{S}$ and every $m \in \mathbf{M}$. □

It is interesting to note that this code has

$$H(\mathbf{M}) = \log v, \qquad H(\mathbf{E}) = \log(\lambda \cdot v \cdot (v - n)/(k - 1)),$$

and $v_S = \gamma \cdot 2^{H(\mathbf{M}) - H(\mathbf{E})}$, where $\gamma = \lambda$ (in Theorem 2.10).

Corollary 3.11. *Suppose there exists a* (v, b, r, k, λ)-BIBD. *Then there is a perfectly non-Cartesian* $\text{AC}(k, v, k \cdot b, k/v, (k - 1)/(v - 1))$.

Proof. This is the case where every group of the group-divisible design has size 1. Note that here we have $v_S = (k - 1)/(v - 1)$, so the bound of Theorem 2.14 is tight.

Corollary 3.12. *Suppose there is a* $\text{TD}(k, \lambda; n)$. *Then there is a perfectly non-Cartesian* $\text{AC}(k, n \cdot k, \lambda \cdot k \cdot n^2, 1/n, 1/n)$.

Consequently, $\varepsilon(k, \alpha, \alpha)$ is $O(k^2/\alpha^2)$ and $v(k, \alpha, \alpha)$ is $O(k^2/\alpha)$, even if we restrict ourselves to perfectly non-Cartesian codes.

These constructions of Theorems 3.5 and 3.10 both have two very nice properties which we have not yet emphasized. First, the encoding strategy in each case is *uniform*: each encoding rule is used with equal probability $1/b$. Second, this encoding strategy yields the stated game values for *any* source distribution.

The final topic we consider is the construction of authentication codes for *uniform* source distributions ($p(s) = 1/k$ for any source state s). This topic was first investigated in [1], where some constructions were given using balanced incomplete block designs. As before, we consider only codes without splitting. The best we could hope for is to attain the bounds $v_I = k/v$ and $v_S = (k - 1)/(v - 1)$. So, we shall study $\text{AC}(k, v, b, k/v, (k - 1)/(v - 1))$; such authentication codes will be called *optimal*.

We have the following characterization of authentication codes which are optimal with respect to the uniform probability distribution on the source states.

Lemma 3.13. *An authentication system is optimal with respect to the uniform probability distribution on the source states if and only if the following properties are*

satisfied:

(i) *For every* $m \in \mathbf{M}$, $\sum_{\{e \in \mathbf{E}: e \in E\}} p(e) = k/v$.
(ii) *For every* $m \neq m'$, $\sum_{\{e \in \mathbf{E}: m, m' \in e\}} p(e) = (k^2 - k)/(v^2 - v)$.

Proof. (i) is given in Theorem 2.1. From Theorem 2.14, $v_S = (k-1)/(v-1)$ if and only if, for every $m \neq m$, we have

$$\sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \cdot p(S = f_e(m)) \Big/ \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) \cdot p(S = f_e(m)) = (k-1)/(v-1).$$

Since the source distribution is uniform, this is equivalent to

$$\sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) \Big/ \sum_{\{e \in \mathbf{E}: m \in M(e)\}} p(e) = (k-1)/(v-1).$$

Using (i), we obtain

$$\sum_{\{e \in \mathbf{E}: m, m' \in M(e)\}} p(e) = (k^2 - k)/(v^2 - v),$$

as desired. □

In many authentication codes, the optimal encoding strategy is to choose every encoding rule with probability $1/b$. If we assume that this encoding strategy is in fact optimal, then the properties above are of a purely combinatorial nature. We have the following:

Theorem 3.14. *An authentication system is optimal with respect to a uniform encoding strategy and a uniform probability distribution on the source states if and only the following properties are satisfied*:

(i) *For every* $m \in \mathbf{M}$, $|\{e \in \mathbf{E}: m \in e\}| = k \cdot b/v$.
(ii) *For every* $m \neq m'$, $|\{e \in \mathbf{E}: m, m' \in e\}| = b \cdot (k^2 - k)/(v^2 - v)$.

This says that the rows of $\mathbf{E}$, considered as unordered sets, form a balanced incomplete block design with parameters (v, b, r, k, λ), where $r = k \cdot b/v$ and $\lambda = b \cdot (k^2 - k)/(v^2 - v)$. So, we can produce optimal authentication codes from BIBDs when the source states are equiprobable.

Using known families of BIBDs, we can obtain many authentication codes for uniform source distributions. For example, using projective geometries, we have the following:

Theorem 3.15. *For any prime power* n, *and any integer* $d \geq 2$, *there is an optimal authentication code for the uniform source distribution on* $n + 1$ *source states, for* $v = (n^{d+1} - 1)/(n - 1)$ and $\lambda = 1$.

References

[1] E. F. Brickell, A few results in message authentication, *Congressus Numerantium*, **43** (1984), 141–154.

[2] E. N. Gilbert, F. J. MacWilliams, and N. J. A. Sloane, Codes which detect deception, *Bell System Tech. J.*, **53** (1974), 405–424.
[3] H. Hanani, On transversal designs, *Math. Centre Tracts*, **55** (1974), 42–52.
[4] G. J. Simmons, A game theory model of digital message authentication, *Congressus Numerantium*, **34** (1982), 413–424.
[5] G. J. Simmons, Message authentication: a game on hypergraphs, *Congressus Numerantium*, **45** (1984), 161–192.
[6] G. J. Simmons, Authentication theory/coding theory, in *Advances in Cryptology: Proceedings of CRYPTO 84*, Lecture Notes in Computer Science, Vol. 196, Springer-Verlag, Berlin, 1985, pp. 411–432.