Information Theoretic Bounds on Authentication Systems in Query Model

Reihaneh Safavi-Naini
School of IT and CS
University of Wollongong
Wollongong, Australia
rei@uow.edu.au

Peter Wild
Information Security Group
Royal Holloway University of London
Egham, UK
P.Wild@rhul.ac.uk

November 20, 2006

Abstract

Authentication codes provide message integrity guarantees in an information theoretic sense within a symmetric key setting. Information theoretic bounds on the success probability of an adversary who has access to previously authenticated messages have been derived by Simmons and Rosenbaum, among others. In this paper we consider a strong attack scenario where the adversary is adaptive and has access to authentication and verification oracles. We derive information theoretic bounds on the success probability of the adversary and on the key size of the code. This brings the study of unconditionally secure authentication systems on a par with the study of computationally secure ones. We characterize the codes that meet these bounds and compare our result with the earlier ones.

Keywords: Unconditional security, authentication system, A-codes, verification oracle.

1 Introduction

Unconditionally secure authentication systems provide message integrity when the adversary's computational power is unknown or unlimited. Unconditional security is particularly important when recent advances in quantum computing and prospect of discovery and realization of efficient algorithms for solving 'hard' problems, is taken into account.

In an *Authentication code (A-code)* [1, 10] authenticated messages (*ciphertexts*) encode states of an information source (referred to as *plaintexts* or *source states*) by a mapping determined by the shared key (also called the *encoding rule*). A-codes are symmetric key systems. The receiver verifies the authenticity of a message using the same key as the sender. In a *spoofing attack of order i* a message-observing adversary observes i authenticated messages transmitted by the sender and then tries to construct a fraudulent message called the *spoofing message*, that will be accepted by the receiver. We do not make any assumptions about the computational power of an adversary.

The performance of an A-code is measured by the probability that the spoofing message is accepted by the receiver. Information theoretic bounds [10, 7, 6] for A-codes give fundamental limits on performance of the codes. Rosenbaum [7] and Pei [6] independently derived a bound on the success probability of attackers in spoofing of order i and employed the bound to derive a lower bound on the key size of A-codes.Extensions and alternative derivations of these bounds are given by a number of authors such as [2, 4].

We extend this analysis by considering adversaries that may be proactive in gathering information. The adversary might obtain information from the sender by having the sender transmit a message corresponding to source state of the adversary's choosing or might obtain information from the receiver by sending a message of the adversary's choosing and observing whether or not the receiver accepts it. Safavi-Naini *et al* [8] have considered A-codes with such an adversary in the context of unconditionally secure digital signature schemes (USDS, Shikata *et al* [9]). This situation, the *query model*, is modelled in terms of an Authentication Oracle (A-oracle) that provides the authenticated message corresponding a to query source state (an A-query) in the same way that the sender would and a Verification Oracle (V-oracle) that provides a response *accept* or *reject* to a query message (a V-query) according as the receiver would or would not accept the message. This terminology parallels that used for schemes relying on computational security. An attack with access to an A-oracle corresponds to an *adaptive chosen plaintext attack* and an attack with access to a V-oracle corresponds to an *adaptive chosen ciphertext attack*.

In this paper we study unconditionally secure A-codes (symmetric key) under the query model and derive information theoretic bounds on the success probability of a query attacker.

1.1 Our results

We consider an adversary who asks exactly i queries, observing the responses of the oracle, and then spoofs. We analyse this via an experiment in which the adversary uses a strategy to choose each query adaptively, taking into account all queries and responses previously observed.

Bounds on success probability. We derive information theoretic bounds on the success probably of the adversary in a general query, response attack model. This can be seen as a generalisation of the Rosenbaum-Pei lower bound [7, 6] for spoofing of order i.

Constructing pure optimal strategies. An adversary's success chance is maximized when he uses an optimal strategy. Finding optimal strategies requires solving a constrained maximisation problem, taking into account all possible sequences of query and responses, and results in an adversary strategy that is represented by a sequence of probability distributions. We show that there always exists an optimal strategy for which each of these probability distributions has the property that there is a unique element with non-zero probability. Such a strategy is called *pure* and we give a recursive algorithm that constructs this pure optimal strategy.

Queries do not decrease success chance. It is known that an adversary's expected chance of spoofing may decrease if the adversary observes a message, compared to his expected chance of spoofing when he spoofs without any observation. A natural question, noting the adversary's control on the choice of the query, is whether it is always beneficial to the adversary to ask a query if possible. We show that as long as there are 'good' queries, asking them will not reduce the average success chance (it may improve it) and so they should be asked. Thus there is no requirement for the adversary to compute the probabilities in order to make a decision on whether to query or not in each particular case. In particular, in the case of verification queries for non-trivial schemes, it is always a better strategy to ask the query and then spoof than to spoof without querying.

Bounds on key entropy. In the case of authentication queries we use the bound on the probability of success to derive a bound on the entropy of the key space. We establish a combinatorial characterisation of the authentication schemes attaining the bound which shows that optimal codes in the authentication query model are also optimal codes in the message observing model.

2 Definitions

A symmetric key authentication system provides integrity guarantees for two parties, referred to as the *sender* and the *receiver*, that share a secret key.

A (symmetric) authentication system consists of two algorithms $\Pi = (\mathsf{Auth}, \mathsf{Ver})$ and is defined over three sets, $\mathcal{S}$, $\mathcal{M}$, and $\mathcal{E}$, called plaintexts(or *source state*), ciphertexts (or *message*) and keys, respectively. The authentication algorithm $\mathsf{Auth}(e, s)$ takes a key e and a plaintext s and generates a ciphertext m. We consider authentication systems *without splitting* in which a key and a message determine a single ciphertext and the algorithm Auth is a one-to-one function. (For authentication systems *with splitting* a pair (e, s) determines a subset of $\mathcal{M}$.) The verification algorithm $\mathsf{Ver}(e, m)$, takes a key e and a ciphertext m and returns a single bit b. $\mathsf{Ver}(e, m)$ is defined in terms of Auth as follows: $\mathsf{Ver}(e, m) = 1$ if $\mathsf{Auth}(e, s) = m$, for some $s \in \mathcal{S}$ and $\mathsf{Ver}(e, m) = 0$, otherwise. We have that for all $e \in \mathcal{E}$ and $s \in \mathcal{S}$, it holds that $\mathsf{Ver}(e, \mathsf{Auth}(e, s)) = 1$. We may also define a decryption function $\mathsf{D}(e, m)$ which satisfies $\mathsf{D}(e, (\mathsf{Auth}(e, s)) = s$ for all $e \in \mathcal{E}$ and $s \in \mathcal{S}$.

The sender and receiver use a probability distribution $p(e)$ over $\mathcal{E}$ to select a secret key. To send a source state $s \in \mathcal{S}$ to the receiver, s is encoded under e to produce message $m = \mathsf{Auth}(e, s)$. The probability distribution $p(e)$ is called the *communicants' strategy* and is assumed to be public. The sender uses the key to authenticate a sequence of plaintexts arising from the source according to a specified probability distribution, also assumed to be public, and transmits the corresponding sequence of ciphertexts to the receiver. The probability distributions on $\mathcal{E}$ and $\mathcal{S}$ together determine a probability distribution on $\mathcal{M}$ given by $p(m) = \sum_{e\in\mathcal{E}, s\in\mathcal{S}:\mathsf{Auth}(e,s)=m} p(e)p(s)$. A ciphertext $m \in \mathcal{M}$ is *valid* for $e \in \mathcal{E}$ if and only if $\mathsf{Ver}(e, m) = 1$

We denote by E, S, M the random variables on sample spaces $\mathcal{E}, \mathcal{S}, \mathcal{M}$, respectively, corresponding to these probability distributions. We assume that the source produces a sequence $s_1, s_2, s_3, \ldots$ of distinct source states so that $p(s|s_1, \ldots, s_j)$, the probability that the next source state is s given that the sequence $s_1, \ldots, s_j$ has arisen so far, is 0 if $s \in \{s_1, \ldots, s_j\}$.

An authentication code can be represented by a matrix with $|\mathcal{M}|$ columns labeled by the messages and $|\mathcal{E}|$ rows labelled by the encoding rules. The entry in row e and column m is $\mathsf{D}(e, m)$ if $\mathsf{Ver}(e, m) = 1$ and 0 otherwise. Alternatively, we may define its *encoding matrix*. This matrix has $|\mathcal{S}|$ columns labelled by the source states and has $|\mathcal{E}|$ rows labelled by the encoding rules. The entry in row e and column s is the message $m = \mathsf{Auth}(e, s)$. The set of elements in row e is denoted $\mathcal{M}(e)$ and has exactly $|\mathcal{S}|$ distinct messages. Thus, for an encoding rule e, a message m is valid if and only if $m \in \mathcal{M}(e)$.

It is assumed that the adversary knows the encoding matrix but does not know the actual secret encoding rule agreed upon by the sender and the receiver.

2.1 Adversaries and success probability

The traditional adversary model for an authentication code is an adversary who has access to i authenticated messages and attempts to construct a forged message (also called a *spoofing message*), that would be accepted by the receiver. The best success probability, P_i, of the adversary in the above attack gives a measure of the security provided by the code.

Simmons [10] derived an information theoretic bound on the success probability of a spoofer when $i = 0$ (an *impersonation attack*). Rosenbaum [7] and Pei [6] independently derived a general form of the bound for $i > 0$.

$$P_i \geq 2^{H(E|M^{i+1})-H(E|M^i)} = 2^{-I(E;M|M^i)} \tag{1}$$

where M^i denotes a random variable associated with the sequence of observed messages $\mathbf{m}^i = m_1 m_2 \ldots m_i$.

The bound can be used to derive an information theoretic bound on the entropy or uncertainty (and hence the length) of the key in terms of the success probabilities $P_0, \ldots, P_i$.

Adversary with Oracle access In Safavi-Naini *et al* [8] the adversary model was strengthened by the addition of access to authentication and verification oracles. This new adversary is adaptive and can ask authentication and verification queries from corresponding oracles. The oracles produce responses to queries in the same way that the legitimate sender and receiver would.

For a given $e \in \mathcal{E}$ an *authentication oracle* $\mathsf{Auth}(e,.)$ (also called a *signature oracle*) takes as input a source state $s \in \mathcal{S}$, computes $m = \mathsf{Auth}(e,s)$ and returns response m.

A *verification oracle*, $\mathsf{Ver}(e,.)$, receives as input a ciphertext $m \in \mathcal{M}$, computes $b = \mathsf{Ver}(e,m) \in \{0,1\}$ and returns response b.

We use q to denote a query (either an A-query s or a V-query m) and use r to denote a response (either m or b). We denote the set of queries by $\mathcal{Q}$ and the set of responses by $\mathcal{R}$. Thus $\mathcal{Q} = \mathcal{S}$ and $\mathcal{R} = \mathcal{M}$ for A-queries and $\mathcal{Q} = \mathcal{M}$ and $\mathcal{R} = \{0,1\}$ for V-queries. Let $\mathbf{x}^i = x_1, x_2 \ldots x_i$ denote a sequence of i elements. We also use $\mathbf{x}^i$ to denote the set $\{x_1, x_2 \ldots x_i\}$. For $j < i$ we say $\mathbf{y}^j$ is a prefix of $\mathbf{x}^i$ if $y_l = x_l, l = 1, \ldots, j$. We use $\mathbf{s}^i$, $\mathbf{m}^i$, $\mathbf{q}^i, \mathbf{r}^i, (\mathbf{q},\mathbf{r})^i$ to denote a sequence of source states, messages, queries, responses, and query and response pairs, respectively. A strategy τ of an adversary is a collection of probability distributions that is used to select the queries and also the final spoofing query. If $(\mathbf{q},\mathbf{r})^j$ is a sequence of query and response pairs then $\tau_{(\mathbf{q},\mathbf{r})^j}(q)$ denotes the probability that q is chosen as the next query (or spoofing message) given that the adversary has asked the sequence $\mathbf{q}^i$ of queries and observed the sequence $\mathbf{r}^i$ of responses. An adversary's strategy is a collection τ of probability distributions $\tau_{(\mathbf{q},\mathbf{r})^j}$.

Let $F_{a,\tau}^{\mathsf{Auth}(e,.),\mathsf{Ver}(e,.)}(i,1)$ be an adversary that has access to both oracles, and uses a strategy τ to adaptively ask i A-queries (receiving the response after each query) and then ask a V-query. Let $F_{v,\tau}^{\mathsf{Ver}(e,.)}(i+1)$ be an adversary that has access to only a verification oracle, and uses a strategy τ to adaptively ask i verification queries (receiving the response after each query) and then ask a further verification query.

Consider the following experiments:

Experiment $\mathbf{Exp}_{\Pi,F_{a,\tau}(i,1)}$
$e \leftarrow \mathcal{E}$
If after asking *exactly* i queries $\mathbf{s}^i$ of $\mathsf{Auth}(e,.)$
and receiving corresponding responses $\mathbf{m}^i$
$F_{a,\tau}^{\mathsf{Auth}(e,.),\mathsf{Ver}(e,.)}$ makes a query m to
the oracle $\mathsf{Ver}(e,.)$ such that the return
$\mathsf{Ver}(e,m) = 1$, and
m had never been returned by
the oracle $\mathsf{Auth}(e,.)$
then return 1 **else** return 0

Experiment $\mathbf{Exp}_{\Pi,F_{v,\tau}(i+1)}$
$e \leftarrow \mathcal{E}$
If after asking *exactly* i queries $\mathbf{m}^i$ of $\mathsf{Ver}(e,.)$
and receiving corresponding responses b^i
$F_{v,\tau}^{\mathsf{Ver}(e,.)}$ makes a query m to
the oracle $\mathsf{Ver}(e,.)$ such that the return
$\mathsf{Ver}(e,m) = 1$, and
m was never asked of
the oracle $\mathsf{Ver}(e,.)$
then return 1 **else** return 0

The output of the experiment is a random variable ($\mathbf{Exp}_{\Pi,F_{a,\tau}(i,1)}$ or $\mathbf{Exp}_{\Pi,F_{v,\tau}(i+1)}$ on $\{0,1\}$) corresponding to the success or failure of an adversary who makes i queries to an oracle before spoofing with a (forged) message m. Each run of the experiment is one instance of the communicants' strategy and the adversary's choices in his attack, and corresponds to a sample point in a probability space where sample points are labelled with a key (the communicants' choice) and the adversary's sequence of query and response pairs $(\mathbf{q},\mathbf{r})^i, (m,b)$, where $\mathbf{r}^i$ is the sequence of responses to the sequence of queries $\mathbf{q}^i$ ($\mathbf{s}^i$ or $\mathbf{m}^i$) and $b = \mathsf{Ver}(e,m)$. Let $p_i^\tau(e,(\mathbf{q},\mathbf{r})^i)$ denote the probability of the sample point with label $e, (\mathbf{q},\mathbf{r})^i$. For the sample points in which the forgery is considered successful the experiment results in 1.

We compute the *advantage* of the forgers in the above experiments as the the probability of the experiment resulting in 1:

$$\mathbf{Adv}_{\Pi,F_{a,\tau}(i,1)} = Pr[\mathbf{Exp}_{\Pi,F_{a,\tau}(i,1)} = 1]$$
$$\mathbf{Adv}_{\Pi,F_{v,\tau}(i+1)} = Pr[\mathbf{Exp}_{\Pi,F_{v,\tau}(i+1)} = 1]$$

This is the average success probability over all keys (with respect to the distribution given by the communicants' strategy) when the adversary uses strategy τ. We use the notation P_i^τ for either of these quantities. This can be written as

$$P_i^\tau = \sum_{q_1\in\mathcal{Q}} \tau(q_1) \sum_{r_1\in\mathcal{R}} p(r_1|q_1) \sum_{q_2\in\mathcal{Q}} \tau_{(\mathbf{q},\mathbf{r})^1}(q_2) \sum_{r_2\in\mathcal{R}} p(r_2|q_2,(\mathbf{q},\mathbf{r})^1)\dots$$

$$\sum_{q_i\in\mathcal{Q}} \tau_{(\mathbf{q},\mathbf{r})^{i-1}}(q_i) \sum_{r_i\in\mathcal{R}} p(r_i|q_i,(\mathbf{q},\mathbf{r})^{i-1}) \sum_{m\in\mathcal{M}} \tau_{(\mathbf{q},\mathbf{r})^i}(m) \sum_{e\in\mathcal{E},\mathsf{Ver}(e,m)=1} p(e|m,(\mathbf{q},\mathbf{r})^i)$$

The *advantage function of* Π is the advantage of a forger with the highest success probability.

$$\mathbf{Adv}_{\Pi,a}(i,1) = \max_{F_{a,\tau}(i,1)} Pr[\mathbf{Exp}_{\Pi,F_{a,\tau}(i,1)} = 1]$$
$$\mathbf{Adv}_{\Pi,v}(i+1) = \max_{F_{v,\tau}(i+1)} Pr[\mathbf{Exp}_{\Pi,F_{v,\tau}(i+1)} = 1]$$

The strategy of a forger with this advantage is called an *optimal strategy*. We write P_i for either of these quantities.

We say that a key e is *consistent* with a query and response pair (q,r) if the following holds: (i) if $(q,r) = (s,m)$ then it holds that $\mathsf{Auth}(e,s) = m$; and (ii) if $(q,r) = (m,b), b \in \{0,1\}$, then it holds that $\mathsf{Ver}(e,m) = b$. Let $\mathcal{E}((\mathbf{q},\mathbf{r})^j)$ denote the set of keys that are consistent with a query response sequence $(\mathbf{q},\mathbf{r})^j$. The conditional probability, $p(e|(\mathbf{q},\mathbf{r})^j)$, that the key is e given the sequence of query and response pairs $(\mathbf{q},\mathbf{r})^j$ is non-zero only if $e \in \mathcal{E}((\mathbf{q},\mathbf{r})^j)$. Similarly, for $m \in \mathcal{M}$, let $\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1)) = \{e \in \mathcal{E}((\mathbf{q},\mathbf{r})^i) : m \in \mathcal{M}(e)\}$. For $e \in \mathcal{E}$ we put: $\gamma(e,m,(\mathbf{q},\mathbf{r})^i) = 1$ if $e \in \mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))$ and $m \notin \mathbf{q}^i \cup \mathbf{r}^i$; and $\gamma(e,m,(\mathbf{q},\mathbf{r})^i) = 0$ otherwise.

In the following we assume $e \in \mathcal{E}$ is chosen by communicants and is unknown to the adversary. We introduce some notation that records probabilities of certain events in the experiments described above. Put

$$p_i^\tau((\mathbf{q},\mathbf{r})^j) = \prod_{l=1}^{j} \tau_{(\mathbf{q},\mathbf{r})^{l-1}}(q_l)p(r_l|q_l,(\mathbf{q},\mathbf{r})^{l-1}) \tag{2}$$

where $p(r_l|q_l,(\mathbf{q},\mathbf{r})^{l-1}) = \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^{l-1}):\mathsf{Ver}(e,q_l)=r_l} p(e|(\mathbf{q},\mathbf{r})^{l-1})$ is the conditional probability that r_l is the response to query q_l given that $\mathbf{r}^{l-1}$ is the sequence of responses to the sequence of queries $\mathbf{q}^{l-1}$. This is the probability in the above experiments of the instances that have $(e,(\mathbf{q},\mathbf{r})^j)$ as the prefix to their label. This probability may be calculated from the communicants' strategy, $p(e)$, the adversary's strategy, τ, and the authentication system Π. This probability satisfies the following recursion.

$$p_i^\tau((\mathbf{q},\mathbf{r})^j) = p_i^\tau((\mathbf{q},\mathbf{r})^{j-1})\tau_{(\mathbf{q},\mathbf{r})^{j-1}}(q_j)p(r_j|q_j,(\mathbf{q},\mathbf{r})^{j-1}). \tag{3}$$

Put

$$P_i^\tau((\mathbf{q},\mathbf{r})^i),(m,1)) = \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i)} p(e|(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(\mathbf{q},\mathbf{r})^i). \tag{4}$$

This is the probability that the spoofing message m is successful given the sequence $(\mathbf{q},\mathbf{r})^i$. Then

$$P_i^\tau((\mathbf{q},\mathbf{r})^i) = \sum_{m\in\mathcal{M}} \tau_{(\mathbf{q},\mathbf{r})^i}(m) P_i^\tau((\mathbf{q},\mathbf{r})^i),(m,1)). \tag{5}$$

Put

$$P_i^\tau((\mathbf{q},\mathbf{r})^j, q_{j+1}) = \sum_{r_{j+1}\in\mathcal{R}} p(r_{j+1}|q_{j+1},(\mathbf{q},\mathbf{r})^j) \sum_{q_{j+2}\in\mathcal{Q}} \tau_{(\mathbf{q},\mathbf{r})^{j+1}}(q_{j+2})$$

$$\cdots \sum_{q_i\in\mathcal{Q}} \tau_{(\mathbf{q},\mathbf{r})^{i-1}}(q_i) \sum_{r_i\in\mathcal{R}} p(r_i|q_i,(\mathbf{q},\mathbf{r})^{i-1}) \sum_{m\in\mathcal{M}} \tau_{(\mathbf{q},\mathbf{r})^i}(m) P_i^\tau((\mathbf{q},\mathbf{r})^i),(m,1)).$$

and

$$P_i^\tau((\mathbf{q},\mathbf{r})^j) \;=\; \sum_{q_{j+1}\in\mathcal{Q}} \tau_{(\mathbf{q},\mathbf{r})^j}(q_{j+1}) P_i^\tau((\mathbf{q},\mathbf{r})^j, q_{j+1}). \tag{6}$$

This is the conditional probability in the above experiments that the instances with labels whose sequences of queries and responses are prefixed by $(\mathbf{q},\mathbf{r})^j)$ output 1.

For each $j = 1,\ldots,i$, the advantage of the adversary may be written

$$P_i^\tau = \sum_{(\mathbf{q},\mathbf{r})^j} p_i^\tau((\mathbf{q},\mathbf{r})^j) P_i^\tau((\mathbf{q},\mathbf{r})^j). \tag{7}$$

We write P_i to denote the maximum of P_i^τ over all strategies τ.

3 Optimal strategies

The adversary's strategy is a collection of probability distributions $\tau_{(\mathbf{q},\mathbf{r})^j}$ that are used to choose queries. We say that $\tau_{(\mathbf{q},\mathbf{r})^j}$ is *pure* if there is a unique query, denoted $q_{(\mathbf{q},\mathbf{r})^j}$, with $\tau_{(\mathbf{q},\mathbf{r})^j}(q_{(\mathbf{q},\mathbf{r})^j}) = 1$ (so that $\tau_{(\mathbf{q},\mathbf{r})^j}(q) = 0$ for all other queries q). We say that a *strategy* τ *is pure* if $\tau_{(\mathbf{q},\mathbf{r})^j}$ is pure for each $(\mathbf{q},\mathbf{r})^j$.

Theorem 3.1 *There always exists a* pure *optimal strategy for an authentication system adversary who has oracle access.*

Proof: We prove the theorem by constructing a pure strategy that is optimal. Let the adversary have i oracle queries, and a single spoofing query. Given any strategy τ of this adversary, we show how to construct a pure strategy τ^i whose advantage is at least that of τ, using i recursive steps. Towards this end, we recursively determine strategies $\tau^0,\ldots,\tau^i$ such that the advantage of an adversary with strategy τ^0 is at least that of an adversary with strategy τ, the advantage, for $j = 1,\ldots,i$, of an adversary with strategy τ^j is at least that of an adversary with strategy τ^{j-1}, and $\tau^j_{(\mathbf{q},\mathbf{r})^l}$ is pure for all $(\mathbf{q},\mathbf{r})^l$, $l = i,\ldots,i-j$. Since τ^i is pure this will establish our claim by taking τ to be an optimal strategy.

Now τ^0 will differ from τ only in the distributions $\tau_{(\mathbf{q},\mathbf{r})^i}$ used to choose the spoofing message. Thus we are considering instances with labels whose sequences of queries and responses have a given prefix $(\mathbf{q},\mathbf{r})^i$. For each $(\mathbf{q},\mathbf{r})^i$ let $m_{(\mathbf{q},\mathbf{r})^i}$ be such that

$$P_i^\tau((\mathbf{q},\mathbf{r})^i),(m_{(\mathbf{q},\mathbf{r})^i},1)) = \max\nolimits_{m\in\mathcal{M}} P_i^\tau((\mathbf{q},\mathbf{r})^i),(m,1)). \tag{8}$$

Put $\tau^0_{(\mathbf{q},\mathbf{r})^i}(m_{(\mathbf{q},\mathbf{r})^i}) = 1$ (and zero otherwise). Thus the strategy $\tau^0_{(\mathbf{q},\mathbf{r})^i}$ is a pure strategy that chooses the spoofing message optimally. For $j < i$ put $\tau^0_{(\mathbf{q},\mathbf{r})^j} = \tau_{(\mathbf{q},\mathbf{r})^j}$ for all $(\mathbf{q},\mathbf{r})^j$.

Hence τ^0 is a strategy that is identical to τ except that each $\tau^0_{(\mathbf{q},\mathbf{r})^i}$ is a special probability distribution with non-zero probability only corresponding to an optimal spoofing query $m_{(\mathbf{q},\mathbf{r})^i}$. Note that from (2) we have $p_i^{\tau}((\mathbf{q},\mathbf{r})^i) = p_i^{\tau^0}((\mathbf{q},\mathbf{r})^i)$ for all $(\mathbf{q},\mathbf{r})^i$. Also note that from (8) we have $P_i^{\tau}((\mathbf{q},\mathbf{r})^i),(m,1)) \leq P_i^{\tau^0}((\mathbf{q},\mathbf{r})^i),(m_{(\mathbf{q},\mathbf{r})^i},1))$ for all $m \in \mathcal{M}$ so that

$$\begin{aligned} P_i^{\tau}((\mathbf{q},\mathbf{r})^i) &= \sum_{m\in\mathcal{M}} \tau_{(\mathbf{q},\mathbf{r})^i}(m) P_i^{\tau}((\mathbf{q},\mathbf{r})^i),(m,1)) \leq \sum_{m\in\mathcal{M}} \tau_{(\mathbf{q},\mathbf{r})^i}(m) P_i^{\tau^0}((\mathbf{q},\mathbf{r})^i),(m_{(\mathbf{q},\mathbf{r})^i},1)) \\ &= P_i^{\tau^0}((\mathbf{q},\mathbf{r})^i),(m_{(\mathbf{q},\mathbf{r})^i},1)) = \sum_{m\in\mathcal{M}} \tau^0_{(\mathbf{q},\mathbf{r})^i}(m) P_i^{\tau^0}((\mathbf{q},\mathbf{r})^i),(m,1)) = P_i^{\tau^0}((\mathbf{q},\mathbf{r})^i). \end{aligned}$$

where the last equality uses equation (5).

Now $P_i^{\tau} = \sum_{(\mathbf{q},\mathbf{r})^i} p_i^{\tau}((\mathbf{q},\mathbf{r})^i) P_i^{\tau}((\mathbf{q},\mathbf{r})^i) \leq \sum_{(\mathbf{q},\mathbf{r})^i} p_i^{\tau^0}((\mathbf{q},\mathbf{r})^i) P_i^{\tau^0}((\mathbf{q},\mathbf{r})^i) = P_i^{\tau^0}$ and the probability of success using strategy τ^0 is at least that using τ.

For $j = 1,\ldots,i$ we recursively define τ^j assuming $\tau^u, u = 0,\cdots j-1$ are defined. Now τ^j will differ from τ^{j-1} only in the distributions $\tau^j_{(\mathbf{q},\mathbf{r})^{i-j}}$. That is, for all $l = 0,\cdots i, l \neq j$, we have $\tau^j_{(\mathbf{q},\mathbf{r})^{i-l}} = \tau^{j-1}_{(\mathbf{q},\mathbf{r})^{i-l}}$ for all $(\mathbf{q},\mathbf{r})^{i-l}$, $l = 0,\cdots j-1$ and so these strategies are pure. In this step $\tau^{j-1}_{(\mathbf{q},\mathbf{r})^{i-j}}$ is replaced by a pure distribution such that $P_i^{\tau^{j-1}} \leq P_i^{\tau^j}$.

Consider instances with labels whose sequences of queries and responses have prefix $(\mathbf{q},\mathbf{r})^{i-j}$.

For each $(\mathbf{q},\mathbf{r})^{i-j}$ let $q_{(\mathbf{q},\mathbf{r})^{i-j}}$ be such that

$$\begin{aligned} & \sum_{r\in\mathcal{R}} p(r|q_{(\mathbf{q},\mathbf{r})^{i-j}},(\mathbf{q},\mathbf{r})^{i-j}) P_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j},(q_{(\mathbf{q},\mathbf{r})^{i-j}},r)) \\ = \; & \max_{q\in\mathcal{Q}} \sum_{r\in\mathcal{R}} p(r|q,(\mathbf{q},\mathbf{r})^{i-j}) P_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j},(q,r)). \end{aligned} \tag{9}$$

Put $\tau^j_{(\mathbf{q},\mathbf{r})^{i-j}}(q_{(\mathbf{q},\mathbf{r})^{i-j}}) = 1$ (and zero otherwise). For $l \neq j$ put $\tau^j_{(\mathbf{q},\mathbf{r})^{i-l}} = \tau^{j-1}_{(\mathbf{q},\mathbf{r})^{i-l}}$ for all $(\mathbf{q},\mathbf{r})^{i-l}$. Then from equation (2) we have $p_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j}) = p_i^{\tau^j}((\mathbf{q},\mathbf{r})^{i-j})$ for all $(\mathbf{q},\mathbf{r})^{i-j}$. Also note that from (9) we have $P_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j},q) \leq P_i^{\tau^j}((\mathbf{q},\mathbf{r})^{i-j},q_{(\mathbf{q},\mathbf{r})^{i-j}})$ for all $q \in \mathcal{Q}$ so that

$$\begin{aligned} P_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j}) &= \sum_{q\in\mathcal{Q}} \tau^{j-1}_{(\mathbf{q},\mathbf{r})^{i-j}}(q) P_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j}),q) \leq \sum_{q\in\mathcal{Q}} \tau^{j-1}_{(\mathbf{q},\mathbf{r})^{i-j}}(q) P_i^{\tau^j}((\mathbf{q},\mathbf{r})^{i-j},q_{(\mathbf{q},\mathbf{r})^{i-j}}) \\ &= P_i^{\tau^j}((\mathbf{q},\mathbf{r})^i,q_{(\mathbf{q},\mathbf{r})^i}) = \sum_{q\in\mathcal{Q}} \tau^j_{(\mathbf{q},\mathbf{r})^{i-j}}(q) P_i^{\tau^j}((\mathbf{q},\mathbf{r})^{i-j},q) = P_i^{\tau^j}((\mathbf{q},\mathbf{r})^{i-j}). \end{aligned}$$

Now

$$P_i^{\tau^{j-1}} = \sum_{(\mathbf{q},\mathbf{r})^{i-j}} p_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j}) P_i^{\tau^{j-1}}((\mathbf{q},\mathbf{r})^{i-j}) \leq \sum_{(\mathbf{q},\mathbf{r})^{i-j}} p_i^{\tau^j}((\mathbf{q},\mathbf{r})^{i-j}) P_i^{\tau^j}((\mathbf{q},\mathbf{r})^{i-j}) = P_i^{\tau^j}$$

and the probability of success using strategy τ^j is at least that using τ^{j-1}.

The strategy τ^i is a pure strategy and is optimal if τ is optimal. We have $\tau^i_{(\mathbf{q},\mathbf{r})^i}(m_{(\mathbf{q},\mathbf{r})^i}) = 1$ and $P_i^{\tau^i}((\mathbf{q},\mathbf{r})^i) = \max_{m\in\mathcal{M}} P_i^{\tau^i}((\mathbf{q},\mathbf{r})^i,(m,1))$ and $m_{(\mathbf{q},\mathbf{r})^i}$ is an optimal spoofing message. Moreover, for each $j = 1,\ldots,i$, we have $\tau^i_{(\mathbf{q},\mathbf{r})^{i-j}}(q_{(\mathbf{q},\mathbf{r})^{i-j}}) = 1$ and

$$\begin{aligned} & \sum_{r\in\mathcal{R}} p_i^{\tau^i}(r|q_{(\mathbf{q},\mathbf{r})^{i-j}},(\mathbf{q},\mathbf{r})^{i-j}) P_i^{\tau^i}((\mathbf{q},\mathbf{r})^{i-j},(q_{(\mathbf{q},\mathbf{r})^{i-j}},r)) \\ = \; & \max_{q\in\mathcal{Q}} \sum_{r\in\mathcal{R}} p_i^{\tau^i}(r|q,(\mathbf{q},\mathbf{r})^{i-j}) P_i^{\tau^i}((\mathbf{q},\mathbf{r})^{i-j},(q,r)). \end{aligned}$$

Thus $\tau^i_{(\mathbf{q},\mathbf{r})^{i-j}}$ is an optimal query. ■

3.1 Bounds on success probability

We consider an adversary that uses strategy τ to make i queries (either A-queries or V-queries) to an oracle and then constructs a spoofing query. Now $p_i^\tau((\mathbf{q},\mathbf{r})^i)$ is the probability that the resulting sequence of query and response pairs is $(\mathbf{q},\mathbf{r})^i$. We write $(Q^\tau, R^\tau)^i$ to denote a random variable that takes the values $(\mathbf{q},\mathbf{r})^i$ with respective probabilities $p_i^\tau((\mathbf{q},\mathbf{r})^i)$. The source distribution and the communicants' strategy determine a distribution $p(m)$ on $\mathcal{M}$. Let $p(m|(\mathbf{q},\mathbf{r})^i)$ be the probability that a message m is transmitted by the sender as the next message given the sequence of query and response pairs $(\mathbf{q},\mathbf{r})^i$.

Theorem 3.2 *Let Π be an authentication system and let P_i^τ be the probability of success of an adversary who spoofs optimally after making i oracle queries using strategy τ. Then*

$$P_i^\tau \geq 2^{H(E|M,(Q^\tau,R^\tau)^i)-H(E|(Q^\tau,R^\tau)^i)} = 2^{-I(E;M|(Q^\tau,R^\tau)^i)}.$$

Moreover, equality holds if and only if, for all $(\mathbf{q},\mathbf{r})^i \in (\mathcal{Q}\times\mathcal{R})^i$ with $p_i^\tau((\mathbf{q},\mathbf{r})^i) \neq 0$ and all $m \in M$ with $p(m|(\mathbf{q},\mathbf{r})^i) \neq 0$, we have $P_i^\tau((\mathbf{q},\mathbf{r})^i,(m,1)) = P_i^\tau$ and $p(m|e,(\mathbf{q},\mathbf{r})^i)$ is constant for all $e \in \mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))$. ■

The proof of Theorem 3.2 is given in the appendix. It is an adaptation of the proof of Theorem 3.1 of Rosenbaum [7]. This bound is analogous to Rosenbaum's bound for a message-observing adversary, with difference being that the distribution associated with the random variable $(Q^\tau, R^\tau)^i$ depends on both the adversary's and the communicants' strategies while in the message observing case it is determined by communicants' strategy and the source state distribution. This means that unlike the case of a message observing adversary where the best success chance is bounded by the quantity $I(E;M|M^i)$ that he cannot change, the case of an adaptive adversary with access to query oracles allows him to influence $I(E;M|(Q^\tau,R^\tau)^i)$ and so have a higher bound on the success chance. The bound depends on the query strategy and applies to any spoofing strategy. For good authentication systems the best spoofing strategy should meet the bound with equality.

As in Rosenbaum [7] we have the following generalisation. Let the values $p^*(e,m,(\mathbf{q},\mathbf{r})^i)$ for $e \in \mathcal{E}$, $m \in \mathcal{M}$ and sequence $(\mathbf{q},\mathbf{r})^i$ of query and response pairs be a joint probability distribution on $\mathcal{E}\times\mathcal{M}\times(\mathcal{Q}\times\mathcal{R})^i$ such that, if $\gamma(e,m,(\mathbf{q},\mathbf{r})^i) = 0$ then $p^*(e,m,(\mathbf{q},\mathbf{r})^i) = 0$ and, for all e and $(\mathbf{q},\mathbf{r})^i$, $\sum_m p^*(e,m,(\mathbf{q},\mathbf{r})^i) = p_i^\tau(e,(\mathbf{q},\mathbf{r})^i)$, the probability that the encoding rule is e and, for strategy τ, the sequence of query and response pairs is $(\mathbf{q},\mathbf{r})^i$. We write M^* to denote a random variable that takes values $m \in \mathcal{M}^*$ with respective probabilities $p^*(m) = \sum_{e,(\mathbf{q},\mathbf{r})^i} p^*(e,m,(\mathbf{q},\mathbf{r})^i)$. The bound in Theorem 3.2 becomes

$$P_i^\tau \geq 2^{H(E|M^*,(Q^\tau,R^\tau)^i)-H(E|(Q^\tau,R^\tau)^i)} = 2^{-I(E;M^*|(Q^\tau,R^\tau)^i)} \tag{10}$$

3.2 Authentication queries

When the adversary has access to an authentication oracle the sample points have labels with sequences of queries and responses of the form $((\mathbf{s},\mathbf{m})^i,(m,b))$ and the probability of output 1 for a given sequence of queries and responses is given by

$$P_i^\tau((\mathbf{s},\mathbf{m})^i),(m,1)) = \sum_{e\in\mathcal{E}(\mathbf{s},\mathbf{m})^i)} p(e|(\mathbf{s},\mathbf{m})^i)\gamma(e,m,(\mathbf{s},\mathbf{m})^i)$$

where $\mathcal{E}((\mathbf{s},\mathbf{m})^i)$ is the set of keys that satisfy $\mathsf{Auth}(e,s_j) = m_j, j = 1,\cdots i$. Theorem 3.2 in its generalized form (expression (10)) can be written as follows.

Theorem 3.3 *Let Π be an authentication system and let P_i^τ be the probability of success of an adversary who uses strategy τ and spoofs optimally after making i oracle queries. Then*

$$P_i^\tau \geq 2^{H(E|M^*,(S^\tau,M^\tau)^i)-H(E|(S^\tau,M^\tau)^i)} = 2^{-I(E;M^*|(S^\tau,M^\tau)^i)}.$$

Moreover, equality holds if and only if, for all $(\mathbf{s},\mathbf{m})^i \in (\mathcal{S}\times\mathcal{M})^i$ with $p_i^\tau((\mathbf{s},\mathbf{m})^i) \neq 0$ and all $m \in M$ with $p^*(m|(\mathbf{s},\mathbf{m})^i) \neq 0$, we have $P_i^\tau((\mathbf{s},\mathbf{m})^i,(m,1)) = P_i^\tau$ and $p^*(m|e,(\mathbf{s},\mathbf{m})^i)$ is constant for all $e \in \mathcal{E}((\mathbf{s},\mathbf{m})^i,(m,1))$. ∎

3.2.1 Bound on the key size

We prove a bound on the key size in terms of the probabilities of success P_j for an adversary that has access to j authentication queries, $j = 0,\dots,i$. Let τ be a strategy such that for all $(\mathbf{s},\mathbf{m})^j$, $j = 0,\dots,i$, we have $\tau_{(\mathbf{s},\mathbf{m})^j}(s) = 0$ whenever $s \in \mathbf{s}^i$. We define $i+1$ adversaries $F_{\tau^j} = F_{a,\tau^j}^{\mathsf{Auth}(e,.),\mathsf{Ver}(e,.)}(i,1)$, $j = 0,\dots,i$, with strategies $\tau^0,\dots,\tau^i$ where $\tau^i = \tau$. For $j = 0,\dots,i$, F_{τ^j} uses strategy τ^j to ask j oracle queries and then uses an optimal spoofing strategy that gives him the best success chance. For $j = 0,\dots,i$ define $\tau^j_{(\mathbf{s},\mathbf{m})^\ell} = \tau_{(\mathbf{s},\mathbf{m})^\ell}$ for all $(\mathbf{s},\mathbf{m})^\ell$, $\ell = 0,\dots,j-1$. Note that $\tau^j_{(s,m)^j}$ is the probability distribution used by F_{τ^j} to select the spoofing message. Also note that for all $u = 1,\dots,j-1$, the probability distribution used by the adversary F_{τ^u} for selecting queries s_ℓ, $\ell = 1,\dots,u$, is the same as the probability distribution used by the adversary F_{τ^j}.

The lower bound on the success probability of F_{τ^j} is,

$$P_j^{\tau^j} \geq 2^{-I(E;M^*|(S^{\tau^j},M^{\tau^j})^j)}$$

where $(S^{\tau^j},M^{\tau^j})^j$ denotes the random variable that takes values $(\mathbf{s},\mathbf{m})^j$ with probability $p_j^{\tau^j}((\mathbf{s},\mathbf{m})^j)$ and M^* is a random variable that has the properties described at the end of section 3.1. Since $p_j^{\tau^j}((\mathbf{s},\mathbf{m})^j) = p_j^\tau((\mathbf{s},\mathbf{m})^j)$ we have $(S^{\tau^j},M^{\tau^j})^j = (S^\tau,M^\tau)^j$.

Let

$$p^*(e,m,(\mathbf{s},\mathbf{m})^j) \quad = \quad p_j^{\tau^j}((\mathbf{s},\mathbf{m})^j)p(e|(\mathbf{s},\mathbf{m})^j)\tau^{j+1}_{(\mathbf{s},\mathbf{m})^j}(s)$$

where $s = \mathsf{D}(e,m)$ and $p^*(e,m,(\mathbf{s},\mathbf{m})^j) = 0$ otherwise. This is the distribution on query and response pairs $(\mathbf{s},\mathbf{m})^j$, keys e, and transmitted messages m, that arises when the source state distribution is $\tau^{j+1}_{(s,m)^j}$. Then $p^*(e,m,(\mathbf{s},\mathbf{m})^j) = 0$ if $\gamma(e,(\mathbf{s},\mathbf{m})^j,m) = 0$.

Now $\tau^{j+1}_{(s,m)^j}(s)$ is the distribution used by $F_{\tau^{j+1}}$ for selection of the $j+1$ query s^{j+1}. The random variable $M_{j+1}^{\tau^{j+1}}$ has distribution satisfying

$$\begin{aligned} p(M_{j+1}^{\tau^j} = m|(\mathbf{s},\mathbf{m})^j) &= \sum_{s\in\mathcal{S}}\ \sum_{e\in\mathcal{E}:\mathsf{Auth}(e,s)=m} \tau^{j+1}_{(\mathbf{s},\mathbf{m})^j}(s)p(e|(\mathbf{s},\mathbf{m})^j) \\ &= \sum_{e\in\mathcal{E}((\mathbf{s},\mathbf{m})^j} \tau^{j+1}_{(\mathbf{s},\mathbf{m})^j}(\mathsf{D}(e,m))p(e|(\mathbf{s},\mathbf{m})^j) \end{aligned}$$

and

$$\begin{aligned} p(M_{j+1}^{\tau^{j+1}} = m) &= \sum_{(\mathbf{s},\mathbf{m})^j} p_{j+1}^{\tau^{j+1}}((\mathbf{s},\mathbf{m})^j)) \sum_{e\in\mathcal{E}((\mathbf{s},\mathbf{m})^j} \tau^{j+1}_{(\mathbf{s},\mathbf{m})^j}(\mathsf{D}(e,m))p(e|(\mathbf{s},\mathbf{m})^j) \\ &= \sum_{(\mathbf{s},\mathbf{m})^j}\ \sum_{e\in\mathcal{E}((\mathbf{s},\mathbf{m})^j} p^*(e,m,(\mathbf{s},\mathbf{m})^j). \end{aligned}$$

Thus $M_{j+1}^{\tau^{j+1}}$ satisfies the conditions on M^* described at the end of section 3.1 and we have

$$P_j^{\tau^j} \geq 2^{-I(E;M_{j+1}^{\tau^{j+1}})|(S^\tau,M^\tau)^j)}$$

where $M_{j+1}^{\tau^{j+1}}$ has the same distribution as responses to query $q_{j+1} = s_{j+1}$ by the adversary $F_{\tau^{j+1}}$. Equivalently, the inequality may be written

$$\log_2 P_j^{\tau^j} \geq -I(E; M_{j+1}^{\tau^{j+1}}|(S^\tau, M^\tau)^j) = -H(E|(S^\tau, M^\tau)^j) + H(E|(S^\tau, M^\tau)^j, M_{j+1}^{\tau^{j+1}}) \tag{11}$$

Theorem 3.4 *Let* Π *be an authentication system. Then*

$$\Pi_{j=0}^i P_j \geq 2^{-H(E)}$$

Proof: For adversaries F_{τ^j}, $j = 0, \ldots, i$, as described above we have

$$\log_2 P_j \geq \log_2 P_j^{\tau^j} \geq -I(E; M_{j+1}^{\tau^{j+1}}|(S^\tau, M^\tau)^j) = -H(E|(S^\tau, M^\tau)^j) + H(E|(S^\tau, M^\tau)^j, M_{j+1}^{\tau^{j+1}}).$$

Note that

$$\begin{aligned} H(E, S_j^\tau|(S^\tau, M^\tau)^{j-1}, M_j^{\tau^j}) &= H(E|(S^\tau, M^\tau)^{j-1}, M_j^{\tau^j}) + H(S_j^\tau|E, (S^\tau, M^\tau)^{j-1}, M_j^{\tau^j}) \\ &= H(S_j^\tau|(S^\tau, M^\tau)^{j-1}, M_j^{\tau^j}) + H(E|(S^\tau, M^\tau)^j) \end{aligned}$$

and since $H(S_j^\tau|E, (S^\tau, M^\tau)^{j-1}, M_j^{\tau^j}) = 0$ we have

$$H(S_j^\tau|(S^\tau, M^\tau)^{j-1}, M_j^{\tau^j}) = H(E|(S^\tau, M^\tau)^{j-1}, M_j^{\tau^j}) - H(E|(S^\tau, M^\tau)^j).$$

Hence

$$\begin{aligned} \log_2(\Pi_{j=0}^i P_j) &\geq (-H(E) + H(E|M_1^{\tau^1})) + (-H(E|(S^\tau, M^\tau)^1) + H(E|(S^\tau, M^\tau)^1, M_2^{\tau^2})) \\ &\quad + \cdots + (-H(E|(S^\tau, M^\tau)^i) + H(E|(S^\tau, M^\tau)^i, M_{i+1}^{\tau^{i+1}})) \\ &= -H(E) + (H(E|M_1^{\tau^1}) - H(E|(S^\tau, M^\tau)^1)) \\ &\quad + \cdots + (H(E|(S^\tau, M^\tau)^{i-1}, M_i^{\tau^i}) - H(E|(S^\tau, M^\tau)^i)) + H(E|(S^\tau, M^\tau)^i, M_{i+1}^{\tau^{i+1}}) \\ &= -H(E) + H(S_1^\tau|M_1^{\tau^1}) + \cdots + H(S_i^\tau|(S^\tau, M^\tau)^{i-1}, M_i^{\tau^i}) + H(E|(S^\tau, M^\tau)^i, M_{i+1}^{\tau^{i+1}}) \\ &\geq -H(E). \end{aligned}$$

■

Let $Pd = \max_i P_i$. Then it follows that $Pd \geq 2^{-\frac{H(E)}{i+1}}$.

4 'Good' queries do not decrease success probability

It is well known that for a message observing adversary, observing an extra message may reduce the success probability of the adversary. For example, Massey [3] defined l-fold secure A-codes as codes that provide perfect protection against spoofing of order zero and one. For such codes with k source states and v messages we have $P_0 = \frac{k}{v}$ and $P_1 = \frac{k-1}{v-1}$. Moreover the probability of success in spoofing of order zero is $\frac{k}{v}$ for each spoofing message m and the probability of success in spoofing of order one is $\frac{k-1}{v-1}$ for each observed message m and spoofing message $m' \neq m$. Thus for these codes observing any message will reduce the success probability of the adversary.

The passive message observing adversary obtains information from valid messages sent across the channel. On the other hand the active querying adversary obtains information from the responses to the queries. Thus, there is a difference: in the latter the adversary can control the amount of information that he receives.

Since adversaries with oracle access have some control on the information that they receive a natural question is whether asking queries is always helpful to such adversaries. For every query, the adversary may calculate his best success chance before and after the query is asked and choose the one that gives him the higher success chance. Of course, the probability of success is unchanged if the adversary repeats a query that has been made before. So we only need to consider the case where the adversary makes distinct queries. We call a query 'good' if it is distinct from previous queries and if asking it will not reduce the success chance of the adversary. Otherwise a query is 'bad'. the success chance of the adversary. In the following we show that the adversary only needs to look for 'good' queries and if there is one then he can be guaranteed that asking the query will not decrease his success chance.

The following theorem gives conditions under which a query is good. It follows that in all but those special cases where no query satisfies these conditions, asking another query is helful and the adversary's probability of success is at least as great with the extra query.

Theorem 4.1 *Let τ be a strategy of an adversary who asks i oracle queries and then constructs a spoofing query for an authentication system Π. Suppose that the adversary modifies the strategy to ask one extra oracle query and then spoof optimally.*

Then an extra query q is good *if it is distinct from previous queries and (i) in the case of verification queries, q is distinct from the unique optimal spoofing message for strategy τ should one exist; and (ii) in the case of authentication queries, there exists an optimal spoofing message $\hat{m}$ for strategy τ such that any key consistent with the previous query and response pairs is not consistent with $(q, \hat{m})$.*

If there exists a good extra query for every sequence of query and response pairs arising from τ then for a suitable modification the adversary's probability of success is at least P_i^τ

Proof: Suppose that the adversary has observed the sequence $(\mathbf{q},\mathbf{r})^i$ of query and response pairs.

The expected success probability of the adversary after another query q_{i+1} is

$$\sum_{r_{i+1}\in\mathcal{R}} p(r_{i+1}|q_{i+1},(\mathbf{q},\mathbf{r})^i)\max_{m\in\mathcal{M}}\left\{\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^{i+1})} p(e|(\mathbf{q},\mathbf{r})^{i+1})\gamma(e,m,(\mathbf{q},\mathbf{r})^{i+1})\right\}$$

where $p(r_{i+1}|q_{i+1},(\mathbf{q},\mathbf{r})^i) = \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^{i+1})} p(e|(\mathbf{q},\mathbf{r})^i)$ is the conditional probability that the response is r_{i+1} given the sequence $(\mathbf{q},\mathbf{r})^i$ and the query q_{i+1}. But

$$p(e|(\mathbf{q},\mathbf{r})^{i+1}) = \frac{p(e|(\mathbf{q},\mathbf{r})^i)}{\sum_{e'\in\mathcal{E}((\mathbf{q},\mathbf{r})^{i+1})} p(e'|(\mathbf{q},\mathbf{r})^i))}.$$

so this is

$$\begin{aligned}
& \sum_{r_{i+1}\in\mathcal{R}} \max_{m\in\mathcal{M}}\left\{\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^{i+1})} p(e|(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(\mathbf{q},\mathbf{r})^{i+1})\right\} \\
\geq\ & \max_{m\in\mathcal{M}}\left\{\sum_{r_{i+1}\in\mathcal{R}}\left[\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^{i+1})} p(e|(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(\mathbf{q},\mathbf{r})^{i+1})\right]\right\} \\
=\ & \max_{m\in\mathcal{M}}\left\{\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i)} p(e|(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(\mathbf{q},\mathbf{r})^i,(q_{i+1},\hat{r}))\right\}
\end{aligned}$$

where $\hat{r} = \mathsf{Ver}(e, q_{i+1})$ or $\mathsf{Auth}(e, q_{i+1})$.

Suppose that q_{i+1} is distinct from $q_1, \ldots, q_i$ and (i) if $q_{i+1} \in \mathcal{M}$ then there exists $\hat{m} \neq q_{i+1}$ with $P_i^\tau((\mathbf{q}, \mathbf{r})^i, (\hat{m}, 1)) = \max_{m \in \mathcal{M}} \left\{ P_i^\tau((\mathbf{q}, \mathbf{r})^i, (m, 1)) \right\}$, and (ii) if $q_{i+1} \in \mathcal{S}$ then under any key $e \in \mathcal{E}((\mathbf{q}, \mathbf{r})^i)$, there exists a message $\hat{m} \in \mathcal{M}$ with $\hat{m} \neq \mathsf{Auth}(e, q_{i+1})$ and $P_i^\tau((\mathbf{q}, \mathbf{r})^i, (\hat{m}, 1)) = \max_{m \in \mathcal{M}} P_i^\tau((\mathbf{q}, \mathbf{r})^i, (m, 1))$.

With $\hat{m}$ so defined we have $\gamma(e, \hat{m}, (\mathbf{q}, \mathbf{r})^i, (q_{i+1}, \hat{r})) = \gamma(e, \hat{m}, (\mathbf{q}, \mathbf{r})^i)$ and so

$$\begin{aligned} \max_{m \in \mathcal{M}} \left\{ \sum_{e \in \mathcal{E}((\mathbf{q}, \mathbf{r})^i)} p(e|(\mathbf{q}, \mathbf{r})^i) \gamma(e, m, (\mathbf{q}, \mathbf{r})^i, (q_{i+1}, \hat{r})) \right\} &= \sum_{e \in \mathcal{E}((\mathbf{q}, \mathbf{r})^i)} p(e|(\mathbf{q}, \mathbf{r})^i) \gamma(e, \hat{m}, (\mathbf{q}, \mathbf{r})^i) \\ &= P_i^\tau((\mathbf{q}, \mathbf{r})^i). \end{aligned}$$

Thus q_{i+1} is a *good* query.

It follows immediately that if there exists a good extra query for every sequence of query and response pairs arising from τ then the adversary's probability of success for a modified strategy that chooses such a query with probability 1 is at least $\sum_{(\mathbf{q}, \mathbf{r})^i} p_i^\tau((\mathbf{q}, \mathbf{r})^i) P_i^\tau((\mathbf{q}, \mathbf{r})^i) = P_i^\tau$ and the result follows. ■

Note that for verification queries the condition for existence of a good query is nearly always satisfied. The exception is when $|\mathcal{M} \backslash \mathbf{q}^i| = 1$ and so the only available query is the spoofing query. As long as $|\mathcal{M} \backslash \mathbf{q}^i| > 1$ there will be a good query and so the success chance of the adversary would not be reduced if that query is asked. In the case of authentication queries, the theorem may not imply the existence of a good query. This is the case if for some sequence of query and response pairs $(\mathbf{q}, \mathbf{r})^i$ arising from τ every optimal spoofing message $\hat{m}$ has the property that for any $q \in \mathcal{S} \backslash \mathbf{q}^i$ there is an $e \in \mathcal{E}((\mathbf{q}, \mathbf{r})^i)$ such that $\mathsf{Auth}(e, q) = \hat{m}$.

Note that the above theorem is helpful as it provides the adversary with a method to determine whether he should ask a query or not: he simply checks whether there are good queries and if there are then he is guaranteed to do better if he asks them.

5 Optimal codes

We consider optimal authentication codes that satisfy the bound $Pd \geq 2^{-\frac{H(E)}{i+1}}$ with equality. We use an argument similar to Rosenbaum [7] and obtain a combinatorial characterization of such authentication codes analogous to that in the message-observing adversary setting. The codes will have minimum number of keys and limit the best success chance of a spoofer with access to i queries to its minimum (that is, satisfy bound 3.2 with equality).

If equality holds then it follows by theorem 3.4 that $P_j = Pd$ for $j = 0, \ldots, i$ and the proof of the theorem shows that $H(E|(S^\tau, M^\tau)^i, M_{i+1}^{\tau^{i+1}}) = 0$ and $H(S_j^\tau|(S^\tau, M^\tau)^{j-1}, M_j^\tau) = 0$ for $j = 1, \ldots, i$. Further, equality holds in Theorem 3.3 and so, for an optimal strategy τ^* (having the properties described above the statement of Theorem 3.4), $p^*(m|e, (\mathbf{s}^*, \mathbf{m})^j)$ is constant for all $e \in \mathcal{E}((\mathbf{s}^*, \mathbf{m})^j, (m, 1))$ and $P_j = P_j^{\tau^*}((\mathbf{s}^*, \mathbf{m})^j, (m, 1))$.

Since $p^*(m|e, (\mathbf{s}^*, \mathbf{m})^j)$ is independent of e it follows that

$$p(e|(\mathbf{s}^*, \mathbf{m})^j, (m, 1)) = \frac{p(e|(\mathbf{s}^*, \mathbf{m})^j)}{\sum_{e' \in \mathcal{E}((\mathbf{s}^*, \mathbf{m})^j, (m,1))} p(e'|(\mathbf{s}^*, \mathbf{m})^j)} \tag{12}$$

and we have $\sum_{e \in \mathcal{E}((\mathbf{s}^*, \mathbf{m})^j, (m,1))} p(e|(\mathbf{s}^*, \mathbf{m})^j) = P_j^{\tau^*}((\mathbf{s}^*, \mathbf{m})^j, (m, 1)) = Pd$.

Using an argument by induction on the number j of query and response pairs gives

$$p(e|(\mathbf{s}^*,\mathbf{m})^j) = \frac{p(e)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} p(e')} \tag{13}$$

for all $j = 0,\ldots,i$. Further, another induction, on the number j of terms in the product, gives

$$\sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} p(e) = \prod_{l=0}^{j-1} P_l \tag{14}$$

for all $j = 0,\ldots,i$. (See the appendix for proofs of these equations.)

Let $\mathbf{s}^{i+1}$ be a sequence of distinct source states and let $e \in \mathcal{E}$. Let $\mathbf{m}^{i+1}$ be the sequence of messages with $m_j = \mathsf{Auth}(e, s_j)$ for $j = 1,\ldots,i+1$. Then $p(e|(\mathbf{s},\mathbf{m})^{i+1}) \neq 0$ and since $H(E|(S^\tau, M^\tau)^i, M^{\tau\, i+1}_{i+1}) = 0$ we have $p(e|(\mathbf{s},\mathbf{m})^{i+1}) = 1$ and $\mathcal{E}((\mathbf{s},\mathbf{m})^i,(m_{i+1},1)) = \{e\}$. Hence $p(e) = \sum_{e'\in\mathcal{E}((\mathbf{s},\mathbf{m})^i, m_{i+1}.1))} p(e') = \prod_{j=0}^{i} P_j$. Thus the distribution $p(e)$ is uniform.

It follows that

$$p(e|(\mathbf{s}^*,\mathbf{m})^j) = \frac{p(e)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} p(e')} = \frac{1}{|\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)|}$$

$$P_j = \sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j,(m,1))} \frac{1}{|\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)|}$$

and $|\mathcal{E}((\mathbf{s},\mathbf{m})^{j+1})| = (\prod_{l=0}^{j} P_l)|\mathcal{E}| = Pd^{j+1}|\mathcal{E}|$. Since $|\mathcal{E}((\mathbf{s},\mathbf{m})^{i+1})| = 1$ we have $|\mathcal{E}| = Pd^{-(i+1)}$ and $|\mathcal{E}((\mathbf{s},\mathbf{m})^j)| = Pd^{j-(i+1)}$. Thus Pd^{-1} is an integer q and $|\mathcal{E}| = q^{i+1}$.

Thus the optimal authentication systems determine a combinatorial structure. We may identify a pair $(e,(s,m))$ satisfying $\mathsf{Auth}(e,s) = m$ with an incidence in a combinatorial design. The above results show that an optimal authentication code corresponds to a combinatorial design in which any j query response pairs $(\mathbf{s},\mathbf{m})^j$ are incident with a constant number q^{i+1-j} encoding rules. Since $H(S_1^\tau|M_1^\tau) = 0$ the authentication code is Cartesian and it follows that it is optimal in the message observing setting also. The authentication systems arising from Reed-Solomon Codes (or Orthogonal Arrays) (see Mitchell *et al* [5]) which are optimal in the message observing adversary setting also provide examples of optimal codes in our query oracle setting.

6 Concluding Remarks

We have given an analysis of authentication systems for an adversary with access to oracle queries. We derived information theoretic bounds on the best success probability of an adversary using a strategy τ to ask i queries and then construct a spoofing query. This bound can be seen as a generalisation of the Simmons and Rosenbaum bound and is derived using the same technique. The adaptive adversary however has the ability to influence the bound through his querying strategy as well as his spoofing strategy while for a message observing adversary the bound is only affected by this latter strategy. We also derived a bound on the key entropy for an adaptive adversary with access to i authentication queries and showed that in this case for an authentication system with probability of deception Pd the key entropy is at least $(i+1)\log Pd$. This is similar to the known result for a message observing adversary.

We gave a combinatorial characterisation of authentication codes that meet the bounds and showed that optimal codes (having the least success probability and the smallest number of keys) correspond to orthogonal arrays. This is analogous to the message observing case.

We gave a result to show that, in the query oracle model, as long as there is always a good query, then asking that query is helpful to the adversary.

7 Appendix

Proof of Theorem 3.2. The following proof is a direct application of the proof used in [7] to the situation where the adversary makes oracle queries instead of observing messages.

We will use Jensen's inequality for *convex functions*. A real function ϕ is convex on the interval (a, b) if $\phi'(r) < \phi'(s)$, for all $a < r < s < b$.

Theorem 7.1 *(Jensen's inequality) [11] Let $w_i, i = 1, \cdots n$ be non-negative numbers such that $\sum_i w_i = 1$. let ϕ be a real function that is convex on the interval (a, b) and let $x_i \in (a, b), i = 1, \cdots n$. Then,*

$$\phi(\sum_i w_i.x_i) \leq \sum_i w_i.\phi(x_i) \tag{15}$$

and equality holds if and only if all x_i are equal.

■

The proof of Theorem 3.2 proceeds as follows.

Let

$$\Psi_{(\mathbf{q},\mathbf{r})^i,m}(e) = \frac{p(e|(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(q,r)^i)}{P_i^{\tau}((\mathbf{q},\mathbf{r})^i),(m,1))}$$

Because $P_i^{\tau}((\mathbf{q},\mathbf{r})^i,(m,1)) = \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(e|(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(\mathbf{q},\mathbf{r})^i)$ it follows that $\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} \Psi_{((\mathbf{q},\mathbf{r})^i,m)}(e) = 1$ and $\Psi_{((\mathbf{q},\mathbf{r})^i,m)}(e)$ is a probability distribution on $\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))$.

Now we have

$$\begin{aligned}
p(m|(\mathbf{q},\mathbf{r})^i) &= \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(m,e|(\mathbf{q},\mathbf{r})^i)\\
&= \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(e|(\mathbf{q},\mathbf{r})^i)p(m|e,(\mathbf{q},\mathbf{r})^i)\\
&= \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(e|(\mathbf{q},\mathbf{r})^i)p(m|e,(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(\mathbf{q},\mathbf{r})^i)\\
&= \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} \Psi_{((\mathbf{q},\mathbf{r})^i,m)}(e)P_i^{\tau}((\mathbf{q},\mathbf{r})^i,(m,1))p(m|e,(\mathbf{q},\mathbf{r})^i)
\end{aligned}$$

Using Jensen's inequality (15) for $\phi(x) = x.\log x$ at $x = p(m|(\mathbf{q},\mathbf{r})^i) = \sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} w_e.x_e$ with $w_e = \Psi_{(\mathbf{q},\mathbf{r})^i,m}(e)$ and $x_e = P(m,(\mathbf{q},\mathbf{r})^i)p(m|e,(\mathbf{q},\mathbf{r})^i)$, we have

$$\begin{aligned}
&p(m|(\mathbf{q},\mathbf{r})^i)\log p(m|(\mathbf{q},\mathbf{r})^i)\\
\leq &\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} \Psi_{(\mathbf{q},\mathbf{r})^i,m}(e)P_i^{\tau}((\mathbf{q},\mathbf{r})^i,(m,1))p(m|e,(\mathbf{q},\mathbf{r})^i)\log[P_i^{\tau}((\mathbf{q},\mathbf{r})^i,(m,1))p(m|e,(\mathbf{q},\mathbf{r})^i)]\\
= &\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(e|(\mathbf{q},\mathbf{r})^i)p(m|e,(\mathbf{q},\mathbf{r})^i)\gamma(e,m,(\mathbf{q},\mathbf{r})^i)\log[P_i^{\tau}((\mathbf{q},\mathbf{r})^i,(m,1))p(m|e,(\mathbf{q},\mathbf{r})^i)]\\
= &\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(e,m|(\mathbf{q},\mathbf{r})^i)\log[P_i^{\tau}((\mathbf{q},\mathbf{r})^i,(m,1))p(m|e,(\mathbf{q},\mathbf{r})^i)]
\end{aligned} \tag{16}$$

We will first show that

$$H(M|(\mathbf{q},\mathbf{r})^i) \geq -\log(P_i^{\tau}((\mathbf{q},\mathbf{r})^i) + H(M|E,(\mathbf{q},\mathbf{r})^i)$$

This is true because using (16) we have,

$$\begin{aligned}
H(M|(\mathbf{q},\mathbf{r})^i) &= -\sum_{m\in\mathcal{M}} p(m|(\mathbf{q},\mathbf{r})^i)\log p((m|(\mathbf{q},\mathbf{r})^i) \\
&\geq -\sum_{m\in\mathcal{M}}\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(e,m|(\mathbf{q},\mathbf{r})^i)\log[P_i^\tau((\mathbf{q},\mathbf{r})^i,(m,1))p(m|e,(\mathbf{q},\mathbf{r})^i)] \\
&= -\sum_{m\in\mathcal{M}} p(m|(\mathbf{q},\mathbf{r})^i)\log P_i^\tau((\mathbf{q},\mathbf{r})^i,(m,1)) \\
&\quad -\sum_{m\in\mathcal{M}}\sum_{e\in\mathcal{E}((\mathbf{q},\mathbf{r})^i,(m,1))} p(e|(\mathbf{q},\mathbf{r})^i)p(m|e,(\mathbf{q},\mathbf{r})^i)\log p(m|e,(\mathbf{q},\mathbf{r})^i)
\end{aligned}$$

Now as the adversary spoofs optimally we have $P_i^\tau((\mathbf{q},\mathbf{r})^i) \geq P_i^\tau((\mathbf{q},\mathbf{r})^i,(m,1))$ and so

$$\begin{aligned}
H(M|(\mathbf{q},\mathbf{r})^i) &\geq -\log P_i^\tau((\mathbf{q},\mathbf{r})^i)\sum_{m\in\mathcal{M}} p(m|(\mathbf{q},\mathbf{r})^i) + H(M|E,(\mathbf{q},\mathbf{r})^i) \\
&= -\log P_i^\tau((\mathbf{q},\mathbf{r})^i) + H(M|E,(\mathbf{q},\mathbf{r})^i).
\end{aligned}$$

The final step uses Jensen's inequality (15) for $\phi(x) = -\log x$.

$$\begin{aligned}
\log P_i^\tau &= \log[\sum_{(\mathbf{q},\mathbf{r})^i} p_i^\tau((\mathbf{q},\mathbf{r})^i)P_i^\tau((\mathbf{q},\mathbf{r})^i)] \\
&\geq \sum_{(\mathbf{q},\mathbf{r})^i} p_i^\tau((\mathbf{q},\mathbf{r})^i)\log P((\mathbf{q},\mathbf{r})^i) \\
&\geq \sum_{(\mathbf{q},\mathbf{r})^i} p_i^\tau((\mathbf{q},\mathbf{r})^i)(H(M|E,(\mathbf{q},\mathbf{r})^i) - H(M|(\mathbf{q},\mathbf{r})^i)) \\
&= H(M|E,(Q^\tau,R^\tau)^i) - H(M|(Q^\tau,R^\tau)^i)
\end{aligned}$$

■

Proof of equation 12 We show that

$$p(e|(\mathbf{s}^*,\mathbf{m})^j,(m,1)) = \frac{p(e|(s^*,m)^j)}{\sum_{e'\in\mathcal{E}((s^*,m)^j,(m,1))} p(e'|(s^*,m)^j)}.$$

Now, for $p^*(e,(\mathbf{s},\mathbf{m})^j,m) \neq 0$ we have

$$p^*(e,(\mathbf{s}^*,\mathbf{m})^j,m) = p(e|(\mathbf{s}^*,\mathbf{m})^j,(m,1))p^*((\mathbf{s}^*,\mathbf{m})^j,m).$$

and

$$p^*(e,(\mathbf{s}^*,\mathbf{m})^j,m) = p^*(m|e,(\mathbf{s}^*,\mathbf{m})^j)p(e|(\mathbf{s}^*,\mathbf{m})^j)p_i^\tau((\mathbf{s}^*,\mathbf{m})^j).$$

Hence

$$p(e|(\mathbf{s}^*,\mathbf{m})^j,(m,1)) = \frac{p^*(m|e,(\mathbf{s}^*,\mathbf{m})^j)p(e|(\mathbf{s}^*,\mathbf{m})^j)p_i^\tau((\mathbf{s}^*,\mathbf{m})^j)}{p^*((\mathbf{s}^*,\mathbf{m})^j,m)}.$$

Now

$$\begin{aligned}
p^*((\mathbf{s}^*,\mathbf{m})^j,m) &= \sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j,(m,1))} p^*(e',(\mathbf{s}^*,\mathbf{m})^j,m) \\
&\quad \sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j,(m,1))} p^*(m|e',(\mathbf{s}^*,\mathbf{m})^j)p(e'|(\mathbf{s}^*,\mathbf{m})^j)p_i^\tau((\mathbf{s}^*,\mathbf{m})^j).
\end{aligned}$$

Thus

$$p(e|(\mathbf{s}^*,\mathbf{m})^j,(m,1)) = \frac{p^*(m|e,(\mathbf{s}^*,\mathbf{m})^j)p(e|(\mathbf{s}^*,\mathbf{m})^j)p_i^\tau((\mathbf{s}^*,\mathbf{m})^j)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j,(m,1))} p^*(m|e',(\mathbf{s}^*,\mathbf{m})^j)p(e'|(\mathbf{s}^*,\mathbf{m})^j)p_i^\tau((\mathbf{s}^*,\mathbf{m})^j)}.$$

But by Theorem 3.3 $p^*(m|e',(\mathbf{s}^*,\mathbf{m})^j)$ is constant for all $e' \in \mathcal{E}((\mathbf{s}^*,\mathbf{m})^j,(m,1))$ so that

$$p(e|(\mathbf{s}^*,\mathbf{m})^j,(m,1)) = \frac{p(e|(\mathbf{s}^*,\mathbf{m})^j)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j,(m,1))} p(e'|(\mathbf{s}^*,\mathbf{m})^j)}.$$

and the result follows. ∎

Proof of equation 13 We show that $p(e|(\mathbf{s}^*,\mathbf{m})^j) = \frac{p(e)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} p(e')}$.

For $j=0$ we have $p(e|(\mathbf{s}^*,\mathbf{m})^0) = p(e)$, so $p(e|(s^*,m)) = \frac{p(e)}{\sum_{e'\in\mathcal{E}((s^*,m)^0)} p(e')}$.

Suppose that $1 \le j \le i$ and $p(e|(s^*,m)^{j-1}) = \frac{p(e)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1})} p(e')}$ for all $e \in \mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1})$. Since $H(S_j^\tau|(S^\tau,M^\tau)^{j-1},M_j^\tau)$ so that $\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1},(m,1)) = \mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1},(s^*,m)) = \mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)$ and because $\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1}) \subseteq \mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)$ we have

$$\begin{aligned}
p(e|(\mathbf{s}^*,\mathbf{m})^j) &= p(e|(\mathbf{s}^*,\mathbf{m})^{j-1},(m,1)) = \frac{p(e|(\mathbf{s}^*,\mathbf{m})^{j-1})}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1},(m,1))} p(e'|(\mathbf{s}^*,\mathbf{m})^{j-1})} \\
&= \frac{\frac{p(e)}{\sum_{f\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1})} p(f)}}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1},(m,1))} \frac{p(e')}{\sum_{f\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1})} p(f)}} \\
&= \frac{p(e)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} p(e')}
\end{aligned}$$

and the result follows. ∎

Proof of equation 14 We show that $\sum_{e\in\mathcal{E}((s^*,m)^j)} p(e) = \prod_{l=0}^{j-1} P_l$.

For $j=1$ we have $P_0 = P_0^{\tau^*}((m,1)) = \sum_{e\in\mathcal{E}((m,1))} p(e) = \sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^1)} p(e)$.

Suppose that $2 \le j \le i+1$ and $\sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1})} p(e) = \prod_{l=0}^{j-2} P_l$.

Now

$$P_{j-1} = \sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1},(m,1))} p(e|(\mathbf{s}^*,\mathbf{m})^{j-1}) = \sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} p(e|(\mathbf{s}^*,\mathbf{m})^{j-1}$$

so that

$$\begin{aligned}
P_{j-1} &= \sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} \frac{p(e)}{\sum_{e'\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^{j-1})} p(e')} \\
&= \sum_{e\in\mathcal{E}((\mathbf{s}^*,\mathbf{m})^j)} \frac{p(e)}{\prod_{l=0}^{j-1} P_l}
\end{aligned}$$

and the result follows. ∎

References

[1] E. N. Gilbert, F. J. MacWilliams and N. J. A. Sloane, ‘Codes which detect deception’, *Bell System Tech. J.* **53**(3) (1974) 405–424.

[2] R. Johannesson and A. Sgarro, ‘Strengthening Simmon’s bound on impersonation’, *IEEE Transactions on Information Theory* **37** (1991), 1182-1185.

[3] J. L. Massey, Cryptography- A Selective survey, *Alta Frequenza*, **LV**(1) (1986), 4-11.

[4] U. Maurer, ‘A unified and generalized treatment of authentication theory’, ”Proceedings of the 13th Symposium on Theoretical Aspects of Computer Science”, Lecture Notes in Computer Science 1046 (1996), 387-398

[5] C. Mitchell, M. Waker and P. Wild, ’The combinatorics of perfect authentication schemes’, *SIAM Journal on Discrete Mathematics* **7** (1994), 102-107.

[6] D. Pei, ‘Information-theoretic bounds for authentication codes and block designs’, *Journal of Cryptology* **8** (1995), 177-188.

[7] U. Rosenbaum, ‘A lower bound on authentication after having observed a sequence of messages’, *Journal of Cryptology* **6** (1993), 135-156.

[8] R. Safavi-Naini, L. McAven and M. Yung, ‘General Group Authentication Codes and Their Relation to ”Unconditionally Secure Signatures”’, Public Key Cryptography 2004, LNCS 2947, pp 231-248.

[9] J. Shikata, G. Hanaoka,Y. Zheng and H. Imai, ’Security notions for unconditionally secure signature schemes’, Eurocrypt 2002, LNCS 2332, pp434-449.

[10] G. J. Simmons, ‘Authentication theory/coding theory’, *Crypto’84* LNCS **196** (Springer–Verlag, 1984) 411–431.

[11] http://planetmath.org/encyclopedia/JensensInequality.html