Constructing optimal authentication codes with perfect multi-fold secrecy

Constructing Optimal Authentication Codes with Perfect Multi-fold Secrecy

Michael Huber
University of Tuebingen
Wilhelm Schickard Institute for Computer Science
Sand 13, D-72076 Tuebingen, Germany
Email: michael.huber@uni-tuebingen.de

***Abstract*—We establish a construction of optimal authentication codes achieving perfect multi-fold secrecy by means of combinatorial designs. This continues the author's work (ISIT 2009, cf. [1]) and answers an open question posed therein. As an application, we present the first infinite class of optimal codes that provide two-fold security against spoofing attacks and at the same time perfect two-fold secrecy.**

I. Introduction

Authentication and secrecy are two crucial concepts in cryptography and information security. Although independent in their nature, various scenarios require that both aspects hold simultaneously. For *information-theoretic* or *unconditional* security (i.e. robustness against an attacker that has unlimited computational resources), authentication and secrecy codes have been investigated for quite some time. The initial construction of authentication codes goes back to Gilbert, MacWilliams & Sloane [2]. A more general and systematic theory of authentication was developed by Simmons (e.g., [3], [4]). Fundamental work on secrecy codes started with Shannon [5].

This paper deals with the construction of optimal authentication codes with perfect multi-fold secrecy. It continues the author's recent work [1], which naturally extended results by Stinson [6] on authentication codes with perfect secrecy. We will answer an important question left open in [1] that addresses the construction of authentication codes with perfect multi-fold secrecy for equiprobable source probability distributions. We establish a construction of optimal authentication codes which are multi-fold secure against spoofing attacks and simultaneously provide perfect multi-fold secrecy. This can be achieved by means of combinatorial designs. As an application, we present the first infinite class of optimal codes that achieve two-fold security against spoofing as well as perfect two-fold secrecy.

The paper is organized as follows: Necessary definitions and concepts from the theory of authentication and secrecy codes as well as from combinatorial design theory will be summarized in Section II. Section III gives relevant combinatorial constructions of optimal authentication codes which bear no secrecy assumptions. In Section IV, we review Stinson's constructions in [6] and recent results from [1]. Section V is devoted to our new constructions.

II. Preliminaries

A. Authentication and Secrecy Codes

We rely on the information-theoretical or unconditional secrecy model developed by Shannon [5], and by Simmons (e.g., [3], [4]) including authentication. Our notion complies, for the most part, with that of [6], [7]. In this model of authentication and secrecy three participants are involved: a *transmitter*, a *receiver*, and an *opponent*. The transmitter wants to communicate information to the receiver via a public communications channel. The receiver in return would like to be confident that any received information actually came from the transmitter and not from some opponent (*integrity* of information). The transmitter and the receiver are assumed to trust each other. Sometimes this is also called an A*-code*.

In what follows, let $\mathcal{S}$ denote a set of k *source states* (or *plaintexts*), $\mathcal{M}$ a set of v *messages* (or *ciphertexts*), and $\mathcal{E}$ a set of b *encoding rules* (or *keys*). Using an encoding rule $e \in \mathcal{E}$, the transmitter encrypts a source state $s \in \mathcal{S}$ to obtain the message $m = e(s)$ to be sent over the channel. The encoding rule is an injective function from $\mathcal{S}$ to $\mathcal{M}$, and is communicated to the receiver via a secure channel prior to any messages being sent. For a given encoding rule $e \in \mathcal{E}$, let $M(e) := \{e(s) : s \in \mathcal{S}\}$ denote the set of *valid* messages. For an encoding rule e and a set $M^* \subseteq M(e)$ of distinct messages, we define $f_e(M^*) := \{s \in \mathcal{S} : e(s) \in M^*\}$, i.e., the set of source states that will be encoded under encoding rule e by a message in M^*. A received message m will be accepted by the receiver as being authentic if and only if $m \in M(e)$. When this is fulfilled, the receiver decrypts the message m by applying the decoding rule e^{-1}, where

$$e^{-1}(m) = s \Leftrightarrow e(s) = m.$$

An authentication code can be represented algebraically by a $(b \times k)$*-encoding matrix* with the rows indexed by the encoding rules, the columns indexed by the source states, and the entries defined by $a_{es} := e(s)$ $(1 \le e \le b,\, 1 \le s \le k)$.

We address the scenario of a *spoofing attack* of order i (cf. [7]): Suppose that an opponent observes $i \ge 0$ distinct messages, which are sent through the public channel using the same encoding rule. The opponent then inserts a new message m' (being distinct from the i messages already sent), hoping to have it accepted by the receiver as authentic. The cases $i = 0$

and $i = 1$ are called *impersonation game* and *substitution game*, respectively. These cases have been studied in detail in recent years (e.g., [8], [9]), however less is known for the cases $i \geq 2$. In this article, we focus on those cases where $i \geq 2$.

For any i, we assume that there is some probability distribution on the set of i-subsets of source states, so that any set of i source states has a non-zero probability of occurring. For simplification, we ignore the order in which the i source states occur, and assume that no source state occurs more than once. Given this probability distribution p_S on $\mathcal{S}$, the receiver and transmitter choose a probability distribution p_E on $\mathcal{E}$ (called *encoding strategy*) with associated independent random variables S and E, respectively. These distributions are known to all participants and induce a third distribution, p_M, on $\mathcal{M}$ with associated random variable M. The *deception probability* P_{d_i} is the probability that the opponent can deceive the receiver with a spoofing attack of order i. The following theorem (cf. [7]) provides combinatorial lower bounds.

Theorem 1: [Massey] In an authentication code with k source states and v messages, the deception probabilities are bounded below by

$$P_{d_i} \geq \frac{k-i}{v-i}.$$

An authentication code is called *t_A-fold secure against spoofing* if $P_{d_i} = (k-i)/(v-i)$ for all $0 \leq i \leq t_A$.

Moreover, we consider the concept of perfect multi-fold secrecy which has been introduced by Stinson [6] and generalizes Shannon's fundamental idea of perfect (one-fold) secrecy (cf. [5]). We say that an authentication code has *perfect t_S-fold secrecy* if, for every positive integer $t^* \leq t_S$, for every set M^* of t^* messages observed in the channel, and for every set S^* of t^* source states, we have

$$p_S(S^*|M^*) = p_S(S^*).$$

That is, the *a posteriori* probability distribution on the t^* source states, given that a set of t^* messages is observed, is identical to the *a priori* probability distribution on the t^* source states.

When clear from the context, we often only write t instead of t_A resp. t_S.

B. Combinatorial Designs

We recall the definition of a combinatorial t-design. For positive integers $t \leq k \leq v$ and λ, a t-(v,k,λ) *design* $\mathcal{D}$ is a pair $(X, \mathcal{B})$, satisfying the following properties:

(i) X is a set of v elements, called *points*,
(ii) $\mathcal{B}$ is a family of k-subsets of X, called *blocks*,
(iii) every t-subset of X is contained in exactly λ blocks.

We denote points by lower-case and blocks by upper-case Latin letters. Via convention, let $b := |\mathcal{B}|$ denote the number of blocks. Throughout this article, 'repeated blocks' are not allowed, that is, the same k-subset of points may not occur twice as a block. If $t < k < v$ holds, then we speak of a *non-trivial* t-design. For historical reasons, a t-(v,k,λ) design with $\lambda = 1$ is called a *Steiner t-design* (sometimes also a *Steiner system*). The special case of a Steiner design with parameters $t = 2$ and $k = 3$ is called a *Steiner triple system STS*(v) of order v. A Steiner design with parameters $t = 3$ and $k = 4$ is called a *Steiner quadruple system SQS*(v) of order v. Specifically, we are interested in Steiner quadruple systems in this paper. As a simple example, the vector space $\mathbb{Z}_2^d$ ($d \geq 3$) with the set $\mathcal{B}$ of blocks taken to be the set of all subsets of four distinct elements of $\mathbb{Z}_2^d$ whose vector sum is zero, is a non-trivial *boolean* Steiner quadruple system SQS(2^d). More geometrically, these SQS(2^d) consist of the points and planes of the d-dimensional binary affine space $AG(d,2)$.

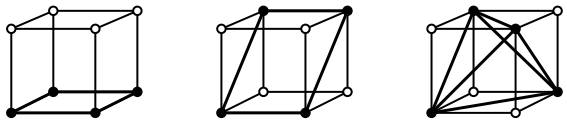

Fig. 1. Illustration of the unique SQS(8), with three types of blocks: faces, opposite edges, and inscribed regular tetrahedra.

For the existence of t-designs, basic necessary conditions can be obtained via elementary counting arguments (see, for instance, [10]):

Lemma 1: Let $\mathcal{D} = (X, \mathcal{B})$ be a t-(v,k,λ) design, and for a positive integer $s \leq t$, let $S \subseteq X$ with $|S| = s$. Then the number of blocks containing each element of S is given by

$$\lambda_s = \lambda \frac{\binom{v-s}{t-s}}{\binom{k-s}{t-s}}.$$

In particular, for $t \geq 2$, a t-(v,k,λ) design is also an s-(v,k,λ_s) design.

It is customary to set $r := \lambda_1$ denoting the number of blocks containing a given point. It follows

Lemma 2: Let $\mathcal{D} = (X, \mathcal{B})$ be a t-(v,k,λ) design. Then the following holds:

(a) $bk = vr$.
(b) $\binom{v}{t}\lambda = b\binom{k}{t}$.
(c) $r(k-1) = \lambda_2(v-1)$ for $t \geq 2$.

For encyclopedic accounts of key results in design theory, we refer to [10], [11]. Various connections of designs with coding and information theory can be found in a recent survey [12] (with many additional references therein).

III. Optimal Authentication Codes

For our further purposes, we summarize the state-of-the-art for authentication codes which bear no secrecy assumptions. The following theorem (cf. [7], [13]) gives a combinatorial lower bound on the number of encoding rules.

Theorem 2: [Massey–Schöbi] If an authentication code is $(t-1)$-fold against spoofing, then the number of encoding rules is bounded below by

$$b \geq \frac{\binom{v}{t}}{\binom{k}{t}}.$$

TABLE I
OPTIMAL AUTHENTICATION CODES WITH PERFECT SECRECY: INFINITE CLASSES

t_A	t_S	k	v	b	Ref.
1	1	$q+1$ q prime power	$\frac{q^{d+1}-1}{q-1}$ $d \geq 2$ even	$\frac{v(v-1)}{k(k-1)}$	[6]
1	1	3	$v \equiv 1 \pmod 6$	$\frac{v(v-1)}{6}$	[1]
1	1	4	$v \equiv 1 \pmod{12}$	$\frac{v(v-1)}{12}$	[1]
1	1	5	$v \equiv 1 \pmod{20}$	$\frac{v(v-1)}{20}$	[1]
2	1	$q+1$ q prime power	q^d+1 $d \geq 2$ even	$\frac{v(v-1)(v-2)}{k(k-1)(k-2)}$	[1]
2	1	4	$v \equiv 2, 10 \pmod{24}$	$\frac{v(v-1)(v-2)}{24}$	[1]

TABLE II
OPTIMAL AUTHENTICATION CODES WITH PERFECT SECRECY: FURTHER EXAMPLES

t_A	t_S	k	v	b	Ref.
2	1	5	26	260	[1]
3	1	5	11	66	[1]
		7	23	253	[1]
		5	23	1.771	[1]
		5	47	35.673	[1]
		5	83	367.524	[1]
		5	71	194.327	[1]
		5	107	1.032.122	[1]
		5	131	2.343.328	[1]
		5	167	6.251.311	[1]
		5	243	28.344.492	[1]
4	1	6	12	132	[1]
		6	84	5.145.336	[1]
		6	244	1.152.676.008	[1]

An authentication code is called *optimal* if the number of encoding rules meets the lower bound with equality. When the source states are known to be independent and equiprobable, optimal authentication codes which are $(t-1)$-fold secure against spoofing can be constructed via t-designs (cf. [6], [13], [14]).

Theorem 3: [DeSoete–Schöbi–Stinson] Suppose there is a t-(v,k,λ) design. Then there is an authentication code for k equiprobable source states, having v messages and $\lambda \cdot \binom{v}{t}/\binom{k}{t}$ encoding rules, that is $(t-1)$-fold secure against spoofing. Conversely, if there is an authentication code for k equiprobable source states, having v messages and $\binom{v}{t}/\binom{k}{t}$ encoding rules, that is $(t-1)$-fold secure against spoofing, then there is a Steiner t-$(v,k,1)$ design.

IV. STINSON'S CONSTRUCTIONS & RECENT RESULTS

Using the notation introduced in Section II-A, we review in Tables I and II previous constructions from [6], [1] for equiprobable source probability distributions. This lists all presently known optimal authentication codes with perfect secrecy.

V. NEW CONSTRUCTIONS

Starting from the condition of perfect t-fold secrecy, we obtain via Bayes' Theorem that

$$p_S(S^*|M^*) = \frac{p_M(M^*|S^*)p_S(S^*)}{p_M(M^*)} = \frac{\sum_{\{e\in\mathcal{E}:S^*=f_e(M^*)\}} p_E(e)p_S(S^*)}{\sum_{\{e\in\mathcal{E}:M^*\subseteq M(e)\}} p_E(e)p_S(f_e(M^*))} = p_S(S^*).$$

It follows

Lemma 3: An authentication code has perfect t-fold secrecy if and only if, for every positive integer $t^* \leq t$, for every set M^* of t^* messages observed in the channel and for every set S^* of t^* source states, we have

$$\sum_{\{e\in\mathcal{E}:S^*=f_e(M^*)\}} p_E(e) = \sum_{\{e\in\mathcal{E}:M^*\subseteq M(e)\}} p_E(e)p_S(f_e(M^*)).$$

Hence, if the encoding rules in a code are used with equal probability, then for every $t^* \leq t$, a given set of t^* messages occurs with the same frequency in each t^* columns of the encoding matrix.

We can now establish an extension of the main theorem in [1]. Our construction yields optimal authentication codes which are multi-fold secure against spoofing and provide perfect multi-fold secrecy.

Theorem 4: Suppose there is a Steiner t-$(v,k,1)$ design, where $\binom{v}{t^*}$ divides the number of blocks b for every positive integer $t^* \leq t-1$. Then there is an optimal authentication code for k equiprobable source states, having v messages and $\binom{v}{t}/\binom{k}{t}$ encoding rules, that is $(t-1)$-fold secure against spoofing and simultaneously provides perfect $(t-1)$-fold secrecy.

Proof: Let $\mathcal{D} = (X, \mathcal{B})$ be a Steiner t-$(v,k,1)$ design, where $\binom{v}{t^*}$ divides b for every positive integer $t^* \leq t-1$. By Theorem 3, the authentication code has $(t-1)$-fold security against spoofing attacks. Hence, it remains to prove that the code also achieves perfect $(t-1)$-fold secrecy under the assumption that the encoding rules are used with equal probability. With respect to Lemma 3, we have to show that, for every $t^* \leq t-1$, a given set of t^* messages occurs with the same frequency in each t^* columns of the resulting encoding matrix. This can be accomplished by ordering, for each $t^* \leq t-1$, every block of $\mathcal{D}$ in such a way that every t^*-subset of X occurs in each possible choice in precisely $b/\binom{v}{t^*}$ blocks. Since every t^*-subset of X occurs in exactly $\lambda_{t^*} = \binom{v-t^*}{t-t^*}/\binom{k-t^*}{t-t^*}$ blocks due to Lemma 1, necessarily $\binom{k}{t^*}$ must divide λ_{t^*}. By Lemma 2 (b), this is equivalent to saying that $\binom{v}{t^*}$ divides b. To show that the condition is also sufficient, we consider the bipartite (t^*-subset, block) incidence graph of $\mathcal{D}$ with vertex set $\binom{X}{t^*} \cup \mathcal{B}$, where $(\{x_i\}_{i=1}^{t^*}, B)$ is an edge if and only if $x_i \in B$ $(1 \leq i \leq t^*)$ for $\{x_i\}_{i=1}^{t^*} \in \binom{X}{t^*}$ and $B \in \mathcal{B}$. An ordering on each block of $\mathcal{D}$ can be obtained via an edge-coloring of this graph using $\binom{k}{t^*}$ colors in such a way that each vertex $B \in \mathcal{B}$ is adjacent to one edge of each color,

and each vertex $\{x_i\}_{i=1}^{t^*} \in \binom{X}{t^*}$ is adjacent to $b/\binom{k}{t^*}$ edges of each color. Specifically, this can be done by first splitting up each vertex $\{x_i\}_{i=1}^{t^*}$ into $b/\binom{k}{t^*}$ copies, each having degree $\binom{k}{t^*}$, and then by finding an appropriate edge-coloring of the resulting $\binom{k}{t^*}$-regular bipartite graph using $\binom{k}{t^*}$ colors. The claim follows now by taking the ordered blocks as encoding rules, each used with equal probability. ■

Remark 1: It follows from the proof that we may obtain optimal authentication codes that provide $(t-1)$-fold security against spoofing and at the same time perfect $(t'-1)$-fold secrecy for $t' \leq t$, when the assumption of the above theorem holds with $\binom{v}{t^*}$ divides b for every positive integer $t^* \leq t'-1$.

As an application, we give an infinite class of optimal codes which are two-fold secure against spoofing and achieve perfect two-fold secrecy. This appears to be the first infinite class of authentication and secrecy codes with these properties.

Theorem 5: For all positive integers $v \equiv 2 \pmod{24}$, there is an optimal authentication code for $k = 4$ equiprobable source states, having v messages, and $v(v-1)(v-2)/24$ encoding rules, that is two-fold secure against spoofing and provides perfect two-fold secrecy.

Proof: We will make use of Steiner quadruple systems (cf. Section II-A). Hanani [15] showed that a necessary and sufficient condition for the existence of a SQS(v) is that $v \equiv 2$ or $4 \pmod 6$ $(v \geq 4)$. Hence, the condition $v \mid b$ is fulfilled when $v \equiv 2$ or $10 \pmod{24}$ and the condition $\binom{v}{2} \mid b$ when $v \equiv 2 \pmod{12}$ in view Lemma 2 (b). Therefore, if we assume that $v \equiv 2 \pmod{24}$, then we can apply Theorem 4 to establish the claim. ■

We present the smallest example:

Example 1: An optimal authentication code for $k = 4$ equiprobable source states, having $v = 26$ messages, and $b = 650$ encoding rules, that is two-fold secure against spoofing and provides perfect two-fold secrecy can be constructed from a Steiner quadruple system SQS(26). Each encoding rule is used with probability $1/650$.

Remark 2: For $v = 26$, the first SQS(v) was constructed by Fitting [16], admitting a v-cycle as an automorphism (*cyclic* SQS(v)). We generally remark that the number $N(v)$ of non-isomorphic SQS(v) is only known for $v = 8, 10, 14, 16$ with $N(8) = N(10) = 1$, $N(14) = 4$, and $N(16) = 1{,}054{,}163$ (cf. [17]). Lenz [18] proved that for the admissible values of v, the number $N(v)$ grows exponentially, i.e. $\liminf_{v\to\infty} \frac{\log N(v)}{v^3} > 0$. For comprehensive survey articles on Steiner quadruple systems, we refer the reader to [19], [20]. For classifications of specific classes of highly regular Steiner quadruple systems and Steiner designs, see, e.g., [21], [22].

ACKNOWLEDGMENT

The author thanks Doug Stinson for an interesting conversation on this topic. The author gratefully acknowledges support of his work by the Deutsche Forschungsgemeinschaft (DFG) via a Heisenberg grant (Hu954/4) and a Heinz Maier-Leibnitz Prize grant (Hu954/5).

REFERENCES

[1] M. Huber, "Authentication and secrecy codes for equiprobable source probability distributions", in *Proc. IEEE International Symposium on Information Theory (ISIT) 2009*, pp. 1105–1109, 2009.
[2] E. N. Gilbert, F. J. MacWilliams and N. J. A. Sloane, "Codes which detect deception", *Bell Syst. Tech. J.*, vol. 53, pp. 405–424, 1974.
[3] G. J. Simmons, "Authentication theory/coding theory", in *Advances in Cryptology – CRYPTO '84*, ed. by G. R. Blakley and D. Chaum, Lecture Notes in Computer Science, vol. 196, Springer, Berlin, Heidelberg, New York, pp. 411–432, 1985.
[4] G. J. Simmons, "A survey of information authentication", in *Contemporary Cryptology: The Science of Information Integrity*, ed. by G. J. Simmons, IEEE Press, Piscataway, pp. 379–419, 1992.
[5] C. E. Shannon, "Communication theory of secrecy systems", *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, 1949.
[6] D. R. Stinson, "The combinatorics of authentication and secrecy codes", *J. Cryptology*, vol. 2, pp. 23–49, 1990.
[7] J. L. Massey, "Cryptography – a selective survey", in *Digital Communications*, ed. by E. Biglieri and G. Prati, North-Holland, Amsterdam, New York, Oxford, pp. 3–21, 1986.
[8] D. R. Stinson, "Combinatorial characterizations of authentication codes", *Designs, Codes and Cryptography*, vol. 2, pp. 175–187, 1992.
[9] D. R. Stinson and R. S. Rees, "Combinatorial characterizations of authentication codes II", *Designs, Codes and Cryptography*, vol. 7, pp. 239–259, 1996.
[10] Th. Beth, D. Jungnickel and H. Lenz, *Design Theory*, vol. I and II, Encyclopedia of Math. and Its Applications, vol. 69/78, Cambridge Univ. Press, Cambridge, 1999.
[11] C. J. Colbourn and J. H. Dinitz (eds.), *Handbook of Combinatorial Designs*, 2nd ed., CRC Press, Boca Raton, 2006.
[12] M. Huber, "Coding theory and algebraic combinatorics", in *Selected Topics in Information and Coding Theory*, ed. by I. Woungang *et al.*, World Scientific, Singapore, 38 pages, 2010 (in press). Preprint at arXiv:0811.1254v1.
[13] P. Schöbi, "Perfect authentication systems for data sources with arbitrary statistics" (presented at EUROCRYPT '86), unpublished.
[14] M. De Soete, "Some constructions for authentication - secrecy codes", in *Advances in Cryptology – EUROCRYPT '88*, ed. by Ch. G. Günther, Lecture Notes in Computer Science, vol. 330, Springer, Berlin, Heidelberg, New York, pp. 23–49, 1988.
[15] H. Hanani, "On quadruple systems", *Canad. J. Math.*, vol. 12, pp. 145–157, 1960.
[16] F. Fitting, "Zyklische Lösungen des Steiner'schen Problems", *Nieuw. Arch. Wisk.*, vol. 11, pp. 140–148, 1915.
[17] P. Kaski, P. R. J. Östergård and O. Pottonen, "The Steiner quadruple systems of order 16", *J. Combin. Theory, Series A*, vol. 113, pp. 1764–1770, 2006.
[18] H. Lenz, "On the number of Steiner quadruple systems", *Mitt. Math. Sem. Giessen*, vol. 169, pp. 55–71, 1985.
[19] A. Hartman and K. T. Phelps, "Steiner quadruple systems", in: *Contemporary Design Theory*, ed. by J. H. Dinitz and D. R. Stinson, Wiley, New York, pp. 205–240, 1992.
[20] C. C. Lindner and A. Rosa, "Steiner quadruple systems – A survey", *Discrete Math.*, vol. 22, pp. 147–181, 1978.
[21] M. Huber, "Almost simple groups with socle $L_n(q)$ acting on Steiner quadruple systems", *J. Combin. Theory, Series A*, 4 pages, 2010 (in press). Preprint at arXiv:0907.1281v1.
[22] M. Huber, *Flag-transitive Steiner Designs*, Birkhäuser, Basel, Berlin, Boston, 2009.