Combinatorial Designs for Authentication and Secrecy Codes
Combinatorial Designs for Authentication and Secrecy Codes

Michael Huber

University of Tuebingen
Sand 13, D-72076 Tuebingen
Germany
michael.huber@uni-tuebingen.de

Full text available at: http://dx.doi.org/10.1561/0100000044
Editorial Scope

Foundations and Trends® in Communications and Information Theory will publish survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians
Foundations and Trends® in Communications and Information Theory, 2008, Volume 5, 6 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328. Also available as a combined paper and online subscription.
Combinatorial Designs for Authentication and Secrecy Codes

Michael Huber

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

Abstract

Combinatorial design theory is a very active area of mathematical research, with many applications in communications and information theory, computer science, statistics, engineering, and life sciences. As one of the fundamental discrete structures, combinatorial designs are used in fields as diverse as error-correcting codes, statistical design of experiments, cryptography and information security, mobile and wireless communications, group testing algorithms in DNA screening, software and hardware testing, and interconnection networks. This monograph provides a tutorial on combinatorial designs, which gives an overview of the theory. Furthermore, the application of combinatorial designs to authentication and secrecy codes is described in depth. This close relationship of designs with cryptography and information security was first revealed in Shannon’s seminal paper on secrecy systems. We bring together in one source foundational and current contributions concerning design-theoretic constructions and characterizations of authentication and secrecy codes.
Contents

1 Introduction 1

1.1 Authentication and Secrecy Model 3
1.2 Combinatorial Designs: A Brief Historical Account 8
1.3 Some Group Theory 10

2 Combinatorial Design Theory 13

2.1 $t$-Designs and Finite Geometries 14
2.2 Latin Squares and Perpendicular Arrays 28
2.3 Authentication Perpendicular Arrays 32
2.4 Splitting Designs and Others 36

3 Applications to Authentication and Secrecy Codes 43

3.1 Secrecy Codes 44
3.2 General Authentication Codes 48
3.3 Authentication Codes with Secrecy 49
3.4 Authentication Codes with Secrecy in the Verification Oracle Model 63
3.5 Authentication Codes with Splitting 67
3.6 Authentication Codes with Arbitration 74
3.7 Other Applications 74
3.8 Synthesis and Discussion 76
4 Appendix

4.1 Multiply Homogeneous Permutation Groups 79

Acknowledgments 83

References 85
Introduction

Authenticity and secrecy are two crucial concepts in cryptography and information security. Concerning authenticity, typically communicating parties would like to be assured of the integrity of information they obtain via potentially insecure channels. Regarding secrecy, protection of the confidentiality of sensitive information shall be ensured in the presence of eavesdropping. 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 can be used to minimize the possibility of an undetected deception. The construction of such codes is of great importance and has been considered by many researchers over the last few decades. Often deep mathematical tools are involved in the constructions, mainly from combinatorics. This close relationship of cryptography and information security with combinatorics has been first revealed in Shannon’s landmark paper “Communication theory of secrecy systems” [182]: a key-minimal secrecy system provides perfect secrecy if and only if the encryption matrix is a Latin square and the keys are used with equal probability. The initial construction
of authentication codes goes back to Gilbert et al. [74], and uses finite projective planes. A more general and systematic theory of authenticity was developed by Simmons (see [183, 184, 185, 186, 187], and [188] for a survey). Further foundational works on authentication and secrecy codes have been carried out by Massey [152] and Stinson et al. [194, 195, 196, 201, 202]. A generalized information-theoretic framework for authentication was introduced by Maurer [157].

The purpose of this monograph is to describe in depth classical and current interconnections between combinatorial designs and authentication and secrecy codes. The latter also include the author’s recent [102, 106, 107] and new contributions (cf. Section 3.4) on multi-fold secure authentication and secrecy codes in various models. Moreover, this issue provides a tutorial overview on the theory of combinatorial designs. These fundamental discrete structures find applications in fields as diverse as error-correcting codes, statistical design of experiments, cryptography and information security, mobile and wireless communications, group testing algorithms in DNA screening, software and hardware testing, and interconnection networks. In particular, the last few years have witnessed an increasing body of work in the communications and information theory literature that makes substantial use of results in combinatorial design theory.

The organization of the monograph is as follows. Section 1.1 introduces the Shannon–Simmons model of information-theoretical authentication and secrecy. We define the important concepts of spoofing attacks and perfect secrecy. A short historical account on combinatorial designs is given in Section 1.2 Since permutation groups often play a crucial role in the construction of combinatorial designs, we introduce basic notions on permutation groups and group actions in Section 1.3 Section 2 provides a tutorial account on combinatorial design theory. We emphasize on the construction of various combinatorial structures including t-designs, finite geometries, Latin squares, orthogonal arrays, perpendicular and authentication perpendicular arrays, splitting t-designs, and others. These combinatorial structures provide essential tools for the construction and characterization of authentication and secrecy codes in the following section. A special notice is placed on examples for each type of combinatorial designs. We also briefly
point to the interplay between $t$-designs and error-correcting codes. Section 3 is devoted to various key applications of combinatorial designs to authentication and secrecy codes. Foundational and recent results concerning the construction and characterization of authentication and secrecy codes are exposed. Starting with Shannon’s classical result, we first deal with secrecy codes in Section 3.1. Authentication codes without any secrecy requirements are considered in Section 3.2. In Section 3.3, codes that offer both authenticity and secrecy are discussed in detail. We distinguish between arbitrary and equiprobable source probability distributions. The advantage of the source states being equiprobable distributed is that the number of encoding rules can be reduced. Section 3.4 is devoted to an extended authentication model, where the opponent can act proactively by having access to a verification oracle. Authentication codes with splitting are considered in Section 3.5. In such a code, several messages can be used to communicate a particular plaintext (non-deterministic encoding). We briefly mention authentication codes that permit arbitration in Section 3.6. In Section 3.7, further recent applications are highlighted which makes substantial use of combinatorial design theory. Finally, we conclude in Section 3.8 with a synthesis of the work and some directions for future research.

1.1 Authentication and Secrecy Model

We rely on the information-theoretical (or unconditional) secrecy model developed by Shannon [182], and by Simmons [183] including authentication. Information-theoretical security means that the security of the model is not dependent on any complexity assumptions and hence cannot be broken given unlimited computational resources. A well-known practical application of such a perfectly-secure system is the Washington–Moscow Hotline (“red telephone”) during the time of the cold war. Modern applications may include protection of digital data where cryptographic long-term security and/or confidentiality is strongly required, e.g., in archiving official documents, notarial contracts, court records, medical data, state secrets, copyright protection as well as further areas concerning e-government, e-health, e-publication, etc.
4 Introduction

The reader may be interested in the area of information-theoretical cryptography [156], long-term secure cryptography [33], post-quantum cryptography [15], and in the broad area of cryptography in general [77, 78, 159, 200].

1.1.1 Basic Preliminaries

We introduce the basic model of information-theoretical authentication and secrecy. Our notation follows, for the most part, that of [152, 195]. Figure 1.1 gives an illustration of the model (cf. [152, 195]).

In this basic 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. Variants of this model will be discussed in Sections 3.4, 3.5, and 3.6.

In what follows, let $S$ denote a set of $k$ source states (or plaintexts), $M$ a set of $v$ messages (or ciphertexts), and $E$ a set of $b$ encoding rules (or keys). Using an encoding rule $e \in E$, the transmitter encrypts a source state $s \in S$ to obtain the message $m = e(s)$ to be sent over the

![Fig. 1.1 Shannon–Simmons authentication and secrecy model.](http://dx.doi.org/10.1561/0100000044)
channel. The encoding rule is an injective function from \( S \) to \( M \), and is communicated to the receiver via a secure channel prior to any messages being sent. For a given encoding rule \( e \in E \), let \( M(e) := \{ e(s) \mid s \in S \} \) denote the set of valid messages. For an encoding rule \( e \) and a set \( M^* \subseteq M(e) \) of distinct messages, we define \( f_e(M^*) := \{ s \in S \mid e(s) \in M^* \} \), i.e., the set of source states that will be encoded under encoding rule \( e \) by a message in \( M^* \). Furthermore, we define \( E(M^*) := \{ e \in E \mid M^* \subseteq M(e) \} \), i.e., the set of encoding rules under which all the messages in \( M^* \) are valid. 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
\[
e^{-1}(m) = s \iff e(s) = m.
\]

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 \( a_{es} := e(s) \) (\( 1 \leq e \leq b, 1 \leq s \leq k \)).

### 1.1.2 Protection Against Spoofing Attacks

We introduce the scenario of a spoofing attack of order \( i \) (cf. \[152\]): Suppose that an opponent observes \( i \geq 0 \) 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 (see, for instance, \[196, 201, 27, 53, 165\]), however less is known for the cases \( i \geq 2 \). In this monograph, we especially 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
\( S \), the receiver and transmitter choose a probability distribution \( p_E \) on \( E \), called an 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 \( M \) 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 by Massey provides combinatorial lower bounds (for the proof, we follow [194, 195]).

**Theorem 1.1 (Massey [152]).** In an authentication code with \( k \) source states and \( v \) messages, for every \( 0 \leq i \leq t \), the deception probabilities are bounded below by

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

**Proof.** Let \( M^* \subset M \) denote a set of \( i \leq t \) distinct messages. We suppose that an opponent observes the \( i \) messages in the channel, and then sends a message \( m \in M \) not in \( M^* \). Let payoff\((m, M^*)\) denote the probability that the message \( m \) would be accepted by the receiver as authentic. Then

\[
\text{payoff}(m, M^*) = \frac{\sum_{e \in E(M^* \cup \{m\})} p(e) \cdot p(S = f_e(M^*))}{\sum_{e \in E(M^*)} p(e) \cdot p(S = f_e(M^*))}.
\]

It follows that

\[
\sum_{m \in M \setminus M^*} \text{payoff}(m, M^*) = k - i.
\]

Hence, there exists some \( m \in M \) not in \( M^* \) such that\( \text{payoff}(m, M^*) \geq (k - i)/(v - i) \). For every set \( M^* \) of \( i \) messages, the opponent can choose such an \( m \). This defines a deception strategy in which the transmitter/receiver can be deceived with probability at least \((k - i)/(v - i)\).

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

We address Shannon’s fundamental idea of perfect secrecy (cf. [182]): An authentication code is said to have perfect secrecy if

\[ p_S(s|m) = p_S(s) \]

for every source state \( s \in S \) and every message \( m \in M \).

That is, the a posteriori probability that the source state is \( s \), given that the message \( m \) is observed, is identical to the a priori probability that the source state is \( s \).

It can easily be shown via Bayes’ Theorem that

\[ p_S(s|m) = \frac{p_M(m|s) \cdot p_S(s)}{p_M(m)} \]

(1.1)

\[ = \frac{\sum_{e \in E|e(s) = m} p_E(e) \cdot p_S(s)}{\sum_{e \in E|e \in M(e)} p_E(e) \cdot p_S(e^{-1}(m))}. \]

(1.2)

Moreover, we introduce the concept of perfect multi-fold secrecy established by Stinson [195], which generalizes Shannon’s perfect (one-fold) secrecy. An alternative definition has been given by Godlewski and Mitchell [75]. Instead of assuming that each encoding rule is used to encode only one message, the situation is extended in a natural way: each encoding rule is used to encode up to \( t_S \) messages for some positive integer \( t_S \). More formally, 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. Obviously, for the case \( t_S = 1 \) this coincides with the definition of perfect secrecy.

When clear from the context, we often only write \( t \) instead of \( t_A \) respectively \( t_S \).

As the encoding rules have to be communicated to the receiver via a secure channel, i.e. \( \log_2 b \) bits for \( b \) encoding rules, we want to minimize the number of encoding rules. With respect to the minimal
number, we will deal with the construction and characterization of *optimal* authentication and secrecy codes in Section 3.

**Remark 1.1.** We note that the term *secrecy code* (sometimes also secrecy system) is customarily used in the above model to describe a cipher that achieves Shannon’s perfect secrecy over *noiseless* channels. This should not be confused with the same expression often used today for describing codes that can achieve both reliable and secure communication over *noisy* channels (also known as wiretap channels). For recent developments on information-theoretic security for noisy channels, we refer to the monograph [142] and the references therein.

### 1.2 Combinatorial Designs: A Brief Historical Account

Combinatorial designs have a long and rich history of work. We briefly highlight three historical examples:

Leonhard Euler considered in 1782 the following problem [64], posed by Catherine the Great according to folklore. This problem came to known as *Euler’s 36 Officers Problem*:

“A very curious question, which has exercised for some time the ingenuity of many people, has involved me in the following studies, which seem to open a new field of analysis, in particular the study of combinations. The question resolves around arranging 36 officers to be drawn from 6 different ranks and also from 6 different regiments so that they are ranged in a square so that in each line (both horizontal and vertical) there are 6 officers of different ranks and different regiments.”
This question asks for finding two orthogonal Latin squares of order 6. Euler correctly conjectured that this was impossible, and a complete proof with an exhaustive search of all Latin squares of order 6 were given in 1900 by Tarry [207, 208]. A short proof is due to Stinson [193].

The Swiss geometer Jakob Steiner posed in 1853 in his classical “Combinatorische Aufgabe” [192] the following question:

“Welche Zahl, \( N \), von Elementen hat die Eigenschaft, dass sich die Elemente so zu dreien ordnen lassen, dass je zwei in einer, aber nur in einer Verbindung vorkommen?”

[Transl.: “For what number, \( N \), of elements is it possible to arrange the elements in triplets, so that every pair of elements is contained in one and only one triplet?”]

Writing \( v, k, \) and \( t \) instead of \( N, 3, \) and \( 2, \) respectively, leads us to the definition of what is now called a Steiner \( t \)-design (or a Steiner system, cf. Definition 2.1). However, there had been earlier work on these combinatorial designs going back to, in particular, Plücker, Woolhouse, and most notably Kirkman.

Thomas Kirkman’s famous 15 Schoolgirl Problem, which he proposed in 1850 in the popular magazine The Lady’s and Gentleman’s Diary [132], states as follows:

“Fifteen young ladies in a school walk out three abreast for seven days in succession: it is required to arrange them daily so that no two shall walk twice abreast.”
This is equivalent to the problem of constructing a Steiner 2-design with parameters $k = 3$ and $v = 15$, having the extra requirement that the set of triples can be partitioned into seven ‘parallel classes’. Kirkman’s problem as well as the more general case for other possible values of $v$ attracted great interest among late 19th and early 20th century mathematicians, including contributions by Burnside, Cayley and Sylvester. However, it was not until 1971 that the general problem was completely resolved by Ray-Chaudhuri and Wilson [173], showing that there exists at least one such design for every $v \equiv 3 \pmod{6}$. For $v = 15$, there are seven different solutions to the problem (up to isomorphism). For all other admissible values $v \geq 21$, the number of solutions remains unknown up to the present.

For a detailed account on the history of combinatorial designs, we refer the interested reader, e.g., to [43, Chap. I.2] and [225].

1.3 Some Group Theory

Often permutation groups play a crucial role in the construction of combinatorial designs. We introduce basic notions on permutation groups and group actions in this section. We will restrict ourselves to finite groups, although most of the concepts also make sense for infinite groups.

Let $X$ be a non-empty finite set. The set $\text{Sym}(X)$ of all permutations of $X$ with respect to the composition

$$x^{gh} := (x^g)^h \quad \text{for} \quad x \in X \quad \text{and} \quad g, h \in \text{Sym}(X)$$

forms a group, called the symmetric group on $X$. If $X = \{1, \ldots, v\}$, then we write $S_v$ for the symmetric group of degree $v$. Clearly, $\text{Sym}(X) \cong S_v$ if and only if $|X| = v$. 
A group $G$ acts (or operates) on $X$, if to each element $g \in G$ a permutation $x \mapsto x^g$ of $X$ is assigned such that

(i) $x^1 = x$ for all $x \in X$ (where $1 = 1_G$ denotes the identity element of $G$),

(ii) $(x^g)^h = x^{gh}$ for all $x \in X$ and all $g, h \in G$.

Evidently, these properties are fulfilled if and only if the map

$$\varphi : g \mapsto (x \mapsto x^g)$$

of $G$ into $\text{Sym}(X)$ is a group homomorphism. In general, any homomorphism $\varphi$ of $G$ into $\text{Sym}(X)$ is said to be an action (or a permutation representation) of $G$ on $X$. If $\ker(\varphi) = 1$ for the kernel of $\varphi$, then $G$ acts faithfully on $X$; in this case, $G$ is a called a permutation group on $X$. If $\ker(\varphi) = G$, then $G$ operates trivially on $X$. The degree of a permutation group is the size of $X$.

**Example 1.1.** The group of symmetries of a three-dimensional cube (cf. Figure 2.3) acts on various sets including the set of 8 vertices, the set of 6 faces, the set of 12 edges, and the set of 4 principal diagonals. Properties (i) and (ii) are clearly satisfied in each case.

Let $G_1$ and $G_2$ be permutation groups acting on the sets $X_1$ and $X_2$, respectively. Then, $G_1$ and $G_2$ are called permutation isomorphic, if there exists a group isomorphism $\sigma : G_1 \rightarrow G_2$ and a bijective map $\tau : X_1 \rightarrow X_2$ with

$$(x^g)^\tau = (x^\tau)^{(g^\sigma)}$$

for all $x \in X_1$ and all $g \in G_1$. Essentially, this means that the groups $G_1$ and $G_2$ are “the same” except for the labeling of the points.

Let $G$ be a group acting on $X$. For $x \in X$, the subgroup

$$G_x := \{ g \in G \mid x^g = x \}$$

denotes the (point-)stabilizer of $x$ in $G$ and the set

$$x^G := \{ x^g \mid g \in G \}$$
is the orbit of $x$ under $G$ (or the $G$-orbit of $x$). For $B \subseteq X$, let

$$G_B := \{ g \in G \mid B^g = B \}$$

be its setwise stabilizer. The order of $G$ is denoted by $|G|$.

A group $G$ acting on $X$ is called transitive on $X$, if $G$ has only one orbit, i.e. $x^G = X$ for all $x \in X$. Equivalently, $G$ is transitive if for any two points $x, y \in X$ there exists an element $g \in G$ with $x^g = y$. For a positive integer $t \leq |X|$, we call $G$ to be $t$-transitive, if for any two injective $t$-tuples $(x_1, x_2, \ldots, x_t)$ and $(y_1, y_2, \ldots, y_t)$ there exists an element $g \in G$ with $x_i^g = y_i$ for all $1 \leq i \leq t$. We say that $G$ is $t$-homogeneous, if it is transitive on the set of all $t$-subsets of $X$. Obviously, $t$-transitive implies $t$-homogeneous.

**Example 1.2.** The symmetric group $S_v$ is $v$-transitive, and the alternating group $A_v$ (i.e., the subgroup of $S_v$ consisting of all even permutations) is $(v - 2)$-transitive in their actions on the set $\{1, \ldots, v\}$ ($v \geq 3$).

We will list all finite multiply homogeneous permutation groups in Appendix 4.1. We note that this classification relies on the Classification of the Finite Simple Groups (CFSG), one of the most powerful tools of modern algebra.

For a detailed treatment of finite group theory and permutation groups, we refer the reader to [5, 38, 41, 61, 117, 118, 139, 223].

Full text available at: http://dx.doi.org/10.1561/0100000044.
References


85
References


References

References


References

References


References

References


References


