Bounds and Constructions for Authentication - Secrecy Codes with Splitting

Marijke De Soete
Seminar of Geometry and Combinatorics
State University of Ghent, Krijgslaan 281, B-9000 Ghent Belgium.

It is the aim to deal with codes having unconditional security, which means that the security is independent of the computing power. Analogously to the theory of unconditional secrecy due to Shannon [12], Simmons developed a theory of unconditional authentication [10]. In this paper we give some new bounds and constructions for authentication/secrecy codes with splitting.

Consider a transmitter who wants to communicate a source to a remote receiver by sending messages through an imperfect communication channel. Then there are two fundamentally different ways in which the receiver can be deceived. The channel may be noisy so that the symbols in the transmitted message can be received in error, or the channel may be under control of an opponent who can either deliberately modify legitimate messages or else introduce fraudulent ones. Simmons [10] showed that both problems could be modeled in complete generality by replacing the classical noisy communications channel of coding theory with a game - theoretic noiseless channel in which an intelligent opponent, who knows the system and can observe the channel, plays so as to optimize his chances of deceiving the receiver. To provide some degree of immunity to deception (of the receiver), the transmitter also introduces redundancy in this case, but does so in such a way that, for any message the transmitter may send, the altered messages that the opponent would introduce using his optimal strategy are spread randomly. Authentication theory is concerned with devising and analizing schemes (codes) to achieve this "spreading".
In the mathematical model there are three participants: a *transmitter*, a *receiver* and an *opponent*. The transmitter wants to communicate some information to the receiver. The opponent wanting to deceive the receiver, can either impersonate the receiver, making him accept a fraudulent message as authentic, or, modify a message which has been sent by the transmitter.
Let S denote the set of k source states, M the set of v messages and E the set of b encoding rules.

S. Goldwasser (Ed.): Advances in Cryptology - CRYPTO '88, LNCS 403, pp. 311-317, 1990.

A *source state* $s \in S$ is the information that the transmitter wishes to communicate to the receiver. The transmitter and receiver will have secretly chosen an *encoding rule* $e \in E$ beforehand. An encoding rule e will be used to determine the message $e(s)$ to be sent to communicate any source state s. In a model with *splitting*, several messages can be used to determine a particular source state. However, in order for a receiver to be able to uniquely determine the source state from the message sent, there can be at most one source state which is encoded by any given message $m \in M$, for a given encoding rule $e \in E$ (this means: $e(s) \neq e(s')$ if $s \neq s'$).

An opponent will play *impersonation* or *substitution*. When the opponent plays impersonation, he sends a message to the receiver, attempting to have the receiver accept the message as authentic. When the opponent plays substitution, he waits until a message m has been sent, and then replaces m with another message m', so that the receiver is misled as to the state of source. More generally, an opponent can observe i (≥ 0) distinct messages being sent over the channel knowing that the same key is used to transmit them, but ignoring this key. If we consider the code as a secrecy system, then we make the assumption that the opponent can only observe the messages being sent. Our goal is that the opponent be unable to determine any information regarding the i source states from the i messages he has observed.

We shall use the following notations. Given an encoding rule e, we define $M(e) = \{e(s)|s \in S\}$, i.e. the set of messages permitted by encoding rule e, and let $|M(e)| = k(e)$. For a set of distinct messages $M' \subset M$ and an encoding rule e, define $f_e(M') = \{s \in S|e(s) \in M'\}$, i.e. the set of source states which will be encoded under encoding rule e by a message in M'. Define also $E(M') = \{e \in E|M' \subseteq M(e)\}$, i.e. the set of encoding rules under which all the messages in M' are permitted.

The following scenario for authentication is investigated. After the observation of i messages $M' \subset M$, the opponent sends a message m' to the receiver, $m' \notin M'$, hoping to have it accepted as authentic. This is called a *spoofing attack of order* i [6], with the special cases $i = 0$ and $i = 1$ corresponding respectively to the impersonation and substitution game. The last games have been studied extensively by several authors (see [2], [5], [10], [13]).

For any i, there will be a probability on the set of i source states which occur. We ignore the order in which the i source states occur, and assume that no source state occurs more than once. Also, we assume that any set of i source states has a non-zero probability of occuring. Given a set of i source states, we define $p(S)$ to be the probability that the source states in S occur.

Given the probability distributions on the source states described above, the receiver and transmitter will choose a probability distribution for E, called an *encoding strategy*. If splitting occurs, then they will also determine a *splitting strategy* to determine $m \in M$, given $s \in S$ and $e \in E$ (this corresponds to non-deterministic encoding). The transmitter/receiver will determine these strategies to minimize the chance that an opponent can deceive them.

Once the transmitter/receiver have chosen encoding and splitting strategies, we can define for each $i \geq 0$ a probability denoted P_{d_i}, which is the probability that the opponent can deceive the transmitter/receiver with a spoofing attack of order i. We denote by $AC(k, v, b)$ an authentication system with k source states, v messages and b encoding rules.

1 Secrecy

Considering the secrecy of a code, we desire no information be conveyed by the observation of the messsages. A code has *perfect L-fold secrecy* (Stinson [14]) if, for every set M_1 of at most L messages observed in the channel, and for every set S_1 of at most $|M_1|$ source states, we have $p(S_1/M_1) = p(S_1)$. This means that observing a set of at most L messages in the channel does not help the opponent to determine the L source states. On the other hand, a code is said to be *Cartesian* ([2], [13]) if any message uniquely determines the source state, independent of the particular encoding rule being used .

2 Bounds on P_{d_i} and b

Bounds on P_{d_0} and P_{d_1} for authentication codes with splitting depending on the entropies of the various probability distributions can be found in [2], [9], [10], [13] and [14]. The most important bounds are given by:

$$P_{d_0} \geq 2^{H(MES)-H(E)-H(M)} = 2^{H(M|ES)+H(S)-H(M)}$$

and for a substitution with secrecy

$$P_{d_1} \geq 2^{-H(E/M)} = 2^{H(M)-H(E)-H(S)+H(M/ES)}.$$

The following bounds for an impersonation, resp. a substitution game are proven in [4]:

$$P_{d_0} \geq min_{e \in E} \frac{k(e)}{v} \text{ (see also [9] [10])},$$

$$P_{d_1} \geq min_{e \in E} \frac{k(e) - max_{s \in S}|e(s)|}{v - max_{s \in S}|e(s)|}.$$

For codes without splitting this results in the known bounds $P_{d_0} \geq k/v$ and $P_{d_1} \geq (k-1)/(v-1)$ ([6], [13], [14]).
These bounds can also be generalized for a spoofing attack of order i [4] to

$$P_{d_i} \geq min_{e \in E} \frac{k(e) - i \cdot max_{s \in S}|e(s)|}{v - i \cdot max_{s \in S}|e(s)|}.$$

An authentication system which achieves equality $\forall i$, $0 \leq i \leq L$, is called *L-fold secure against spoofing* (this is a generalization of the definition for codes without splitting, see [6], [14]).

The number of keys is basically influenced by the following two aspects: (i) the distribution on the source states and (ii) the secrecy of the code. In [4] we obtain the following bound:
If a code achieves perfect L-fold secrecy and is $(L-1)$-fold secure against spoofing, then

$$b \geq \frac{v \cdot (v - max_{s \in S}|e(s)|) \cdots (v - (L-1) \cdot max_{s \in S}|e(s)|)}{L!}.$$

Analogously as for codes without splitting [14], we define an *optimal L-code* to be a code which achieves perfect L-fold secrecy, which is $(L-1)$-fold secure against spoofing and which meets equality in the foregoing formula.

3 Constructions for authentication codes with arbitrary source distribution

3.1 Authentication codes derived from partial geometries

A (finite) *partial geometry* (PG) is an incidence structure $\mathcal{G} = (P, B, I)$ in which P and B are disjoint (nonempty) sets of objects called *points* and *lines* resp., and for which I is a symmetric point-line incidence relation satisfying the following axioms:

1. Each point is incident with $1 + t$ lines ($t \geq 1$) and two distinct points are incident with at most one line.

2. Each line is incident with $1 + s$ points ($s \geq 1$) and two distinct lines are incident with at most one point.

3. If x is a point and L a line not incident with x, then there are exactly α, ($\alpha \geq 1$) points $x_1, x_2, \ldots, x_\alpha$ and α lines $L_1, L_2, \ldots, L_\alpha$ such that $x\ I\ L_i\ I\ x_i\ IL$, $i = 1, 2, \ldots, \alpha$.

Partial geometries were introduced by R. C. Bose. The partial geometries with $\alpha = 1$ are the *generalized quadrangles* (GQ).

There holds $|P| = (s+1)(st+\alpha)/\alpha$, $|B| = (t+1)(st+\alpha)/\alpha$, $\alpha(s+t+1-\alpha) | st(s+1)(t+1)$ and $(s+1-2\alpha)t \leq (s-1)(s+1-\alpha)^2$ (and dually). We remark that the dual incidence structure $G' = (P', B', I')$, $P' = B$, $B' = P$, $I' = I$, is a partial geometry with parameters $t' = s$, $s' = t$ and $\alpha' = \alpha$.
Further information about PG and GQ can be found in [7].

1. From a generalized quadrangle of order (s, t), $s, t > 1$, we can define the following two authentication codes without splitting [3].

   - *A GQ of order (s,t) defines a cartesian $AC(t+1, (t+1)s, ts^2)$ which is 0-fold secure against spoofing and for which $P_{d_1} = 1/s$.*
   - *If the GQ contains a regular point, the foregoing code can be improved to an $AC(t+1, (t+1)s, (t+1)s^2)$ which is 0-fold secure against spoofing, which has perfect 1-fold secrecy, and for which $P_{d_1} = 1/s$.*

2. A PG with parameters $s, t \geq 1$, $\alpha > 1$ defines an $AC(t+1, (t+1)s, (t+1)st(s+1-\alpha))$ code which has 0-fold security against spoofing and which has perfect 1-fold secrecy [4].

3. A *spread* of a PG $\mathcal{G}$ is a set $\mathcal{R}$ of lines of $\mathcal{G}$ such that each point of $\mathcal{G}$ is incident with a unique line of $\mathcal{R}$. Hence there holds $|\mathcal{R}| = (st + \alpha)/\alpha$.
Let $\mathcal{G}$ be a PG with parameters $s, t > 1$, $\alpha \geq 1$, containing a spread $\mathcal{R}$. Then we can define the following authentication codes.

   - *For $\alpha > 1$, $\mathcal{G}$ defines an optimal 1-code with splitting* [4].
   - *For $\alpha = 1$, $\mathcal{G}$ defines an optimal 1-code without splitting* [3].

3.2 Authentication codes derived from designs

Consider an *affine resolvable BIB-design.* This is a 2-(v, k, λ) design $\mathcal{D} = (P, B, I)$ for which there exists a partition of $B = B_1 \cup B_2 \ldots B_r$ of the block set, $|B_i| = n$, such that each point occurs exactly once in the blocks of any set B_i, $1 \leq i \leq r$ and any two blocks of different sets have exactly μ, $\mu > 0$, points in common [1]. There holds $|B| = rn$, $|P| = kn$, $\lambda = r(k-1)/(nk-1)$ and $k = \mu n$.

In [4] we construct the following authentication code with splitting:

An affine resolvable design $\mathcal{D}$ defines an $AC(n, kn, (r-1)n^2)$ which is 0-fold secure against spoofing, which has 1-fold secrecy, and for which $P_{d_1} = \lambda/(r-1)$.

References

[1] Beth T., Jungnickel D., Lenz H., *Design Theory.* Wissenschaftsverlag Bibliografisches Institut Mannheim, 1985.

[2] Brickell E. F., *A few results in message authentication.* Proc. of the 15th Southeastern Conf. on Combinatorics, Graph theory and Computing, Boca Raton LA (1984), 141–154.

[3] De Soete M., *Some Constructions for Authentication / Secrecy codes*, Proceedings of Eurocrypt'88, Davos, L.N.C.S., to appear.

[4] De Soete M., *New Bounds and Constructions for Authentication / Secrecy Codes with Splitting.* In preparation.

[5] Gilbert E. N., MacWilliams F. J., Sloane N. J. A., *Codes which detect deception.* Bell Sys. Techn. J., Vol.53–3 (1974), 405–424.

[6] Massey J. L., *Cryptography - A Selective Survey.* Proc. of 1985 Int. Tirrenia Workshop on Digital Communications, Tirrenia, Italy, 1985, Digital Communications, ed. E. Biglieri and G. Prati, Elsevier Science Publ. 1986, 3–25.

[7] Payne S. E., Thas J. A., *Finite generalized quadrangles.* Research Notes in Math. #110, Pitman Publ. Inc. 1984.

[8] Shrikhande S. S., *Affine resolvable balanced incomplete block designs: a survey.* Aequat. Math. 14 (1976), 251–269.

[9] Simmons G. J., *Message Authentication: A Game on Hypergraphs.* Proc. of the 15th Southeastern Conf. on Combinatorics, Graph Theory and Computing, Baton Rouge LA Mar 5–8 1984, Cong. Num. 45 (1984), 161–192.

[10] Simmons G. J., *Authentication theory / Coding theory.* Proc. of Crypto'84, Santa Barbara, CA, Aug.19–22, 1984, Advances in Cryptology, ed. R. Blakley, Lect. Notes Comp. Science 196, Springer 1985, 411–432.

[11] Simmons G. J., *A natural taxonomy for digital information authentication schemes.* Proc. of Crypto '87, Santa Barbara, CA, Aug 16–20, 1987, to appear in Advances in Cryptology, ed. C. Pomerance, Springer Verlag, Berlin.

[12] Shannon C. E., *Communication Theory of Secrecy Systems.* Bell Technical Journal, Vol.28 (1949), 656–715.

[13] Stinson D. R., *Some Constructions and Bounds for Authentication Codes.* J. Cryptology 1 (1988), 37–51.

[14] Stinson D. R., *A construction for authentication / secrecy codes from certain combinatorial designs.* Crypto '87, Santa Barbara, CA, Aug 16–20, 1987, to appear in J. Cryptology.