Authentication Codes, from Error-Correcting Codes

Henk van Tilborg

SPCodingSchool
January 2015
Campinas, Brazil

1 Definitions and introduction

In an ***authentication code*** (Gilbert e.o. 1974, Simons 1992), a sender S replaces a message by a "codeword" so that the intended receiver R can recover the message and check that it indeed came from this sender.

So, secrecy is not necessarily an issue here.

Sender and receiver share a common key.

The main goal is:

unconditional security

Unconditional security means:

- No mathematical assumptions like the difficulty of factoring large numbers or taking a discrete logarithm.
- The adversary is assumed to have unlimited computer power.
- So, also symmetric cryptosystems are out, because they are vulnerable to an exhaustive key search attack.

Aspects that will come up are:

- large data files
- multiple use
- secrecy.

More formally:

Definition 1.1: *An authentication code is a triple* $(\mathcal{M}, \mathcal{K}; \mathcal{C})$ *and a collection of mappings* $f_k : \mathcal{M} \times \mathcal{K} \to \mathcal{C}$, $k \in \mathcal{K}$, *such that for each fixed* k *in* $\mathcal{K}$ *the mapping* f_k *is one-to-one.*

$$f_k(m) = \underline{c}. \tag{1}$$

If privacy is no issue, one often makes the authentication code "systematic". In this case, f_k is of the form

$$f_k(m) = (m; \tau_k(m)). \tag{2}$$

Here τ is a mapping from $\mathcal{M} \times \mathcal{K}$ to some set $\mathcal{T}$ and $\tau_k(m)$ is called the ***tag***.

The assumption is that the adversary knows the authentication code, but not the key.

"Unconditional security"does not mean that an adversary can not fool the system. If an adversary guesses a correct c, so $c = f_k(m)$ for some m under the right key k, she will get message m accepted by R as authentic from S.

A toy example:
S wants to send a single bit of information (a "yes" or a "no") to R by means of a word of length 2. S and R have 4 possible keys available. They make use of the following scheme:

key \ codeword	00	01	10	11
1	0	1	–	–
2	1	–	0	–
3	–	0	–	1
4	–	–	1	0

message

So, message 1 will be sent as codeword 11 under key 3.

There are two types of attack.

Impersonation The adversary sends a codeword $\hat{c}$ and hopes that it gets accepted.

Substitution The adversary replaces a transmitted codeword c by a different $\hat{c}$ and hopes that it gets accepted.

	00	01	10	11
1	0	1	–	–
2	1	–	0	–
3	–	0	–	1
4	–	–	1	0

The probability of a successful *impersonation attack* is given by $P_I = 1/2$.

The probability of a successful *substitution attack* is also given by $P_S = 1/2$.

But there is unconditional secrecy!

The best that an impersonator can do is transmit a codeword that for the largest fraction of keys will be accepted.

Assuming a uniform distribution on $\mathcal{K}$ and $\mathcal{M}$ one has:

$$P_I \geq \frac{|\mathcal{M}|}{|\mathcal{C}|}. \tag{3}$$

because for each key k exactly $|\mathcal{M}|$ of the $|\mathcal{C}|$ possible codewords represent an authenticated message $f_k(m)$.

For a substitution attack you can do not better than transmit a codeword that for the largest fraction of the keys will be accepted that are possible, given the transmitted codeword.

Again assuming a uniform distribution on $\mathcal{K}$ and $\mathcal{M}$, one has:

$$P_S \geq \frac{|\mathcal{M}| - 1}{|\mathcal{C}| - 1}. \tag{4}$$

because when a codeword is observed, at least $|\mathcal{M}| - 1$ of the remaining $|\mathcal{C}| - 1$ codewords are still authentic.

The maximum of the probabilities P_I and P_S is often called the probability of successful deception: $P_D = \max\{P_I,\ P_S\}$.

Without proof (Johansson 1994) we quote

$$P_D \geq \frac{1}{\sqrt{|\mathcal{K}|}}. \tag{5}$$

This bound is called the *square root bound*. Authentication codes meeting this bound are called perfect.

We also cite:

Theorem 1.2: *A necessary condition for an authentication code to be perfect is that*

$$|\mathcal{M}| \leq \sqrt{|\mathcal{K}|} + 1. \tag{6}$$

This implies that keys are twice as long as messages are.

2 Projective plane construction

This construction is due to E.N. Gilbert, F.J. MacWilliams, and N.J.A. Sloane (1974) and makes use of projective planes.

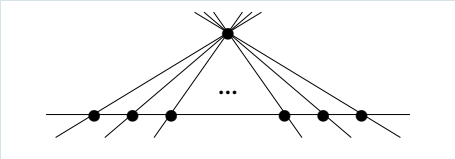
Definition 1.3: *A projective plane is a pair $(\mathcal{P}, \mathcal{L})$ where $\mathcal{P}$ is a finite set of elements, called points, and $\mathcal{L}$ a finite collection of subsets of $\mathcal{P}$, called lines, with the following properties:*

PP0: *There are at least four points, no three of which lie on the same line.*

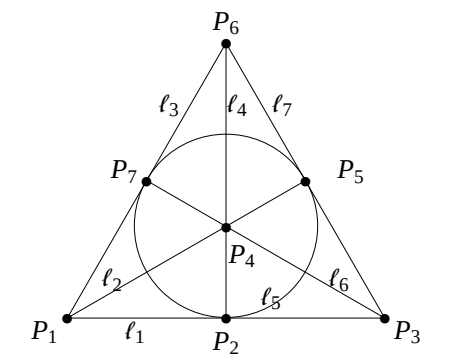
PP1: *For every pair of points there is a unique line going through them.*

PP2: *Every pair of lines intersect in a unique point.*

Requirement PP0 is there to avoid the trivial construction:



The best known example of a projective plane is Fano's plane:



It is an easy exercise to prove the following theorem:

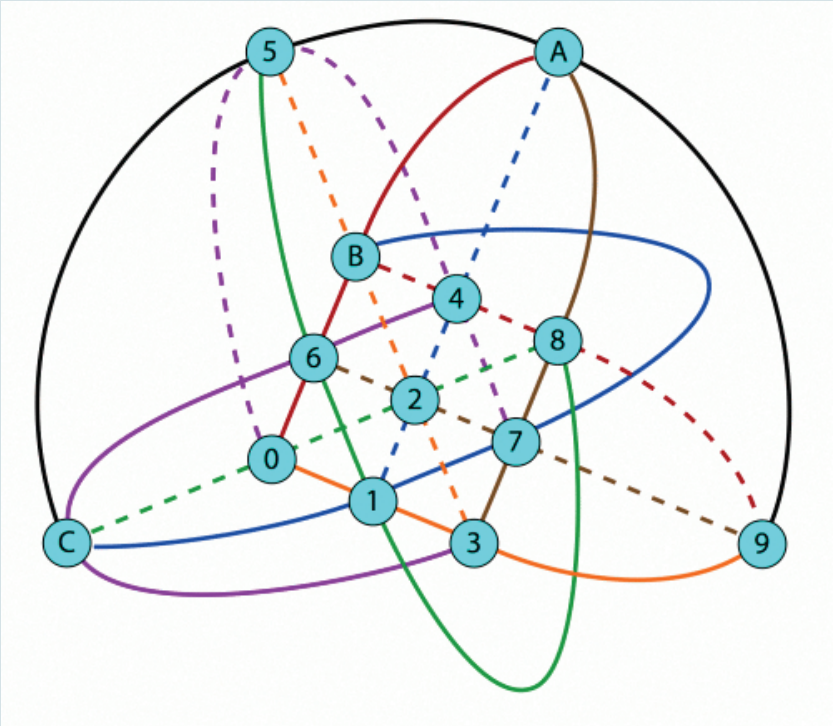
Theorem 2.1: *If $(\mathcal{P}, \mathcal{L})$ is a projective plane then a constant n exists, called the order of the projective plane, such that:*

PP3: *Every line contains exactly $n + 1$ points.*

PP4: *Every point lies on exactly $n + 1$ lines.*

PP5: $|\mathcal{P}| = |\mathcal{L}| = n^2 + n + 1$.

And here is a projective plane of order 3:

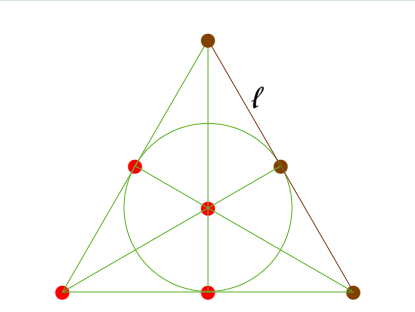


Construction 2.2: *Let $(\mathcal{P}, \mathcal{L})$ be a projective plane of order n and fix l in $\mathcal{L}$. Then an authentication code $(\mathcal{M}, \mathcal{K}; \mathcal{C})$ is defined by*

$\mathcal{M}$ *consists of the points on* l,
$\mathcal{K}$ *consists of all the points not on* l,
$\mathcal{C}$ *consists of all lines except for* l,
$f_k(m) = c$, *the line* c *through* k *and* m.

Note that

$$|\mathcal{M}| = n + 1, \quad |\mathcal{K}| = n^2, \quad |\mathcal{C}| = n^2 + n.$$



Theorem 2.3: *The projective plane construction defines a perfect authentication code, i.e.*

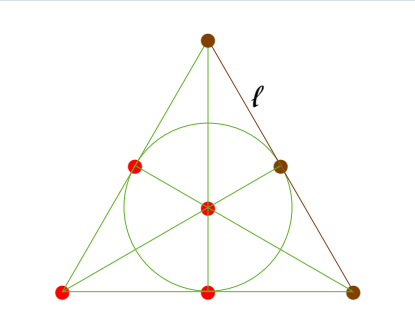
$$P_I = P_S = P_D = \frac{1}{n}.$$

Also, (6) is met with equality: $|\mathcal{M}| = n + 1 = \sqrt{n^2} + 1 = \sqrt{|\mathcal{K}|} + 1.$

Indeed, an impersonator can do no better than to select a line $\hat{l}$ as codeword that contains as many points outside l (keys) as possible. But this number is n and is independent of the choice of $\hat{l}$.

A similar argument holds for a substitution attack.

Projective planes exist for all orders n that are a prime power (use the finite field of this size).



The receiver intersects the received line c with l and retrieves message m as unique point of intersection.

The receiver checks if the secret key k lies on the received line c.

Note that there is no secrecy!

Note also that you can use the same key only once, because the intersection of two transmitted codewords, c and $\hat{c}$, made with the same key, would be this key!

3 The affine scheme

The following systematic authentication code has the same drawback as the projective geometry construction: keys are twice as long as messages.

Theorem 3.1: *Let* $\mathcal{M} = \mathcal{T} = GF(q)$, $\mathcal{K} = \{(a,b) \mid a,b \in GF(q)\}$ *and* $\tau : \mathcal{M} \times \mathcal{K} \to \mathcal{T}$ *defined by*

$$\tau_{a,b}(m) = a.m + b.$$

Then $(\mathcal{M}, \mathcal{K}; \mathcal{T})$ *is a systematic authentication code with*

$$|\mathcal{M}| = |\mathcal{T}| = q, \quad |\mathcal{K}| = q^2.$$

Moreover, $(\mathcal{M}, \mathcal{K}; \mathcal{T})$ *is perfect, i.e.* $P_I = P_S = P_D = \frac{1}{q}$.

4 Making A-codes from EC-codes I

To obtain authentication codes with a better ratio between key length and message length, compromises will have to be made.

T. Johansson (Ph.D. thesis, 1994) gives a construction of authentication codes by means of shift register sequences. They are easier to implement than authentication codes based on projective geometry.

The same author, together with G.A. Kabatianskii and B. Smeets, shows a year earlier how error-correcting codes (EC-codes) can be used to make authentication codes (A-codes).

A q-ary EC-code $\mathcal{C}$ has parameters $(n, |\mathcal{C}|, d_H)_q$ if it consists of $|\mathcal{C}|$ words of length n over $GF(q)$ such that different words in $\mathcal{C}$ have Hamming distance at least d_H.

In this, so called, q-twisted construction, the EC-code must satisfy a special property.

Assumption 4.1: *EC-code C has the additional property that*

$$\underline{c} \in C \quad \Rightarrow \quad \forall_{\lambda \in GF(q)}[\underline{c} + \lambda \underline{1} \in C].$$

This restriction is less serious than it may look:

- Any linear code containing the all-one vector meets this requirement.
- The vector $\underline{1}$ may be replaced by any other weight n word (this results in an equivalent authentication code).

This property allows us to define an equivalence relation on C.

$$\underline{c} \sim \underline{c}' \text{ if and only if } \underline{c} - \underline{c}' = \lambda \underline{1} \text{ for some } \lambda \in GF(q). \tag{7}$$

Ingredients of the A-code

Let M be a subcode of the EC-code C, containing one representative from each equivalence class. Set M has cardinality $|C|/q$ and will correspond to the set $\mathcal{M}$ of messages in the A-code that we construct. Messages will now be denoted by $\underline{m}$.
Note that message $\underline{m}$ may still have redundancy, so it can be more efficiently represented by an integer in $\mathbb{Z}_{|\mathcal{M}|}$.

Let $GF(q) = \{\alpha_1, \alpha_2, \ldots, \alpha_q\}$.

Define the set V of q-ary vectors of length nq by

$$V = \{\underline{v}^{(\underline{m})} = (\underline{m} + \alpha_1\underline{1},\ \underline{m} + \alpha_2\underline{1},\ \ldots,\ \underline{m} + \alpha_q\underline{1}) \mid \underline{m} \in M\}. \qquad (8)$$

The set of keys $\mathcal{K}$, of the A-code is given by the coordinate set of the vectors in V. So $|\mathcal{K}| = nq$.

The authenticator $\tau_k(\underline{m})$ of message $\underline{m}$ under key k simply is given by the k-th coordinate of $\underline{v}^{(\underline{m})}$.

Example 4.2: Let C be the quaternary $[4, 3, 2]$ EC-code, that is the dual code of the repetition code. It satisfies Assumption 4.1.

$\mathcal{M}$ contains 4^2 messages, one vector from each set $\{\underline{c},\ \underline{c}+\underline{1},\ \underline{c}+\alpha\underline{1},\ \underline{c}+\alpha^2\underline{1}\}$, where $\underline{c} \in C$ and $\alpha^2 = 1 + \alpha$.

Here, we take M as the set of 16 codewords in C with last coordinate equal to 0.

$$M \ = \ \{(c_1, c_2, c_3, c_4) \in C \mid c_4 = 0\}.$$

The third symbol in each word in M is the sum of the first two, because each codeword is orthogonal to $(1, 1, 1, 1)$.

So, the first two symbols form a nice representation of the words in M.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	$\underline{c} = \underline{m}$							$\underline{v}^{(\underline{m})}$							
0	0	0	0	1	1	1	1	α	α	α	α	α^2	α^2	α^2	α^2
0	1	1	0	1	0	0	1	α	α^2	α^2	α	α^2	α	α	α^2
0	α	α	0	1	α^2	α^2	1	α	0	0	α	α^2	1	1	α^2
0	α^2	α^2	0	1	α	α	1	α	1	1	α	α^2	0	0	α^2
1	0	1	0	0	1	0	1	α^2	α	α^2	α	α	α^2	α	α^2
1	1	0	0	0	0	1	1	α^2	α^2	α	α	α	α	α^2	α^2
1	α	α^2	0	0	α^2	α	1	α^2	0	1	α	α	1	0	α^2
1	α^2	α	0	0	α	α^2	1	α^2	1	0	α	α	0	1	α^2
α	0	α	0	α^2	1	α^2	1	0	α	0	α	1	α^2	1	α^2
α	1	α^2	0	α^2	0	α	1	0	α^2	1	α	1	α	0	α^2
α	α	0	0	α^2	α^2	1	1	0	0	α	α	1	1	α^2	α^2
α	α^2	1	0	α^2	α	0	1	0	1	α^2	α	1	0	α	α^2
α^2	0	α^2	0	α	1	α	1	1	α	1	α	0	α^2	0	α^2
α^2	1	α	0	α	0	α^2	1	1	α^2	0	α	0	α	1	α^2
α^2	α	1	0	α	α^2	0	1	1	0	α^2	α	0	1	α	α^2
α^2	α^2	0	0	α	α	1	1	1	1	α	α	0	0	α^2	α^2

1	2	3	4	5	6	7	8	9	19	11	12	13	14	15	16
	$\underline{c} = \underline{m}$							$\underline{v}^{(\underline{m})}$							
0	0	0	0	1	1	1	1	α	α	α	α	α^2	α^2	α^2	α^2
0	1	1	0	1	0	0	1	α	α^2	α^2	α	α^2	α	α	α^2
0	α	α	0	1	α^2	α^2	1	α	0	0	α	α^2	1	1	α^2
0	α^2	α^2	0	1	α	α	1	α	1	1	α	α^2	0	0	α^2
1	0	1	0	0	1	0	1	α^2	α	α^2	α	α	α^2	α	α^2
1	1	0	0	0	0	1	1	α^2	α^2	α	α	α	α	α^2	α^2
1	α	α^2	0	0	α^2	α	1	α^2	0	1	α	α	1	0	α^2
1	α^2	α	0	0	α	α^2	1	α^2	1	0	α	α	0	1	α^2
α	0	α	0	α^2	1	α^2	1	0	α	0	α	1	α^2	1	α^2
α	1	α^2	0	α^2	0	α	1	0	α^2	1	α	1	α	0	α^2
α	α	0	0	α^2	α^2	1	1	0	0	α	α	1	1	α^2	α^2
α	α^2	1	0	α^2	α	0	1	0	1	α^2	α	1	0	α	α^2
α^2	0	α^2	0	α	1	α	1	1	α	1	α	0	α^2	0	α^2
α^2	1	α	0	α	0	α^2	1	1	α^2	0	α	0	α	1	α^2
α^2	α	1	0	α	α^2	0	1	1	0	α^2	α	0	1	α	α^2
α^2	α^2	0	0	α	α	1	1	1	1	α	α	0	0	α^2	α^2

The authenticator consists of one symbol. For instance, message $(1, \alpha)$ being the first two symbols of the word $(1, \alpha, \alpha^2, 0)$ will get α as authenticator when key 13 is used.

Since each symbol occurs four times in each row of the table, we may conclude that $P_I = 1/4$.

Suppose on the other hand that $(1, \alpha; \alpha)$ is an intercepted codeword. From the table an opponent can conclude that the key that has been used is among keys 2, 7, 12 and 13.

Looking at the corresponding columns, we observe that for each message the same authenticator is used at most twice in these columns. So the best the opponent can do is transmit a message with an authenticator that occurs twice and obtain $P_S = 1/2$. For instance, he can send $(1, 1; 1)$.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	$\underline{c} = \underline{m}$							$\underline{v}^{(\underline{m})}$							
0	0	0	0	1	1	1	1	α	α	α	α	α^2	α^2	α^2	α^2
0	1	1	0	1	0	0	1	α	α^2	α^2	α	α^2	α	α	α^2
0	α	α	0	1	α^2	α^2	1	α	0	0	α	α^2	1	1	α^2
0	α^2	α^2	0	1	α	α	1	α	1	1	α	α^2	0	0	α^2
1	0	1	0	0	1	0	1	α^2	α	α^2	α	α	α^2	α	α^2
1	1	0	0	0	0	1	1	α^2	α^2	α	α	α	α	α^2	α^2
1	α	α^2	0	0	α^2	α	1	α^2	0	1	α	α	1	0	α^2
1	α^2	α	0	0	α	α^2	1	α^2	1	0	α	α	0	1	α^2
α	0	α	0	α^2	1	α^2	1	0	α	0	α	1	α^2	1	α^2
α	1	α^2	0	α^2	0	α	1	0	α^2	1	α	1	α	0	α^2
α	α	0	0	α^2	α^2	1	1	0	0	α	α	1	1	α^2	α^2
α	α^2	1	0	α^2	α	0	1	0	1	α^2	α	1	0	α	α^2
α^2	0	α^2	0	α	1	α	1	1	α	1	α	0	α^2	0	α^2
α^2	1	α	0	α	0	α^2	1	1	α^2	0	α	0	α	1	α^2
α^2	α	1	0	α	α^2	0	1	1	0	α^2	α	0	1	α	α^2
α^2	α^2	0	0	α	α	1	1	1	1	α	α	0	0	α^2	α^2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\underline{m}_6$	1	1	0	0	0	0	1	1	α^2	α^2	α	α	α	α	α^2	α^2
$\underline{m}_7$	1	α	α^2	0	0	α^2	α	1	α^2	0	1	α	α	1	0	α^2

Subtract $1 \times \underline{1}$ from the first row and $\alpha \times \underline{1}$ from the second row.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\underline{m}_6 - 1 \times \underline{1}$	0	0	1	1	1	1	0	0	α	α	α^2	α^2	α^2	α^2	α	α
$\underline{m}_7 - \alpha \times \underline{1}$	α^2	0	1	α	α	1	0	α^2	1	α	α^2	0	0	α^2	α	1

Note that each $\binom{1}{\alpha}$ in columns $\{2, 7, 12, 14\}$ is mapped to $\binom{0}{0}$ after subtracting $1 \times \underline{1}$ from the first row and $\alpha \times \underline{1}$ from the second row. This amounts to identical coordinates in the message parts $\underline{m}_6$ resp. $\underline{m}_7$. There can not be more that 2 such places because $d = 2$.

In general, there cannot be more that $n - d$ identical coordinates between two codewords. So, a particular tag will not occur more than $n - d$ times in the restriction of a different row to these n columns.

Theorem 4.3: *Let C be a EC-code with parameters $(n, |C|, d)$, satisfying (7). Let M represent a subcode of representatives of all equivalence classes.*

Then the triple $(\mathcal{M} = M, \mathcal{K} = \{1, 2, \ldots, nq\}; \mathcal{T} = GF(q))$ and mapping $\tau_k(\underline{m}) = (v^{(\underline{m})})_k$ defines a systematic A-code with

$$|\mathcal{M}| = |C|/q, \quad |\mathcal{K}| = nq, \quad |\mathcal{T}| = q,$$

and

$$P_I = 1/q, \quad P_S = 1 - d/n.$$

The cardinalities above are obvious.

Also, it is obvious that $P_I = 1/q$, because in

$$v^{(\underline{m})} = (\underline{m} + \alpha_1\underline{1},\ \underline{m} + \alpha_2\underline{1},\ \ldots,\ \underline{m} + \alpha_q\underline{1})$$

each value in $GF(q)$ occurs equally often!

General Class 4.4: *Take for C the extended q-ary $[q, k+1, q-k]$ Reed Solomon code.*

Then the q-twisted A-code, obtained from C, has parameters:

$$|\mathcal{M}| = q^k, \qquad |\mathcal{K}| = q^2, \qquad P_I = \frac{1}{q}, \qquad P_S = \frac{k}{q}.$$

Note that with this construction $|\mathcal{M}|$ can be much larger than $|\mathcal{K}|$ (as opposed to $|\mathcal{M}| \leq \sqrt{|\mathcal{K}|} + 1$).

Also, P_I satisfies (3) with equality, because $P_I = \frac{1}{q} = \frac{|\mathcal{M}|}{|\mathcal{C}|}$. Indeed, $|\mathcal{C}| = q|\mathcal{M}|$, because every message m can have any value in $GF(q)$ as authenticator.

The price we pay is in P_S.

We have $P_S = \frac{k}{q}$, as opposed to the bound $\frac{|\mathcal{M}|-1}{|\mathcal{C}|-1} \approx \frac{1}{q}$.

The method explained in this section is certainly not the only way to make A-codes from EC-codes. However, it does have the additional property that

each impersonation has the same probability of success

(here $1/q$).

An *I-equitable code* is a systematic A-code with the additional property that

$$P_I = \frac{|\{k \in \mathcal{K} \mid \exists_{m \in \mathcal{M}}[\tau(m,k) = \tau]\}|}{|\mathcal{K}|}, \quad \text{for all } \tau \in \mathcal{T}. \tag{9}$$

Note that the construction above satisfies (9).