Generalized coset schemes for the wire-tap channel: application to biometrics

Gérard Cohen
ENST, 46 rue Barrault
75634 Paris 13 France
e-mail: cohen@enst.fr

Gilles Zémor
ENST, 46 rue Barrault
75634 Paris 13 France
e-mail: zemor@enst.fr

I. Introduction

A traditional commitment consists of sending or publishing $y = f(\mathbf{b})$ where f is a one-way function and $\mathbf{b}$ is a binary vector that is supposed to remain hidden until it is disclosed. Checking that $f(\mathbf{b}) = y$ forbids one from disclosing a vector different from the one that was committed to. The *fuzzy commitment* problem arises when one wishes the protocol to accept not only the original $\mathbf{b}$, but also any vector $\mathbf{b}' = \mathbf{b}+\mathbf{e}$ where $\mathbf{e}$ is a vector of small Hamming weight. This problem appears typically in biometrical contexts when $\mathbf{b}$ encodes, for example, a fingerprint. Successive biometrical measures of the same finger will always tend to differ slightly. To solve this problem Juels and Wattenberg introduced the following fuzzy commitment scheme. Choose a random secret vector $\mathbf{s}$ and encode it as a codeword $\mathbf{c}$ of some fixed, publicly known code C. Then publish (in practice that may mean write on a smartcard)

$$\mathbf{W} = \mathbf{c} + \mathbf{b} \qquad \text{together with } \ f(\mathbf{c})$$

where $f(\mathbf{c})$ is the image of $\mathbf{c}$ by some cryptographic hash function. When $\mathbf{b}+\mathbf{e}$ is submitted to the protocol, it adds it to $\mathbf{W}$, yielding a noisy version $\mathbf{c}+\mathbf{e}$ of $\mathbf{c}$. The vector $\mathbf{c}+\mathbf{e}$ is then submitted to a decoding algorithm which yields $\mathbf{c}$. Validity is checked by computing $f(\mathbf{c})$ and comparing it to the stored value.

In the idealized setting of [1], the "biometric" vector $\mathbf{b}$ is assumed to be uniformly distributed among vectors of a given length. In that case, the published vector $\mathbf{W}$ yields no information on the secret codeword $\mathbf{c}$ or the original secret $\mathbf{s}$ However, in practice the distribution of $\mathbf{b}$ may be far from uniform and in that case the vector $\mathbf{W}$ is liable to leak undesirable partial knowledge of $\mathbf{c}$, and hence of $\mathbf{s}$, to an unauthorized third party (hereafter "the eavesdropper").

Our present contribution to this problem is threefold. We start with an information-theoretic approach and model the situation by involving wire-tap channel models. This means that we consider the eavesdropper to have access to a very noisy version of the secret codeword, while the correct protocol yields access to a less noisy version of the secret codeword. We would like to insure no leakage of secret information to the eavesdropper: this problem can be remodeled as that of maximizing the amount of information that can be reliably transmitted through the less noisy "channel" with maximum "equivocation", i.e. insuring that the eavesdropper gets essentially no information on the secret data.

Next, we generalize Wyner's coset coding scheme [2] to the case when both the main channel and the wiretap channel are noisy. We then prove that this scheme, i.e. the use of linear codes, achieve the Shannon-capacity of the system. However, this scheme goes only half-way to providing a practical "zero-leakage" fuzzy commitment protocol, because to achieve capacity we need to involve random codes with an unacceptable decoding complexity.

Finally, we present a complete solution to the original problem by making use of LDPC codes, with low decoding complexity. The price to pay is a reduction of the system's original capacity to a subcapacity that we evaluate.

II. A biometric identification scheme with zero information leakage

The information leakage problem can be modeled as a wire-tap problem as follows. Alice's secret $\mathbf{s}$ is first encoded as a codeword $\mathbf{c}$ of some code C and transmitted to Bob over an additive channel that adds noise $\mathbf{e}$ to $\mathbf{c}$, and over an additive wire-tap channel to the eavesdropper Eve who receives $\mathbf{c}+\mathbf{b}$. The problem is to make the secret $\mathbf{s}$ as big as possible while ensuring that Eve gets zero information on $\mathbf{s}$.

The modified fuzzy commitment (or biometric identification) scheme is as follows.

- Choose an error-correcting code C_1, a code C_2 obtained by randomly choosing its parity-check matrix $\mathbf{H}_2$, and a one-way hash function f. Choose a random codeword $\mathbf{c} \in C_1$ whose syndrome for code C_2 equals $\mathbf{s}$.
- publish $\mathbf{W} = \mathbf{c} + \mathbf{b}$, $f(\mathbf{s})$.
- Authentication phase. Check that $\mathbf{b}'$ is close to $\mathbf{b}$ ($d(\mathbf{b}, \mathbf{b}') \leq e$) : this means decode $\mathbf{W} + \mathbf{b}'$ to its closest codeword $\mathbf{z}$ of C_1 and compute its syndrome for C_2, $\sigma_2(\mathbf{z}) = \mathbf{H}_2\,{}^t\mathbf{z}$.
  Finally check that $f(\sigma_2(\mathbf{z})) = f(\mathbf{s})$.

Our hypothesis on the wire-tap (or biometric) channel is that there exists as set T of typical biometric vectors $\mathbf{b}$ such that the probability $P(\mathbf{b} \notin T)$ decreases exponentially with the codelength n, and such that the distribution of $\mathbf{b}$ conditional to $\mathbf{b} \in T$ is very close to uniform. We have therefore $|T| \approx 2^{H(\mathbf{b})}$ where $H(\mathbf{b})$ is the entropy of the biometric vector. This condition on $\mathbf{b}$ includes in particular discrete memoryless channels and also many channels with memory.

Under this condition we can guarantee that, whatever the structure of code C_1, as long as the sum of both the redundancies of C_1 and C_2 does not exceed the biometric entropy $H(\mathbf{b})$, Eve obtains zero information. The code C_1 can therefore, either be chosen to achieve the capacity of the main channel (for $\mathbf{e}$), but this will require a decoder with non-polynomial (and non-practical) complexity, or be chosen to reach some subcapacity but to come together with a practical decoding algorithm (e.g. an LDPC code).

References

[1] A. Juels and M. Wattenberg, "A fuzzy commitment scheme", in 6th ACM Conference on Computer and Communications Security, pp. 28–36, ACM Press, 1999.

[2] A. Wyner "The wire-tap channel", BSTJ 54 , 1355-1387 (1975).