Non-Malleable Codes and Coset Coding

Alain Patey

Télécom ParisTech
Morpho (SAFRAN Group)
Identity and Security Alliance (The Morpho and Télécom ParisTech Research Center)

Joint work with H. Chabanne, G. Cohen, J.-P. Flori

Oct. 11^{th}, 2012. Journées C2

Outline

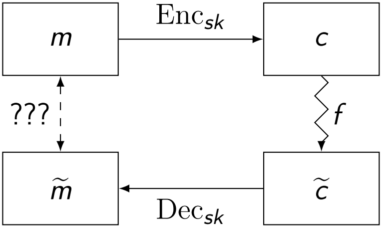
1. General framework
2. Definitions
3. Non-Malleable Codes from the Wire-Tap Channel
4. Secure Network Coding
5. Non-Malleable Codes w.r.t. Linear Tampering Functions

Outline

Non-Malleability – Cryptography

Non-malleability: cryptographic property introduced by Dolek et al. in 1991 [DDN91].

A cryptographic scheme is non-malleable if a decrypted tampered ciphertext reveals no information about the original plaintext.

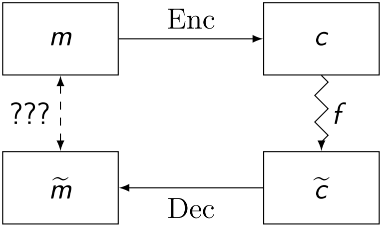


Non-Malleability – Coding Theory

This principle was transposed to coding theory by Dziembowski et al. in 2010 [DPW10].

For a coding scheme to be non-malleable, a decoded tampered codeword should

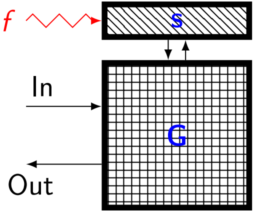
- either be corrected
- or reveal no information about the original message.



This is known as the tampering experiment.

Tamper proofness and read proofness

Such constructions can be used to protect a system where computations are performed in tamper and read proof circuit, but they depend on a secret state which is stored in read proof only memory.



The first step of the computation is then to decipher or decode the encrypted or encoded secret state.

Algorithmic tamper proofness

- Such algorithmic tamper proofness was first studied by Gennaro et al. in 2004 [GLM+04]. Basically, their solution was to store the secret state together with a signature, so that, if the secret state is tampered with, the signature check fails and the system aborts.

$$\langle G, s\rangle \rightarrow \langle G^{\mathrm{Sign},\mathrm{Check},(sk,pk)}, (s, \mathrm{Sign}(s, sk))\rangle$$

- Much influenced by this work, Dziembowski et al. [DPW10] proposed to use non-malleable codes to transform such a system.

$$\langle G, s\rangle \rightarrow \langle G^{\mathrm{Enc},\mathrm{Dec}}, \mathrm{Enc}(s)\rangle$$

Outline

Tampering functions

A tampering function is a function

$$f : \mathbb{F}_2^n \to \mathbb{F}_2^n \ .$$

Some families of particular interest:

1. $\mathcal{F}_{all} = \mathbb{F}_2^{n\,\mathbb{F}_2^n}$ the set of all tampering functions;
2. $\mathcal{F}_{bit} = (f_1, \ldots, f_n)$ the set of bitwise independent tampering functions;
3. $\mathcal{F}^{lin}$ the set of linear functions

Non-malleability with regard to a family is defined as non-malleability against each function in that family.

Formal definition I

- A coding scheme is a couple $(\mathrm{Enc}, \mathrm{Dec})$ where
  - $\mathrm{Enc} : \mathbb{F}_2^k \to \mathbb{F}_2^n$ is a randomized encoding procedure
  - $\mathrm{Dec} : \mathbb{F}_2^n \to \mathbb{F}_2^k \cup \{\perp\}$ is the associated deterministic decoding procedure
- The tampering experiment induces a probability distribution for every message m and tampering function f denoted by $\mathrm{Tamper}_f(m)$.
- The idea of non-malleability is that an attacker should not gain more information by tampering the codeword than by having just black box access to the circuit.

Formal definition II

Definition (Non-malleability)

A coding scheme $(\mathrm{Enc}, \mathrm{Dec})$ is said to be non-malleable with regard to a family $\mathcal{F}$ of tampering functions iff, for every $f \in \mathcal{F}$, there exists a probability distribution $\mathcal{D}_f$ over $\mathbb{F}_2^k \cup \{\perp, \textsf{same}\}$, such that for every message m, the two following distributions are indistinguishable:

$$\mathrm{Tamper}_f(m) \approx \left\{ \begin{array}{c} \widetilde{m} \leftarrow \mathcal{D}_f \\ \mathrm{Output} \begin{cases} m \text{ if } \widetilde{m} = \textsf{same} \\ \widetilde{m} \text{ otherwise} \end{cases} \end{array} \right\}$$

In particular, the distribution $\mathcal{D}_f$ does not depend on the message m.

Relation to classical models

- An error correcting code should be able to correct errors introduced by tampering functions.
  Error correction implies non-malleability: the associated distribution is $\mathcal{D}_f =$ **same** for every tampering function.
- An error detecting code should either return the unmodified codeword, or a special symbol $\perp$.
  Error detection implies non-malleability if the probability of error detecting is independent of the source message.

(Im)possibility results

Theorem (Impossibility)

There exists no code non-malleable with regard to the family $\mathcal{F}_{all}$.

For example, for any coding scheme $(\mathrm{Enc}, \mathrm{Dec})$, there is a function f which associates to a codeword $c = \mathrm{Enc}(m)$ the codeword $\widetilde{c} = \mathrm{Enc}(m+1)$

Theorem (Possibility)

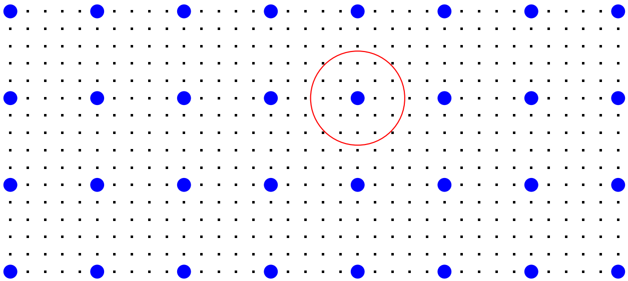
For any family $\mathcal{F}$ such that $\log\log \#\mathcal{F} < n$, there exists a non-malleable code.

The upper bound on the size of the family is to be compared with
$\log\log \#\mathcal{F}_{all} = n + \log n$.
(The proof is not constructive)

Outline

Linear codes

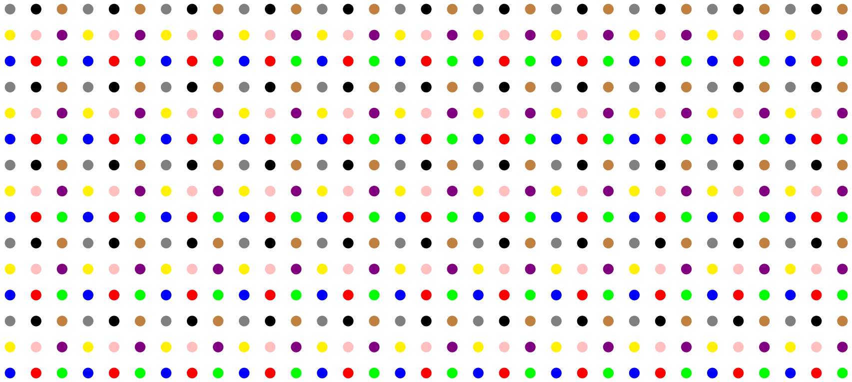
Error correction can typically be done by finding the closest codeword to the received tampered codeword.



For a linear code whose generating matrix is G and parity check is H, a bitstring x is a codeword iff $H^t x = 0$. Otherwise the value $H^t x = e$ is called the syndrome of x.

Linear coset coding

The idea of linear coset coding is to encode each message as a coset of a linear code.



Decoding a codeword can then be done by computing its syndrome.

Formal Definition

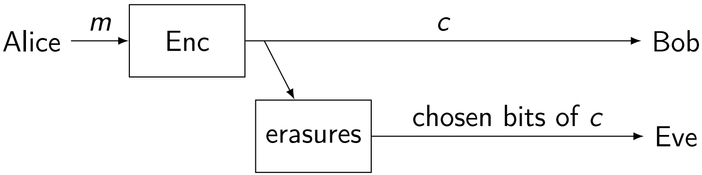
Definition (Linear Coset Coding)

Given: C a $[n, n-k, d]$ linear code with a $k \times n$ parity-check matrix H
Encode: $m \in \mathbb{F}_2^k \mapsto_R c \in \mathbb{F}_2^n$ s.t. $Hc = m$
Decode: $c \in \mathbb{F}_2^n \mapsto m = Hc$

The Wire-Tap Channel II

In 1984, Ozarow and Wyner introduced a second version of the Wire-Tap Channel [OW84].

- Alice wants to transmit a message to Bob without Eve getting any information.
- Both channels are noiseless
- But Eve can only get a given number of bits on hers.



Bitwise Independent Tampering, LCC and WTC

- First, we consider non-malleability w.r.t. bit-wise independent tampering functions, i.e. functions $f : x \mapsto f(x) = (f_1(x_1), \ldots, f_n(x_n))$ where $f_i \in \{\mathbf{0}, \mathbf{1}, \mathbf{keep}, \mathbf{flip}\}$.
- Both NMC and WTC problems can be solved using linear coset coding.
- In particular, for the second version of the Wire-Tap Channel, if we denote by $d^\perp$ the dual distance of the linear code used, and if Eve has only access to less than $d^\perp$ bits of information, then she gains absolutely no information on the message.
- Efficient implementations using LDPC codes.

Which version for bitwise tampering?

Using a linear coset coding, we can not be protected against functions in $\mathcal{F}_{err}$, i.e. with bit functions only **keeping** of **flipping** bits. Indeed if m is the original message and an error e is added, then

$$\widetilde{c} = c + e \ ,$$

and the decoded message is nothing but

$$\widetilde{m} = H^t c + H^t e = m + H^t e \ .$$

So we must include some bit functions setting bits to **0** or **1**. This can be naturally seen as the erasures of the second version of the Wire-Tap Channel.

NMC from the WTC

Result from [CCFP11]

Theorem (Non-Malleability of LCC wrt bit-wise independent tampering functions)

Let $\mathcal{F}$ be a family of bitwise independent tampering functions such that $\forall f = (f_1, \ldots, f_n) \in \mathcal{F}, \# \{i \mid f_i = \mathbf{0} \text{ or } f_i = \mathbf{1}\} \geq D$.
Let C be a $[n, n-k]$ MDS linear code such that $k < D$.
Then a linear coset coding using C is non-malleable w.r.t. $\mathcal{F}$.

From bit-wise to linear tampering

- Linear functions $f : x \mapsto A.x + B$ with $A \in \mathbb{F}_2^{n\times n}, B \in \mathbb{F}_2^n$
- Bit-wise independent functions can easily be described by linear functions, with a diagonal matrix A:

$f = (\textbf{keep},\textbf{flip},\textbf{0},\textbf{1})$. $\forall x \in \mathbb{F}_2^4, f(x) = A.x + B$ with

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Starting from this observation, can we adapt the theorem of last slide to a result of non-malleability w.r.t. linear functions ? How is adapted the condition on the number of **0**,**1** ? On a condition on the rank of A ?

Outline

1. General framework
2. Definitions
3. Non-Malleable Codes from the Wire-Tap Channel
4. **Secure Network Coding**
5. Non-Malleable Codes w.r.t. Linear Tampering Functions

Secure Network Coding (SNC)

Introduced by Cai and Yeung [CY02], see also [CC11].

- Network represented as a directed acyclic graph
- Single source node sends a message m through the network
- User nodes are at the end of the paths in the graph
- Message is encoded before being sent through the network
- Inner nodes transmit linear combinations of the packets that they receive

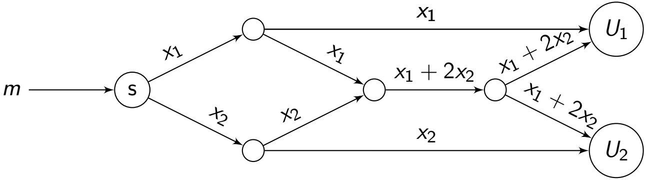


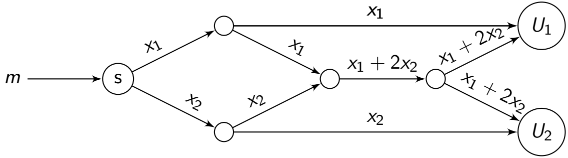
Figure: A Secure Butterfly Network over $\mathbb{F}_3$ $(x_1 \in_R \mathbb{F}_3, x_2 = m - x_1)$

Security of SNC

Security Requirements:

- The user nodes can recover the original message m from the packets that they received
- For any subset of edges that we allow the adversary to obtain, the adversary gets no information on m.

Usually, adversary is allowed to access any subset (but only one at a time) of up to μ edges.

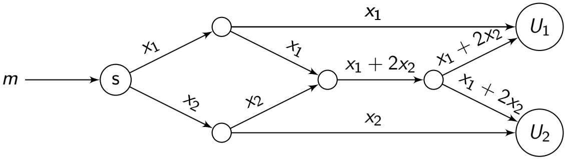


In this example, we satisfy the security requirements (with μ=1):

- User nodes can recover the message. For instance U_2 subtracts the packets he received to obtain his output.
- If an adversary accesses $\mu = 1$ edge, he learns no information on m

SNC Using Linear Coset Coding I

- The source node wants to send a k-symbol message to the user nodes
- It uses Linear Coset Coding with a $k \times n$ parity-check matrix H
- n symbols are sent over the network
- The intermediate nodes send a pre-determined linear combination of the elements they receive



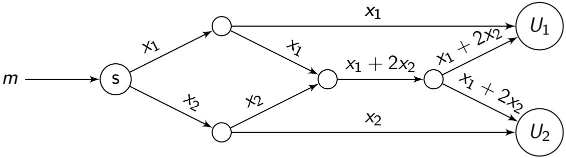
Here, $k = 1$, $n = 2$ and $H = \begin{pmatrix} 1 & 1 \end{pmatrix}$

SNC Using Linear Coset Coding II

Security Result from [ERS07]:

Theorem (Security of SNC using LCC)

A SNC based on LCC based on a MDS code with a $k \times n$ parity-check matrix H, such that no linear combination of $\mu \leq n - k$ packets sent over edges belongs to the space spanned by the rows of H, is secure against an adversary who can observe μ edges.



Here, the linear combinations are $\begin{pmatrix} 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 \end{pmatrix}$. None of them belongs to the span of $H = \begin{pmatrix} 1 & 1 \end{pmatrix}$.

Outline

Back to NMC

- Transposition: the linear combinations performed by the intermediate nodes in the SNC models are viewed as linear tampering functions performed by the adversary in the NMC model
- Furthermore, the decoding procedure needs to be taken into account, since the adversary does not observe the result of his tampering
- Roughly, the adversary observes $HAx + HB$ and the considerations on the rank of the linear combinations in the SNC model are transposed to conditions on the rank of HA

NMC from SNC

Main result from [CCP12]

Theorem (Non-Malleability of LCC wrt linear tampering functions)

Let C be a $[n, n-k]$ MDS linear code, with a $k \times n$ parity-check matrix H. Let $\mathcal{F}^{lin} \subset \mathbb{F}_2^{n\mathbb{F}_2^n}$ be a family of linear tampering functions such that $\forall f : x \mapsto A.x + B \in \mathcal{F}^{lin}$,

1. $rank(HA) \leq n - k$
2. $span(rows\ of\ HA) \cap span(rows\ of\ H) = \{0\}$

Then a LCC using C is non-malleable w.r.t. $\mathcal{F}^{lin}$.

Remarks

- Bit-wise independent tampering functions satisfying the first theorem (NMC+bit-wise) satisfy this theorem (NMC+linear)
- For the same LCC, the class of linear functions considered in the theorem of [ERS07] (SNC+linear) is included in the class of functions considered in this theorem (NMC+linear). (Indeed, the adversary in the SNC model can in particular apply the decoding algorithm)
- The reciprocal property is not true. For instance LCC are non-malleable wrt to $f : x \mapsto x + c$ where c is a codeword, since the tampered codewords are always corrected during the decoding procedure. This function does not satisfy this theorem.

Conclusion

- Parallels between Non-Malleable Codes and known models in coding theory
- NMC wrt bitwise/linear tampering functions built with standard tools
- Perspectives: non-linear tampering, other codes ...

That's all, folks !

Thank you for your attention.
Questions ?

A few words about my PhD

- Title: "Secure Distributed Biometric Matching"
- Goal: design privacy-preserving biometric identification/matching protocols
- Tools: Secure Multi-Computation (Garbled circuits, oblivious transfer. . .), Homomorphic Encryption
- Applied to: Euclidean distance, Hamming distance, scalar product, comparison. . .

References I

Ning Cai and T. Chan.
Theory of secure network coding.
Proceedings of the IEEE, 99(3):421 –437, march 2011.

H. Chabanne, G. Cohen, J. Flori, and A. Patey.
Non-malleable codes from the wire-tap channel.
In *Information Theory Workshop (ITW), 2011 IEEE*, pages 55 –59, oct. 2011.

Hervé Chabanne, Gérard D. Cohen, and Alain Patey.
Secure network coding and non-malleable codes: Protection against linear tampering.
In *ISIT*, pages 2546–2550. IEEE, 2012.

References II

Ning Cai and R.W. Yeung.
Secure network coding.
In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, page 323, 2002.

Danny Dolev, Cynthia Dwork, and Moni Naor.
Non-malleable cryptography (extended abstract).
In *STOC*, pages 542–552. ACM, 1991.

Stefan Dziembowski, Krzysztof Pietrzak, and Daniel Wichs.
Non-malleable codes.
In Andrew Chi-Chih Yao, editor, *ICS*, pages 434–452. Tsinghua University Press, 2010.

References III

Salim Y. El Rouayheb and Emina Soljanin.
On wiretap networks ii.
In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pages 551 –555, june 2007.

Rosario Gennaro, Anna Lysyanskaya, Tal Malkin, Silvio Micali, and Tal Rabin.
Algorithmic tamper-proof (ATP) security: Theoretical foundations for security against hardware tampering.
In Moni Naor, editor, *TCC*, volume 2951 of *Lecture Notes in Computer Science*, pages 258–277. Springer, 2004.

Lawrence H. Ozarow and Aaron D. Wyner.
Wire-tap channel II.
In *EUROCRYPT*, pages 33–50, 1984.