Doctoral School on Engineering Sciences
Università Politecnica delle Marche

Extended summary

Cryptography and Physical Layer Security: The Role of Channel Coding

Curriculum: "Ingegneria Elettronica, Elettrotecnica e delle Telecomunicazioni"

Author

Marco Bianchi

Tutor(s)

Franco Chiaraluce

Date: 27-02-2013

Abstract.

Channel coding theory deals with the information that could be obtained through the observation of a sequence of symbols added with channel noise.
The aim of this research is to present two applications of channel coding techniques able to guarantee security both in a traditional cryptographic manner, and by adopting a different, physical layer based, approach.
The main application of channel coding, and the first application for which it has been developed, is to obtain an efficient communication by overcoming the errors induced by a physical channel, however it could also be used to distinguish users of a communication system, hence it has also important cryptographic and security purposes: when an authorized user is able to exploit some error correction capability, through proper channel coding techniques, and a malevolent user is not, only the authorized one is able to obtain the plain message.
For what concerns cryptography, the first public key code-based cryptosystem was derived by Robert J. McEliece in 1978, it is based on binary Goppa codes, it is still unbroken after more than 30 years, and the crucial point is that no polynomial-time algorithm is known to recover the

secret message in the McEliece Cryptosystem even using quantum computers, so it is a valid candidate for post quantum cryptography. The same is not true for other worldwide implemented public key cryptographic solutions or signature schemes as RSA, DSA, ECDSA.

The original McEliece Cryptosystem has two main advantages and one main disadvantage with regard to RSA: the advantages are the post quantum security and the very small computational complexity, while the main disadvantage is the great key size that prevented its usage in most of real-life applications.

We will show how to reduce the key size by order of magnitude with respect to the original proposal in such a way to make it comparable with that of RSA, while keeping the computational complexity lower than RSA's one.

On the other hand, channel coding can be used, at the physical layer, to distinguish among the users of a communication system through the intrinsic difference of the physical channels connecting one sender to one or many receivers.

Channel coding and other proper techniques, as frame scrambling and concatenation, could be used to minimize the SNR gap, between the legitimate user and a possible attacker, that is needed to make the first able to receive the correct message and the latter unable to recover sufficient information.

Keywords. Channel Coding, Post quantum cryptography, McEliece Cryptosystem, Security Gap, Physical Layer Security.

1 Problem statement and objectives

The research has been focused on developing practical solution for the following two kind of security: code based cryptography and physical layer security.

State-of-the-art public key cryptography that is implemented in all the security standards underlying e-commerce and e-banking, such as RSA, DSA, ECDSA are based on the prime factorization problem or the discrete logarithm problem, both of them are considered very hard to solve using non-quantum algorithms, that is using algorithms able to run on classic, non-quantum, computers.

They were cryptanalyzed for tens of years, by a huge number of researchers all over the world, without considerably meaningful improvements, except for some natural parameters revision.

However, Shor, in 1994, introduced an algorithm [1] able to solve the integer factorization problem and the discrete logarithm problem, on which standard public key cryptography is based, in polynomial time with respect to the dimension of the public key.

It is not possible to make a precise anticipation about the date on which a fully working large scale quantum computer will be available, however in the last few years, a lot of improvement were made.

Several research groups obtained the first prototypes having desirable features such as machines able to handle decoherence effects and able to work at room temperature[2], computers based on single atom qubits and with silicon technology that is the same as that used for standard computers.

Moreover the D-Wave System, in 2011, sold the first commercial, large-scale, 128 qubits-quantum computer to Lockheed Martin. It was not a fully working quantum machine but it is able to implement quantum annealing at physical level to increase computation speed and it is able to solve large and difficult optimization problems [3].

These results pose a reasonable date, for the day we will be forced to change all our cryptographic primitives, in less than ten years.

The McEliece cryptosystem [4] is a public key cryptosystem based on coding theory rather than on number theory, and the problem of decoding a random linear code is NP-completed. It is very unlikely that a polynomial time algorithm exists for solving a NP-complete problem, even if the attacker uses a quantum computer, in fact, no algorithm exists for breaking the McEliece cryptosystem.

This cryptosystem, in its original version, is based on Goppa codes[5], that are a large family of hard-decodable linear codes, and the main disadvantage of such a system, with respect to RSA (related to the choice of the codes), is the large key size that prevented its usage in almost all practical application.

On the other hand, it is extremely fast both in encryption phase and in the decryption one, so it could be a valid candidate for post-quantum cryptography.

In the past, a lot of variants were proposed to decrease the key size, the most efficient ones required a change of the secret linear code in two possible ways: through the adoption of Maximum-Distance-Separable codes, as the Generalized Reed Solomon codes, or through the adoption of codes that can be described by matrices having a compact representation.

However no significant change was ever proposed in the main parts of the system, this fact resulted in weaker schemes: the structure of the secret code is not strongly hidden in the McEliece cryptosystem, that is, the public code is a permuted version of the secret one and this fact was used to mount attacks against version of the cryptosystems based on highly structured codes (such as GRS [6] or dyadic Goppa codes).

One of the aims of this thesis is to develop secure and practical variants of the original McEliece cryptosystem, not only with regards to the code choice, but also in such a way to modify the system, making the secret code strongly disguised in the public key. This allowes to reconsider the usage of GRS codes as well as to improve the security and performance of Quasi-Cycic Low Density Parity Check Codes (QC-LDPC), that were already proposed by the author of this thesis and the research group of Università Politecnica delle Marche[7][8].

For what concerns physical layer security, it is already known that given two different channels between a sender (Alice) and an authorized user (Bob), and from the same sender and an unauthorized user, or eavesdropper (Eve), when the channel of Eve is worse than that of Bob, from the SNR point of view, there is a certain amount a secret information that Bob is able to recover, while Eve is not. However the maximum amount of secret information is proven to be reachable under ideal and asymptotic conditions, such as the perfect knowledge of the channel from Bob and Alice point of view, or the usage of linear codes able to reach the channel capacity.

Such ideal assumptions do not fit with real scenarios, so we studied a practical framework to implement physical layer security using real-life, finite-length codes and some further techniques, namely frame scrambling, concatenation and, if a feedback channel is available between Alice and Bob, Hybrid Automatic Repeat reQuest (HARQ). This analysis could be (and it has actually been) the basis for the study of a physical layer security implementation in real life scenarios such as the key distribution and renewal in wireless networks without any user intervention; similar techniques could also be used for others applications such as Near Field Communication (NFC) commerce.

2 Research planning and activities

Code-based cryptosystem could be attacked by using general decoding algorithms, the more effective ones are Information Set Decoding algorithms (ISD). These algorithms are not polynomial in the input size, so a proper choice of the system parameters can always avoid any attack, so the proposed variants were tested and designed to resist the more recent version of ISD algorithms [9][10].

Moreover, for the structured codes (GRS) variant, we tested that the proposed modification made the system able to resist all the known attacks, so, in this case, we considered not only the ISD threat, but also key retrieving attacks based on the knowledge of the structure of the secret code. The more dangerous one is that based on Distinguisher [11] and will be analyzed in the next section.

For what concerns QC-LDPC codes, a secure variant was already proposed [7], but, given the probabilistic nature of the iterative decoder needed for such codes, no algorithmic method was known to design the code parameters, that is, for each security level, long simulations were needed to establish the proper choice of code length, dimension and density. We proposed an analytic way to estimate the error correction capability of QC-LDPC codes in the particular considered channel (that is the channel were a fixed number of errors is introduced) through adapting some well known techniques able to compute the decoding waterfall threshold for ideal infinite LDPC codes under Bit Flipping decoding [12].

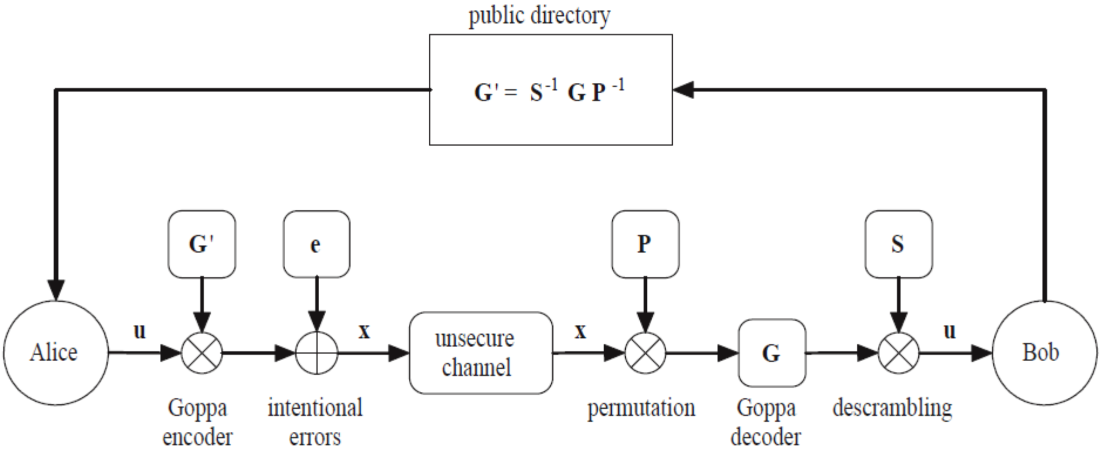

Figure 3.1: The original McEliece Cryptosystem

Moreover, this procedure made us able to establish the best choice of system parameters under the assumption of regular codes, so new optimum public key sizes were obtained.

Switching the context to physical layer security, we carried out an analysis under the assumption of AWGN channel, able to estimate the security level without recurring to theoretical limits, by adopting the Security Gap.

The Security Gap is the difference between SNR values of Bob and Eve able to made the bit error probability of Bob enough small (in the example case we choose 10-5) and that of Eve enough high (0.4 in the examples).

We carried out the same analysis also for the case with scrambled and concatenated frames, and for all the other proposed scenarios (negative Security Gap and feedback channel available).

Both for the cryptographic proposals and for the physical layer security analysis, the theoretical results were proven through extended computer simulation using C++ software written by the author of this thesis and by the research group of Università Politecnica delle Marche.

3 Analysis and discussion of main results

The original McEliece cryptosystem is depicted in Fig. 3.1, where G is the generator matrix of the secret code.

The proposed variants deeply change the nature of the public key by substituting the permutation matrix P with a proper scrambling (or transformation matrix) Q that has to choose in accordance with the secret code. The introduction of such a matrix makes the public code no longer permutation equivalent to the secret one, in particular they have (in general) different weight spectrum, and different correction capability, moreover it implies an error propagation effect to the receiver side that has to be compensated by the increased correction capability of the proposed secret codes with regards to the original Goppa ones.

The structure of Q for the GRS case is the following Equation (1),

$$\mathbf{Q}=\mathbf{R}+\mathbf{T}=\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}^T \bullet \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + \mathbf{T} \qquad (1)$$

where **T** is a sparse scrambling matrix having density equal to *m* and a_i and b_i are element of the field where the code is defined.

With such a choice a distinguisher attack able to recover the secret [] key could be performed when *m*=1 or when the rank of **R** is very low. We proved that it is not the case when the density of **T** is properly chosen and the results in Table 3.1 are referred to a secure parameters set.

For what concerns QC-LDPC codes, the recursion probabilities function describing the error probabilities in each step of the iterative decoding process were derived, they are quite involved and do not fit in the spirit of this extended summary, so we will give only the final simple equation for the number of residual bit errors at the *l*-th iteration:

$$q_l = t - t \cdot f^b\left(q_{l-1}\right) + \left(n-t\right) \cdot g^b\left(q_{l-1}\right) \qquad (2)$$

where $f^b(x)$ is the probability that a message originating from a variable node is correct if the codeword contains *x* errors, $g^b(x)$ is the probability that the message is wrong, *n* is the code length and *t* is the weight of the intentional error.

QC codes can be compactly represented by one row of their parity check matrix or by one columns of their generator matrix so their resulting public keys are very small in comparison with other solutions.

In Table 3.1 two not already mentioned parameters are reported: *k* that is the code dimension, and the operations needed for decrypting one information bit.

Table 3.1: Parameters comparison

			Security	Dec. Operations	Key size
	n	*k*	level	tions per bit	(bytes)
RSA	1024	1024	80	$2^{19.5}$	256
RSA	3072	3072	128	2^{22}	384
Goppa	1632	1269	80	2^{13}	57581
Goppa	2960	2288	128	2^{18}	1537536
GRS	4599	3483	128	2^{21}	431892
QC-LDPC (not optimized)	24576	18432	80	$2^{10.8}$	2304
QC-LDPC	16384	12288	100	$2^{7.7}$	1538
QC-LDPC	28672	21504	128	$2^{8.5}$	2688

For the physical layer security framework, different combination were considered under AWGN channel.

In particular the plain systematic coded transmission (both under LDPC or BCH coding) was not able to reach a sufficient security level by adopting small security gaps. It is due to the fact that the degradation of decoding performance induced by the channel varies quite slowly with the channel deterioration.

For this reason we resorted to the scrambling of the coded bits: in a way similar to what happens in the McEliece cryptosystem, the descrambling operation makes the decrypted word wrong in roughly half of its bits when just one bit is wrong after the decoding phase.

Following the same approach we used also the frame concatenation: in this way a single error after the decoding phase could afflict not only half of the frame bits, but also half of the bits in the frames involved in the concatenation process, obtaining a typical "all-or - nothing" effect, that is, all the frame are correctly decrypted or no information is gained.

Examples are given in Figures 4.1 and 4.2 where P_e^E indicates Eve's error probability, the code rate is about 2/3, the code dimensions are 1576 for the LDPC code and 1354 for the BCH code in Figure 4.1, while the code in Figure 4.2 is a 2047,1354 BCH, perfect scrambling means the best theoretical achievable scrambling effect, *w* indicates the weight of one row of a real scrambling matrix, and *L* is the number of concatenated frames.

The effect of a feedback channel is depicted in Figure 4.3 where no concatenation is supposed and Bob is allowed to request and receive at most 2 retransmission of an erred frame, while Eve is not able to require anything. This has the effect of making the error probability curves not monotone, in fact Eve can take advantage of the frame retransmissions, but not as much as Bob. So when Bob requires two retransmissions with very high probability, also Eve takes advantage of those requests, while, when Bob start to request (on average) just one retransmission, Eve cannot use the two retransmissions for her purposes since they involve (on average) not the same frames that Eve would require.

The same happens when Bob passes from one to zero retransmission, so we have to local maxima for Eve's curves.

Eve's error probability however remains too low for a null or negative security gap, so we have to rise it up, for example recurring to frame concatenation. As an example, for the parameters in Figure 4.3, 170 frame are needed to be concatenated together for reach a reasonable security level.

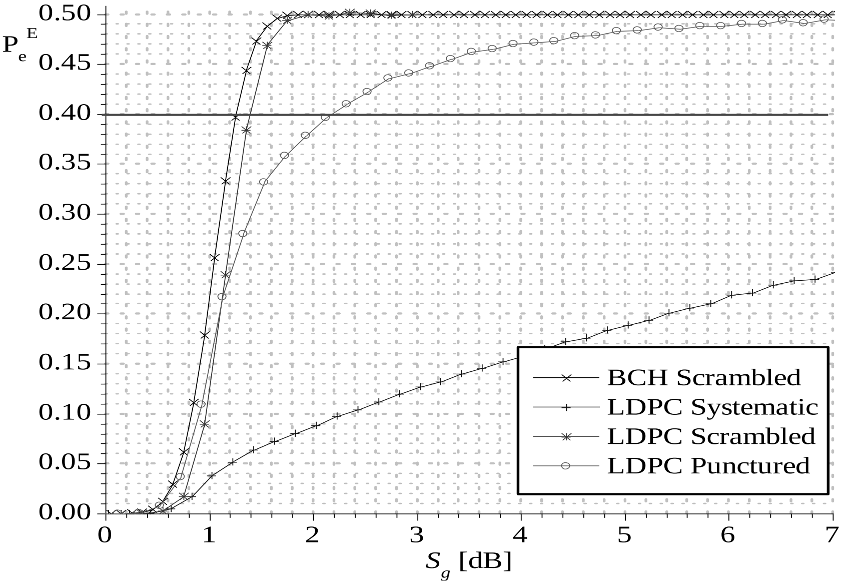

Figure 4.1: Security gap needed for different code choices

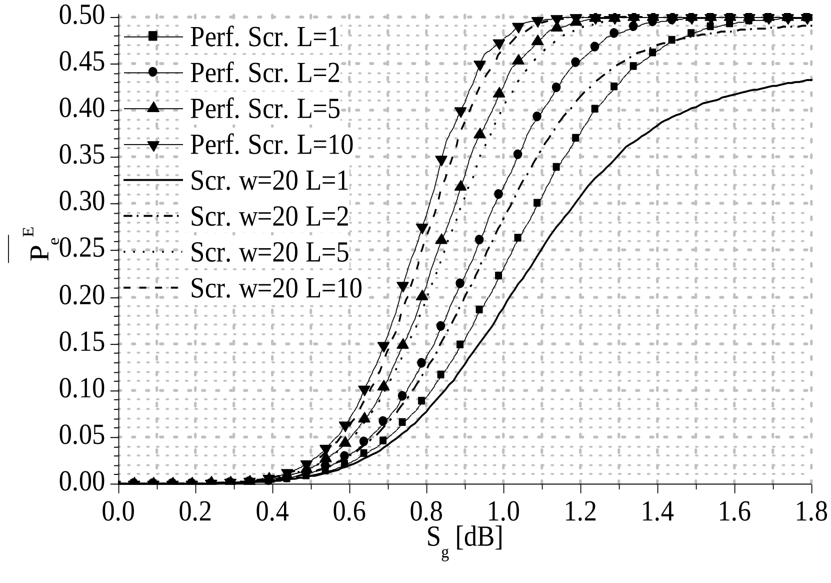

Figure 4.2: Effect of scrambling and concatenation

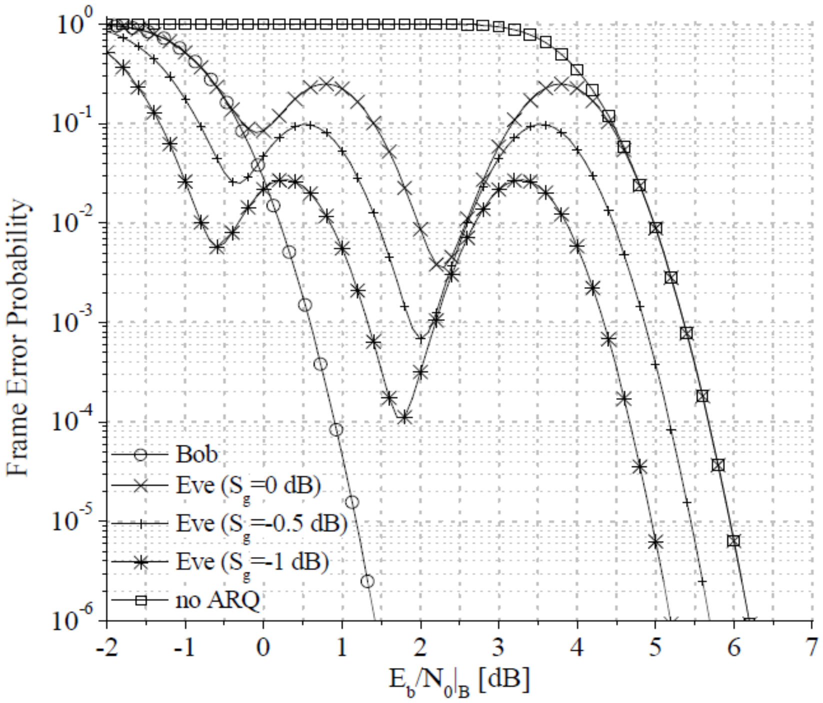

Figure 4.3: Effect of the feedback channel Between Bob and Eve

4 Conclusions

We analyzed two applications of channel coding for security issues, the first one was the study of a new variant of the McEliece cryptosystem based on a different transformation matrix, with respect to the original permutation one, able to change the secret code structure in such a way to allow the reintroduction of codes with high error correction capability such as GRS codes and the introduction of QC-LDPC codes instead of Goppa ones.

The higher error correction capability and the use of structured matrices, made possible the reduction of the public key size up to almost 1000 times with respect to the original system, obtaining keys less than 10 times greater than RSA ones, so practically usable.

The second one regarded a physical layer security framework. We proposed the adoption of widely used linear codes, such as BCH and LDPC, in such a way to minimize the security gap between the authorized and unauthorized users, needed to ensure security to the transmission in a real-life scenario.

Future works in this fields could regard the proof of security for the code-based cryptosystem and a further key size reduction by resorting to irregular QC-LDPC codes, while the study of the physical layer security framework could be extended to other channels (as the Rayleigh channel) and could be the basis for a working prototype of this security paradigm for automatic key distribution and renewal in wireless networks.

References

[1] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 5, no. 26, p. 1484–150, 1997.

[2] T. van der Sar, Z. H. Wang, M. S. Blok, H. Bernien, T. H. Taminiau, D. M. Toyli, D. A. Lidar, D. D. Awschalom, R. Hanson, and V. V. Dobrovitski, "Decoherence-protected quantum gates for a hybrid solid-state spin register," Nature, vol. 484, pp. 82–86, Apr. 20 2012.

[3] M. W. Johnson, et al ..., and G. Rose, "Quantum annealing with manufactured spins," *Nature*, vol. 473, pp. 194–198, May 2011.

[4] R. J. McEliece, "A public-key cryptosystem based on algebraic coding theory," *DSN Progress Report*, no. 114, pp. 42–44, 1978.

[5] V. D. Goppa, "A new class of linear error-correcting codes," *Probl. Peredachi Inf.*, vol. 6, no. 3, pp. 24–30, Sep. 1970.

[6] V. M. Sidelnikov and S. O. Shestakov, "On insecurity of cryptosystems based on Generalized Reed Solomon codes," *Discrete Math. Appl.*, vol. 2, no. 4, pp. 439–444, 1992.

[7] M. Baldi, "LDPC codes in the McEliece cryptosystem: Attacks and countermeasures,"*Proc. NATO Science for Peace and Security Series-D: Information and Communication Security*, vol. 23, pp. 160–174, 2009.

[8] M. Baldi, F. Chiaraluce, R. Garello, and F. Mininni, "Quasi-cyclic lowdensity parity-check codes in the McEliece cryptosystem," *ICC '07 IEEE International Conference on Communication*, pp. 951 –956, Jun. 2007.

[9] C. Peters, "Information-set decoding for linear codes over fq," *IACR Cryptology ePrint Archive*, vol. 2009, p. 589, 2009. [Online]. Available: http://eprint.iacr.org/2009/589

[10] D. J. Bernstein, T. Lange, and C. Peters, "Smaller decoding exponents: ball-collision decoding," IACR Cryptology ePrint Archive, vol. 2010, pp. 585, 2010. [Online]. Available: http://eprint.iacr.org/2010/585

[11] V. Gauthier, A. Otmani, and J.-P. Tillich, "A distinguisher-based attack on a variant of McEliece's cryptosystem based on Reed Solomon codes," *CoRR*, vol. abs/1204.6459, 2012. [Online]. Available: http://arxiv.org/abs/1204.6459

[12] R. G. Gallager, "Low-Density Parity-Check Codes," Cambridge, MA, USA: *The M.I.T. Press*, 1963.