Code-based cryptography

Raphael Overbeck[1] and Nicolas Sendrier[2]

[1] EPFL, I&C, LASEC.
[2] INRIA Rocquencourt, projet SECRET.

1 Introduction

In this chapter, we consider the theory and the practice of code-based cryptographic systems. By this term, we mean the cryptosystems in which the algorithmic primitive (the underlying one-way function) uses an error correcting code $\mathcal{C}$. This primitive may consist in adding an error to a word of $\mathcal{C}$ or in computing a syndrome relatively to a parity check matrix of $\mathcal{C}$.

The first of those systems is a public key encryption scheme and it was proposed by Robert J. McEliece in 1978 [48]. The private key is a random binary irreducible Goppa code and the public key is a random generator matrix of a randomly permuted version of that code. The ciphertext is a codeword to which some errors have been added, and only the owner of the private key (the Goppa code) can remove those errors. Three decades later, some parameter adjustment have been required, but no attack is known to represent a serious threat on the system, even on a quantum computer.

Similar ideas have been used to design other cryptosystems. Among others, let us mention some public key systems, like the Niederreiter encryption scheme [52] or the CFS signature scheme [14], and also identification schemes [73, 76], random number generators [19, 30] or a cryptographic hash function [3]. Some of the most important of those proposals are reviewed in §2.

As for any class of cryptosystems, the practice of code-based cryptography is a trade-off between security and efficiency. Those issues are well understood, at least for McEliece's scheme. Even though, no practical application of code-based cryptography is known to us. This might partly be due to the large size of the public key (100 kilobytes to several megabytes), but maybe also to a lack of publicity in a context were alternative solution were not urgently needed. Anyway, apart from the key size that we already mentioned, the McEliece encryption scheme has many strong features. First, the security reductions are tight (see [38] for instance). Also, the system is very fast, as both encryption and decryption procedures have a low complexity.

We will discuss in details the two aspects of security in §3 and §4. The first security assumption is the hardness of decoding in a random linear code [6]. This is an old problem of coding theory for which only exponential time solutions are known [4]. The second security assumption, needed only for public key systems, is the indistinguishability of Goppa codes [66]. Though it is not as old, in this form, as the first one, it relates to old problems of algebraic coding theory and is believed to be valid.

We will conclude this chapter with some practical aspects, first on the implementation, then on the key size issue, and we finish with a key point for the practicality of McEliece and related systems: how to efficiently construct a semantically secure (IND-CCA2) variant.

2 Cryptosystems

The first cryptosystem based on coding theory was a public key encryption scheme, presented in 1978 by McEliece [48]. Nearly all subsequently proposed asymmetric cryptographic schemes based on coding theory have a common disadvantage: the large memory requirements. Several other schemes followed, as the identification scheme by Stern [73], hash functions [3], random number generators [19] and efforts to build a signature scheme. The latter however all failed (compare [79], [32], [1] and [74]), until finally in 2001 Courtois, Finiasz and Sendrier made a promising proposal [14]. However, even if the latter is not broken, it is not suited for standard applications since besides the public key sizes the signing costs are large for secure parameter sets.

In 1986, Niederreiter proposed a knapsack-type PKC based on error correcting codes. This proposal was later shown to have a security equivalent to McEliece's proposal [42]. Among others, Niederreiter estimated GRS codes as suitable codes for his cryptosystem which were assumed to allow smaller key sizes than Goppa codes. Unfortunately, in 1992 Sidelnikov and Shestakov were able to show that Niederreiter's proposal to use GRS codes is insecure. In the following a couple of proposals were made to modify McEliece's original scheme (see e.g. [27], [26], [28], [70] and [35]) in order to reduce the public key size. However, most of them turned out to be insecure or inefficient compared to McEliece's original proposal (see e.g. [54] or [38]). The most important modifications for McEliece's scheme are the conversions by Kobara and Imai in 2001. These are CCA2-secure, provably as secure as the original scheme [37] and have almost the same transmission rate as the original system.

The variety of possible cryptographic applications provides sufficient motivation to have a closer look at cryptosystems based on coding theory as an serious alternative to established PKCs like the ones based on number theory. In this section we will concentrate on the most important cryptographic schemes based on coding theory.