Journal of Symbolic Computation 64 (2014) 67–87

Contents lists available at ScienceDirect

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

Computational aspects of retrieving a representation of an algebraic geometry code

Irene Márquez-Corbella [a], Edgar Martínez-Moro [b,c], Ruud Pellikaan [d], Diego Ruano [e]

[a] Grace INRIA Saclay Île-de-France, Laboratoire d'Informatique (LIX) UMR 7161 X-CNRS, 1 rue Honoré d'Estienne d'Orves, Campus de l'École Polytechnique, 91120 Palaiseau, France
[b] Institute of Mathematics, University of Valladolid, Castilla, Spain
[c] Department of Mathematics and Statistics, Eastern Kentucky University, USA
[d] Department of Mathematics and Computing Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[e] Department of Mathematical Sciences, Aalborg University, 9220 Aalborg Øst, Denmark

ARTICLE INFO

Article history:
Received 20 December 2012
Accepted 29 October 2013
Available online 7 December 2013

Keywords:
Public key cryptosystem
Code-based cryptography
Algebraic geometry codes
Gröbner basis

ABSTRACT

Code-based cryptography is an interesting alternative to classic number-theoretic public key cryptosystem since it is conjectured to be secure against quantum computer attacks. Many families of codes have been proposed for these cryptosystems such as algebraic geometry codes. In Márquez-Corbella et al. (2012) – for so called very strong algebraic geometry codes $\mathcal{C} = C_L(\mathcal{X}, \mathcal{P}, E)$, where $\mathcal{X}$ is an algebraic curve over $\mathbb{F}_q$, $\mathcal{P}$ is an n-tuple of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and E is a divisor of $\mathcal{X}$ with disjoint support from $\mathcal{P}$ – it was shown that an equivalent representation $\mathcal{C} = C_L(\mathcal{Y}, \mathcal{Q}, F)$ can be found. The n-tuple of points is obtained directly from a generator matrix of $\mathcal{C}$, where the columns are viewed as homogeneous coordinates of these points. The curve $\mathcal{Y}$ is given by $I_2(\mathcal{Y})$, the homogeneous elements of degree 2 of the vanishing ideal $I(\mathcal{Y})$. Furthermore, it was shown that $I_2(\mathcal{Y})$ can be computed efficiently as the kernel of certain linear map. What was not shown was how to get the divisor F and how to obtain efficiently an adequate decoding algorithm for the new representation. The main result of this paper is an efficient computational approach to the first problem, that is getting F. The security status of the McEliece public key cryptosystem using algebraic geometry codes is still not completely settled and is left as an open problem.

E-mail addresses: irene.marquez-corbella@inria.fr (I. Márquez-Corbella), edgar@maf.uva.es (E. Martínez-Moro), g.r.pellikaan@tue.nl (R. Pellikaan), diego@math.aau.dk (D. Ruano).

0747-7171/$ – see front matter
http://dx.doi.org/10.1016/j.jsc.2013.12.007

1. Introduction

McEliece (1978) introduced the first public key cryptosystem (PKC) based on error-correcting codes. The security of this scheme is based on the hardness of the decoding of random linear codes, or equivalently the problem of finding a minimum-weight codeword in a large linear code without any visible structure. This property makes the scheme of McEliece an interesting candidate for post-quantum cryptography. Another advantage consists of its fast encryption and decryption procedures. So one might hope that it is suitable for constrained devices like RFID tags or sensor networks, see Eisenbarth et al. (2009) for further results related to this issue. However, it has one important disadvantage: its low encryption size compared to its large key size. This does not mean that code-based cryptography is inherently inefficient. There have been many attempts on how to reduce the key size while keeping the same level of security, see for example Baldi et al. (2008), Berger et al. (2009), Biasi et al. (2012), Gaborit (2005), Misoczki and Barreto (2009), Misoczki et al. (2012), Monico et al. (2000). There are other public-key primitives based on the theory of error-correcting codes like signature schemes (Courtois et al., 2001), stream ciphers (Gaborit et al., 2007) or hash functions (Augot et al., 2005).

The principle of the McEliece cryptosystem is as follows:

Key generation: Given $\mathcal{C}$ an $[n, k, d]$ linear code defined over $\mathbb{F}_q$ with an efficient bounded distance decoding algorithm which corrects up to $t \leqslant \lfloor \frac{d-1}{2} \rfloor$ errors. Let

(1) G be a generator matrix of $\mathcal{C}$,
(2) S be an arbitrary nonsingular matrix of size $k \times k$,
(3) P be an arbitrary permutation matrix of size $n \times n$.

Let $G' = SGP$. Then the *McEliece public key* and the *McEliece private key* are given respectively by

$$\mathcal{K}_{\text{pub}} = \left(G', t\right) \quad \text{and} \quad \mathcal{K}_{\text{secret}} = (G, S, P).$$

Encryption: Suppose we want to send a message $\mathbf{m} \in \mathbb{F}_q^k$ using the public key (G', t). First, we choose a random error vector $\mathbf{e}' \in \mathbb{F}_q^n$ with Hamming weight at most t, and then, we compute the ciphertext $\mathbf{y}' = \mathbf{m}G' + \mathbf{e}'$.

Decryption: Using the private key (G, S, P) the receiver first computes

$$\mathbf{y} := \mathbf{y}'P^{-1} = \mathbf{m}G'P^{-1} + \mathbf{e}'P^{-1} = \mathbf{m}SG + \mathbf{e}.$$

Since SG is also a generator matrix of the code $\mathcal{C}$, he can apply the decoding algorithm for $\mathcal{C}$ to find $\mathbf{m}S$ and finally obtain the plaintext $\mathbf{m}$ from $\mathbf{m}SS^{-1}$.

McEliece proposed to use a $[1024, 524, 101]$ binary Goppa code. These parameters, however, do not attain the promised security level. We have mainly two different ways of cryptanalyzing the McEliece cryptosystem. There are also some side-channel attacks (Avanzi et al., 2011; Shoufan et al., 2010; Strenzke et al., 2008) but they are beyond the scope of this article.

(1) **Generic decoding attacks:** The best known technique for addressing the general decoding problem in cryptology is *Information Set Decoding* (ISD). The first approach to this method was introduced in Prange (1962). The variants which are used today are derived mainly from the algorithms of Stern (1989) and Lee and Brickell (1988). See Canteaut and Chabaud (1998), Peters (2011) and the reference therein, for recent improvements which were presented independently. Bernstein et al. (2008) presents the first successful attack on the original parameters of the McEliece scheme that required just under 8 days. More recent results (Becker et al., 2012; Bernstein et al., 2011; Finiasz and Sendrier, 2009; May et al., 2011) provide asymptotic improvements. Note that ISD, though much more efficient than a brute-force search, still needs exponential time in the code length. Therefore, more efficient generic attacks make the use of larger codes in the McEliece scheme necessary.