A New Structural Attack for GPT and Variants

Raphael Overbeck

GK Electronic Commerce,
TU-Darmstadt,
Department of Computer Science,
Cryptography and Computer Algebra Group
overbeck@cdc.informatik.tu-darmstadt.de

Abstract. In this paper we look at the Gabidulin version of the McEliece cryptosystem (GPT) and its variants. We propose a new polynomial time attack, which recovers an alternative private key. Our attack is applicable to all variants proposed so far and breaks some of them completely.

Keywords: public key cryptography, code based cryptography, rank distance codes, Gabidulin codes.

1 Introduction

The security of cryptosystems based on error correcting codes is connected to the hardness of the general decoding problem. In 1991 Gabidulin, Paramonov and Tretjakov proposed a variant of the McEliece scheme (GPT) [7] using *rank distance* codes instead of hamming distance codes. Smaller public-key sizes have been proposed for GPT than for the original McEliece cryptosystem, as general decoding algorithms are much slower for the rank metric than for the hamming-metric.

Gibson developed two structural attacks for the GPT cryptosystem (see e.g. [4] and [8]) and proved the parameter sets proposed in [7] and [4] to be insecure. A drawback of Gibson's attacks is, that they have exponential runtime if the secret key is carefully chosen. There were several attempts to modify the GPT cryptosystem, in order to avoid structural attacks, but most of these variants rely on security assumptions very similar to the ones for the original proposal (see [2] and [11]).

In this paper we build a new structural attack on the GPT cryptosystem. Unlike Gibson's attacks it has polynomial runtime, breaks the original GPT cryptosystem from [7] completely and is applicable to all GPT variants proposed so far.

The paper is structured as follows: First we give a short introduction to rank distance codes. Then we present the GPT cryptosystem and its Niederreiter variant. Finally we show how to attack the GPT cryptosystem.

E. Dawson and S. Vaudenay (Eds.): Mycrypt 2005, LNCS 3715, pp. 50–63, 2005.

2 Rank Distance Codes

Rank distance codes were presented by Gabidulin in 1985. They are linear codes over the finite field $\mathbb{F}_{q^m}$ for q (a power of a) prime and $m \in \mathbb{N}$. As their name says they use the concept of rank distance.

Definition 1. *Let $x = (x_1, \cdots, x_n) \in \mathbb{F}_{q^m}^n$ and $b_1, \cdots, b_m$ a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. We can write $x_i = \sum_{j=1}^m x_{ij} b_j$ for each $i = 1, \cdots, n$ with $x_{ij} \in \mathbb{F}_q$. The rank norm $\|x\|_r$ of x is defined as the rank of the matrix $(x_{ij}) \in \mathbb{F}_{q^m}^{n \times m}$.*

The rank norm of a vector $x \in \mathbb{F}_{q^m}^n$ is uniquely determined (independent of the choice of basis) and induces a metric, called *rank distance*.

Definition 2. *An (n,k)-code $\mathcal{G}$ over a finite field $\mathbb{F}$ is a k-dimensional subvectorspace of the vector space $\mathbb{F}^n$. We call the code $\mathcal{G}$ an (n,k,d) rank distance code if $d = \min_{x,y \in \mathcal{G}} \|x - y\|_r$. The matrix $G \in \mathbb{F}^{k \times n}$ is a generator matrix for the (n,k) code $\mathcal{G}$ over $\mathbb{F}$, if the rows of G span $\mathcal{G}$ over $\mathbb{F}$. The matrix $H \in \mathbb{F}^{n \times (n-k)}$ is called check matrix for the code $\mathcal{G}$ if it is the right kernel of G. The code generated by $H^\top$ is called dual code of $\mathcal{G}$ and denoted by $\mathcal{G}^\perp$.*

In [9] Ourivski and Johansson presented an algorithm which solves the general decoding problem in $\mathcal{O}\left((m\frac{d-1}{2})^3 q^{(d-3)(k+1)/2}\right)$ operations over $\mathbb{F}_q$ for (n,k,d) rank distance codes over $\mathbb{F}_{q^m}$. A special class of rank distance codes are the *Gabidulin codes* for which an efficient decoding algorithm exists [4]. We will define these codes by their generator matrix.

Definition 3. *Let $k \leq n \leq m \in \mathbb{N}$ and $g \in \mathbb{F}_{q^m}^n$ be a vector s.t. the components g_i, $i = 1, \cdots, n$ are linearly independent over $\mathbb{F}_q$. The (n,k,d) Gabidulin code $\mathcal{G}$ is the rank distance code with generator matrix*

$$G = \begin{pmatrix} g_1 & g_2 & \cdots & g_n \\ g_1^q & g_2^q & \cdots & g_n^q \\ \vdots & & \ddots & \vdots \\ g_1^{q^{k-1}} & g_2^{q^{k-1}} & \cdots & g_n^{q^{k-1}} \end{pmatrix} \in \mathbb{F}_{q^m}^{k \times n}. \tag{1}$$

An (n,k) Gabidulin code $\mathcal{G}$ corrects $\lfloor \frac{n-k}{2} \rfloor$ errors and has a minimum distance of $d = n - k + 1$. The dual code of an (n,k) Gabidulin code is an $(n, n-k)$ Gabidulin code (see [4]). The vector g is said to be a *generator vector* of the Gabidulin code $\mathcal{G}$ (it is not unique). Error correction based on the *right Euclidean division algorithm* takes $\mathcal{O}\left(d \log_2^2 d + dn\right)$ operations over $\mathbb{F}_{q^m}$ for (n,k,d) Gabidulin codes [4].

Throughout this paper we will use the following notation. We write $\mathcal{G} = \langle G \rangle$ if the (n,k)-code $\mathcal{G}$ over the field $\mathbb{F}$ has the generator matrix G. If the rows of a $(n-k) \times n$ matrix M span $\mathcal{G}^\perp$ we write $G^\perp = M$. We will identify $x \in \mathbb{F}^n$ with $(x_1, \cdots, x_n), x_i \in \mathbb{F}$ for $i = 1, \cdots, n$. For any (ordered) subset $\{j_1, \cdots j_m\} =: J \subseteq \{1, \cdots n\}$ we denote the vector $(x_{j_1}, \cdots, x_{j_m}) \in \mathbb{F}^m$ with x_J. Similarly, for a $k \times n$ matrix M we denote by $M_{.J}$ the submatrix consisting of the columns corresponding to the indices of J and write $M_{J'.} = \left(\left(M^\top\right)_{.J'}\right)^\top$ for any (ordered) subset J' of $\{1, \cdots, k\}$. Block matrices will be given in brackets.