Cryptanalysis of McEliece's Public-Key Cryptosystem

Valery I. Korzhik
Department of Communication Theory
Leningrad Electroengineering Institute of Communications
Mojka 66
Leningrad, 191065, USSR

and

Andrey I. Turkin
Computer Department
Gorky Polytechnical Institute
Nizhnii Novgorod, USSR

Abstract: An approach is proposed for the cryptanalysis of the well-known version of McEliece's public-key cryptosystem that is based on a new iterative optimization algorithm for decoding an arbitrary linear code. The algorithm provides guaranteed correction of all error patterns with Hamming weight less than d/2, where d is the minimum distance of the code, and has time complexity about $O(n^3)$ where n is the block length. The approach is illustrated by the cryptanalysis of McEliece's system when a (63, 36) binary code with d = 11 is the underlying linear code.

1. INTRODUCTION

We consider the well-known version [1] of McEliece's public-key cryptosystem. This cryptosystem is based on the generator matrix G of an (n, k) linear error-correcting code with minimum distance d having an efficient decoding algorithm that corrects all patterns of [(d -- 1)/2] or fewer errors (e. g., a Goppa code). In McEliece's system, a k-bit plaintext message **m** is transformed to a binary cryptogram **c** as

$$\mathbf{c} = \mathbf{m}\,G' + \mathbf{e}$$

where + denotes bitwise modulo-two addition, The $k \times n$ matrix G' is obtained from G in the manner

$$G' = S\,G\,P,$$

where S is a nonsingular $k \times k$ binary matrix and P is an $n \times n$ permutation matrix. The binary n-tuple **e** is an "error pattern" of Hamming weight at most [(d -- 1)/2] that is randomly chosen for each

D.W. Davies (Ed.): Advances in Cryptology - EUROCRYPT '91, LNCS 547, pp. 68-70, 1991.

message **m** that is encrypted. Knowledge of the individual matrices S, G and P allows the intended recipient to compute

$$c P^{-1} = (m S) G + e P^{-1},$$

then to recover the code word (**m** S) by use of the decoding algorithm, and finally to complete the decryption by computing (**m** S) S^{-1} = **m**. However, if the individual matrices S, G and P are not known but only their product G' is known, as is the case for the cryptanalyst, then the crucial difficulty is to determine the error pattern **e**.

2. A NEW APPROACH TO CRYPTANALYSIS OF McELIECE'S SYSTEM

One approach open to the cryptanalyst of the McEliece system is to find k coordinates in which the matrix G' contains a nonsingular $k \times k$ matrix and where **e** contains only zeroes. In [1] it is shown how to choose the parameters k and t for a Goppa code of given length n to maximize the cryptanalyst's effort for this attack.

A second approach for the cryptanalyst is to exploit the fact that the matrix G' may itself turn out to be equivalent to the generator matrix of another Goppa code, in which case knowledge of this matrix alone suffices to permit correction of the error pattern **e**. The probability of such a "fortunate event" (from the cryptanalyst's standpoint) is calculated in [1] and shown to be negligibly small.

We propose a third approach for the cryptanalyst based on the iterative optimization algorithm reported in [2]. This algorithm results in guaranteed correction of error patterns with weight at most $[(d - 1)/2]$ for an arbitrary linear code with minimum distance d. The cryptanalyst can use this algorithm on the cryptogram **c** to obtain the code word **m** G', after which he simply recovers the message **m** by inverting any k columns of G' that contain a nonsingular matrix.

The main features of this efficient (i.e., polynomial time) algorithm for decoding an arbitrary linear code up to its guaranteed error-correcting limit are the following:

(1) the embedding of a discrete set of binary-valued vectors into the continuous vector space $\mathbf{R}^n$;

(2) the reduction of the decoding problem to the problem of finding extrema in $\mathbf{R}^n$ under the constraints determined by the code; and

(3) an iterative procedure for finding the desired extremum when the Hamming weight of the error pattern is at most $[(d - 1)/2]$.

We remark that the existence of this efficient iterative decoding algorithm, a complete description and proof of which will be given in [2], does not

contradict the fact that the problem of decoding an arbitrarily given binary word to the nearest code word in an arbitrary linear code is known to be NP-complete [3], because the mentioned algorithm requires that the given binary word be at distance at most $[(d - 1)/2]$ from a code word in order to guarantee decoding to the nearest code word. We remark also that this algorithm does not give a solution to the related NP-complete problem of integer programming.

3. CRYPTANALYSIS FOR THE (63, 26) CODE

To show the applicability of the iterative decoding algorithm to cryptanalysis does not require that its general capabilities be proved; it suffices to show that cryptanalysis based upon it succeeds in many cases. To show this for McEliece's system, we considered the case where G is the generator matrix of a (63, 36) Bose-Chaudhuri-Hocquenghem (BCH) code with minimum distance $d = 11$. This matrix was transformed into G' by means of a randomly chosen 36×36 nonsingular matrix S and a 63×63 randomly chosen permutation matrix P. The error pattern **e** was a randomly selected 63-tuple of Hamming weight 5. In 100 trials of the cryptanalytic procedure described above, the correct message **m** was found every time. The cryptanalysis required about 10^6 operations, compared to the approximately $5 \cdot 10^6$ operations that would be required in an exhaustive attack.

In general, for an arbitrary (n, k) code, the attack reported here requires about $20 \cdot n^3$ operations. In particular, for the (1024, 654) BCH code with $d = 65$ that was recommended in [1], cryptanalysis on an IBM PC requires about 60 hours. The conclusion in [1] about the excellent security afforded by this system appears now to have been premature.

(During Eurocrypt '91, the above described cryptanalytic attack on the McEliece public-key cryptosystem was demonstrated on an IBM PC.)

REFERENCES

[1] C. M. Adams and H. Meijer, *Security-related comments regarding McEliece's public-key cryptosystem*, **IEEE Trans. Info. Th.**, vol. IT-35, pp. 454-457, March 1989.

[2] A. I. Turkin and V. I. Korzhik, *The practically-optimal decoding algorithm for arbitrary linear codes over a BSC with polynomial time complexity*, to be presented at the IEEE Intl. Symp. Info. Th., Budapest, June 1991.

[3] E. R. Berlekamp, R. J. McEliece and H. C. A. van Tilborg, *On the inherent intractability of certain coding problems*, **IEEE Trans. Info. Th.**, vol. IT-24, pp. 384-386, May 1978.