Security-Related Comments Regarding McEliece's Public-Key Cryptosystem

Carlisle M. Adams
Henk Meijer

Department of Computing and Information Science
Queen's University, Kingston, Canada

August 1986

Abstract

The optimal values for the parameters of the McEliece public key cryptosystem are computed. Using these values improves the cryptanalytic complexity of the system and decreases its data expansion. Secondly it is shown that the likelihood of the existence of more than one trapdoor in the system is very small.

Keywords: Public key cryptosystems, Goppa codes.

1. Introduction

McEliece [1978] has introduced a public-key cryptosystem which is based on algebraic coding theory. In this system, the receiver first constructs an easily-solvable linear error-correcting Goppa code C with generator matrix G and then transforms this matrix into G', a generator matrix for a seemingly difficult-to-solve linear code C'. The matrix G' is the public key of this system -- a message is encrypted by multiplying it with G' and adding errors to the resulting codeword. The legitimate receiver can recover the message by using a decoding algorithm for the original code C.

In the first part of this paper we compute the optimal number of errors that should be introduced in the encryption algorithm; in the second part we comment on the likelihood of finding transformations that will map the code C' into the code C, or into some other easily-solvable Goppa code. The results indicate that introducing the optimal number of errors yields very high security for this cryptosystem and that there is, with high probability, only one transformation from C' into an easily-solvable Goppa code (this is the transformation known to the receiver).

We will begin this paper with a short description of McEliece's system. We continue in Sections 3 and 4 with an analysis of two attacks which may be considered by an intended eavesdropper and close, in the final section, with some concluding remarks on our results.

This work was partially supported by Grant #A0282 from the Natural Sciences and Engineering Research Council.

C. Pomerance (Ed.): Advances in Cryptology - CRYPTO '87, LNCS 293, pp. 224-228, 1988.

2. McEliece's System

McEliece's public-key cryptosystem can be briefly described as follows:

1) The receiver constructs an easily-solvable, binary, error-correcting Goppa code C which has a $(k \times n)$ generator matrix G and an error-correcting capability of t errors (note that G is necessarily of full rank).

2) The matrix G is transformed by

$$G' = SGP \tag{1}$$

where S is a $(k \times k)$ invertible scrambling matrix and P is an $(n \times n)$ permutation matrix. The $(k \times n)$ matrix G' is then a generator matrix for an apparently arbitrary linear code C' (i.e. one for which a fast algorithm for correcting errors is not known).

3) G' is published as the encryption key; the sender encrypts a k-bit message vector $\underline{m}$ into n-bit ciphertext vector $\underline{c}$ by

$$\underline{c} = \underline{m}\, G' \oplus \underline{e}, \tag{2}$$

where $\underline{e}$ is an n-bit error vector of weight t chosen by the sender.

4) The receiver, knowing that

$$\begin{aligned} \underline{c} &= \underline{m}\, G' \oplus \underline{e} \\ &= \underline{m}\, SGP \oplus \underline{e} \end{aligned}$$

computes

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

and uses a decoding algorithm for the original code C to remove the error vector $\underline{e}\, P^{-1}$ and recover the vector $\underline{m}$ S. The sender's message is then easily found by

$$\underline{m} = (\underline{m}S)\, S^{-1}.$$

The private key for this system, therefore, consists of the three matrices G, S, and P.

This paper is mainly concerned with equations (1) and (2) above. We begin by calculating the optimal weight of the error vector $\underline{e}$ in equation (2), where a weight is "optimal" if it yields maximum security for this system. We then go on to consider equation (1). From (1) we have

$$G = S^{-1} G' P^{-1}. \tag{3}$$

We will compute the expected number of matrices S_i and P_i such that the code C_i with generator matrix

$$G_i = S_i G' P_i$$

is an easily-solvable Goppa code.

3. Parameters k and t.

As noted in [Adams 1985] and [McEliece 1978], there are several ways of attacking McEliece's cryptosystem. Of the known attacks, the one which follows has the lowest complexity. We will show how a suitable choice of parameters k and t will maximize this complexity and thus strengthen the algorithm against this attack.

Recall equation (2):

$$\underline{c} = \underline{m}\, G' \oplus \underline{e}.$$

Since $\underline{m}$ is a k-bit vector, we can reduce this to

$$\underline{c}_k = \underline{m}\, G'_k \oplus \underline{e}_k,$$

where $\underline{c}_k$ denotes any k components of $\underline{c}$ (ie. $\underline{c}_k = c_{i_1}, c_{i_2}, ..., c_{i_k}$), $\underline{e}_k$ denotes the corresponding k components of $\underline{e}$, and G'_k is the square matrix consisting of columns $i_1, i_2, ..., i_k$ of G'. Thus we have

$$\underline{c}_k \oplus \underline{e}_k = \underline{m}\, G'_k$$

or, if G'_k is invertible,

$$(\underline{c}_k \oplus \underline{e}_k)\, (G'_k)^{-1} = \underline{m}. \tag{4}$$

Note that if the k components of $\underline{e}_k$ are all zero, (4) reduces to

$$\underline{c}_k\, (G'_k)^{-1} = \underline{m}$$

and an enemy can recover the sender's message without decoding (since $\underline{c}$ and G' are known).

The work factor for this attack can be calculated as follows. The error vector $\underline{e}$ is an n-bit vector with t ones and n-t zeros. Therefore, the probability of choosing (without replacement) k zero components from $\underline{e}$ is

$$p = \binom{n-t}{k} \Big/ \binom{n}{k}.$$

Note that the enemy must, on average, make 1/p attempts before being successful and, for each attempt, must invert the (k × k) submatrix G_k. Assuming that matrix inversion requires k^a steps (see, for example, [Bunch, Hopcroft 1974] and [Pan 1978]), this gives a total expected work factor for this attack of

$$w = k^a \binom{n}{k} \Big/ \binom{n-t}{k} \text{ steps.} \tag{5}$$

From [Berlekamp 1973] or [McEliece 1977] it can be seen that for $n = 2^i$, n, k, and t are related by

$$k = 2^i - it = n - it.$$

Therefore, for n = 1024 (as suggested in [McEliece 1978]), we have k = 1024 - 10t. It can easily be shown by exhaustive search that for values of a between 2 and 3 the value of (5) is maximal for t = 37. For a = 3, the value of (5) is approximately $2^{84.1}$ for t = 37, while for t = 50 (the value proposed in [McEliece 1978]) (5) has the value $2^{80.7}$.

Moreover the lower value of t increases the value of k from 524 to 654 and therefore reduces the data expansion of the cryptosystem.

4. Trapdoors

Brickell [1985] has shown that iterated knapsack cryptosystems (proposed in [Merkle, Hellman 1978]) can be broken. The idea of the proof is that the public (difficult) knapsack can be transformed by the enemy into one of several easy knapsacks; finding the receiver's original easy knapsack is not necessary. We can examine McEliece's system in this light by estimating the likelihood of there being several transformations from the public key G' into an easily-solvable decoding problem.

To do this we define an equivalence relation R on the set of binary $(k \times n)$ matrices of full rank as follows:

A R B if and only if there exists a $(k \times k)$ invertible matrix S
and an $(n \times n)$ permutation matrix P such that $A = S\ B\ P$.

If we call [A] the equivalence class induced by R containing A then it is clear that the private matrix G of McEliece's system is in the equivalence class [G'] of the public matrix G'. However, if there are other Goppa code generator matrices in the equivalence class [G'], a Brickell-like attack on this cryptosystem may be feasible.

If we assume that Goppa code generator matrices are evenly distributed over the set of all $(k \times n)$ matrices of full rank, we can calculate the expected number EXP of Goppa code generator matrices in an equivalence class of R by

$$EXP = \#G/\#C \qquad (6)$$

where #G is the number of Goppa code generator matrices for a given n and k and #C is the number of equivalence classes of R. To our knowledge, the above assumption has never been mentioned in the open literature (likely due to the fact that Goppa code generator matrices are not, in general, recognizable as such).

The values #G and #C have been computed in [Adams 1985]. For Goppa codes with error-correcting capability $t = 50$ and dimension $k = 524$ (the parameters suggested in [McEliece 1978]), #G is less than or equal to the number of irreducible polynomials of degree 50 over $GF(2^{10})$ (for $n = 2^{10} = 1024$) and #C is roughly equal to the total number of binary (524×1024) matrices of full rank divided by the average size of an equivalence class. Substituting the calculated values into (6) we have

$$EXP < 2^{504}/2^{500000} << 1. \qquad (7)$$

It can be shown that for $t = 37$ and $k = 654$ (from the previous section) the value of EXP is even smaller.

Given the above assumption, then, equation (7) shows that we expect that an arbitrary equivalence class of R does not contain a generator matrix for a Goppa code. From the construction of the cryptosystem, however, we know that the equivalence class of G' contains the receiver's private matrix G; therefore, we conclude from (7) that G is the only Goppa code generator matrix in [G']. Thus, the only transformation from G' to an easy generator matrix is the original transformation (3) chosen by the receiver and a Brickell-like attack against this system will be unsuccessful.

5. Conclusions

We conclude that McEliece's public-key cryptosystem appears to be fairly secure. We have shown that the lowest complexity cryptanalytic attack yet proposed has a work factor of roughly 2^{84} steps -- this is significantly higher than that of DES and compares very favourably with that of the RSA system. Furthermore, it seems that an attack similar in nature to Brickell's attack on the Knapsack cryptosystem will be unsuccessful.

References

Adams, C.M. (1985), Examination and Analysis of McEliece's Public-Key Cryptosystem, M.Sc. Thesis, Department of Computing and Information Science, Queen's University, Kingston.

Berlekamp, E.R. (1973), Goppa Codes, IEEE. Transactions on Information Theory, Vol. IT-19 #5 (Sept.).

Brickell, E.F. (1985), Breaking Iterated Knapsacks, Advances in Cryptology: Proceedings of Crypto 84, Blakley, G.R., Chaum, D. (Editors), Springer-Verlag, Berlin.

Bunch, J., Hopcroft, J.E. (1974), Triangular Factorization and Inversion by Fast Matrix Multiplication, Mathematics of Computation, Vol. 28; 125.

McEliece, R.J. (1977), The Theory of Information and Coding (Volume 3 of the Encyclopedia of Mathematics and its Applications), Addison-Wesley, Reading, Mass.

McEliece, R.J. (1978), A Public-Key Cryptosystem Based on Algebraic Coding Theory, DSN Progress Report (Jan, Feb), Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif.

Merkle, R., Hellman, M. (1978) Hiding Information and Signatures in Trapdoor Knapsacks, IEEE. Transactions on Information Theory, Vol. IT-24 #5 (Sept.)

Pan, V. (1978), Strassen's Algorithm is not Optimal, the 19th Annual Symposium on the Foundations of Computer Science.