МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 1999
MATHEMATICS AND EDUCATION IN MATHEMATICS, 1999
Proceedings of Twenty Eighth Spring Conference of
the Union of Bulgarian Mathematicians
Montana, April 5–8, 1999

SOME SECURITY-RELATED COMMENTS REGARDING McELIECE'S PUBLIC-KEY CRYPTOSYSTEM*

Yuri Borissov

Adame and Mejer [1] have computed the optimal values of the parameters used in McEliece's public-key cryptosystem by exhaustive search. In this paper we show that these values can be found by logarithmic search. The same technique can be applied to compute the optimal values of the parameters used in the scheme which outstands an attack based on the parity-check matrix.

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

1) The receiver constructs a binary error-correcting Goppa code $\mathbf{C}$ with a $(k \times n)$ generator matrix G and error-correcting capability of t errors.

2) He/she transforms the matrix G into

$$G^{'} = SGP,$$

where S is a $(k \times k)$ invertible scrambling matrix and P is a $(n \times n)$ permutation matrix. The $(k \times n)$ matrix $G^{'}$ is a generator matrix for an apparently arbitrary linear code $\mathbf{C}^{'}$ (i.e. for which a fast decoding algorithm is not known [5]).

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

$$\mathbf{c} = \mathbf{m}G^{'} + \mathbf{e}, \tag{1}$$

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

4) The receiver, knowing that

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

computes

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

and uses a decoding algorithm for original code $\mathbf{C}$ in order to remove the error-vector $\mathbf{e}P^{-1}$ and to recover the vector $\mathbf{m}S$. Finally the sender's message is found by $\mathbf{m} = (\mathbf{m}S)S^{-1}$. The private keys for this scheme,therefore, are the matrices G, S and P.

*This research was partially supported by the Bulgarian NSF under Contract I-803/98.

McEliece has proposed an attack on his cryptosystem which can be briefly described as follows [3]. Since $\mathbf{m}$ is a $k-$bit vector, we can reduce equation (1) to

$$\mathbf{c}_k = \mathbf{m}G_k^{'} + \mathbf{e}_k,$$

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

$$\mathbf{c}_k + \mathbf{e}_k = \mathbf{m}G_k^{'}$$

or, if $G_k^{'}$ is invertible,

$$(\mathbf{c}_k + \mathbf{e}_k)(G_k^{'})^{-1} = \mathbf{m}. \tag{2}$$

Note that if $\mathbf{e}_k = \mathbf{0}$, (2) reduces to

$$\mathbf{c}_k(G_k^{'})^{-1} = \mathbf{m} \tag{3}$$

and an opponent can recover the sender's message without decoding (since $\mathbf{c}$ and $G^{'}$ are public). If $d_H(\mathbf{c}_k(G_k^{'})^{-1}G, \mathbf{c}) \leq t$, then the opponent can claim that $\mathbf{c}_k(G_k^{'})^{-1}$ is true $\mathbf{m}$ [6].

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

$$p_{n,k,t} = \binom{n-t}{k} / \binom{n}{k}.$$

The opponent must, on average, make $1/p_{n,k,t}$ attempts before being successfull and, for each attempt, must invert the $(k \times k)$ submatrix $G_k^{'}$. Assuming that matrix inversion requires k^3 steps, this gives a total expected work factor for this attack of

$$W_G(t) = k^3 \binom{n}{k} / \binom{n-t}{k}. \tag{4}$$

From [4] and [8] we know that for $n = 2^i$ parameters n, k and t are related by

$$k \geq 2^i - it.$$

Equality will be used in this paper (see also [1] and [2]). Therefore, for $n = 1024$ (as suggested in [3]), we have $k = 1024 - 10t$. Adams and Mejer [1] have shown by exhaustive search that the value of $W_G(t)$ is maximal for $t = 37$. Here we obtain the same result using logarithmic search.

Optimal value of *t*. The following equation is obvious

$$\binom{n}{k} / \binom{n-t}{k} = \frac{n!}{(n-k)!} \frac{(n-t-k)!}{(n-t)!}. \tag{5}$$

Let $V_G(t) = \dfrac{W_G(t+1)}{W_G(t)}$. Using (5) in case $n = 1024$ and $k = 1024 - 10t$, we get:

$$V_G(t) = \frac{[1024-10(t+1)]^3}{(1024-10t)^3}\frac{[9(t+1)]!}{[10(t+1)]!(1024-t-1)!}\frac{(10t)!(1024-t)!}{(9t)!} =$$

$$\left(1-\frac{10}{1024-10t}\right)^3 \frac{(9t+9)(9t+8)\dots(9t+1)}{(10t+10)(10t+9)\dots(10t+1)}(1024-t) =$$

$$\left(1-\frac{10}{1024-10t}\right)^3 \frac{1024-t}{10t+10}\frac{9t+9}{10t+9}\frac{9t+8}{10t+8}\cdots\frac{9t+1}{10t+1} =$$

$$\left(1-\frac{10}{1024-10t}\right)^3 \frac{1024-t}{10t+10}\left(1-\frac{1}{9/t+10}\right)\left(1-\frac{1}{8/t+10}\right)\cdots\left(1-\frac{1}{1/t+10}\right).$$

Therefore $V_G(t)$ decreases when t increases. By straightforward calculations we obtain that $V_G(36) = 1.0020477$, while $V_G(37) = 0.9736544$. This means that the value of $W_G(t)$ is maximal for $t = 37$.

An attack on the scheme based on the parity-check matrix. Let $H^{'}$ be the parity-check matrix of the code $\mathbf{C}^{'}$. $H^{'}$ is an $(n-k \times n)$ binary matrix of rank $n-k$. It can be efficiently computed using generator matrix $G^{'}$ (see [7]). Furthermore the well known fact that syndrome of the error-vector is equal to syndrome of the transmitted vector gives the equation

$$H^{'}\mathbf{e}^T = H^{'}\mathbf{c}^T = \mathbf{s}. \tag{6}$$

Let $H^{'}_{n-k}$ be a square submatrix of $H^{'}$ consisting of columns $i_1, i_2, \dots, i_{n-k}$. Assuming $e_i = 0$ for $i \notin \{i_1, i_2, \dots, i_{n-k}\}$, (6) can be rewritten as

$$H^{'}_{n-k}\mathbf{e}^T_{n-k} = \mathbf{s},$$

where $\mathbf{e}_{n-k} = (e_{i_1}, e_{i_2}, \dots, e_{i_{n-k}})$.

Therefore, if $H^{'}_{n-k}$ is invertible

$$\mathbf{e}^T_{n-k} = (H^{'}_{n-k})^{-1}\mathbf{s}$$

and, if the weight of $\mathbf{e}_{n-k}$ is $\leq t$, then true error-vector $\mathbf{e}$ was found. Further on, knowing $\mathbf{e}$, the opponent can use (3) to recover the message $\mathbf{m}$ without decoding.

Since $H^{'}$ and $\mathbf{s}$ have to be computed only once, ignoring these computations, the work factor of just described attack is

$$W_H = (n-k)^3\binom{n}{k}/\binom{n-t}{k}. \tag{7}$$

Note that factor k^3 from (4) is replaced by $(n-k)^3$ in (7) since we must invert an $(n-k \times n-k)$ matrix.

Let us apply the same technique as in Section 2 to find the optimal value of t used in the scheme which outstands the above attack. Let $V_H(t) = \dfrac{W_H(t+1)}{W_H(t)}$. Using (5) after some noncomplex computations (again in case $n = 1024$ and $k = 1024 - 10t$) we get:

$$V_H(t) = \left(1 + \frac{1}{10t}\right)^3 \frac{1024 - t}{10t + 10} \left(1 - \frac{1}{9/t + 10}\right) \left(1 - \frac{1}{8/t + 10}\right) \ldots \left(1 - \frac{1}{1/t + 10}\right).$$

From this it is not difficult to see that $V_H(t)$ decreases when t increases, and by straightforward calculations we obtain $V_H(40) = 1.0136889 > 1$, while $V_H(41) = 0.09864933 < 1$. This means that the value of $W_H(t)$ is maximal for $t = 41$.

Note also, that since $W_G(t) > W_H(t)$ (when $n = 1024$ and $t \leq 51$) the attack based on the parity-check matrix is more efficient than the attack based on generator matrix.

REFERENCES

[1] C. M. Adams, H. Mejer, Security Related Comments Regarding McEliece PKC, CRYPTO-87.

[2] M. Alabbadi, S. B. Wicker, Combined Data Encryption and Reliability Using McEliece's Public-key Cryptosystem, *Proceedings of the International Symposium on Information Theory and its Applications*, Sydney, Australia, 20–24 November 1994.

[3] R. J. McEliece, A PKC based on Algebraic Coding Theory, *DSN Progress Report* (Jan. Feb. 1978), Jet Propulsion Laboratory.

[4] E. Berlekamp, Goppa Codes, *IEEE Trans. on IT-19*, **5**, Sept. 1973, 590-595.

[5] E. Berlekamp, R. McEliece, H. C. A. van Tilborg, On the inherent intractibility of certain coding problem, *IEEE Trans. Inform. Theory IT-24* **3**, (1978), 384–386.

[6] P. J. Lee, E. F. Brickell, An Observation on the Security Of McEliece's Public-key Cryptosystem, In *Advances in Cryptology-Eurocrypt'88 Proceedings*, 1989, Springer-Verlag, 275–280.

[7] W. W. Peterson, Error-Correcting Codes, MIT Press, Cambridge, MA, 1961.

[8] R. J. McEliece, The Theory of Information and Coding, (Encyclopedia of Maths and its Applications, vol. 3), Addison-Wesley, Reading, Mass.

Yuri Borissov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
G. Bonchev str., block 8
1113 Sofia, Bulgaria

НЯКОИ КОМЕНТАРИ ОТНОСНО СИГУРНОСТТА НА КРИПТОСИСТЕМАТА НА McELIECE

Юри Борисов

Adams и Mejer [1] поставят и решават задачата за намиране на оптимални стойности на параметрите на публично-ключовата криптосистема на McEliece, използвайки изчерпващо търсене. В настоящата статия показваме, че тези стойности могат да се намерят с логаригмично търсене. Същата техника може да се приложи и за пресмятане на оптималните стойности на параметрите на тази схема при разглеждане на въпроса за криптоустойчивостта ѝ относно атака, базираща се на проверочната матрица на използвания линеен код.