STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **LV**, Number 1, 2010

ENCRYPTION SYSTEM OVER THE SYMMETRICAL GROUP OF ORDER N

STELIAN FLONTA, LIVIU-CRISTIAN MICLEA, ENYEDI SZILÁRD

Abstract. The encryption systems with public keys are related to algebraic structures, which have to ensure, by means of their computational properties or their dimensions, a high level of security for the generated keys, which are secret. The ElGamal algorithm is defined over the group of remainder classes modulo n, which is a suitable structure for an encryption algorithm. This paper chooses another structure for which the ElGamal algorithm is to be applied to. This structure is the symmetrical group of order n. Properties which ensure the opportunity of choosing this structure are: the cardinal of the group is high enough, the operation of composition of the permutations is simple from a computational point of view and every permutation can be decomposed uniquely into a product of disjunctive cycles.

1. Introduction

The main goal of this paper is to develop an encryption system. Among the secondary objectives, which are necessary in order to achieve the main goal, the following stand out: elaborating a method for encoding / decoding a message by means of a permutation and presenting a method for choosing the permutation for the key generation. The paper starts with a paragraph that outlines different results of the research community. It continues with the presentation of the algorithm for encoding / decoding a message by the help of a permutation and the description of the cipher of type ElGamal. In the end, a study regarding the security perspective of the system is conducted and the author's contributions are pointed out.

Received by the editors: October 10, 2009.

2010 *Mathematics Subject Classification.* 94A60, 11T71.

1998 *CR Categories and Descriptors.* E.3 [**DATA ENCRYPTION**]: Subtopic – *Public key cryptosystems*;

Key words and phrases. ElGamal over the permutations group, discrete logarithm problem.

2.Related Work

The encryption systems are conceived starting from the properties of certain mathematical structures. Algorithms with public keys highlight those properties that ensure the generation of some keys sufficiently "secure" in a relatively simple way from computational point of view. Starting from the problem of the discrete logarithm over a group, an idea is presented in [9, p. 107] [8, p. 294] [1, ch. 12, p. 2] [7, p. 2]. A structure can be chosen to ensure a high level of difficulty for the discrete logarithm's determination from computational point of view. The problem of the discrete logarithm can be stated using different group structures. There are versions where the group is formed over the remainder classes or over the set of points which belong to the elliptical curves. Starting from these structures, the ElGamal and ElGamal over the elliptical curves encryption systems have been defined. The paper [13, p. 218] proposes a group structure formed over the set of points which belong to the conical curves. In this case, a conic is considered and the operation is defined so that a group structure is obtained over which the ElGamal system is implemented.

Another approach, which is described in [11, p. 473], concerns the mathematical structure over which the encryption system is defined. In the research community, some papers present hybrid systems obtained by combining some encryption systems with the ElGamal system [12, p. 436]. Other existing papers analyze the properties of the encryption function, proving that it is a homomorphism [3, p. 645]. Different approaches of the systems, which are based on the discrete logarithm problem [10, p. 210] [6, p. 2] [2, p. 9], outline the properties related to the method of calculus. There are limited possibilities for choosing a structure for which the problem of the discrete logarithm would be difficult to solve. Another solution [5, p. 445, tome III] presents an ElGamal type cryptographic primitive, with the key divided among the modulo n remainder classes. This differs from the present paper in that we propose an ElGamal type algorithm without a key divided over the n order symmetrical group. The paper [4, p. 226-234] proposes an encryption system based on discrete logarithm in symmetric group of n order problem. The differences between [4, p. 226-234] and the model developed in this paper consist in the coding mode of the message and the encryption algorithm. These differences will be presented in detail furthermore in the paper. The properties of the symmetrical group of order n recommend choosing it in the case of implementing an algorithm of ElGamal type. In this situation, the structure is a finite group which is not commutative.

3. ElGamal over the permutation group

This system will be defined in the following paragraph and it will be noted EGPGP. In order to construct a sufficiently secure system, it is necessary to choose a group with a great number of elements which allows the generation of some keys comparable with the ones frequently used, that is the ones of length 512 bits, 1024 bits or 2048 bits. Such a group is S_n, which has $n!$ elements. Applying the Stirling formula, which approximates $n!$, an estimation can be made for the cardinal of the group S_{128} and S_{1024}, so:

$128! \approx \left(\frac{128}{e}\right)^{128} \cdot \sqrt{2\pi 128} \geq 47^{128} \cdot 2^4 \geq 32^{128} \cdot 2^4 = 2^{644}$

$1024! \approx \left(\frac{1024}{e}\right)^{1024} \cdot \sqrt{2\pi 1024} \geq (2^8)^{1024} \cdot 2^6 = 2^{8198}$.

Another important aspect to be able to elaborate ElGamal over S_n is the codification of the information so that the product between the coded message m and the permutation h to be possible. From this results the necessity for the message m to be coded in a permutation, meaning that to every message m one and only one permutation will be associated. The association must be made using an easily computable bijective function. It is also important that the inverse of this function should be easily computed, this property being necessary in the decryption stage of the algorithm.

We consider A a set of symbols having the cardinal $n-k$. A message, of fixed length k, is $m = \alpha_1\alpha_2...\alpha_k$ where $\alpha_j \in A, \forall j = \overline{1,k}$.

Each symbol is encoded by a bijective function $f : A \to \{k+1, ..., n\}$. For each message m, a function $g_m : \{\alpha_1, \alpha_2, ..., \alpha_k\} \to \{1, ..., k\}$, $g_m(\alpha_j) = j, \forall j = \overline{1,k}$ is defined. This function associates to each symbol from the message m a number which means the symbol's position in the message. It is obvious the fact that the function g_m is bijective. We define the function

$$h : \{1, ..., n\} \to A, h(x) = \begin{cases} g_m^{-1}(x), x \in \{1, ..., k\} \\ f^{-1}(x), x \in \{k+1, ..., n\} \end{cases}$$

For each $\alpha_j \in m, j = \overline{1,k}$, we chose $i_1^j, i_2^j, ..., i_{(l-1)_j}^j, i_{l_j}^j \in \{1, ..., k\}$ such that $h(i_1^j) = h(i_2^j) = ... = h(i_{(l-1)_j}^j) = h(i_{l_j}^j) = \alpha_j$ and is attached the cycle

$$c_j = (i_1^j, i_2^j, ..., i_{(l-1)_j}^j, i_{l_j}^j) \in S_n, \;\; i_1^j \leq i_2^j \leq ... \leq i_{(l-1)_j}^j \leq i_{l_j}^j.$$

Through this process we attach to each symbol $\alpha_j, \forall j = \overline{1,k}$, which appears in the message, a disjoint cycle of the other cycles. The length of this cycle is $s+1$, where s is the number of times the symbol appears within the message. The permutation $p(m) = \prod c_j$ with the property that the cycles are all disjoint two by two. They are also longer than or equal to two. The reverse process through which the message m of $p(m)$ is determined is described in the following lines.

All disjoint cycles of $p(m)$, longer than or equal to two, are paired, and for each cycle the symbols $(i_1, i_2, , .., i_{l-1}, i_l)$, $\quad i_1 \leq i_2 \leq ... \leq i_{l-1} \leq i_l$ are determined such that $\alpha_{i_1} = \alpha_{i_2} = ... = \alpha_{i_{l-1}} = h(\alpha_{i_l})$. The message m recovery is realized by merging the symbols in increasing order of indices.

Of course, the cycles are permutations from the group where EGPGP is applied to, that is S_n. In practice, this can be S_{128} or S_{256} and the standard length of a message can be 32 characters or 64 characters. This choice is determined by the structure of the ASCII code. Starting from a standard length of the message, this idea can be generalized. If the standard length of the message is k, then the chosen group is S_{k+96}. We will come back to these possibilities with detailed explanations later in this paper. In [4, p. 230] the message m is an integer number. The permutation's determination, which is associated, is made using the representation in factorial base system, which is a positional system. In this paper the message m is formed by concatenation of alphanumeric symbols. The encoding through a permutation is made using a permutation's decomposition property in a unique mode in a product of disjoint cycles, making abstraction of order, and a numeric code for the used symbols, for example the ASCII code. Therefore the coding mode is totally different from [4, p. 230]. Further, a version of EGPGP over the group S_n will be presented. This variant is different from [1, ch. 12, p. 2] and [4, p. 226-234] by the mode of the keys generation, of encryption and decryption.

The steps of this algorithm are:

Key generation

A permutation $g \in S_n$ is considered such that the problem of the discrete logarithm is difficult to solve and $ordg = r$ is determined. Also the numbers $x_1, x_2 \in Z_{|\mathrm{H}|}$ are chosen such that $(x_1, r) = 1$, that is the two numbers are prime between them, where $H = \langle g \rangle$ and $h_1 = g^{x_1}, h_2 = g^{x_2}$ are computed. Also a number $s < n$ is chosen such that s divides the number x_2. From the extended Euclid algorithm the numbers α, β are determined such that $x_1\alpha + r\beta = 1$. The public key is $\{h_1, h_2, s\}$ and the secret key is $\{x_2, \alpha\}$. The elements g, r, x_1, β are secret, but they are not keys, consequently they will not be transmitted to the user of the secret keys.

Message encryption

If we desire the encryption of message m, with a maximum length k, than we determine $p(m)$, which is the associated permutation. Then $y \in Z_{|\mathrm{H}|}, t \in S_n$ are chosen, where t is a cycle of order s and $c_1 = t \cdot {h_1}^y$and $c_2 = p(m) \cdot {h_2}^y$ are computed. The encrypted message is (c_1, c_2).

Message decryption

The first step for decrypting the message (c_1, c_2) is to compute

$$\frac{c_2}{(c_1)^{x_2\alpha}} = \frac{p(m)\cdot(g^{x_2})^y}{t^{x_2\alpha}\cdot((g^{x_1})^y)^{x_2\alpha}} =$$

$$= \frac{p(m)\cdot g^{x_2y}}{e\cdot g^{yx_2(-\beta r+1)}}\frac{p(m)\cdot g^{x_2y}}{(g^{(-\beta r+1)})^{yx_2}} = \frac{p(m)\cdot g^{x_2y}}{(g)^{yx_2}} = p(m) \ .$$

In the next step, message m is determined from $p(m)$ using the decoding algorithm [1, 8].

4. Choosing the Permutation

A few observations regarding the actual implementation of the EGPGP cipher are necessary. In order to generate efficient keys, a permutation g is needed for which the problem of the discrete logarithm is difficult to solve. This is true if the order of g is sufficiently high. A way of generating such a permutation is based on the theorem of decomposing a permutation into disjunctive cycles and on the way of computing the order of the permutation. From these results, one deduces that a permutation has a great order if it is the product of some disjunctive cycles whose the least common multiple of lengths is maximum with the restriction that their sum is constant. The choice of some disjunctive cycles from S_n is reduced to choosing a disjunctive subset of the set $\{1, ..., n\}$. Every subset $\{i_{1,}i_{2,...,}i_s\}$ generated the cycle $(i_{1,}i_{2,...,}i_s)$. Consequently, an optimization problem can be stated: Determine the numbers $k_{1,}k_2, ..., k_j \in Z^*$, prime among them two by two, such that

$$\begin{cases} k_1k_2...k_j \to \max \\ k_1 + k_2 + ... + k_j = n \end{cases} .$$

A solution to this problem corresponds to a maximum order permutation which can also be written as a product of disjunctive cycles that have the lengths $k_{1,}k_2, ..., k_j \in Z^*$.

5. Conclusions

In the ElGamal version $({Z_q}^*, \cdot)$, which is the classical version, there are computed multiplications, additions, exponentiations, modulo q inversions. If q is a great number there are adequate resources. The ElGamal over $(S_n, \cdot)$ version assumes smaller exponents and the product and the inversion of the permutation are simple operations. This is another reason for which ElGamal over $(S_n, \cdot)$is preferred. In the process of key generation for the ElGamal over $(S_n, \cdot)$ system it is very important to be able to choose permutations of high order to ensure the resistance against breaking by means of "brute force". A solution to this problem corresponds to a permutation of maximum order which is written as a product of cycles that have the lengths $k_{1,}k_2, ..., k_j \in Z^*$. Practically, n is given, $k_{1,}k_2, ..., k_j \in Z^*$ are chosen, prime among them, such that $k_1 + k_2 + ... + k_j = n$ and then the set $\{1, ..., n\}$ is decomposed in sets with respectively $k_{1,}k_2, ..., k_j$ elements and after that the cycles obtained from these sets are multiplied. The product obtained this way is the permutation g that can be used for key generation. The next computations, using the sum respectively the product of the prime numbers smaller or equal to 97 and the approximation of the factorial number, using the Stirling formula, shows that in $(S_{1096}, \cdot)$ exist permutations which have the order greater or equal to 2^{128}. Also, the order of the group is very large 2^{8773}.

- $2 + 3 + 5 + 7 + 11 + 13 + 17 + 19 + 23 + 29 + 31+$
  $+37 + 41 + 43 + 47 + 53 + 59 + 61 + 67 + 71+$
  $+73 + 79 + 83 + 89 + 97 = 1060$
- $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 37 \cdot 41\cdot$
  $\cdot 43 \cdot 47 \cdot 53 \cdot 59 \cdot 61 \cdot 67 \cdot 71 \cdot 73 \cdot 79 \cdot 83 \cdot 89 \cdot 97 \approx$
  $\approx 2,3 \cdot 10^{36} \approx 2^{128}$
- $$1096! \geq 2^{8773}$$

Consequently the group $(S_{1096}, \cdot)$ has at least the order 2^{8773}. Also, from this group, a permutation of order 2^{128} can be easily chosen.

We developed an encryption system of ElGamal type over the symmetrical group of order n, and we compared this system with the cipher defined over the modulo n remainder classes. Also, we have presented an algorithm for encoding / decoding a message by using permutations and a method for choosing the permutation for the key generation.

References

[1] A. Atanasiu, *Cryptography*, course notes, Bucharest, 2009, http://www.galaxyng.com/adrian_atanasiu/cript.htm

[2] B. Chevallier-Mames, P. Paillier and D. Pointcheval, *Encoding-free ElGamal encryption without random oracles* , Public Key Cryptography - PKC 2006, pp. 91-104.

[3] L. Chen, Y. Xu, W. Fang and C. Gao, *A New ElGamal-Based Algebraic Homomorphism and Its Application*, 2008 ISECS International Colloquium on Computing, Communication, Control and Management, Guangzhou, 2008, pp. 643-648.

[4] J. N. Doliskani, E. Malekian and A. Zakerolhosseini, *A Cryptosystem Based on the Symmetric Group Sn*, IJCSNS International Journal of Computer Science and Network Security, vol. 8 No. 2, 2008, pp. 226-234

[5] S. Flonta and L. Miclea, *An extension of the El Gamal encryption algorithm, Proceedings of 2008 IEEE International Conference on Automation, Quality and Testing, Robotics, Cluj-Napoca, AQTR 2008, pp. 444-446.*

[6] M. P. Jhanwar and R. Barua, *A Public Key Encryption In Standard Model Using Cramer-Shoup Paradigm*, Cryptology ePrint Archive, 2008.

[7] A. Mahalanobis, *The discrete logarithm problem in the group of non-singular circulant matrices*, Cryptology ePrint Archive, 2009.

[8] A. Menezes, P. Oorschot and S. Vanstome, *Handbook of Applied Cryptography*, CRC Press, 1996.

[9] V. V. Patriciu, M. Ene-Pietroșanu, I. Bica and J. Priescu *Electronic Signatures and Security in Informatics*, Editura All, Bucureşti 2006.

[10] R. Schmitz, *Public Key Cryptography - A Dynamical Systems Perspective*, 2008 Second International Conference on Emerging Security Information, Systems and Technologies, Cap Esterel, 2008, pp. 209-212.

[11] S. H. Paeng, K. C. Ha, J. H. Kim, S. Chee and C. Park, *New public key cryptosystem using finite non-abelian groups*, Crypto 2001 (J. Kilian, ed.), LNCS, vol. 2139, Springer-Verlag, 2001, pp. 470-485.

[12] Q. Bing-cheng, Y. Yang-zin, Z. Xi-min and C. Yin-dong, *Iterative Composite Encryption Algorithm Based on Tea and Elgamal*, 2009 WRI World Congress on Computer Science and Information Engineering, Los Angeles, 2009, pp. 435-438.

[13]Dalu Zhang , Min Liu , Zhe Yang,*Zero-Knowledge Proofs of Identity Based on ELGAMAL on Conic*, IEEE International Conference on E-Commerce Technology for Dynamic E-Business, CEC-East'04, Beijing, 2004, pp. 216-223.

Technical University of Cluj-Napoca, Cluj-Napoca, Romania
E-mail address: Stelian.FLONTA@aut.utcluj.ro, Liviu.Miclea@aut.utcluj.ro, Szilard.Enyedi@aut.utcluj.ro