http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

International Technology Journal 5 (3): 520-523, 2006
ISSN 1812-5638

The Subset Search Algorithm: A New Attack on the Discrete Logarithm Problem over Finite Fields and Elliptic Curves

Kaqish Malek
Department of Computer Information Systems, Al-Ahliyya Amman University Al Salt Road,
P.O. Box: 183, Amman 19328, Jordan

Abstract: The security of the new information society that exchange data via modern intelligent communication systems is nowadays very essential because of public connectivity and the related threads (espionage or sabotage etc.). Many security product and specially cryptographic systems (e.g., RSA, ElGamal, Diffie-Hellman, Elliptic Curves cryptosystems, etc.) was invented to encrypt and decrypt data, where the security of such cryptosystems are based on the apparent intractability of solving some number theoretic problems (e.g., The Discrete Logarithm Problem, Integer Factorization Problem, Diffie-Hellman Problem, Quadratic Residuosity Problem, Knapsack problem, etc.). Such problems are generally considered as being difficult to solve if the associated parameters are carefully chosen. The Discrete Logarithm Problem (DLP) on finite fields can be defined as followed: If we assume Z_p (denotes the set of integers {0, 1, 2, …., p - 1}, where addition and multiplication are performed modulo p) is a finite cyclic group of order p, where α a generator of Z_p and $\beta \epsilon Z_p$, then the Discrete Logarithm of β to the base α, denoted $\log_a \beta$, is the unique integer x, $0 \le x \le p-1$, such that $\beta = \alpha^x$. Many years this problem was studied but no known polynomial-time algorithm for solving the Discrete Logarithm Problem (DLP) has been found. This study introduce a new attack on the Discrete Logarithm Problem over finite fields- ($F_{2^f}^*, F_{g^f}^*, F \ge 1$, p prime) and elliptic curves groups; this attack is more significant on elliptic curves groups, where the group size is much more smaller compared to finite fields groups because there is no known sub-exponential algorithm for computing the discrete logarithm problem on elliptic curves unlike the discrete logarithm problem over finite fields. This attack is similar to Shanks Baby-step Giant-step algorithms but contains some differences, e.g., the new algorithm requires 50% less memory usage and thus the discrete logarithm can be found faster. The Shank Baby-step Giant-step algorithm is consider as the one of the best algorithms of solving the DLP over elliptic curves, this fact give the new attack a cryptographic importance.

Key words: Cryptography, cryptanalysis, cryptosystems, discrete logarithm problem, elliptic curves, shanks baby-step giant step algorithm

INTRODUCTION

Since Internet, computer and communications technology are one of the fastest technology in today's world, it is very important to have suitable security products and security systems that meet all security need and wishes of the different user's.

Many standards organizations (e.g., Internet Architecture Board (IAB), Internet Engineering Task Force (IETF), etc.) did specify a huge set of security protocols, algorithms and applications that provide security services which meets that needs for data privacy and secure communication.

Although not all users needs and wishes can be achieved in one single mechanism; however, we can note that Cryptography underlies most of the security mechanisms. Cryptographic techniques or generally Cryptography is the science of data encryption and decryption.

The Codebreaker (Kahn, 1967) describes a 4000 years historical overview from its limited use by the Egyptians until it major use in the twentieth century. Today Cryptography enables you to securely store sensitive information or transmit it across insecure networks so that it cannot be read by anyone except the intended recipient. By using a powerful tool such as encryption we gain privacy, authenticity, integrity and limited access to data.

Cryptographic systems can be divided in private key cryptography (also known as conventional cryptography systems) and public key cryptography.

Private key cryptography, also known as secret-key or symmetric-key encryption, has an old history and is based on using one key for encryption and decryption. In the 1960s many modern private key cryptographic systems where developed based on Feistel cipher, e.g., Data Encryption Standard (DES), Triple Data Encryption standards (3DES), Advanced Encryption

Standard (AES), The International Data Encryption Algorithm (IDEA), Blowfish, RC5, CAST, etc.

Diffie and Hellman *et al.* (1976) published a revolutionary concept of public-key cryptography based on two keys (Public and Private key) that solved many weaknesses and problems in private key cryptography. Upon this, many public key cryptographic systems was invented (e.g., RSA (Rivest *et al.*, 1978), ElGamal (ElGamal, 1985), Diffie-Hellman *et al* (1976) key exchange, elliptic curves (Koblitz, 1987), etc.). The security of such Public key cryptosystems often based on apparently difficult mathematical number theory problems like the discrete logarithm problem over finite fields and over elliptic curves, the integer factorization problem, the Diffie-Hellman Problem, etc.

This study discuss a new attack on the discrete logarithm problem, a discrete logarithm based cryptosystem is only secure if discrete logarithms in the underlying group is difficult.

PROBLEM DEFINITION

Let p be a prime number, then Z_p denotes the set of integers {0, 1, 2, …., p - 1}, where addition and multiplication are performed modulo p. It is well-known that there exists a non-zero element $\alpha \epsilon Z_p$ such that each non-zero element in Z_p can be written as a power of α; such an element α is called a generator of Z_p. A group is called cyclic if such an element α exists.

A Field is a nonempty set F of elements with two operations "+" (called addition) and "•" (called multiplication) satisfying the following axioms: for all a, b, c ∈ F,

(i) F is closed under + and •, i.e., a + b and a •b are in F;
(ii) Commutative laws: a + b = b + a, a • b = b • a;
(iii) Associative laws: (a + b) + c = a + (b + c), a • (b • c) = (a • b) • c;
(iv) Distributive law: a • (b + c) = a • b + a • c.

Furthermore, two distinct identity elements 0 and 1 (called the additive and multiplicative identities, respectively) must exist in F satisfying

(v) a + 0 = a for all a ∈ F;
(vi) a •1 = a and a •0 = 0 for all a ∈ F;
(vii) For any a in F, there exists an additive inverse element (a) in F such that a + (a) = 0;
(viii) For any a ≠ 0 in F, there exists a multiplicative inverse element α^{-1} in F such that $\alpha \bullet \alpha^{-1} = 1$.

Finite field of prime order p or prime power $q = p^f (f \geq 1)$ is commonly denoted F_q or GF (q) (for Galois field) and because Z_m is a field if and only if m is a prime, we denote the field Z_m by F_m. This is called a prime field.

For simplicity I will only consider the Discrete Logarithm Problem in Z_p, which can be defined as follows: If we assume Z_p is a finite cyclic group of order p, where, α a generator of Z_p and $\beta \epsilon Z_p$, then the discrete logarithm of β to the base α, denoted $\log_\alpha \beta$, is the unique integer x, $0 = x = p-1$, such that $\beta = \alpha^x$ (Menezes *et al.*, 1999).

The Discrete Logarithm Problem (DLP) in Z_p has been the object of much study. The problem is generally regarded as being difficult if field parameters are carefully chosen. In particular, there is no known polynomial-time algorithm for the DLP. There are two types of algorithms for solving the discrete logarithm problem. Special-purpose algorithms (Denny *et al.*, 1998) attempt to exploit special features of the prime p. In contrast, the running times of general-purpose algorithms which depend only on the size of p, for example exhaustive search, Shank's Baby step Giant step (Cohen, 1993), Pohlig Hellman (Pohlig *et al.*, 1978) (for group with small prime factors), Pollard's rho algorithm (Pollard, 1978), Index Calculus method (Enge *et al.*, 2000; Gaudry, 1999) (efficient only in certain groups), Number field Sieve (Gordon, 1993), etc. The fastest general-purpose algorithms known for solving the discrete logarithm problem over finite fields are based on a method called the index-calculus and the best current algorithm known for the DLP is the number field sieve. To thwart these attacks, p should have at least 150 digits and p - 1 should have at least one "large" prime factor, but it is important to notice that security is relative regarding time and what is secure today may need to be improved tomorrow. The utility of the discrete logarithm problem in a cryptographic setting is that finding discrete logs is (probably) difficult, but the inverse operation of exponentiation can be computed efficiently (e.g., the square-and-multiply method), this property is known in number theory as trapdoor or one-way function, there is no proof existence of one way function, but it is widely believed.

SUBSET SEARCH ALGORITHM

The new attack introduced in this paper can computes discrete logarithm problem in arbitrary finite cyclic group, but for simplicity I will only consider the discrete logarithm problem in Z_p. The SUBSET SEARCH algorithm for finding the discrete logarithm x in finite fields Z_p (where, α generator of a cyclic group Z_p of order p, $\alpha^x = \beta$, $\beta \epsilon Z_p$) is based on the following observation:

Instead of computing x in:

$$\alpha^x = \beta$$

We consider:

$$\alpha^{xi}\ \alpha^{x} = \alpha^{j}$$

$x_i \in Z_p$ Can be set any random number, but for faster calculation of the discrete logarithm problem we set

$$x_i = -i*m \text{ Or } x_i = i*m \text{ where, } i \in \{0,1,..,m\}$$

We build an array A of group elements.
Search in array A for

$$j \text{ and } \alpha^{j}$$

If j, α^j found in table A, then:

$$x \equiv (j-x_1) \bmod (p-1) \text{ Or } x \equiv (j+x_1) \bmod (p-1)$$

Proposed algorithm: The Subset search algorithm for computing the discrete logarithm x can be described as followed:

Step 1: Choose maximum array length m (equals a Mersenne-number $M_n = 2^n-1$) that fit in memory.

A Mersenne List L_n (n>1) of length 2^n-1 with sorted elements have the following properties:

(i) Each Mersenne list has a middle element. His index is at 2^{n-1}.
(ii) All element right (left) from the middle element build also Mersenne list L_{n-1}
(ii) All 2^{n-1} Elements right (left) from the middle element are bigger (smaller) than the middle element

The maximum number of searching step in a Mersenne List L_n requires maximum n steps.

Step 2: Construct a one dimension array A of length m,

with entry $(\overline{p} * \alpha^j + j)$ for $0 = j < m$, $\overline{p} < p$, $\overline{p}$ is prime, $j_0 = 0, j_1 = 1, ..., j_{m-1} = m-1$.

A: $\overline{p}\alpha^{j_0}+j_0$	$\overline{p}\alpha^{j_1}+j_1$	$\overline{p}\alpha^{j_{m-2}}+j_{m-2}$	$\overline{p}\alpha^{j_{m-1}}+j_{m-1}$
0	1	m-2	m-1

Note: if we choose j_x randomly then Array A elements can also be set any arbitrary group elements, but for faster calculation of the discrete logarithm problem we choose the above parameters.

Step 3: Sort array A by α^j.

Notes: 1: One can also use a hash table instead of sorted array to implement the array lookup.
2: The discrete logarithm x can be found faster if we choose larger m.

Step 4: Check array A, if $\alpha^x = \alpha^{jx}$ then return $x = j_x$

Step 5: For i = 1 to m-1 do the following:
Step 5.1: Set $x_1 = i*m$
Step 5.2: Compute α_{x1} and α^{-x1}
Step 5.3: Compute $\alpha_{x12} = \alpha^x * \alpha_{x1}$ and $\alpha_{x22} = \alpha^x * \alpha_{-x1} \bmod (p-1)$
Step 5.4: Check array A, if $\alpha_{x12} = \alpha_{jx}$ then return $x = j_x - x_1 \bmod (p-1)$
Step 5.5: Check array A, if $\alpha_{x22} = \alpha_{jx}$ then return $x = j_x + x_1 \bmod (p-1)$

We now consider algorithm runtime, this algorithm require storage for O(m) group elements, The array takes O(m) multiplications to construct and O(m*log (m)) comparison to sort. After having constructed the array, step 5 takes $O(\sqrt{p})$ array look-ups. Thus the algorithm can be implemented to run in $O(\sqrt{p})$ time.

Example: The following sketch was made with Mathematica 5, it shows calculation time comparison (sec) between Shanks Baby step Giant step (Fig. 1) and the Subset search algorithm for (Fig. 2) solving the discrete logarithm x of 100 randomly chosen $\beta = \alpha^x$, where p = 897579869, α = 3, m = 100000 in Shanks algorithm, m = 200000 in Subset search algorithm (50% less memory).

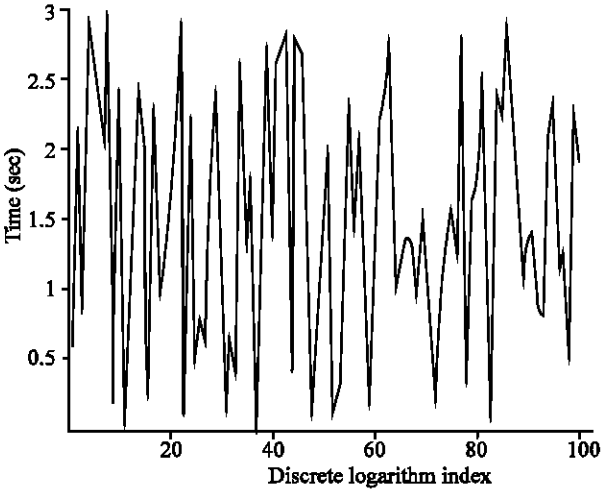

Fig. 1: Subset search algorithm

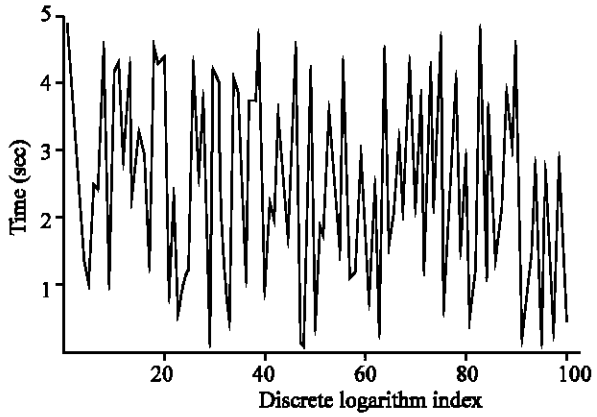

Fig. 2. Shanks Baby step Giant step algorithm

CONCLUSION

This study we briefly discussed the security of public-key cryptographic systems that are based on the discrete logarithm problem, however; they will no longer be secure if the corresponding hard problem is solved in the future.

Some threads come from general purpose attacks. The new attack described above have several advantages compared with other general purpose attacks, first it can be applied on finite fields and elliptic curves groups, this attack is of more importance on elliptic curves groups, because elliptic curves group size is much more smaller compared to finite fields groups and where elliptic curve encryption is widely used especially on limited memory devices such as smart cards, additionally the new attack use 50% less memory usage compared with Shanks baby step giant step algorithm and it can easily be implemented. Nevertheless the calculation of the discrete logarithm x in F^*_p where p have at least 150 digits and p-1 have at least one "large" prime factor, will stay hard.

The best known algorithm for finding the discrete logarithm in finite field groups is the index calculus method, unfortunately it can not be transformed on elliptic curves groups, the best known algorithms for any other finite abelian group (such as elliptic curves groups) is the Pollard (1978) is described in the present study techniques. Shanks Baby step Giant step algorithm (Cohen, 1993), or the new attack, all these attacks are close to the best what cryptanalysis till now achieved on finite fields and elliptic curves.

Many cryptosystems (e.g., based on knapsack problem) have been broken; our assumption about the intractability of the discrete logarithm problem may change in the future due to mathematic insights that will allow the definition of new attacks that may solve the Discrete Logarithm Problem in reasonable time.

REFRRENCES

Cohen, H., 1993. A Course in Computational Algebraic Number Theory of Graduate Texts in Mathematics. Springer-Verlag.

Denny, T. and D. Weber, 1998. The Solution of McCurley's Discrete Log Challenge. Advances in Cryptology- CRYPTO '98, Lecture Notes in Computer Science, colume 1462, Springer-Verlag, pp: 458-471.

Diffie, W. and M. Hellman, 1976. New directions in cryptography. IEEE Transactions on Information Theory, 22: 644-654.

Gaudry, P., 1999. A variant of the Adleman-DeMarrais-Huang algorithm and its application to small genera, Laboratoire d' inforatique Preprint, LIX/RR/99/04.

Gordon, D., 1993. Discrete logarithms in GF(p) using the number field sieve. SIAM J. Discrete Mathematics, 6: 124: 138.

ElGamal, T., 1985. A public-key cryptosystem and a signature scheme based on discrete logarithms. IEEE Transactions on Information Theory, 31: 469-472.

Enge, A. and P. Gaudry, 2000. A General Framework for Subexponential Discrete Logarithm Algorithms, Manuscript, (Feb, 2000), pp: 19.

Kahn, D., 1967. The Code breakers: the Comprehensive History of Secret Communication from Ancient to the Internet, Published.

Koblitz. N., 1987. Elliptic curve cryptosystems. Mathematics of Computation, 48: 203-209.

Menezes, A., P. van Oorscot and S. Vanstone, 1999. Handbook of Applied Cryptography, CRC Press.

Rivest, R.L., A. Shamir and L. Adleman, 1978. A method for obtaining digital signatures and public key cryptosystems. Commun. of the ACM, 21: 120-126.

Pollard, J., 1978. Monte Carlo methods for index computation mod p. Mathematics of Computation, 32: 918-924.

Pohlig, S. and M. Hellman, 1978. An improved algorithm for computing logarithms over GF(p) and its cryptographic significance. IEEE Transaction on Information Theory, 24: 106-110.

Pomerance, C., J.W. Smith and R. Tuler, 1988. A pipeline architecture for factoring large integers with the quadratic sieve algorithm, SIAM J. Comput., 17: 387-403.

Shor, P.W., 1997. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comput. Available at http://www.research.att.com/shor., 26: 1484-1509.