ISSN No. 0976-5697

Volume 3, No. 4, July- August 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Review of Key Length SelectionFormula for Elliptic Curve Cryptosystems

Tito Waluyo Purboyo[*1] and Budi Rahardjo[2]
School of Electrical Engineering and Informatics
Institut Teknologi Bandung
Bandung, Indonesia
[*1]titowaluyo@yahoo.com
[2]br@paume.itb.ac.id

Abstract: Cryptography is an important tool in e-commerce because it allows the cryptographic protection of electronic information in an e-commerce. The effectiveness of this protection depends on many variables such as length of the key, cryptographic protocols and the selection of passwords. This paper specifically discusses the selection of the keylength for elliptic curve cryptosystems. Selection of keylength is an important factor in determining the security level of the cryptography used. In this paper we discuss the determination of the key length for public key cryptography based on a set of parameters which are explicitly formulated. Another factor that determines the key length is also discussed in this paper such as the effort and costs incurred for the attack against elliptic curve cryptosystem.

Keywords: The Key length, Level of Security, Elliptic Curve Cryptosystems, Cryptographic Key Sizes, Cryptography.

I. INTRODUCTION

Key length expressed in number of bits contained in a specific cryptographic key or a related arithmetic structure. If no adequate explanation, the relationship between key length and security can be confusing. As an illustration, the key length 80, 160 and 1024 are clearly different key length. The new key length can be said to be equivalent if the 80 length of the key is a symmetric key length, the 160 length of the key is a hash length and the 1024 length of the key is a number of bits in an RSA modulus. The relationship between the seemingly different key lengths can be explained by the fact that the symmetric encryption key length and the B-bit cryptographic hash with the 2B-bit security have roughly the same [2].

The following data on the symmetric key length, RSA and ECC has a comparable level of security (Table 1).

Table 1. Key Length Comparison [4, 7-9]

RSA Key Length (bit)	ECC Key Length (bit)	Symmetric Key Length (bit)
1024	160	80
2048	224	112
3072	256	128
7680	384	192
15360	512	256

From the Table 1 above shows that the key length is usually a power of 2. It is not for mathematical reasons or security but because the data are usually highly processed and stored in a suitable multiple of 8 bits (byte), 32 bit (word) and 64 bit (block) and so on.

The strength of a cryptosystem depends on the algorithm and key, especially depending on the length of the key. Let us assume that an algorithm is robust, we will discuss how well the importance of key length for symmetric key systems and asymmetric key systems [4, 6].

Now we will discuss why the key length is crucial in cryptography.

A good cryptographic system is a cryptographic system based on the known algorithms (not suppressed). This means that the algorithm is considered adequate to be announced to the public, because only a knowledge of the key mechanisms that will work to decrypt the encrypted data.

Assume that an algorithm is robust, meaning that there is no known weaknesses in the algorithms that can be exploited by attackers, and that is practically impossible to reconstruct the plaintext message without knowing the key. The only way for the attacker / eavesdroppers to reconstruct the plaintext message is to find the key. Assume also that the key can only be found by trying all the possibilities that exist, namely to find the key by brute force.

In a brute force attack, the attacker tries all possible keys and each key is used with the cipher textuntil the plaintext messages are found. The attacker does not try to find the key by certain logic, but based on its ability to generate all the possible keys that should be attempted until the discovery of a suitable key. All encryption algorithms have a weakness for crack brute force attacks. To reduce this gap then we made a longer key. If the bit length of a lock longer then the key space is greater. Therefore the chance that the key should be attempted by an attacker is up.

Thus, the cryptographic key size is very influential. The longer the key, the harder it is to crack an encrypted data block. The longer the key bit difficult to try all possible keys grows exponentially. Increasing the key length for a bit to make the good guy job to be a little more difficult, but making the bad guy to work twice as hard (because the number of possible keys to be doubled).

At the time of deciding to use a key length specified in the use of cryptographic algorithms, one must consider the two exchanges (tradeoffs). Long keys can provide better security. Short keys can provide better efficiency. Therefore, it is

important to determine the optimal key length by considering the possibility of whether a key can be guessed, on the other hand should also consider the influence of key length of time required to encrypt the plaintext and the time required by the recipient to decrypt the cipher text.

Computers in today's era are usually used to try all possible keys in a brute force attack. Until the mid-1990s, individual attackers could not find a computer capable of trying all the keys in a relatively short time. Practical brute force attacks can only be done by an entity that has a lot of money as large companies or governments. But now, highly capable computers can be found widely. To accelerate the discovery of a key, an attacker put together a group of computers working in parallel in an organized way to crack the key. Each machine is assigned a key block to be tested. If there are n computers to crack the key, then in theory the key will be found in 1 / n if use only one computer. This makes a brute force attack against the key can be performed by an attacker with low financial capability (but they must have very many friends to be able to do so).

In reality, not just a brute force attack to try every possible key, but the attacker must be able to determine when the correct key was found. Humans may be able to recognize plaintext generated as the plaintext message if there are certain words used in everyday language. But if the plaintext has been compressed before it is encrypted, then the work to determine that the resulting plaintext is the original message will be more difficult. The original message can also be a number of data which led to the determination of the original message becomes more difficult. And if the computer is used to try the key, it would require additional software to recognize the semantics of the original message. The bottom line is there something that needs to complete a brute force attack the plaintext knowledge of what to expect and how to distinguish the results of many plaintext.

Moore's Law describes a prediction that the power of the computer will form the doubling every 18 months. Increased computer power has led to a brute force attack to find the key in the relatively short time. Challenge to find the key should be made more difficult because of the increase in computer power. The way to do that is by increasing the number of possible keys that must be tried by increasing the number of bits used in the key. If the computer becomes m times faster, the key length must be added as $\log_2 m$ bits. For example, 56-bit key length is no longer sufficient to DES (a symmetric key cryptography) as to the year 1975, because in 2000 it computer to 1000 times faster than in 1975. To provide the same protection, the key length should be $\log_2 (1000)$, or approximately 10 bits longer is 66 bits [1, 2].

There are two ways to defeat the cryptographic algorithm. The first way is to look for weaknesses in the cryptographic algorithm itself while the second way is to do a brute force attack [4].

In Section II we will discuss the topic on the level of security and other topics related to the level of security.

II. THE LEVEL OF SECURITY

A. Generic Attack:

For a symmetric cryptosystem, attacks generic (generic Attack) is defined as an attack in which the key is reconstructed from plaintext-cipher text pair is known. Plaintext and cipher text in the case of block ciphers can be composed of several blocks that are not too long. In the cryptographic literature, such an attack is called a known plaintext attack on the (known plaintext attack). It is assumed that the plaintext-cipher text pair can be used to reconstruct a unique key. For asymmetric cryptosystems, generic attack is defined as an attack where the private key is reconstructed from a known public key [2].

Generic attack does not include attacks where the attacker has access to other data that can only be generated by an unknown key, as in linear cryptanalysis and differential cryptanalysis against block ciphers. We will only review the generic attack because it is generally believed that the generic attack is closest to the situation in real life. Another reason is that generic attack against block ciphers is an offensive attack with the lowest business costs [2].

B. Level of Security:

If a generic attack on a symmetric cryptosystem with a λ-bit key length requires more expensive than the cost required to perform the exhaustive key search, then the cryptosystem that has a security level λ. Exhaustive key search for keys with a length of λ-bit keys may include the total amount of 2λ-1 as a different key, which for the worst case could reach the keys 2λ. In general, a cryptographic system offers the security level λ if a generic attack that had expected to require as much effort 2λ-1 (the approximation)[2]. How an attack attempt (attack effort) can be measured will be explained in the followingsection.

C. The cost of an Attack Effort:

Security level refers explicitly to attempt an attack (attack effort) and do not refer to the time it takes to make it happen.

All the attacks discussed here can be fully parallelized (fully parallelizable). Assume that an attack can be realized within d days at a device at a cost of c dollars. Furthermore, for a number w, the attack can be realized in d/w day at a cost of cw dollar device. This means that an approach to measure the attack effort is obtained by multiplying the time required by the cost of equipment (equipment cost). The cost of an attack effort measured in dollar days, so businesses attacks mentioned above will cost dollar days dc. Examples of fully parallelizable attack is exhaustive key search, because the search space can be any number divided by the number of processors that can work independently on a range of search which are assigned to processors [2].

D. Relationship Between Security and Security Level:

To determine whether a cryptographic system offer adequate security or protection, is not always useful to apply the definition of security levels on the length of symmetric key cryptosystems. The amount of time and money needed to realize an attack is always the smaller enterprises all the time because computers become faster and cheaper. Furthermore, the amount of protection offered by a certain security level is constant is also reduced. It is also related to the development of

the cryptanalysis of the time affect the security level of a cryptographic system is not due to the less cost for development of methods of attack but the attack itself. Furthermore, the security level definition includes a constant of proportionality (the approach taken constant 2λ-1) so that means can vary from one system to another. So the term adequate protection is dependent on the application that can still be subjective [2].

E. Defining Adequate Protection:

Data Encryption Standard (DES) is a symmetric cryptosystem with 56-bit key length, published in 1977 by the U.S. Department of Commerce. Although there is doubt about the security level of this DES, but because there is no generic attack were found better than the exhaustive key search of DES is still believed to have a security level 56 [2].

Because DES is widely adopted, it held a consensus in 1982, the first year in which to do a review of the DES is DES still offers adequate protection for a number of commercial applications. For this reason, the adequate protection is defined as the security offered by DES in 1982. By ignoring the constant of proportionality (2λ-1), adequate protection is synonymous with security level 56 in the year 1982 [2].

F. The Costs to Solve DES:

Exhaustive key search attack against DES in 1980 (of course with the value of money and technology in 1980) takes 2 days (on average) on the device (device) which was built at a cost of approximately 50M USD. The design used was fully parallelizable as described above. So, in 1980, DES can be solved at a cost of about 100M dollar days (2 days multiplied by 50M USD) [1].

Furthermore, public key cryptosystems are briefly discussed in chapter III.

III. PUBLIC KEY CRYPTOSYSTEM

In public-key cryptosystem, the receiver R has a private key (a secret) and the corresponding public key which anyone can access them, including S. The sender S uses R's public key to encrypt a message to be sent to R, and R using a private key to decrypt the encrypted message. If the private key can be derived from the public key, it means the system can be solved. The contents of the public key and private key, and how difficult they are used to solve the system depends on public-key cryptosystem is used. For the purposes of cryptanalysis and historical reasons, we distinguish three types of cryptosystems, namely the classical asymmetric systems, discrete logarithm systems subgroup and elliptic curve systems [1].

The systems will be discussed in this paper is the elliptic curve systems.

A. Elliptic Curve Systems:

Elliptic curve systems were first introduced by Koblitz and Miller independently in 1985 as mentioned in [9]. The idea is [1] g generator that generates a subgroup of the group of points on an elliptic curve E over the field to the (finite field) GF (p). Number q of the subgroup generated by the generator g is a prime number and a private key k is in the range [1, q-1].

Security of the EC system based on the difficulty of calculating discrete logarithms in the subgroup generated by g. Discrete logarithms in this subgroup can be computed if discrete logarithms in a group of points on elliptic curves over the field until the (finite field) can be calculated. This issue is known as the elliptic curve discrete logarithm problem (Elliptic Curve discrete logarithm problem or ECDLP). There was no known method to solve ECDLP better than to solve problems in all the cyclic subgroups and combine the results. Therefore, the difficulty of ECDLP depends on the size of the largest prime divisor of the group order of points on elliptic curves (whose value is close to p). For this reason, the p, E and q are usually chosen such that the size of p and q close enough. Therefore, the security of EC systems depends on the size of q, and the size of the EC key refers to the bit length of the subgroup size q. The number of bits required to store the public key EC key length greater than q EC, because EC public key contains p, e, g and y [1].

B. Elliptic Curve Discrete logarithm problem (ECDLP):

To clarify the discussion of the elliptic curve discrete logarithm problem (ECDLP), review the definition of ECDLP are given points P and Q in the group, find a number k so that kP = Q.

An example used in the finite field is a Galois Field ECDLP (GF) of the polynomial, GF(2^m). Element of GF(2^m) is a polynomial of the form

$$a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \cdots + a_2x^2 + a_1x + a_0;\ a_i \in [0,1].$$

The coefficients $a_i$is an element of integers modulo 2. The elements of GF (2m) can be expressed as vectors of the form

$$(a_{m-1}, a_{m-2}, \dots, a_2, a_1, a_0).$$

In order for GF(2^m) can be defined in complete, irreducible polynomials need to be used. This polynomial of degree m irreducible and cannot be factored into polynomials of degree less than m with coefficients 0 or 1. The main coefficients must be equal to 1 i.e.

$$x^m + f_{m-1}x^{m-1} + f_{m-2}x^{m-2} + \cdots + f_2x^2 + f_1x + f_0;\ f_i \in [0,1].$$

For example, review the GF(2^4). We determine the irreducible polynomial of the form f (x) = x4 + x +1. In the sum of the elements of GF(2^4), the coefficients of the corresponding rank sum modulo 2 as shown as follows:

$$(x^2+1) + (x^3+x^2+x) = (1 mod 2)x^3 + (2 mod 2)x^2 + (1 mod 2)x + (1 mod 2) = x^3 + x + 1.$$

Reduction also applies to the same thing.

For the multiplication of elements of GF(2^4), the multiplication of these elements must be also an element in GF(2^4).

$$(x^2+1)(x^3+x^2+x) = (1 mod 2)x^5 + (1 mod 2)x^4 + (2 mod 2)x^3 + (1 mod 2)x^2 + (1 mod 2)x = x^5 + x^4 + x^2 + x.$$

It turned out that the above is not the result of multiplying the elements of GF(2^4) due to a greater degree than m-1 or 3. In order to store the product on top of being an element of GF(2^4) it is used to reduce the irreducible polynomial multiplication result, namely:

$$(x^5 + x^4 + x^2 + x) mod (x^4 + x + 1) = x^4.$$

How to make the last the same as that used in the way integers are:

$$(x^5 + x^4 + x^2 + x) = (x^4 + x + 1)x + x^4.$$

The general form for elliptic curves over GF (2m) is

$$y^2 + xy = x^3 + ax^2 + b.$$

Conditions must be met in order for an elliptic curve over not having a double root is $b \neq 0$.

Now we will calculate Q = 2P. The slope of the elliptic curve over GF(2^m) can be calculated as follows:

$$s = -(\partial F/\partial x)/(\partial F/\partial y) = -(3x^2 + 2ax - y)/(-2y - x)$$
$$= -(3x^2 + 2ax - y)/(-2y - x)$$
$$= -\frac{x^2 + 2x^2 + 2ax - y}{-2y - x}.$$

Since all the coefficients must be reduced modulo 2, is obtained

$$s = -(x^2 - y)/(-x)$$
$$= -(-x + \frac{y}{x})$$
$$= x - \frac{y}{x}.$$

Minus sign can be ignored because (-1mod2) = (1mod2) = 1 so the last equation becomes

$$s = x + \frac{y}{x}.$$

Negative from the point (x, y) is the point (x, x + y). The coordinates of the point q = 2P can be calculated using the following formula [3]:

$$s = x_p + \frac{y_p}{x_p}$$
$$x_Q = s^2 + s + a$$
$$y_Q = x_P + (s + 1)x_Q.$$

The most efficient method known to attack the EC system is a method of Pollard rho [1].

Because p and q are assumed the same order with respect [1], then the costs of operating comparable group $(\log_2 (q))^2$. From the estimates obtained there provided that the 109-bit EC system with p = 2109 takes 18,000 years to the PC (equivalent to one year at 18,000 PC or equivalent with 8 MMY [1].

In chapter IV, we will discuss how to determine the length of the key EC (Elliptic Curve) to formulate an inequality that contains parameters relating to the EC key length.

IV. DETERMINING THE EC (ELLIPTIC CURVE) KEY LENGTH

In this chapter we will discuss about how to determine the key length of a formula that consists of parameters that have been defined previously. EC system used is 109 bit EC system with p = 2109 [1].

Four main considerations used to determine the selection of a cryptographic key size, namely:

a. Life Span, the expectations which the information will be protected.
b. Security Margin, the level of an attack will succeed expressed in years (the year when the information is expected to be protected).
c. Computing environment, the expected changes of the computing resources available to the attacker.
d. Cryptanalysis, cryptanalysis developments that are expected to occur [1].

The formula to obtain the EC key length can be obtained by first understanding the assumptions and definitions used.

A. Some Definitions:

The following definitions are derived from [1].

a. Definition I. Margin security (security margin) s is defined as the year in which the user will trust the DES (Data Encryption Standard).Default setting I. The default setting for s is s = 1982.
b. Definitions II. The variable m> 0 is defined as the number of months it takes (on average) to increase the processor speed and double the memory size is doubled.Default setting II. The default setting for m is m = 18.
c. Definition III. The value t = 0 or t = 1 defines how m should be interpreted:If t = 1, means the amount of computing power and RAM are obtained by someone for every dollar expected to double every m months.If t = 0, means the amount of computing power and RAM is expected to double every m months, regardless of price.Default setting III. The default setting for t is t = 1.
d. Definition IV. Variable b> 0 is defined as the number of years in which the budget is expected attacker (attacker) to be doubled.Default setting IV. The default setting for b is b = 10.
e. Definition V. Number r> 0 is defined as the number of months (on average) which is expected to affect the development of cryptanalysis of classical asymmetric system be twice as effective, i.e., r months from now is expected that the attacks on classical asymmetric system requires half the cost of computing the current business this.Default settings V. The default setting for r is r = 18.
f. Definition VI. Number c≥ 0 is defined as the number of months (on average) which affect the development of EC cryptanalytic be twice as effective. If c = 0 means no cryptanalytic progress towards the expected EC.Default setting VI. The default setting for c is c = 0.
g. Definition VII. Number P> 0 is defined as the price of a PC in the U.S. dollar which is equipped with at least 64 MB of RAM. PC equipped with a 450 MHz Pentium II processor, motherboard & hardware communication. Default Setting VII. The default setting for P is P = 100.

B. Infeasible Number of Mips-Years (IMY):

For example, the key length must be determined to achieve the security margin at least until the year y. DES is required to solve $5*10^5$ Mips-Years is equal to 0.5 MMY [1]. MMY-Mips = Million Years. IMY computing offers a number of acceptable levels of security in the year s (Definition I). Based on the definition of I-IV, mean that in the year y, i.e. y-s years later, the number of computingIMY (y) = $5*10^5 * (2^{12(y-s)/m}) * (2^{t(y-s)/b})$ Mips-Years offers an acceptable level of security.IMY (y) means that infeasible number of Mips-Years for the year y.

Factors that affect the IMY are described as follows.

a. Factor of $2^{12(y-s)/m}$ based on the expectations of the speed s processors in the year s to year y (Definition I and II), and
b. Factor of $2^{t(y-s)/b}$ reflects the expectations of an increase in budget attacker (Definition I, III & IV).

Value IMY (y) is generated is used to obtain / lower key sizes that offer an acceptable level of security until the year y.

C. Run Time Convention:

Computing resources (computing power) are measured in Mips-Years is defined as the amount of computation that can be done in one year by a DEC computer VAX 11/780.

Conversion used in the calculation is one year equivalent PC with 450 Mips-Years. PC here meant PC with a Pentium II 450MHz [1].

D. Inequality of EC (Elliptic Curve)Key Length:

If the EC key length is selected such that

$$\frac{2^{u/2} * u^2}{IMY(y) * C} \geq \frac{2^{109/2} * 109^2}{2{,}2 * 10^6}$$

where

C = 1 (if c = 0 then be taken C = 1, while if≠c0 then take $C = 2^{12(y-1999)/c}$)

y = 2010

t = 1

s = 1982 (default)

b = 10 (default)

m = 18 (default)

$$IMY(y) = 5 * 10^5 * 2^{12(y-s)/m} * 2^{t(y-s)/b} \ Mips - Years$$

$$IMY(2010) = 5 * 10^5 * 2^{12(2010-1982)/18} * 2^{1(2010-1982)/10} \ Mips - Years$$

$$= 5 * 10^5 * 416.127{,}66146 * 6{,}96440451 \ Mips - Years$$

$$= 1.449.040.680.348{,}40739 \ Mips - Years$$

Then the security of EC systems until at least y / minimal "cost and computationally equivalent" to the security offered by DES in year s [1].

By substituting the value of IMY (2010) and C = 1 (assuming c = 0) to the inequality, we obtain

$$2^{u/2} * u^2 \geq 1{,}99363503351803476 * 10^{26}.$$

The last inequality is very difficult to solve analytically, so we use Maple to solve it.

Calculation of IMY and all coefficients on the inequality of EC key length are performed using spreadsheet software.

Then we obtain the solution

u1 = 145.9728390 and u2 = 145.9934650 +17.44293557 i (complex numbers).

So the derived EC key length for the year 2010 amounted to 146.

For 2015 and 2050, the key length is obtained by the same manner.

EC key length data for the years 2010 to 2050 could be seen in the Table 2.

Table 2. TheEC Key Length for c=0

Year (x)	EC Key Length (y) (c=0)
2010	146
2015	154
2020	161
2025	169
2030	176
2035	184
2040	191
2045	198
2050	206

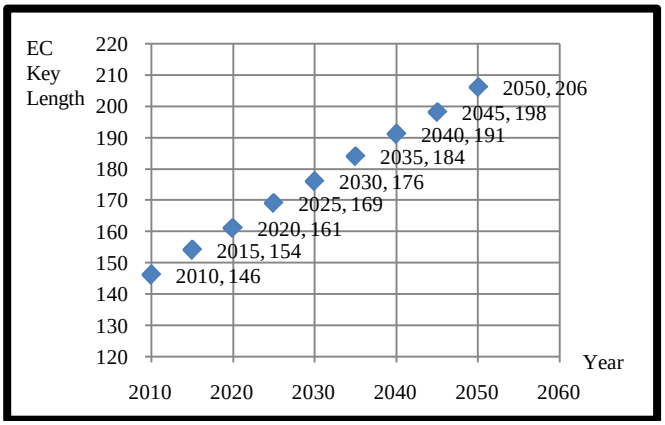

Figure 1. The Graph of EC Key Length for c=0

So thatthe calculationbecomesmuch simpler, you can use linearinterpolation. Takethe points(2010.146) and (2050, 206)andsubsequentlydeterminedtheequation ofa linethrough these two points.

As an example, the gradient line through the two points is g, then g = (206-146) / (2050-2010) or g = 60/40 or g = 1.5.

Equation of the line is y-146 = 1.5 (x-2010) or y = 1.5 (x-2010) +146.

So for 2050, the derived key length y = 1.5 (2050-2010) or y = 206 +146.

In the above explanation it is assumed that no development of cryptanalysis, so we can take c = 0. If it is assumed that the development of cryptanalysis to be twice as effective in 18 months, then take c = 18.

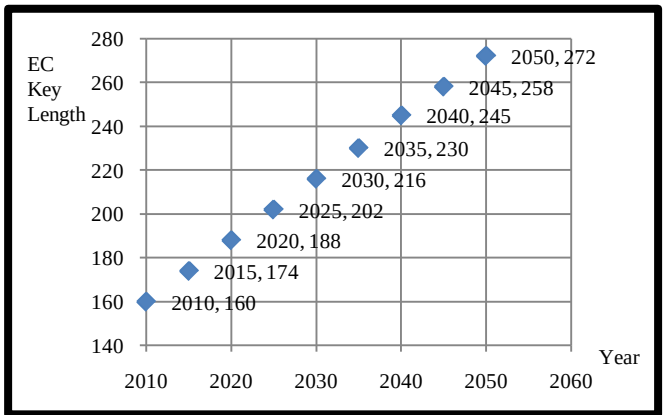

Figure 2. The Graph of EC Key Length for c=18

To the assumption c = 18 months, the calculations performed in the same manner and the results of the calculations can be seen in the Figure 2.

Table 3. The EC Key Length for c=18.

Year (x)	EC Key Length (y) (c=18)
2010	160
2015	174
2020	188
2025	202
2030	216
2035	230
2040	245
2045	258
2050	272

V. CONCLUSION

For 2010,the EC key length calculation method proposed by Lenstra/Verheul in [1] generate EC key length equal to the EC key length recommended by NIST (National Institute of Standards and Technology) [5].

For 2050, the methods of Lenstra/Verheul generate EC key length 272, while the NIST recommended EC key length 384, as can be seen in [5].

The results of calculation of EC key length for c = 0 and c = 18 can be seen in Table 4.

Table 4. Comparison of EC Key Length for c = 0 and c = 18

Year (x)	EC Key Length (y) (c=0)	EC Key Length (y) (c=18)
2010	146	160
2015	154	174
2020	161	188
2025	169	202
2030	176	216
2035	184	230
2040	191	245
2045	198	258
2050	206	272

VI. REFERENCES

[1] A. K. Lenstra,E. R Verheul, "Selecting Cryptographic Key Sizes", Journal of Cryptology VOL. 14 No. 4, 255-293, Sringer-Verlag, 2001.

[2] A. K. Lenstra, "Key Lengths: Contribution of The Handbook of Information Security", Lucent Technologies and Technische Universiteit Eindhoven, NJ USA, 2004.

[3] V. Hassler, Security Fundamentals for E-Commerce, Artech House, Inc., 2001.

[4] A. Janko, "How to Apply Strong Cryptographic Controls", ISSA (Information Systems Security Association) Journal, 2006.

[5] http://www.keylength.com, accessed on July 2012.

[6] W. Stalling, Cryptography and Network Security Principles and Practices, 5th Edition, Pearson Education, Inc., USA, 2011.

[7] R. Soram, M. Khomdram, "Juxtaposition of RSA and Elliptic Curve Cryptosystem", IJCSNS International Journal of Computer Science and Network Security VOL. 9 No.9, September 2009.

[8] R. Shanmugalakshmi, M. Prabu, "Research Issues on Elliptic Curve Cryptography and Its applications", IJCSNS International Journal of Computer Science and Network Security VOL.9 No.6, June 2009.

[9] F. Rodriguez-Henriquez, N. A. Saqib, A. Diaz-Perez, "Ultra Fast Parallel Implementation of Elliptic Curve Point Multiplication over GF(2^m)", Elsevier Science, 2003.