ISSN:0975-9646

Navpreet Kaur et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4) , 2014, 5404-5407

Authenticated Diffie-Hellman Key Exchange Algorithm

Navpreet Kaur[1], Ritu Nagpal[2]
[1]*M.Tech(CSE),GJUS&T(Hisar)*
[2]*Asst. Professor, Dept. of Computer Science & Engg., GJUS&T(Hisar)*

Abstract-The ability to distribute cryptographic keys has been a challenge for centuries. The Diffie-Hellman was the first practical solution to the problem .However, if the key exchange takes place in certain mathematical environments, the key exchange become vulnerable to a specific Man-in-Middle attack, first observed by Vanstone. This paper is an effort to solve a serious problem in Diffie-Hellman key exchange, that is, Man-in-Middle attack. In this paper we have used RSA algorithm along with Diffie-Hellman to solve the problem. We explore the Man-in-Middle attack, analyse the countermeasures against the attack.

Index Terms-Cryptography, Diffie-Hellman, Man-in-Middle attack, primality testing.

INTRODUCTION

Cryptography and encryption/decryption methods fall into two broad categories- symmetric and public key. In symmetric cryptography, sometimes called classical cryptography, parties share the same encryption/decryption key. Therefore, before using a symmetric cryptography system, the users must somehow come to an agreement on a key to use. An obvious problem arises when the parties are separated by large distances which is commonplace in today's worldwide digital communications. If the parties did not meet prior to their separation, how do they agree on the common key to use in their crypto system without a secure channel? They could send a trusted courier to exchange keys, but that is not feasible, if time is a critical factor in their communication.

The problem of securely distributing keys used in symmetric ciphers has challenged cryptographers for hundreds of years. If an unauthorized user gains access to the key, the cryptographic communication must be considered broken. Amazingly, in 1977, Whitfield Diffie and Martin Hellman published a paper in which they presented a key exchange protocol that provided the first practical solution to this dilemma. The protocol, named the Diffie-Hellman key exchange (or key agreement) protocol in their honour, allows two parties to derive a common secret key by communications over an unsecured channel, while sharing no secret keying material at prior .

Before conducting the key exchange using the Diffie-Hellman protocol, the parties must agree on a prime number that defines the mathematical environment in which the key exchange will take place. If the prime number is large enough, a brute force attack to find the secret key becomes infeasible. However, if the two parties agree on certain prime numbers, an active adversary can compromise their communication.

This paper investigates the Diffie-Hellman protocol and the difficulty of the discrete logarithm problem the protocol relies on. We then analyse the man-in-middle attack described above by developing an algorithm to conduct the attack. We then consider methods to defend against the attack and demonstrate their effectiveness.

BACKGROUND AND REVIEW

Before beginning a discussion of the Diffie-Hellman protocol and the man-in-middle attack, we investigate and present some basic definitions and theorems. This information is available in any standard algebra text, such as Fraleigh's Abstract Algebra, or discrete mathematics text, such as Rosen's Discrete Mathematics and Its Applications. It is assumed the reader is familiar with common mathematical, logical, and set notation.

NUMBER THEORY

If a and b are integers and $a \neq 0$, we say that a divides b if there is an integer c such that b = ac . When a divides b we say that a is a factor of b and that b is a multiple of a. The notation $a \backslash b$ denotes a divides b . Given two integers a and b, both non-zero, the largest integer d such that d\a and d\b is called the greatest common divisor of a and b. The greatest common divisor of a and b is denoted by *gcd(a,b)*. The integers a and b are relatively prime, if their greatest common divisor is one.

Every positive integer greater than one is divisible by at least two integers, itself and 1. If these are its only factors, integer is called a prime. A positive integer that is greater than 1, and not prime, is called composite. The Fundamental Theorem of Arithmetic states that every positive integer greater than one can be written uniquely as a product of two of more primes, where the prime factors are written in order of non-decreasing size. Given a positive integer, n , let the prime factorization of n be denoted by

$$n=\prod_{i=1}^{k} P_i^{ai}$$

In some situations, we care only about the remainder of an integer when it is divided by some specified positive integer, denoted by m. If a and b are integers, then a is congruent to b modulo m if m divides a -b. We use the notation $a \equiv b \pmod m$ to indicate that a is congruent to b modulo m.

The great French mathematician *Pierre de Fermat (1601–1655)* demonstrated that the congruence $a^{p-1} = 1 \pmod p$ holds when p is a prime, and this gives us a theorem that

will prove crucial in our analysis of the man-in-the-middle attack.

Fermat's Theorem: If $a \epsilon Z$ and p is a prime not dividing a, then p divides a^{p-1}, that is,

$$a^{p-1} \equiv 1 \pmod p .$$

Euler gave a generalization of Fermat's theorem, but we must first define Euler's Totient Function. Commonly referred to as Euler's Phi Function, the function gives the number of integers less than or equal to n which are relatively prime to n , and is denoted by $\phi(n)$. It is not hard to show that, if

$$n=\prod_{i=1}^{k} P_i^{\alpha i} \quad , then$$

$$\phi(n)=n \prod_{i=1}^{k} P_i^{\alpha i} (1-1/P_i)$$

Euler's Theorem: If $a \epsilon Z$ and is relatively prime to n, then $a^{\phi(n)}- 1$ is divisible by n, that is,

$$a^{\phi(n)} \equiv 1 \pmod n.$$

GROUP THEORY

A **group** <G,*> is a set *G*, closed under a binary operation *, such that the following axioms are satisfied:

Associativity: For all $a, b, c \epsilon G$ $a*(b*c)=(a*b)*c$

Identity: There is an element *e* in *G* such that *for all* $a \epsilon Z$

$a*e=e*a=a$

Inverse: Corresponding to each $a \epsilon G$, there exist an element a' such that

$a*a'=a'*a=e$

If the set *G* has a finite number of elements. In this case, the number of elements is called the ***order*** of *G*, denoted by $|G|$.

If *n* is a prime *p*, then the set $Z^*_p=Z_p-\{[0]_p\}$ forms a group under multiplication modulo *n* . It is a necessary requirement to remove the zero class because zero has no inverse under multiplication. $<Z^*_p,.>$,is the multiplicative group of the set of congruence classes of prime integers. The Diffie-Hellman key exchange protocol sets this group as the environment for the key agreement.

Sub Group

Let G be a group and H be its subset. The subset H is called subgroup of G if following conditions are satisfied.

1. if a, b ϵ H, the product ab also belongs to H.
2. e(identity element of G) belongs to H.
3.if a ϵ H its inverse also belongs to H.

Lagrange's Theorem: *Let H be a subgroup of a finite group G.*

Then the order of H is a divisor of the order of G.

This powerful theorem makes the attack we will analyse later possible.

We know the order of $<Z^*_p,.>$ is *p* -1. The two properties mentioned above tell us that any subgroup of $<Z^*_p,.>$ will also be cyclic and the order of the subgroup will be a divisor of *p* -1.

Groups of Prime Order

A group g is called a group of prime order if it is:

1. A cyclic group having a prime number as its order.
2. isomorphic to the quotient of group of integers by a subgroup generated by aprime.
3. a simple abelian group.
4. adittive group of finite prime field.

The number of distinct subgroups of a group are either 0 or congruent to 1(mod p).

Let G be a group and H, K be its subgroups each of order p, where p is a prime then,

$H \cap K = \{e\}$ or $H=K$.

PRIMALITY TESTING

Suppose a large integer is given. Sometimes we want to determine whether the number is composite or prime. A ***primality test*** is an algorithm used for this, that is, for determining whether an input number is prime. Primality tests can be deterministic and probabilistic. Deterministic primality tests prove with certainty whether a number is prime or composite. Probabilistic primality tests tell us a number is composite or *probably* prime.

Miller-Rabin Primality Test

The Miller-Rabin Primality Test is an efficient probabilistic algorithm to test for primality based on the idea of strong pseudoprimes. Consider an odd composite number *n* and $n -1 = d .2^s$ with *d* odd. *n* is a ***strong pseudoprime*** if either $a^d \equiv 1 \pmod n$ or $a^{d.2r} \equiv -1 \pmod n$ with $r = 0,1,...s -1$. The Carmichael numbers are Fermat pseudoprimes for every base. However, a composite number can only be a strong pseudoprime to at most one quarter of all bases.

The algorithm is as follows:

Choose a random integer $a \epsilon [2, n - 2]$. *If* $a^d \neq 1 \pmod n$ *and* $a^{d.2r} \neq -1 \pmod n$ *for all* $r = 0,1,...s -1$, *then a is called a witness and n is composite*. Otherwise, *n* is a strong probable prime to base *a*. If $n > 9$ and is odd composite, the probability that the algorithm will fail to produce a witness for *n* is $<1/ 4$. The probability that we fail to find a witness after *k* iterations is $<1/ 4^k$. We can make this probability as small as we desire with a large number of iterations. For instance, if we wanted to ensure the probability of calling a composite number a prime is less than 10^{-6}, we must compute 10 iterations or more.

The Miller-Rabin test is very fast and has a complexity of $O((\log n)^3)$. Of course, because it is probabilistic, there is a chance of the test returning a number as prime when it is in fact composite. The Miller-Rabin test offers us both speed, as compared to other primality tests, and the ability to control the probability of error and will be our tool of choice.

DIFFIE-HELLMAN AND THE DISCRETE LOGARITHM

THE DIFFIE-HELLMAN PROTOCOL

The Diffie-Hellman protocol provided the first practical solution to the key distribution problem, allowing two parties, never having met in advance or shared keying material, to establish a shared secret by exchanging messages over an open channel. The key can then be used to encrypt subsequent communications using a symmetric key cipher. The security rests on the intractability of the Diffie-Hellman problem and the related problem of

computing discrete logarithms. We will call the two parties conducting the key exchange "Alice" and "Bob."

Protocol steps:

1. A prime number p and generator α *of* $\mathbf{Z}^*$ p $(2 \leq \alpha \leq p - 2)$ are selected and published.
2. Alice chooses a random secret x, $1 \leq x \leq p - 2$, and sends Bob $\alpha^x \bmod p$.
$A \rightarrow B$: $\alpha^x \bmod p$
3. Bob chooses a random secret y, $1 \leq y \leq p - 2$, and sends Alice $\alpha^y \bmod p$.
$B \rightarrow A$: $\alpha^y \bmod p$
4. Bob receives α^x and computes the shared key as $K = (\alpha^x)^y \bmod p$.
5. Alice receives α^y and computes the shared key as $K = (\alpha^y)^x \bmod p$.

if the prime number used is large enough, no computing power available today can exhaust the key space. For instance, most applications recommend 1024-bit primes. This correlates to a number of about 300 digits and makes searching the key space one by one infeasible.

THE DISCRETE LOGARITHM

Eve has more information than just the fact that the key resides in the interval (1, p -1). Because the exchange occurs over an open channel, Eve knows α^x and α^y as well. If $\beta \equiv \alpha^x \pmod p$ and $\gamma \equiv \alpha^y \pmod p$, then p ,α ,β and γ are known. All Eve has to do is solve $\alpha^x \equiv \beta \pmod p$ for x or $\alpha^y \equiv \gamma \pmod p$ for y. Once x or y are known, Eve simply raises α^x to y or α^y to x and arrives at the secret key K . However, if p is large, solving $\alpha^x \equiv \beta \pmod p$ for x in general is considered difficult. The problem of finding x if α^x is known as the ***discrete logarithm problem*** (DLP), often abbreviated $x = L_\alpha(\beta)$.

The difficulty of solving the DLP yields useful cryptosystems. Diffie-Hellman key exchange protocol, El Gamal encryption system, and the Digital Signature Algorithm all rely on the difficulty of solving the DLP. In 2005, a 168 digit prime (556 bits) discrete logarithm was computed, setting a record at that time. The record factorization up to then was 200 digits (663 bits).

MAN-IN-THE-MIDDLE ATTACK

THEORY BEHIND THE ATTACK

Wiener and Van Orschot noted that, if certain primes are used, a potentially fatal protocol attack on the Diffie-Hellman key exchange protocol becomes possible. The idea is based on forcing the parties to agree on a shared key that resides in a subgroup of the cyclic group Z^*_p. If the order of the subgroup is small enough, an adversary can exhaustively search the subgroup, retrieve the secret key, and eavesdrop on the communication of Alice and Bob.

For instance, consider the case when the prime used for the key exchange is of the form $p = 2q + 1$, where q is a prime. Then, $\alpha^q = \alpha^{(p-1)/2}$.

Claim: $\alpha^{(p-1)/2}$ is an element of order two.

Proof: By Fermat's little theorem, $\alpha^{p-1} = 1 \bmod p$. So $\alpha^{(p-1)/2}$ must be +1 or -1. But if $\alpha^{(p-1)/2} = 1$ then α must have order $(p - 1)/2$. This is a contradiction, because α is a primitive root of Z^*_p and must be of order p -1. So $\alpha^{(p-1)/2} = -1$ and is an element of order two.

If Alice and Bob respectively send each other unauthenticated messages and α^x, and α^y active intruder may substitute $(\alpha^x)^q$ for the first, and $(\alpha^y)^q$ for the second. When Alice receives $(\alpha^y)^q$ and computes $(\alpha^{qy})^x$ and when Bob receives $(\alpha^x)^q$ and computes $(\alpha^{qx})^y$, they will arrive at only one of two possible values, +1 and -1. The intruder can then try both possible keys and gain access to Alice and Bob's secret communications. Obviously, if Alice and Bob demonstrate vigilance, they will agree in advance to suspect any key agreement that arrives at +1 or -1.

We can generalize the situation if Alice and Bob use a prime number of the form $p = Rq + 1$, where R is a small integer and q is again a large prime.

Claim: $\alpha^{(p-1)/R}$ *is an element of order* R.

Proof: Raising $\alpha^{(p-1)/R}$ to consecutive powers, starting with 0, we get:
$(\alpha^{(p-1)/R}) = 1, (\alpha^{(p-1)/R})^2, (\alpha^{(p-1)/R})^3, \ldots, (\alpha^{(p-1)/R}) = \alpha^{p-1} = 1$

This produces a list of R different values. Continuing after R ,
$(\alpha^{(p-1)/R})^{(R+1)} = (\alpha^{(p-1)/R})^R . (\alpha^{(p-1)/R}) = 1.(\alpha^{(p-1)/R})$,
$(\alpha^{(p-1)/R})^{(R+2)} = (\alpha^{(p-1)/R})^R . (\alpha^{(p-1)/R})^2 = 1. (\alpha^{(p-1)/R})^2, \ldots\ldots,$
$(\alpha^{(p-1)/R})^{(R+n)} = (\alpha^{(p-1)/R})^R . (\alpha^{(p-1)/R})^n = 1.(\alpha^{(p-1)/R})^n$

For $n < R$, the results are in the original list.

For $n \geq R$, we can write $R + n = R + kR + m$ *with* $0 \leq m \leq R - 1$ *and* $m, k \in \mathbf{Z}$.
$(\alpha^{(p-1)/R})^{(R+n)} = (\alpha^{(p-1)/R})^{(R+kR+m)} = (\alpha^{(p-1)/R})^R . (\alpha^{(p-1)/R})^{kR} . (\alpha^{(p-1)/R})^m =$
$1.1^k . (\alpha^{(p-1)/R})^m = (\alpha^{(p-1)/R})^m$

Because $0 \leq m \leq R - 1$, this is in our original list and $\alpha^{(p-1)/R}$ is of order R .

So, if the prime Alice and Bob agree to use is of the form $p = Rq + 1$, Eve can force them to agree on a key in a subgroup of Z^*_p of order R by replacing α^x and α^y with $(\alpha^x)^q$ and $(\alpha^y)^q$. Even if Alice and Bob are vigilant, the key can take any of R values and the generalized attack poses a significant threat to an unauthenticated key exchange using the Diffie-Hellman protocol.

COUNTERMEASURES AGAINST THE ATTACK

To prevent this potentially fatal protocol attack, Alice and Bob have several options. The easiest method is to force authentication prior to the key exchange.

Authentication

The attack we have discussed is not the only man-in-the-middle attack Diffie -Hellman is vulnerable to. The

Appendix details another attack, if no authentication occurs prior to the key exchange. To combat these attacks, a variation of Diffie-Hellman that ensures authentication can be used. An example of such a variation is the Station-to-Station protocol (STS). STS is a three-pass variation of the basic Diffie-Hellman protocol that allows the establishment of a shared secret key between two parties with mutual entity authentication and mutual explicit key authentication. The STS employs digital signatures. A digital signature of a message is a number dependent on some secret known only to the signer; and, additionally, on the content of the message being signed. The STS protocol is frequently employed with the RSA signature scheme.

To employ an RSA signature scheme, public and private key pairs must first be generated.

RSA signature scheme key generation steps:

1. Generate two large distinct random primes p and q, each roughly the same size
2. Compute $n = pq$ and $\phi = (p-1)(q-1)$
3. Select a random integer e, $1 < e < \phi$, such that $\gcd(e,\phi) = 1$
4. Use the extended Euclidean algorithm to compute the unique integer d, $1 < d < \phi$ such that $ed = 1 \pmod{\phi}$
5. The user's public key is (n, e) and the user's private key is d

Each user should generate a public and private key

If we let E denote a symmetric encryption algorithm, and $S_A(m)$ denote Alice's signature on m, the protocol is as follows

Set up:

a. A prime number p and generator α of $Z^*_p (2 \le \alpha \le p-2)$ are selected and published
b. Alice selects RSA public and private signature keys (n_A, e_A) and d_A (Bob selects analogous keys). Assume each party has access to authentic copies of the other's public key.

Actions:

a. Alice generates a secret random x, $1 \le x \le p-2$, encrypts the message with its signatures and sends to Bob $E_{SA}(a^x \bmod p)$.
   $A \rightarrow B : E_{SA}(a^x \bmod p)$.
b. Bob decryptes message using public key of Alice.
c. Bob generates a secret random y, $1 \le y \le p-2$, and computes the shared key $K=(a^x)^y \bmod p$. Bob encrypts the message using its signatures and sends to Alice $E_{SB}(a^y \bmod p)$.
   $B \rightarrow A: E_{SB}(a^y \bmod p)$
d. Alice computes the shared key $k = (a^y)^x \bmod p$, decrypts the encrypted data, and uses Bob's public key to verify the received value as the signature on the hash of the clear text
   Upon successful verification, Alice and Bob accepts k that is actually shared with Bob, and sends Bob an analogous message.

Eve cannot alter the original exponentials without triggering a failure during Alice and Bob's key agreement. This precludes the man-in-middle attack we have focused on and defends Alice and Bob's key exchange against several other possible active man-in-middle attacks.

CONCLUSIONS AND FUTURE WORK

This thesis investigated and analyzed a particular man-in-the-middle attack on the Diffie-Hellman key exchange protocol. We created an algorithm to carry out the attack and demonstrated how it is constrained by the primality test used by the attacker. In particular, if the Miller-Rabin primality test is used, the algorithm's complexity is $O((\log N)^3)$ with N being the input prime number. We showed that prime numbers of the form $p=Rq+1$ with R bounded are common with small primes but become increasingly rare as larger numbers are considered. In fact, with low bit primes such as 128 bits, a reasonably-sized R will give an attacker a good chance of the prime being of the desired form. However, when large primes such as 1024 and 2048 bits are considered, a very large value of R is required to give an attacker a reasonable chance of conducting the attack. We demonstrated how two techniques, authentication and prime order subgroups, can prevent the attack. In fact, it appears industry has begun to adopt the prime order subgroup technique to defend against the attack. It is possible that analyzing the given prime number, capturing the required messages, altering those messages, and forwarding the messages to the intended recipients will be too time-consuming. This would obviously alert the parties of possible compromise. In addition, it may be possible to alter the attack to compromise communications that are authenticated and render several Diffie-Hellman variants such as the STS protocol vulnerable.

REFERENCES

[1] A. J. Menezes, P. C. van Oorshot, and S. A Vanstone, *Handbook of Applied Cryptography*. CRC Press, New York, New York, 1997.
[2] P. C. van Oorschot and M. J. Wiener, On Diffie-Hellman Key Agreement with Short Exponents. *EUROCRYPT'96*, LNCS 1070, Springer-Verlag, 1996, pp. 332–343.
[3] S. Singh, *The Code Book*. Doubleday, 1999.
[4] J. B. Fraleigh, *A First Course in Abstract Algebra.* Addison-Wesley, San Francisco, CA, 7th Edition, 2002.
[5] K. H. Rosen, *Discrete Mathematics and Its Applications*. McGraw Hill, San Francisco, CA, 6th Edition, 2007.
[6] R. Crandall and C. Pomerance, *Prime Numbers: A Computational Perspective*. Springer, New York, NY, 2001.
[7] A. L. Atkin and F. Morain, Elliptic Curves and Primality Proving. *Res. Rep.* 1256, INRIA, June 1990.
[8] M. Agrawal, N. Kayal, N. Saxena, PRIMES is in P. *Annals of Mathematics 160.* 2004.
[9] C. Pomerance and H.W. Lenstra, Primality testing with Gaussian periods, preprint.
[10] W. Diffie and M. E. Hellman, *New Directions in Cryptography.* IEEE IT- 22, 1976, pp. 644–654.
[11] W. Trappe and L. Washington, *Introduction to Cryptography with Coding Theory.* Pearson, Upper Saddle River, NJ, 2nd Edition, 2006.
[12] S. Pohlig and M. Hellman, An Improved Algorithm for Computing Logarithms over GF(p) and its Cryptographic Significance. *IEEE Transactions on Information Theory,* 24, 1978, pp. 106–110.
[13] T. Agoh, On Sophie Germain Primes. *Tatra Mt. Math.* Publ. 20, 2000.
[14] P. Stanica, private communication, 2009.
[15] Internet Engineering Task Force (IETF) Request for Comment (RFC) 2631, June 1999.