

Weak Keys in Diffie-Hellman Protocol

Aniket P. Kate (04305001) Prajakta S. Kalekar (04329008)
Deepti Agrawal (04329020)
Indian Institute of Technology, Powai, Mumbai -400076

November 15, 2004

Contents

1	Introduction	3
2	Emergence of public key system	3
3	Diffie-Hellman Protocol	3
4	Basics of Abstract Algebra	3
5	Known attacks on DHP	5
5.1	Trivial Attacks	5
5.2	Subgroup Confinement Attack	5
5.3	Composite Order Subgroup Attack	6
6	DHP over general linear group	6
6.1	Modulus condition and solution of the DHP	7
6.2	Conjugate class and the Diffie-Hellman conjecture	8
7	DHP over general extension field	10
8	DHP over elliptic curves	10
9	Conclusion	10
10	Our Observations	10
10.1	Weak key in Diffie Hellman	10
10.2	Comparison between the Composite Order Subgroup attack and the Man-in-the-middle attack	11
10.3	Vulnerabilities in DHP over extension fields	11
11	Appendix	11
11.1	Fermat's Little Theorem	11
11.2	Discrete Logarithm Problem	11
11.3	Safe Primes	11

Abstract

The Diffie-Hellman key agreement protocol (also called exponential key agreement) was developed by Diffie and Hellman in 1976 and published in the ground-breaking paper “New Directions in Cryptography”. The protocol allows two users to exchange a secret key over an insecure medium without any prior secrets. The protocol is vulnerable to various kinds of attacks, some of which have been discussed in this report.

1 Introduction

The success of the Diffie-Hellman protocol lies in the difficulty of solving Discrete log Problem (Refer to 11.2). According to the Diffie-Hellman Conjecture, the Diffie-Hellman Problem (DHP) of finding the shared secret key reduces to solving the DLP, which is a sub-exponential problem.

2 Emergence of public key system

3 Diffie-Hellman Protocol

The protocol has two system parameters p and g . They are both public and may be used by all the users in a system. Parameter p is a prime number and parameter g (usually called a generator) is an integer less than p , which is capable of generating every element from 1 to $p-1$ when multiplied by itself a certain number of times, modulo the prime p . Suppose that Alice and Bob want to agree on a shared secret key using the Diffie-Hellman key agreement protocol. They proceed as follows:

1. First, Alice generates a random private value a and Bob generates a random private value b .
2. Then they derive their public values using parameters p and g and their private values.
3. Alice's public value is $g^a \bmod p$ and Bob's public value is $g^b \bmod p$. They then exchange their public values.
4. Finally, Alice computes $k(ab) = (g^b)^a \bmod p$, and Bob computes $k(ba) = (g^a)^b \bmod p$. Since $k(ab) = k(ba) = k$, Alice and Bob now have a shared secret key k .

4 Basics of Abstract Algebra

Abstract algebra Abstract algebra is the set of advanced topics of algebra that deal with abstract algebraic structures rather than the usual number systems. The most important of these structures are groups, rings, and fields.

Group A group G is a finite or infinite set of elements together with a binary operation (called the group operation) that together satisfy the four fundamental properties of closure, associativity, the identity property, and the inverse property. The operation with respect to which a group is defined is often called the “group operation,” and a set is said to be a group “under” this operation. Elements A, B, C, \dots with binary operation between A and B denoted AB form a group if

1. Closure: If A and B are two elements in G , then the product AB is also in G .
2. Associativity: The defined multiplication is associative
3. Identity: There is an identity element I (a.k.a. $1, E$, or e)
4. Inverse: There must be an inverse or reciprocal of each element. Therefore, the set

must contain an element such that for each element of G .

A group is a monoid each of whose elements is invertible. **Example:** The set of integers under usual addition is a group.

Ring A ring (in the mathematical sense) is a set S together with two binary operators and (commonly interpreted as addition and multiplication, respectively) satisfying the following conditions:

1. Additive associativity
2. Additive commutativity
3. Additive identity
4. Additive inverse
5. Multiplicative associativity
6. Left and right distributivity

Example: The set of all n -square matrices over real numbers is a ring R with respect to addition and multiplication.

Field A field is any set of elements that satisfies the field axioms for both addition and multiplication and is a commutative division algebra. A field with a finite number of members is known as a *finite field or Galois field*. Because the identity condition is generally required to be different for addition and multiplication, every field must have at least two elements. Examples include the complex numbers, rational numbers, and real numbers, but not the integers, which form only a ring. **Examples** include the complex numbers, rational numbers, and real numbers, but not integers, which form only a ring.

Irreducible Polynomial A polynomial is said to be irreducible if it cannot be factored into nontrivial polynomials over the same field. **For Example,** in the finite field $GF(2)$, $1 + x + x^2$ is irreducible, whereas $1 + x^2$ is not, as it can be factored as $(1 + x)(1 + x)$

Extension Field A field K is said to be an extension field (or field extension, or extension), denoted K/F of a field F if F is a subfield of K . **For Example,** the complex numbers are an extension field of the real numbers, and the real numbers are an extension field of the rational numbers.

General Linear Group In abstract algebra, the general linear group of degree n over a field F (written as $GL(n, F)$) is the group of n -by- n invertible matrices with entries from F , with the group operation that of ordinary matrix multiplication. (This is indeed a group because the product of two invertible matrices is again invertible.) If F is a finite field of order q , then we sometimes write $GL(n, q)$ instead of $GL(n, F)$. If the field is R (the real numbers) or C (the complex numbers), the field is sometimes omitted when it is clear from the context, and we write $GL(n)$. $GL(n, F)$ and subgroups of $GL(n, F)$ are important in the development of group representations, and also arise in the study of spatial symmetries and symmetries of vector spaces in general, as well as the study of polynomials. If F is a finite field with q elements, then $GL(n, F)$ is a finite group with $(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{n-1})$ elements. This can be shown by counting the possible columns of the matrix: the first column can be anything but the zero column; the second column can be anything but the multiples of the first column, etc.

Minimal Polynomial Minimal polynomial of a matrix is the polynomial in of smallest degree n such that

$$p(A) = \sum_0^n c_i A_i = 0 \tag{1}$$

The minimal polynomial divides any polynomial q with $q(A)=0$, in particular, it divides the characteristic polynomial. If the characteristic polynomial factors as

$$\text{char}(A)(x) = (x - \lambda_1)^{n_1} \dots (x - \lambda_k)^{n_k} \quad (2)$$

then its minimal polynomial is

$$p(x) = (x - \lambda_1)^{m_1} \dots (x - \lambda_k)^{m_k} \quad (3)$$

where $1 \leq m_i \leq n_i$

For Example, let

$$\mathbf{A} = \begin{pmatrix} 1 & 51 \\ 1 & 1 \end{pmatrix} \quad (4)$$

Then its minimal polynomial is $x^2 + 51x + 3$

5 Known attacks on DHP

5.1 Trivial Attacks

Simple Exponents If one of k and l can be easily determined, the protocol can be broken.
For Example

- **$k = 1$ or $l = 1$**
If k or l equals 1 then $g^k = g$ or $g^l = g$ which any observant attacker will be able to detect. Thus k and l should not be chosen as 1.
- **$k = p-1$ or $l = p-1$**
If k or l equals $p-1$ then $g^{p-1} = 1$ (Refer to 11.1) and hence the shared secret key will be equal to 1, which any observant attacker will be able to detect. Thus k and l should not be chosen as $p-1$.

Simple Substitution Attack The attack is mounted as follows :

- Oscar intercepts g^x and g^y and replaces them with 1.
- Both Alice and Bob compute the same shared secret key which equals one.
If the computer program does not realize that g^x , g^y and g^{xy} cannot equal 1, the protocol is vulnerable.

5.2 Subgroup Confinement Attack

The generator g in the Diffie Hellman protocol is a primitive root of the prime p , i.e. the order of the group generated by g is equal to $p-1$. If the selected prime p is such that $p-1$ has several small prime factors, then some values between 1 and $p-1$ do not generate groups of order $p-1$, but of subgroups of smaller orders. Hence, within the group of order $p-1$ there are subgroups of smaller orders. If the public parameter of either Alice or Bob lies within one of these small subgroups, then the shared secret key would be confined to that subgroup. If the order of the

subgroup is small enough, the intruder may launch a brute force attack to determine the exact value of the shared secret key.

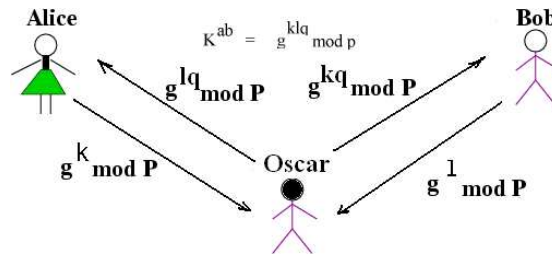
Example : Let $p = 19$ and $g = 2$
 Then the group generated by g is
 $(2, 4, 8, 16, 13, 7, 14, 9, 18, 17, 15, 11, 3, 6, 12, 5, 10, 1)$
 Now, Let $k = 2$, $A = 2^2 = 4$
 Subgroup generated by $A = S_A = (4, 16, 7, 9, 17, 11, 6, 5, 1)$
 Let $l = 3$, $B = 2^3 = 8$
 Sub-group generated by $B = S_B = (8, 7, 18, 11, 12, 1)$
 $K_{ab} = 2^6 \text{ mod } 19 = 7$

It can be clearly seen from the example above that the shared secret key K_{ab} lies in the intersection of the subgroups generated by k and l .

The **Solution** to counter this kind of an attack is to choose a **Safe Prime** (Refer to Safe Primes11.3)

5.3 Composite Order Subgroup Attack

Even if the selected prime is a Safe Prime, there exists a small subgroup of order 2. The attacker can exploit this fact to launch an attack as follows. Oscar can intercept the messages g^k and g^l and exponentiate them by q . (He will replace g^k by g^{kq} and g^l by g^{lq} .) The secret key will be g^{klq} which allows Oscar to find this value by exhaustive search. This is done by noting that the order of $g^q = g^{\frac{(p-1)}{2}}$ is 2^1 which implies that the secret key can only take one of two values. Hence, Oscar can use a brute force search (only two elements to try) in order to determine what the shared secret key is.



6 DHP over general linear group

We shall denote the general linear group and the algebra of $n \times n$ matrices over a finite field K as $GL_n(K)$ and $M_n(K)$ or simply as GL_n and M_n respectively, whenever the field is known from the context or is irrelevant. The minimal polynomial of a matrix A shall be denoted as $h(A, x)$.

Definition 1: DHP over GL_n . A matrix A in GL_n and matrices $B = A^k$ and $C = A^l$ are given for some unknown positive integers k, l ; $\text{ord } A$. Determine the matrix $A^{kl} = B^l = C^k$. The matrix A^{kl} is called the shared key of the DH protocol. The triple (A, B, C) shall be called the public data of the DHP ■

¹the subgroup generated by $g^{\frac{(p-1)}{2}}$ is $(g^{\frac{(p-1)}{2}} = p-1, (g^{\frac{(p-1)}{2}})^2 = 1)$

6.1 Modulus condition and solution of the DHP

It is of interest to solve this problem for large k, l . For convenience we shall assume that $2n \leq k, l \leq \text{ord } A$ in this matrix case throughout this paper. The matrix logarithm problem is that of solving for k from B and is analyzed in [4]. Clearly, solution of the logarithm leads to the solution of the DH problem above. We present an approach for solving the DHP which makes use of the fact that $GL_n \subset M_n$. Denote

$$\begin{aligned} h_c(x) &= \text{lcm}(h(A, x), h(C, x)) \\ h_b(x) &= \text{lcm}(h(A, x), h(B, x)) \end{aligned}$$

Proposition 1. There exist polynomials $f(x), g(x)$ with $\deg f < \deg h_c, \deg g < \deg h_b$ such that

$$\begin{aligned} B &= f(A) \\ C^k &= f(C) \end{aligned}$$

and

$$\begin{aligned} C &= g(A) \\ B^l &= g(B) \end{aligned}$$

Conversely if polynomials f, g satisfy above conditions then $f(x) = x^k \pmod{h_c(x)}$ and $g(x) = x^l \pmod{h_b(x)}$.

Proof. Let $f(x) = x^k \pmod{h_c(x)}$ and $g(x) = x^l \pmod{h_b(x)}$. Then these polynomials satisfy the above properties.

Conversely, if $B = f(A)$ and $C^k = f(C)$ (respectively $C = g(A)$ and $B^l = g(B)$) then $x^k - f(x)$ (resp. $x^l - g(x)$) is an annihilating polynomial of both A and C and hence is divisible by $h_c(x)$. Clearly, since $\deg f < \deg h_c$ it follows that $f(x) = x^k \pmod{h_c(x)}$. Case for g can be proved similarly.

Remark 1. Note that the numbers k, l are not given as data in the DHP hence the polynomials f, g cannot be computed from their definitions above. However the above result guarantees existence of f, g which can be used to express the shared key. These polynomials f or g can be found from the relations $B = f(A)$ and $C = g(A)$. In general these equations may not have unique solutions f, g . It is shown below that unique f, g satisfying these equations exist for a restricted class of triples (A, k, l) . Solutions to these equations nevertheless exist, even though non-unique in general, and have important implications for PKC.

Definition 2: The modulus condition. The triple (A, k, l) with A in $GL_n(K)$ and $2n < k, l < \text{ord } A$, is said to satisfy the *modulus condition* if any one of the following conditions hold

$$\begin{aligned} \text{C1: } & x^k \pmod{h(A, x)} = x^k \pmod{h_c(x)} \\ \text{C2: } & x^l \pmod{h(A, x)} = x^l \pmod{h_b(x)} \end{aligned}$$

■ **Theorem 1.** The following statements hold

1. There exists a polynomial $f(x)$ with $\deg f(x) < \deg h(A, x)$ which satisfies $B = f(A)$ and $C^k = f(C)$ iff (A, k, l) satisfies the modulus condition C1. Such a polynomial is unique.
2. There exists a polynomial $g(x)$ with $\deg g(x) < \deg h(A, x)$ which satisfies $C = g(A)$ and $B^l = g(B)$ iff (A, k, l) satisfies the modulus condition C2. Such a polynomial is unique.

Proof. Only the first item is proved as the second item follows by similar reasoning. Let (A, k, l) satisfy condition C1 and choose $f(x) = x^k \bmod h(A, x) = x^k \bmod h_c(x)$. Then f satisfies the required conditions. This proves sufficiency.

Conversely, let f be a polynomial of $\deg f < \deg h(A, x)$ which satisfies $B = f(A)$ and $C^k = f(C)$. Then $x^k - f(x)$ is annihilating for both A and C , hence divisible by their minimal polynomials. Hence $x^k - f(x)$ is also divisible by $h_c(x)$. Hence f is the unique polynomial which equals $x^k \bmod h(A, x) = x^k \bmod h(C, x)$ since $\deg f(x) < \deg h(A, x) \leq \deg h_c(x)$. This proves the necessity.

Remark 2. Note that the equation $B = f(A)$ (resp. $C = g(A)$) always has a unique solution f (resp. g) of degree less than that of $h(A, x)$. These equations are linear systems over the field K of fixed size n^2 equations in d unknowns where d is the degree of $h(A, x)$ for any k (resp. l). The shared key C^k (resp. B^l) is then obtained as $f(C)$ (resp. $g(B)$) for triples (A, k, l) satisfying the modulus condition C1 (resp. C2).

6.2 Conjugate class and the Diffie-Hellman conjecture

Consider the public data A, B, C of the DH protocol. The problem to be addressed now is to decide whether the triple (A, k, l) satisfies the modulus condition purely from the public data. This is possible in a special class defined below.

Definition 3: **A.** triple (A, k, l) is said to belong to the *conjugate class* relative to k if $h(A, x) = h(B, x)$ and relative to l if $h(A, x) = h(C, x)$. ■

Theorem 3. The DH conjecture is false for triples (A, k, l) belonging to the conjugate classes relative to any one of k or l .

Proof. Consider the public data (A, B, C) of the DHP. If the triple (A, k, l) belongs to the conjugate class relative to l , then it clearly satisfies the modulus condition C1. The knowledge that this is so is obtained only from A and C which belong to the public data. The polynomial f is now solved from the equation (1) and the shared key equals $C^k = f(C)$. Hence the shared key is computed purely from the public data. The computation of f moreover does not imply computation of k or l due to theorem 2. The case of conjugate class relative to k follows similarly.

Examples

In this section we present examples which illustrate the above theory for solving the DHP for matrices. The parameters used in these problems are of very small sizes and by no means realistic. Moreover the condition $2n \nmid k, l$ is used only for one of k, l as this is sufficient to illustrate the results.

Example 1. Consider the field be F_53 and A in GL_2 given by

$$A = \begin{bmatrix} 1 & 51 \\ 1 & 1 \end{bmatrix}$$

Let $k = 3, l = 53$ then

$$A^3 = B = \begin{bmatrix} 48 & 51 \\ 1 & 48 \end{bmatrix}$$

$$C = A^{53} = \begin{bmatrix} 1 & 2 \\ 52 & 1 \end{bmatrix}$$

The shared key is

$$A^{53*3} = \begin{bmatrix} 48 & 2 \\ 52 & 48 \end{bmatrix}$$

The minimal polynomials are $h(A, x) = h(C, x) = x^2 + 51x + 3$. Now the polynomial Solution of the linear system $B = f(A)$ gives $f(x) = x + 47$. It is easy to see that $A^{533} = f(C)$. In this example the exponent $l = 53$ is of the form p^j for $j = 1$. Where p is the field characteristic.

Example 2. In this example $h(A, x) = h(C, x)$ is satisfied for exponents k, l which are not of the form pj . Let the field be F_{13} . The matrices A, B and C are given respectively as

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$$

$$B = A^3 = \begin{bmatrix} 8 & 0 \\ 0 & 5 \end{bmatrix}$$

$$C = A^{11} = \begin{bmatrix} 7 & 0 \\ 0 & 2 \end{bmatrix}$$

The shared key A^{kl} is given by

$$A^{3*11} = \begin{bmatrix} 5 & 0 \\ 0 & 8 \end{bmatrix}$$

Now solving the linear system $B = f(A)$ gives the polynomial $f(x) = 2x + 4$. Then it can be seen that $f(C) = A^{kl}$.

Example 3. This example shows that modulus condition is satisfied even if $h(A, x) \neq h(C, x)$. Let the field be F_7 . The matrix A is chosen as

$$A = \begin{bmatrix} 0 & 5 \\ 1 & 0 \end{bmatrix}.$$

Let $k = 7$ and $l = 3$. Then

$$B = \begin{bmatrix} 0 & 2 \\ 6 & 0 \end{bmatrix}.$$

$$C = \begin{bmatrix} 0 & 4 \\ 5 & 0 \end{bmatrix}.$$

The shared key A^{kl} is

$$A^{7*3} = \begin{bmatrix} 0 & 3 \\ 2 & 0 \end{bmatrix}$$

Solving the linear system $B = f(A)$ gives $f(x) = 6x$. Then the shared key can be computed as $A^{21} = 6C$. Here $h(A, x) = x^2 + 2$ and $h(C, x) = x^2 + 1$. Since $k = 7$, we get $6x = x^7 \pmod{h(A, x) = x^7 \pmod{h(C, x)}}$. In the next section we show ways to extend results of this section to extension fields F_{p^m} , which lead to the solution of the DHP for an analogous class of triples such as those satisfying the modulus condition and for the conjugate class above.

7 DHP over general extension field

The conditions derived above for General Linear Group also hold for Extension Fields. For extension field (F_{p^m}) , the values of k and l satisfying the conjugate class condition are of the form

$p^r \pmod{n}$

For Example Assume extension field of prime field 2 over irreducible polynomial $x^3 + x + 1$

Let g be the generator of the extension field.

Hence, $g^3 + g + 1 = 0$

Take $k = 6$ and $l = 2$

Now,

$$A = g^k = g^6 = g^2 + 1 = f(g) \quad B = g^l = g^2$$

Shared key is

$$g^{12} = g^7 \cdot g^5 = g^5 = g^2 + g + 1$$

Also,

$$f(B) = f(g^2) = g^4 + 1 = g^2 + g + 1$$

8 DHP over elliptic curves

9 Conclusion

10 Our Observations

10.1 Weak key in Diffie Hellman

If either of the parties **Alice** or **Bob**, selects its private key as $\frac{(p-1)}{2}$, then the public parameter for that party would be $p - 1$. Hence, if the public parameter is $p - 1$, an intruder can

safely assume that the private key is $\frac{(p-1)}{2}$. Hence the private key $\frac{(p-1)}{2}$ is a weak key and should not be used.

10.2 Comparison between the Composite Order Subgroup attack and the Man-in-the-middle attack

On the surface, the Composite Order Subgroup attack (Section refsec:compOrderSubGrp) may appear similar to a simple Man-in-the-middle attack. However, the Composite Order Subgroup attack is a more sophisticated attack. In the regular Man-in-the-middle attack, the intruder would have to share a secret key each with Alice and Bob. Using these secret keys, every message being sent by Alice to Bob would first have to be decrypted by the intruder using the secret key shared with Alice, and would then have to be encrypted using the secret key shared with Bob. Hence, in the Man-in-the-middle attack, the intruder would have to be an active listener. On the other hand, in the Composite Order Subgroup attack, once a shared secret key has been established between Alice and Bob, the intruder needs to determine the key using brute-force technique (As mentioned in Section 5.3). Once the intruder knows the shared secret, he can passively listen to the messages being exchanged between Alice and Bob.

10.3 Vulnerabilities in DHP over extension fields

(To be written again)

11 Appendix

11.1 Fermat's Little Theorem

If p is prime and p is not a factor of a , then
 $a^{p-1} = 1 \pmod{p}$

11.2 Discrete Logarithm Problem

The problem of finding r such that $g^r = d$, where d and g are elements in a given group. The discrete logarithm problem has a sub exponential complexity.

11.3 Safe Primes

Safe primes are prime numbers of the form $p = 2q + 1$ where q is prime. Such primes have various cryptographic advantages.

References

- [1] W. Diffie and M. Hellman. New Directions in Cryptography. *IEEE Trans. on Information Theory*, 22:644–654, 1976.
- [2] A. A. Kalele and V. R. Sule. On the Diffie-Hellman problem over GL_n . pages 1–14, 2004.
- [3] R. Lidl and G. Pilz. *Applied Abstract Algebra*. Springer-Verlag, 1st edition edition, 1984.
- [4] A. J. Menezes and Yi-Hong Wu. The discrete logarithm problem in gl_n . *ARS Combinatoria*, 47:23–32, 1998.
- [5] Jean-Francois Raymond and Anton Stiglic. Security issues in the diffie-hellman key agreement protocol. *IEEE Trans. on Information Theory*, pages 1–17, 1998.
- [6] William Stallings. *Cryptography and Network Security*. Pearson Education, 3rd edition edition, 2003.