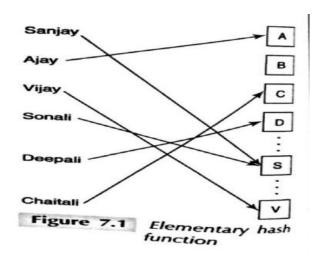
Module 2

Cryptographic Hash

2.1 INTRODUCTION

- ➤ <u>Definition:</u> A hash function is a deterministic function that maps an input element from a larger (possibly infinite) set to an output element in a much smaller set.
- The input element is mapped to a <u>hash value</u>.
- For example, in a district-level database of residents of that district, an individual's record may be mapped to one of 26 hash buckets.
- Each hash bucket is labelled by a distinct alphabet corresponding to the first alphabet of a person's name.
- ➤ Given a person's name (the input), the output or hash value is simply the first letter of that name (Fig. 7.1).
- Hashes are often used to speed up insertion, deletion, and querying of databases.

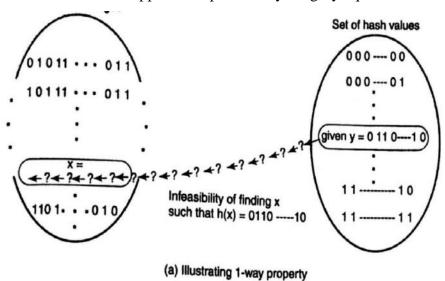


➤ In the example above, two names beginning with the same alphabet map to the same hash bucket and result in a collision.

2.2 PROPERTIES

7.2.1 Basics

- A cryptographic hash function, h(x), maps a binary string of arbitrary length to a fixed length binary string.
- \triangleright The properties of h are as follows:
 - 1. One-way property. Given a hash value, y (belonging to the range of the hash function), it is computationally infeasible to find an input x such that b(x) = y
 - 2. Weak collision resistance. Given an input value x1, it is computationally infeasible to find another input value x2 such that h(x1) = h(x2)
 - 3. **Strong collision** *resistance*. It is computationally infeasible to find two input values x1 and no x2 such that h(x1)=h(x2)
 - 4. *Confusion* + *diffusion*. If a single bit in the input string is flipped, then each bit of the hash value is flipped with probability roughly equal to 0.5.



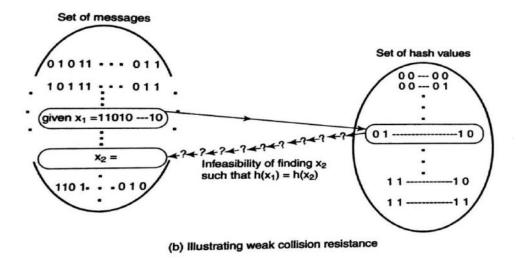


Figure 7.2 Properties of the cryptographic hash

- There is a subtle difference between the two collision resistance properties.
- In the first, the hash designer chooses x1 and challenges anyone to find an x2, which maps to the same hash value as of x1. This is a more specific challenge compared to the one in which the attacker tries to find and x2 such that h(x1) = h(x2).
- In the second challenge, the attacker has the liberty to choose x1.

2.2.2 Attack Complexity

Weak Collision Resistance

- How low long would it take to find an input, x, that hashes to a given value y?
- Assume that the hash value is w bits long. So, the total number of possible hash values is 2^w
- brute force attempt to obtain x would be to loop through the following operations

```
do
{
    generate a random string, x'
    compute h(x')
}
while (h(x') != y)
return (x')
```

 \triangleright assuming that any given string is equally likely to map to any one of the 2^W hash values, it follows that the above loop would have to run, on the average, 2^{W-1} times before finding an x' such that h(x') = y.

A similar loop could be used to find a string, x2, that has the same hash value as a given string x1.

Strong Collision Resistance

- ➤ A Brute-force attack on strong collision-resistance of a hash function involves looping through the program in Fig. 7.4.
- ➤ Unlike the program that attacks weak collision resistance, this program terminates when the hash of a newly chosen random string collides with any of the previously computed hash values.

```
// S is the set of (input string, hash value) pairs
// encountered so far

notFound = true
while (notFound)
{
    generate a random string, x'
    search for a pair (x, y) in S where x = x'
    if (no such pair exists in S)
    {
        compute y' = h(x')
        search for a pair (x, y) in S where y = y'
        if (no such pair exists in S)
            insert (x', y') into S
        else
            notFound = false
    }
}
return (x and x') // these are two strings that have
    // the same hash value
```

Figure 7.4:program to attack strong collision resistance.

THE BIRTHDAY ANALOGY

- Attacking strong collision resistance is analogous to answering the following:
- ➤ "What is the minimum number of persons required so that the probability of two or more in the, group having the same birthday is greater than 1/2?"
- ➤ It is known that in a class of only 23 random individuals, there is a greater than 50% chance that: the birthdays of at least two persons coincide (a "Birthday Collision").
- This statement is referred, to as the Birthday Paradox.

THE BIRTHDAY ATTACK

- ➤ The following idea, first proposed by Yuval illustrates the danger in choosing hash lengths less than 128 bits.
- A malicious individual, Malloc, wishes to forge the signature of his victim, Alka, on a fake document, F.
- F could, for example, assert that Alka owes Malloc several million rupees.
- ➤ Malloc does the following:
 - 1. He creates millions of documents, Fl, F2,.....Fm, etc. that are, for all practical purposes, "clones" of F.
 - 2. This is accomplished by leaving an extra space between two words, etc.
 - 3. If there are 300 words in F, there are 2300 ways in which extra spaces may be left between words.
 - 4. He computes the hashes, h(F1), h(F2), ... h(Fm) of each of these documents.
 - 5. He creates an innocuous document, D one that most people would not hesitate to sign. (For example, it could espouse an environmental cause relating to conservation of forests.)
 - 6. He creates millions of "clones" of D in the same way he cloned F above.
 - 7. Let D1, D2, ... be the cloned documents of D.
 - 8. He computes the hashes, h(D1), h(D2), ... h(Dm) of each of the cloned documents.
 - 9. Malloc asks Alka to sign the document D, and Alka obliges.
 - 10. Later Malloc accuses Alka of signing the fraudulent document
 - 11. the digital signature is obtained by encrypting the hash value of the document using the private key of the signer.
 - 12. Thus, Alka's signature on Dj, is the same as that on Fi,.
 - 13. Hence, at a later point in time, Malloc can use Alka's signature on Dj), to claim that she signed the fraudulent document, F.,.

2.3 CONSTRUCTION

2.3.1 Generic Cryptographic Hash

- The input to a cryptographic hash function is often a message or document.
- ➤ To accommodate inputs of arbitrary length, most hash functions (including the commonly used MD-5 and SHA-1) use iterative construction as shown in Fig. 7.5.
- > C is a compression box.
- > It accepts two binary strings of lengths **b** and **w** and produces an output string of length **w**.
- > Here, b is the block size and w is the width of the digest.
- > During the first iteration, it accepts a pre-defined initialization vector (IV), while the top input is the first block of the message.
- In subsequent iterations, the "partial hash output" is fed back as the second input to the C-box.
- The top input is derived from successive blocks of the message.
- This is repeated until all the blocks of the message have been processed.
- > The above operation is summarized below:
- \rightarrow h, = C (IV, m₁) for first block of message
- \rightarrow **hi** = C (**h**_{i-1}.**m**_i) for all subsequent blocks of the message

C = Compression function

S = Multiplexor

IV = Initialization vector

m_i = ith block of message m

h_i = Hash value after ith iteration

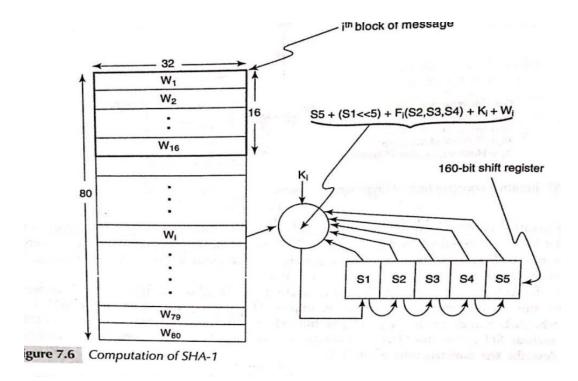
7.5 Iterative construction of cryptographic hash

Figure 7.5 Iterative construction of cryptographic hash

- The above iterative construction of the cryptographic hash function is a simplified version of that proposed by **Merkle and Damgard.**
- ➤ It has the property that if the compression function is collision-resultant, then the resulting hash function is also collision-resultant.
- ➤ MD-5 and SHA-1 are the best known examples. MD-5 is a 128-bit hash, while SHA-1 is a 160-bit hash.

2.3.2 Case Study: SHA-1

> SHA-1 uses the iterative hash construction of Fig. 7.5.



```
initialize the shift register, S1 S2 S3 S4 S5
for each block of the (message + pad + length field) {
    create the 80-word array [using Eq. (7.2)]
    for i = 1 to 80 {
        temp \leftarrow S5 + (S1 << 5) + F_i(S2, S3, S4) + K_i + W_i
        S5 \leftarrow S4
        S4 \leftarrow S3
        S3 \leftarrow S2 >> 2
        S2 \leftarrow S1
        S1 \leftarrow temp
}
```

$$F_i$$
 (S2, S3, S4) = (S2 \land S3) \lor (\sim S2 \land S4),
 F_i (S2, S3, S4) = S2 \oplus S3 \oplus S4,
 F_i (S2, S3, S4) = (S2 \land S3) \lor (S2 \land S4) \lor (S3 \land S4),
 F_i (S2, S3, S4) = S2 \oplus S3 \oplus S4
 F_i (S2, S3, S4) = S2 \oplus S3 \oplus S4
 F_i (S2, S3, S4) = S2 \oplus S3 \oplus S4

- The message is split into blocks of *size 512 bits*.
- The length of the message, expressed in binary as a 64 bit number, is appended to the message.
- ➤ Between the end of the message and the length field, a pad is inserted so that the length of the (message + pad + 64) is a multiple of 512, the block size.
- > The pad has the form: 1 followed by the required number of 0's.

Array Initialization

- Each block is split into 16 words, each 32 bits wide.
- These 16 words populate the first 16 positions, W1, W2W16, of an array of 80 words.

The remaining **64 words** are obtained from :

$$W_{i} = W_{i-3} \oplus W_{i-8} \oplus W_{i-14} \oplus W_{i-16} \quad 16 < i \le 80$$

> This array of words is shown in Fig. 7.6.

Hash Computation in SHA 1

- A 160-bit shift register is used to compute the intermediate hash values (Fig. 7.6).
- > It is initialized to a fixed pre-determined value at the start of the hash computation.
- ➤ We use the notation S1, S2, S3, S4, and S5 to denote the five 32-bit words making up the shift register.
- The bits of the shift register are then mangled together with each of the words of the array in turn.
- The mangling is achieved using a combination of the following Boolean operations: +, v, ~, ^, XOR ROTATE.

2.4 APPLICATIONS AND PERFORMANCE

2.4.1 Hash-based MAC

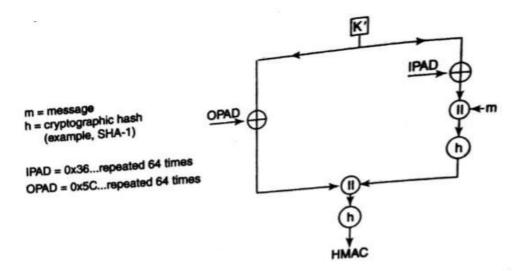
MAC

- ➤ MAC is used as a message integrity check as well as to provide message authentication.
- It makes use of a common shared secret, k, between two communicating parties.
- > The hash-based MAC that we now introduce is an alternative to the CBC-MAC.
- > The cryptographic hash applied on a message creates a digest or digital fingerprint of that message.
- > Suppose that a sender and receiver share a secret, k.
- ➤ If the message and secret are concatenated and a hash taken on this string, then the hash value becomes a fingerprint of the combination of the message, m and the secret, k.
- \rightarrow MAC = h (m/|k)
- The MAC is much more than just a *checksum* on a message.
- ➤ It is computed by the sender, appended to the message, and sent across to the receiver.

- ➤ On receipt of the **message** + **MAC**, the receiver performs the computation using the common secret and the received message.
- It checks to see whether the MAC computed by it matches the received MAC.
- A change of even a single bit in the message or MAC will result in a mismatch between the computed MAC and the received MAC.
- In the event of a match, the receiver concludes the following:
- > (a) The sender of the message is the same entity it shares the secret with thus the MAC provides source authentication.
- > (b) The message has not been corrupted or tampered with in transit thus the MAC provides verification of message integrity.
- > Drawbacks:
- An attacker might obtain one or more message—MAC pairs in an attempt to determine the MAC secret.
- First, if the hash function is one-way, then it is not feasible for an attacker to deduce the input to the hash function that generated the MAC and thus recover the secret.
- ➤ If the hash function is collision-resistant, then it is virtually impossible for an attacker to suitably modify a message so that the modified message and the original both map to the same MAC value.

HMAC

- There are other ways of computing the hash MAC other than this method using HMAC.
- Another possibility is to use key itself as the Initialization Vector (IV) instead of concatenating it with the message.
- ➤ Bellare, Canetti, and Krawczyk proposed the HMAC and showed that their scheme is re against a number of subtle attacks on the simple hash-based MAC.
- Figure 7.7 shows how an HMAC is computed given a key and a message.



7.7 Computation of an HMAC

- The key is padded with O's (if necessary) to form a 64-byte string denoted K' and XORed with a constant (denoted IPAD).
- > It is then concatenated with the message and a hash is performed on the result.
- > K' is also XORed with another constant (denoted OPAD) after which it is prepended to the output of the first hash.
- > Once again hash is then computed to yield the HMAC.
- As shown in Fig. 7.7, HMAC performs an extra hash computation but provides greatly enhanced security.

2.4.2 Digital Signatures

- The same secret that is used to generate a MAC on a message is the one that is used to verify the MAC.
- Thus the MAC secret should be known by both parties the party that generates the MAC and the party that verifies it.
- A digital signature, on the other hand, uses a secret that only the signer is privy to.
- An example of such a secret is the signer's private key.
- A crude example of an RSA signature by A on message, m, is $E_{A,pr}(m)$
- where A.pr is A's private key.
- The use of the signer's private key is a fundamental aspect of signature generation.
- Hence, a message sent together with the sender's signature guarantees not just integrity and authentication but also non-repudiation, i.e., the signer of a document

cannot later deny having signed it since she alone has knowledge or access to her private key used for signing.

- The verifier needs to perform only a public key operation on the digital signature (using the signer's public key) and a hash on the message.
- The verifier concludes that the signature is authentic if the results of these two operations tally,

$$E_{A.pu}$$
 $(E_{A.pr}(h(m))) \stackrel{?}{=} h(m)$

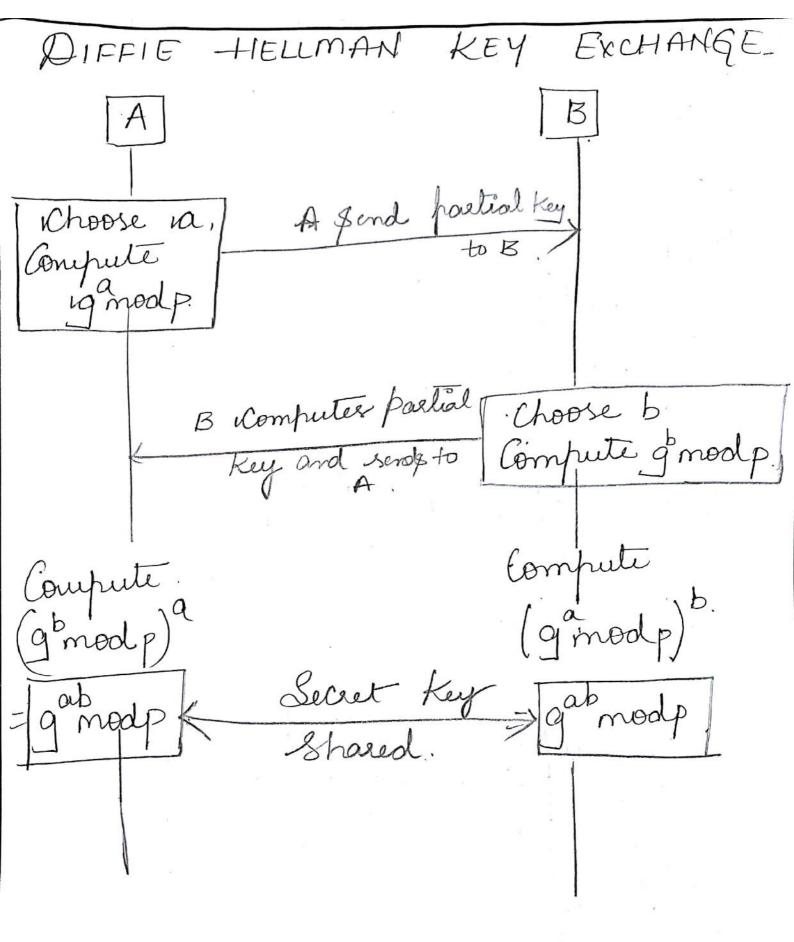
Question Bank (module 2-chapter 2)

- 1. Explain generic hash computation and HMAC.
- 2. Define hashing Explain the properties of hashing with a neat figure.
- 3. Explain SHA-1 computation with a neat illustration.
- 4. Explain weak and strong collision resistance.
- 5. Explain digital signature.
- 6. Explain birthday analogy and birthday attack

DISCRETE LOGARITHM AND ITS APPLICATIONS. INTRODUCTION. Consider the finite, multiplicative group $(Z_p^*, *_p)$ where p is prime. - Let ig be the igenerator of the group. igmodp, igmodp, ... gp-modp. Let x be an element in {0,1,--P-13. The function: y= gr (modp) onentiation ruth Base of and moduler - The Inverse Operation is o $|\chi = \log_q y(\text{med}_p)|$ Discrete logouthm

MODULE 2 - Chapter 3

Example. > Let p=131 ig = 2 KEY EXCHANGE. * DIFFIE -- HELLMAN PROTOCOL. reed to eaguer whom a shared secret for the obweation of their Current Session. > In 1976, Diffie and hellman proposed the violea of a private Key and Coversponding public Key, 1) A chooser a vandom integer a, 1<a < p-1, computer the partial Key gamodpand sends to B. B chooser is random integer b, 1xbxp-1, computer the partial Key ig mod p and sends to A. (gamodp) mod p = gab modp 4) On the receipt of B's mag, A computer (gbmodp) modp = gabmodp.



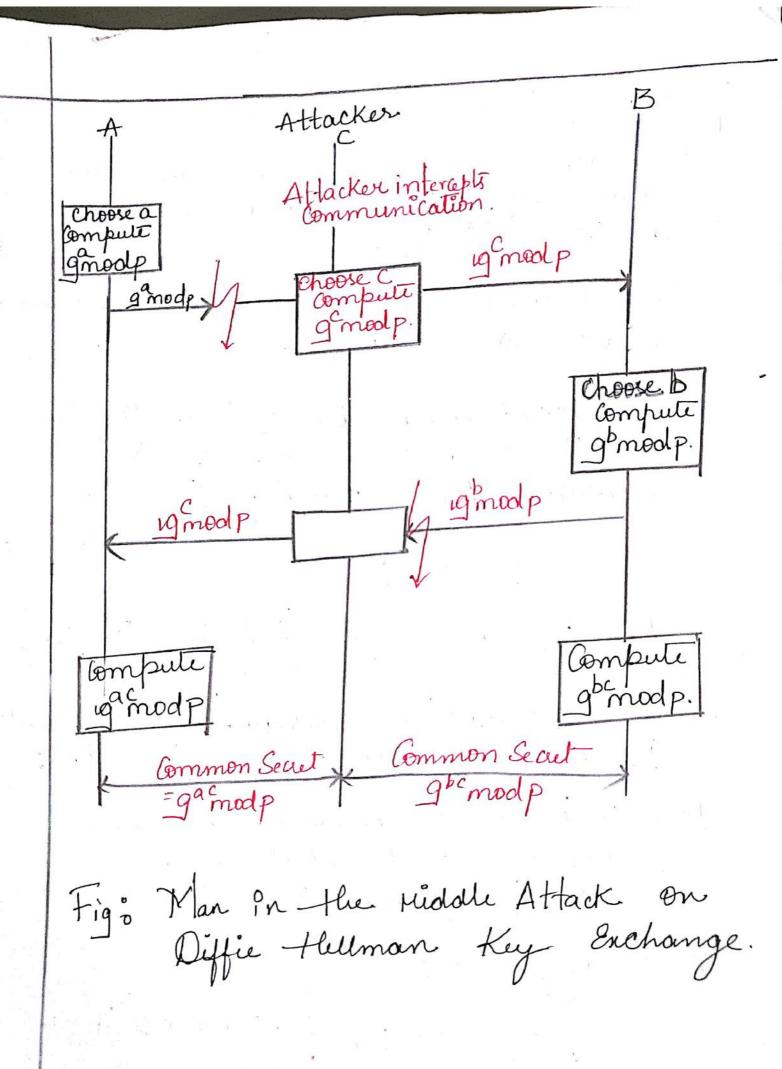
Nagashree. C Asst Professor, Department of CSE,SVIT

Example: Compute Diffic Hellman partial Keyr and Secret Keyr. uthere a=24, b=17, · 9 = 2 and p=131. i) A computer faction Key. = gamod p. = 2 mod 131 2). B Computes partial Key: = ighmodp. = 2 mod (3) 3) A computer Secret Key rafter receiving B'\$ partial Key. = (ighmodp) B's partial key. = (72)24 mod 131 B computer Secret Key: (gmodp) = 46 med 131 = 113

ATTACKS The partial Keys, igamode and gemode Nove sent in clear An Eaverdropper nith the Knowledge of the partial keys and public parameters (\$ and g). Ideduce the Common Secret gabmod p, idenued by A of B. This publim is referred to is Computational Diffie Hellman problem. MAN IN THE MIDDLE ATTACK ON DIFFIE - HELLMAN KEY EXCHANGE. - An vattacker, ic chooser an integer c and computer gic mod p. Cothen interrupts A's merrage to B, substitutes it with gemod p and sends this instead to B. C valso untercepts Bs merrage to A sending gemod p ûnstead. - After the message transfer

B Computer > (gcmod p) mod p → gbc mod p.

hetile A Computer, $(ig^{c} mod p)^{a} mod p = ig^{c} mod p.$ il valso computer the two Secrets -> igac mod p and y bc mod p. A and B might—think that they have in secure channel for to comminication by encrypting all messages. But A Sharer the Secret ganodpmith - B shower the Secret go modp with C. - Every Subsequent merage encypted by A and untended for B Can be idecrypted by C. Similarly Every message from B to A Can be idetrypted by C. This is a classic Example of an active "Man in the Hiddle Stack".



EL GAMAL ENCRYPTION. El gamal encryption user va large prime number p and generator ig in (Z_p^*, p) . An Elgamal primate Key, is an integer a, I < a < p-1. The Coverfooding public Key is the truplet (p, 19, x) rehere & is the encryption Key Calculated of X = g mod p. - Let (p, 19, x) be the kublic Key of - 10 Encrypt a mersage to be sent to A, B idser the following: 1) B chooses a random number r, 1< r < p-1 such that r is relatively prime to \$-1 2) B Computer: C1 = gmodp C2 = (m * X) modp

3) B sends the Ciphertest $C = [C_1, C_2]$ to A. Decuption At Aside * A user its primate key is to ideceypt and obtain plaintent m., (c, -a) * C2 mod p. * ELGAMAL SIGNATURES. > Let ia be the primate key of A. -> Let (p, 1g, x) he the public Ky of A. -> To sign ia mersage m, A doer the following: i) She computer the hash h(m) of the 2) She Chooses 10 random number 7, 1/2/2/1, such That r is relatively prime to \$-1.

3) She Computer X= ig mod p 4) She Computer y= (h(m) - ax) & mod (p-1) The Signature is the pair (x,y). * Signature revification user X, 10 prove Elgamal Signature: Consider step 4 Egn y=(h(m) - ax) r mod p-1 y = (h(m) -ax) 1 mod p-1 Ty=(h(m)-ax)+K(p-1), where K is an integer Rawing Both Bider to power of 9
and reducing Modulo P. J. Equal

19 = 9 16 mod p. [Fermals] gry = gh(m) 1 modp. - [hersem]

where $X = h(m | l | g^{mod} p)$ and Y = (r + ax) mod q.

where $X = g^{mod} p$.

V be random rumbur 1≤√≤q-1.

'KROBLEMS ON ELGAMAL ENCRYPTION. a. A Block of plaintent message m=3, how to be encrypted, Assume P=11, 1g=2, vecipients primate Key=5, Sender Chooser Vandom inleger 7=7. Verform Encuption & Decuption. Step1: p=11, y=2 Recipients primate Key, 10=5. Compute public Key of receiver: X=gmodp. x=2 mod 11 x = 32 mod 11 X=10 Step 2: Compute C, and C2 [Sendue hour to compute] C1=g modp

$$C_{1} = g^{8} \mod p \qquad [r=+]$$

$$= 2^{\frac{1}{2}} \mod 11$$

$$= 128 \mod 11$$

$$C_{1} = 7$$

$$C_{2} = m * x \mod p \qquad [m=3]$$

$$= 3 * 10^{\frac{1}{2}} \mod 11$$

$$C_{2} = 8$$

$$C = [7, 8]$$

$$C = [7, 8]$$

$$Step 3: Decuypt$$

$$m = C_{1}^{-2} * C_{2} \mod p$$

$$= 7^{\frac{1}{2}} * 8 \mod 11$$

$$= (7_{1}^{-1})^{\frac{1}{2}} * 8 \mod 11$$

$$= 8^{\frac{1}{2}} * 8 \mod 11$$

= 8⁵ * 8 mod 11

Take 56 mod 11

= 1

Equivalent #01

Scanned by CamScanner

Substitute

7 = 8 governe.

7×3=21 modinx

7x 5=35mod11x

i not

1 Equal to

hence

Continue

Q.
$$p = 23$$
, $g = 11$, $10 = 6$, $r = 3$, $m = 10$.

Step1: $k = g \mod p$

$$= 11 \mod p$$
 $k = 9$

Step3: $(empute C_1, C_2)$
 $C_1 = ug \mod p$

$$= (10 * g \mod 23)$$
 $C_1 = 20$

Step3: $(empute C_1) = (10 * g \mod 23)$
 $(empute C_2) = (10 * g \mod 23)$
 $(empute C_3) = (10 * g \mod 23)$
 $(empute C_4) = (10 * g \mod 23)$
 $(empute C_4)$

Scanned by CamScanner

Rublic Key Cryptography and RSA Step 1: choose two large krume numbers Steps: Compute the modulur step3; Compute Euler totient function Q(n)=(p-1)*(q-1) Step 4: Choose the enceyption key & such that 19cd (e, p(n)) =1 Steps: Compute idecryption key id ble mod $\phi(n) = 1$ id=emodp(n) e is called public Key id is Called private key.

Nagashree, C Asst Professor, Department of CSE,SVIT Step 7: Encyption:

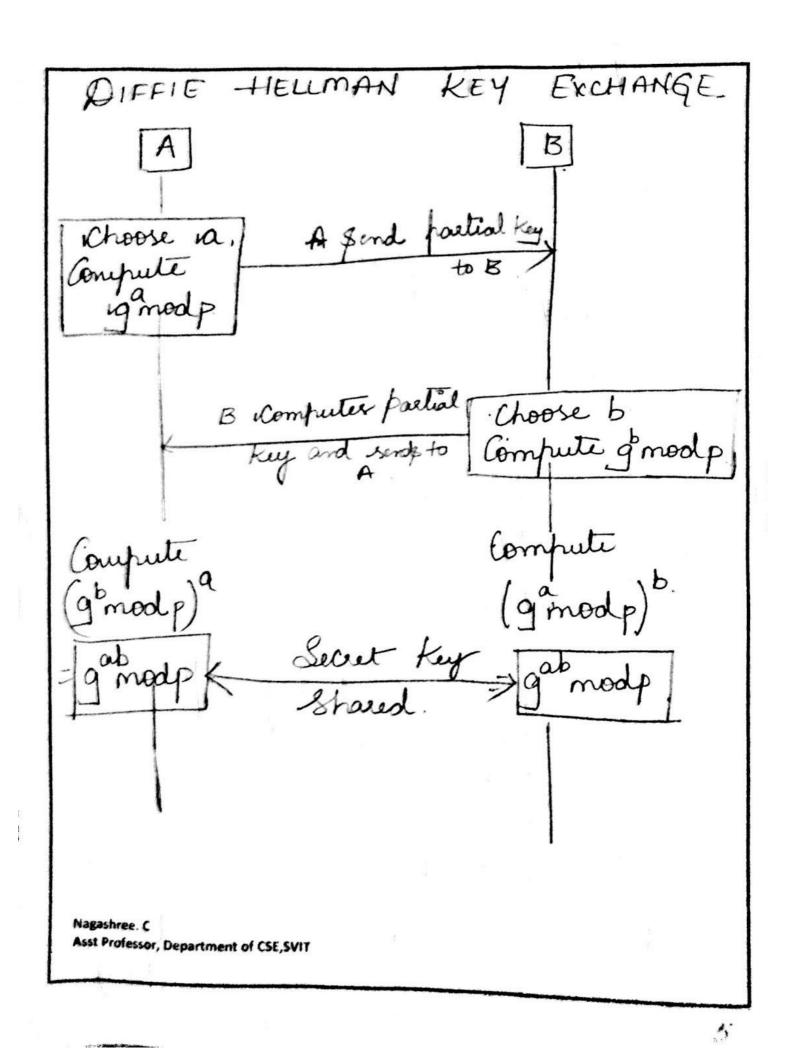
Ci = minodn

Step 7: Decryption:

my= c; modn Example: suppose RSA prime numbers are \$=3, 9=11, e=3., M=00111011 Solution; Step 1: Compute modulus n m= pxq steps: Compute $\phi(n)$ $\phi(n) = (p-1)*(q-1)$ = (3-1) * (11-1) D(n)= 20 Nagashree. C Asst Professor, Department of CSE,SVIT

> step 3: Compute encyption the igcd (e, φ(n))=1 igcd (3, 20) =1 e=3= public Key. 7 Step4: Compute Deceyption Key id= e mod q(n) = 3 mod 20 Sty 6: Decemption 2 Steps: Enception: Ri= mimodn | Mi= Cimedn. Block 1 Block Z NOTE: Plain Level M of divided into Block 6 bits as number of buts viequired to represen M=33 requilier 6

Decryption. Encyption m;= C; mod n C; = m; mod n my=001110, N=7 M1= 5 mod 33. m= 14. 14 med 33 C, = 5 m1 = 14 m2 = C2 mod n. icz= m² mod n m = 000011, $m_z = 3$ > C2 Computed 1 27 substitute C, id & nuclee im = Comed n C2 = memodn = 27 mod n. = 3 mod n C2 = Q7 =(27 mod 33 x 27 mod 33) mod 33. Nagashree. C Asst Professor, Department of CSE,SVIT



partial Compute Diffic Hellman Secret Keyr. Keyr and ulhere a=24, b=17, · 19 = 2 and \$= 131. i) A computer partial Key. : igamed p = 2 mod 131 2). B Computes partial Key: = ighmodp = 2 med 131 3) A Computer Secret Key rafter receiving B'4 partial Key. = (ngt mode) Bis fortial key. $= (72)^{24} \mod 131$ & B computer Secret Key: (gmodp) = 46 med 131 = 13