Module 2 Cryptographic Hash

2.1 INTRODUCTION

- Definition: 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 *hash value*.
- 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.
- > 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)
 - 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.



(a) Illustrating 1-way property



(b) Illustrating weak collision resistance

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 x_1 .

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 ł generate a random string, x' compute h(x')while (h(x') = y)urn (x')

> 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)
ł
   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.
 - 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:
- > \mathbf{h} , = C (IV, \mathbf{m} ₁) for first block of message

> $hi = C (h_{i-1}.m_i)$ for all subsequent blocks of the message





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 <u>Merkle and Damgard.</u>
- 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.





$1_1(02, 05, 07) = (52 \times 55) \vee (-52 \times 54),$	$1 \le i \le 20$
F_i (S2, S3, S4) = S2 \oplus S3 \oplus S4,	21 < i < 10
F_i (S2, S3, S4) = (S2 \land S3) \lor (S2 \land S4) \lor (S3 \land S4),	$41 \le i \le 60$
$F_i (S2, S3, S4) = S2 \oplus S3 \oplus S4$	$61 \le i \le 80$

- > 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.
- $\succ MAC = h (m||k)$
- > 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

MODULE 2 - Chapter 3 DISCRETE LOGARITHM AND ITS APPLICATIONS. INTRODUCTION. Consider the finite, multiplicative group (Zp*, *p) rehere p is prime. - Let ig be the generator of the group. ig modp, ig modp, g^p-modp. Let x be an element in {0, 1, -- P-1]. The function: y= gr (modp) -> Modular onintiation with Base 19 and moduler Inverse - The operation is o $\chi = \log_{q} y(med_{p})$ Discrete logarithm Scanned by CamScanner

Example. > Let p=131 ig = 2 KEY EXCHANGE. * DIFFIE - HELLMAN PROTOCOL. A f B that need to raquee upon a shared scoret for the iduration of their current Session. > In 1976, Diffie and hellman proposed the idea of a primate Key and Covverponding public Key, 1) A chooser a random integer a, 1<a < p-1, computer the partial Key gemodp and sends to B. B chooser a random integer b, 1×b×p-1, computer the partial Key 1g^bmod p and sends to A. s) On the receipt of A's msg, B Computer (gamodp) mod p = gab mod p 4) On the receipt of B's msg, A computer (gbmodp) mod p = gab mod p.

Scanned by CamScanner

EXCHANGE_ DIFFIE -HELLMAN KEY В A A gend postial key to B Khoose a, Compute ignodp. B computer partial Choose b Key and serols to Compute g modp. Compute Compute (gmodp)^b. g mod p) Key Secret ==== gab modp modp Shared

Nagashree. C Asst Professor, Department of CSE,SVIT

Example: Compute Diffic Hellman partial Keyr and Secret Keyr. uhere a=24, b=17, ·19=2 and p=131. is A computer partial Key. = gramod p. = 2 mod 131 = 46 partial Key". 27. B Computes = igemed P. = 2 mod [3] 3) A computer Secret Key after receiving B'& partial Key. = (rgp modp) " B's portial key. $= (72)^{24} \mod 131$ B computes secret Key: (grodp)b = 46 med 131 =113

ATTACKS The partial Keys, gamodp and gbmodp Dure 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 as Computational Diffie Hellman problem. MAN IN THE MIDDLE ATTACK ON DIFFIE - HELLMAN KEY EXCHANGE. - An vattacker, i chooser an intéger c and computer gir mod p. C then interrupts A's meriage to B, substitutes it with genod p and substitutes this instead to B. C valso intercepter Bs menage to A sending ig mod p ünstead. - After the message transfer B computer } (gernod p) mod p $\rightarrow g^{bc} \mod p$.

hihile A Computer, (ig'mod p) anod p = ig mod p. iC ialso the two Secrets Computer -> ig mod p and - igbc mod p. A and B might think that they have a scrube channel for f comminication by encrypting all messages. But A Sharer the Secut gamod printh C, - B shares the Secret go modp with C. - Every Subsequent message encryptid by A and intended for B land be idecrypted by C. Similarly Every message from B to A Van be idecrypted by C. This is a classic Example of an vacture "Man in the Hiddle Attack".

B Attacker -A Attacker intercepts Communication. Choose a compute gnoolp ig mool p Choose C Compute geneolp. g mode. Choose B Compute gemoolp. 1 gmodp 19 mod P Compute g^{bc}modp. Compute ig mod P Common Secret Common Secret goc modp Figö Man in the Hiddle Attack on Diffie Hellman Key Exchange.

EL GAMAL ENCRYPTION. Et gamal encryption user a large prime number p and generator ig in $(\mathbb{Z}_p^*, \mathbb{P})$. An Elgamal primate Key, is an indege integer 12, 1×2×p-1. The Coverfonding public Key is The truplet (p, 19, 12) rehere & is the encryption Key calculated X = g'med p. - Let (p, 1g, x) be The public Key of A. - 10 Encrypt a mersage to be sent to A, B idder the following: 1) B chooses a random number r, I<r<p>1<r<p>K p-1 such that r is velatively prime to p-1 2) B Computer: $C_1 = g^modP$ $C_2 = (m * \chi^m) modP$

3) B sends the apportent $C = [C_1, C_2] + A$. Debugtion At Apide * A user its primate Key is to ideceypt and obtain plaintent m.; $(v_1^{-a}) \star C_2 \mod p$ * ELGAMAL SIGNATURES. > Let a be the primate Key of A. -> Let (pigix) be the public KygA. > To sign a menage m, A doer the plowing: i) She computer the hash h(m) of the mersage. 2) She chooses a random number r, 1<r< p-1, such That r is relatively prime to p-1.

3) She Computer X= ig mod p \$? She computer y= (h(m) - ax) & mod (p-1) 5) The Signature is the pair (x, y). * Signature relification user X, 10 proue Elgamal Signature: Consider step 4 Egn y=(h(m) - ax) r mod p-1 $y = (h(m) - ax) \prod_{x} \mod p - 1$ Ty=(h(m)-ax) + K(p-1), neherek ur anmliger Raining Both Bider to powerfg and reducing Modulo P. 19^{Ty} = 9^h(m) -ax. 19^{Ty} = 9^h(m) -ax. 19^{Ty} = 9^h(m) -ax. 19^{Ty} = 10^{to} from the second p. (Fermilis) $g^{y} = g^{h(m)} - mod P$. - hersem]

$$g^{ax} \xrightarrow{yy} = g^{h(m)} \mod p.$$

$$10 \cdot x^{2} \times x^{2} = g^{h(m)} \mod p. [since \\ K = g^{a} \mod p. \\ \chi = h(m \text{ II. } g^{a} \mod p) \text{ and } \\ \underline{Y} = (r + ax) \mod q. \\ \text{where } X = g^{a} \mod p. \\ Y = and \text{ for an and } \\ I \leq Y \leq q^{-1}. \end{cases}$$

F

TROBLEMS ON ELGAMHL ENCRYPTION. Q. - A Block of plaintext message m=3, har to be encrypted, Assume P= 11, 1g = 2, recipients primate Key = 5, Sender Chooser Vandom inleger 7=7. Verform Encuption & Decuption. Step1: p=11, y=2 Recipients primate Key, 1a=5. Compute public Key of receiver: X=gmodp. $K = 2 \mod 11$ x = 32 mod 11 K = 10Step 2: Compute C, and C2 [Benderharto Compute] C1=g modp

$$C_{1} = g^{T} \mod P \qquad [r = +]$$

$$= 2^{T} \mod 11$$

$$= 128 \mod 11$$

$$C_{2} = m \times K \mod P \qquad [m = 3]$$

$$= 3 \times 10^{T} \mod 11$$

$$C_{2} = 8$$

$$C = [7, 8] \qquad T \times 3 = 21 \mod 11 \times 10^{T} \mod 11$$

$$C_{2} = 8$$

$$C = [7, 8] \qquad T \times 3 = 21 \mod 11 \times 10^{T} \mod 11 \times 10^{T} \mod 11$$

$$= (7, -1)^{5} \times 8 \mod 11$$

$$= (7, -1)^{5} \times 8 \mod 11$$

$$= (7, -1)^{5} \times 8 \mod 11$$

$$= (8^{5} \times 8 \mod 11$$

$$= 1$$

$$[Equivalent \neq 0 1]$$

$$= 1$$

$$Equivalent \neq 0$$

$$= 1$$

$$Equivalent \neq 0$$

Q.
$$p = 23$$
, $g = 11$, $a = 6$, $r = 3$, $m = 10$.
step1: $k = g^{a} \mod p$.
 $= 11^{a} \mod p$
Steps: $(mpute C_{1}, C_{2}$
 $C_{1} = g^{a} \mod p$
 $= 11^{a} \mod 23$
 $C_{1} = 20$
 $C_{2} = (m * k \mod p)$
 $= (10 * q^{3} \mod 23)$
 $C_{1} = 20$
 $C_{3} = 22$
Steps: $Decupt$:
 $m = C_{1}^{-a} * C_{2} \mod p$.
 $= 20^{-6} * 22 \mod 23$
 $= (20^{-1})^{-6} * 22 \mod 23$
 $= (15)^{-6} * 22 \mod 23$
 $m = 10$
 $m = 10$

MODULE-2 Rublic Key Cryptography and RSA RSA Step1: choose two darge kume numbers pand og Steps: Compute the modulur m= p*2 steps: Compute Euler totient function Q(n) = (p-1) * (q-1)Step 4: Choose the enception key e such that $|qcd(e, \phi(n)) = 1|$ Steps: Compute decryption key of the mod $\phi(n) = 1$ W= e mod p(n) e is called public Key id is called primate they. Nagashree, C Asst Professor, Department of CSE,SVIT

Scanned with CamScanner

step 6: Encryption: Ci = mimodn step 7: Decryption: m= c; modn Example: suppose RSA frime numbers are p=3, q=11, e=3, M=00111011Solution: Step1: Compute modulus n m= prg N=3*11 n=33 steps: compute $\phi(n)$ $\phi(n) = (p-1)*(q-1)$ = (3-1)*(11-1) 2+10 \$(n)= 20 Nagashree. C Asst Professor, Department of CSE,SVIT

2

> step 3: Compute encyption the igcd (e, p(n))=1 igcd (3,20)=1 C=3= public Key. 7 Stept: Compute Decemption Key nd = e mod q(n) = 3 mod 20 |d=7| Step 6: Decemption > Steps: Enception: Ki=minodn Mi=Cimedn. m=10011101 Block 1 Block 2 NOTE: plain Level M ip divided into Block 6 bits as number of bits required to represent Nagashree. C m=33 requilier 6 Asst Professor, Department of CSE,SVIT

Scanned with CamScanner

Decryption. Enception mg=C; mod n C:=mimod n my=001110, d=7 m1= 5 mod 33. m= 14. K1= 14 mod 33 C1 = 51 $m_1 = 14$ $m_2 = C_2 \mod n$. $\mathcal{L}_2 = m_2 \mod n$ m = 000011, $m_1 = 3$ > C2 computed is 27 substitute C, id & nuclee im = c mod n C2 = me mod n = 27 mod n. = 3° mod n $C_2 = QT$ = (27 mod 33 x 27 mod 33) mod 33. m_= 3 Nagashree. C Asst Professor, Department of CSE,SVIT

Scanned with CamScanner

DIFFIE HELLMAN KEY EXCHANGE B A A send poetial key Choose a, Compute ignodp B computer partial choose b key and sends to Compute gemoodp (grodp)b. (oupute (g^bmodp)^q g modp K Secret Key Jgab mode Nagashree. C Asst Professor, Department of CSE, SVIT

5

Example. Compute Diffic Hellman partial " Secret Keyr. Keyr and ulhere a= 24, b=17, ·19=2 and \$=131. 17 A computer partial Key: = 19 mod p = 2 mod 131 = 46 27. B Computes partial Key: = ig mod p = 2 mod 131 3) A computer Secret Key rafter receiving B'& partial Key. = (gomodp) B's postial Key. $= (72)^{24} \mod 131$ & B computer secut Key: (gmodp)b = 46 mod.131 = 113

Scanned with CamScanner