Data Integrity and Chosen Ciphertext Attacks

Cryptography and Protocols

Andrei Bulatov

Data Integrity

- Privacy is not the same as integrity!!!
- If we encrypt data with a CPA-secure scheme, does it mean that we also protect its integrity?
- NO
- Suppose we encrypt message $P = P_1 \ldots P_n$ with the PRF-based CPA secure scheme, so that ciphertext is $\langle r, f_s(r) \oplus P \rangle$
- The attacker flips the last bit of the ciphertext making Bob to believe that the message sent is $P_1 \ldots P_{n-1}\overline{P_n}$

Chosen Ciphertext Attacks

- In a chosen ciphertext attack Eve is allowed to ask for encryptions of chosen plaintexts, and for decryptions of chosen ciphertexts
- The login problem

  Suppose that a server and a client share a secret PIN, I, that was chosen at random $0 \leq I \leq 10^4$ (13 bits)

  They also share a secret key k

  Protocol:

  the client sends encrypted I

  the server decrypts and check if the PIN is correct

  if PIN is incorrect the server aborts the communication

- Can the adversary learn the PIN?

Chosen Ciphertext Attacks (cntd)

- **Lemma**

  There exists a CPA-secure scheme (K,E,D) such that if the client and the server use (K,E,D) in this protocol, Eve that sits on the communication channel can learn the PIN after at most 13 sessions.

- **Proof**

  (K,E,D):

  - K chooses key k uniformly at random from $\{0,1\}^n$

  - E to encrypt $P \in \{0,1\}^{n/2}$ encodes each bit of P as follows: 0 is encoded with 00 and 1 with 11. Then use the standard PRF-based scheme $\langle r, f_k(r) \oplus P\rangle = \langle r, C\rangle$

  - D to decrypt sets $\tilde{P} = C \oplus f_k(r)$ and then decode 00,01,10 to 0, and 11 to 1

Chosen Ciphertext Attacks (cntd)

- The encryption scheme is valid, meaning $D(E(P)) = P$, and CPA-secure (exercise)
- Property of (K,E,D):

  Eve can flip any bit of P to 0

  To flip i-th bit flip $2i$-th bit of C
- Now, to find the PIN Eve flips each of the 13 bits in turn and watches the response of the server.
- If corrupted PIN is rejected, the corresponding bit is 1 otherwise it is 0

Message Authentication Schemes

- A Message Authentication Scheme (MAC) consists of 2 algorithms (Sign, Ver)
- There is a key shared between the signer and the verifier (Alice and Bob).
- Alice sending a message P computes $s = \mathrm{Sign}_k(P)$ called a signature or tag. Then she sends (P,s) to Bob.
- Bob accepts the pair (P, s) only if $\mathrm{Ver}_k(P, s) = 1$

- Security of MACs is defined in terms of chosen message attacks (CMA)
- Let n be the key length, m the message length, and t the tag length

Chosen Message Attacks

- CMA secure MAC
- A pair (Sign,Ver), $\mathrm{Sign}: \{0,1\}^n \times \{0,1\}^m \to \{0,1\}^*$

  $\mathrm{Ver}: \{0,1\}^n \times \{0,1\}^m \times \{0,1\}^* \to \{0,1\}$ is a (T,ε)-secure MAC if

  - For any P and k: $\mathrm{Ver}_k\big(P, \mathrm{Sign}_k(P)\big) = 1$ (validity)
  - For any Eve of time complexity at most T in the following game:
    - choose k uniformly at random
    - give Eve access to black boxes Sign_k and Ver_k
    - Eve wins if she comes up with a pair $\langle P', s' \rangle$ such that

      (a) P' is not one of the messages that Eve gave to Sign_k

      (b) $\mathrm{Ver}_k(P', s') = 1$

    Eve wins with probability at most ε

Construction of a MAC

- The following are not CMA secure MACs
  - a CPA-secure scheme
  - a checksum or cyclic redundancy code

- The following is a MAC

  A pair of algorithms (Sign,Ver) that use a PRF $\{f_k\}$

  $Sign_k(P) = f_k(P)$

  $Ver_k(P, s) = 1 \Leftrightarrow f_k(P) = s$

Security of MAC

- **Theorem**

  The MAC defined on the previous slide is CMA secure

- **Proof**

  Suppose that the scheme is not secure. This means there is Eve that wins in the game with probability > ε.

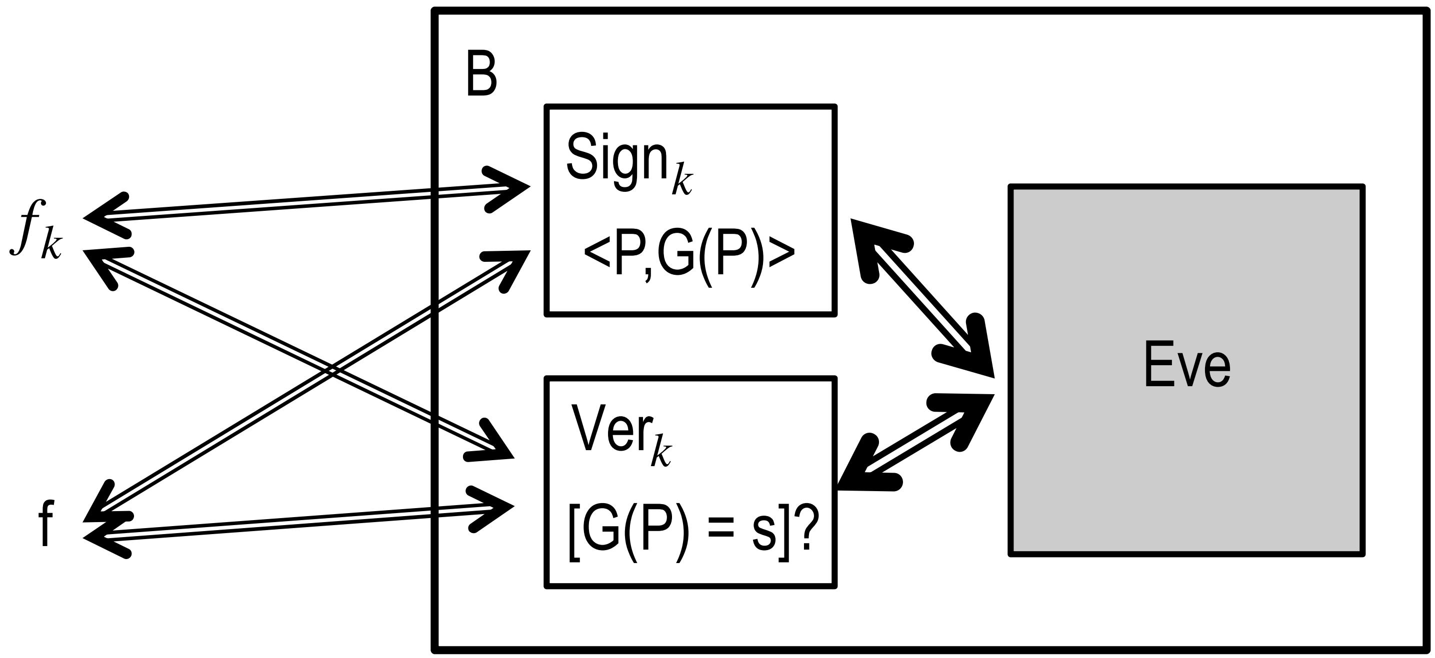
  We use this Eve to construct an algorithm that distinguishes between $\{f_k\}$ and a random function

  As before we construct an ideal MAC that uses a random function instead of a pseudorandom

  Such ideal MAC is unbreakable: When Eve comes up with a guess $\langle P', s' \rangle$, she has never asked about the value of the function on P'. However, this value is random, so the probability to guess s' correctly is $1/2^{|s'|}$

Security of MAC (cntd)

- A distinguisher B works with a function $G \in \{f_k, f\}$ and constructed as follows



Security of MAC (cntd)

- Distinguisher B works as follows, given a function $G \in \{f_k, f\}$

  - run Eve

  - when Eve requests for $\mathrm{Sign}_k(P)$ return $\langle P, G(P)\rangle$

  - when Eve requests for $\mathrm{Ver}_k(P, s)$ return 1 if $G(P) = s$, return 0 otherwise

  - when Eve makes her guess $\langle P', s'\rangle$

    output Game 1 if $G(P') = s'$ (Eve wins)

    output Game 2 otherwise

- Analysis

  |Pr[B wins Game 1] – Pr[B wins Game 2]|

  = |Pr[Eve wins with real MAC] – Pr[Eve wins with ideal MAC]|

  $> \varepsilon - 1/2^n$

Authentication and CCA

- Revisit the login problem. There are 3 ways to authenticate and encrypt PIN
- Encrypt and then Authenticate (EtA)

  Compute $C = E_k(\mathrm{PIN})$ and $t_C = \mathrm{Sign}_{k'}(C)$. Send $\langle C, t_C \rangle$

  (IPSec style)
- Authenticate and then Encrypt (AtE)

  Compute $t_P = \mathrm{Sign}_{k'}(P)$ and then $E_k(P, t_P)$

  (SSL style)
- Encrypt and Authenticate (E&A)

  Compute $C = E_k(\mathrm{PIN})$ and $t_P = \mathrm{Sign}_{k'}(P)$. Send $\langle C, t_P \rangle$

  (SSH style)
- WEP style is don't authenticate

Authentication and CCA (cntd)

- **Theorem**

  1. If (K,E,D) is a CPA secure SES and (Sign,Ver) is a CMA secure MAC, then the probability that poly-time Eve guesses the PIN after seeing polyniomially many interactions of the EtA protocol is less than 1/9999

  2. There is a CPA-secure SES such that for any CMA secure MAC, Eve can learn PIN after 13 sessions of AtE protocol

  3. There is a CMA secure MAC such that for any CPA secure SES, Eve can learn PIN after 1 session of E&A protocol

Practical MACs

- Hash functions:

  Let h be a hash function

  $$\mathrm{Sign}_k(P) = h(k||P)$$

  $$\mathrm{Ver}_k(P, s) = 1 \Leftrightarrow h(k||P) = s$$

- Weaknesses of hash functions: collision attacks, length extension attacks
- HMAC-h (h is a hash function)

  $$\mathrm{Sign}_k(P) = h(k \oplus a||h(k \oplus b||P))$$

  a, b are constants
- UMAC
- MAC using block ciphers (to be considered later)