# Identity-Based Cryptography

Ratna Dutta

Associate Professor

Department of Mathematics

Indian Institute of Technology, Kharagpur

ACM Summer School on Cryptology Research, ISI-Kolkata

June 4-22, 2018

# Public Key Cryptosystem

- Public keys are used for encryption and digital signature verification.
- Private keys are used for decryption and digital signature generation.
- Public keys are accessible to all parties.
- Private keys are to be kept secret.
- How to associate entities with their respective public keys?

- An attacker may present a harmful key as the public key of a victim.
- Before using a public key, one should verify that the key belongs to the claimed party.

# Public Key Certificates

- There is a trusted Certification Authority (CA).
- CA issues public-key certificates to parties.
- A certificate contains a public key, some identifying information of the party to whom the key belongs, a period of validity.
- The certificate is digitally signed by the CA.
- Key compromise and/or malicious activities may lead to revocation of certificates.
- The CA maintains a list of revoked certificates.

# Use of Public Key Certificates

- Alice wants to send an encrypted message to Bob.
- Alice obtains Bob’s public-key certificate.
- Alice verifies the signature of the CA on the certificate.
- Alice confirms that Bob’s identity is stored in the certificate.
- Alice checks the validity of the certificate.
- Alice ensures that the certificate does not reside in the revocation list maintained by the CA.
- Alice then uses Bob’s public key for encryption.

## Problems of Public Key Certificates

- A trusted CA is needed.
- Every certificate validation requires contact with the CA for the verification key and for the revocation list.

# Identity-Based Public Keys

- Alices identity (like e-mail ID) is used as her public key.
- No contact with the CA is necessary to validate public keys.
- A trusted authority is still needed: Private-Key Generator (PKG) or Key-Generation Center (KGC).
- Each party should meet the PKG privately once (registration phase).
- **Limitation:** Revocation of public keys may be difficult.

# Historical Remarks

- Shamir (Crypto 1984) introduces the concept of identity-based encryption (IBE) and signature (IBS). He gives a concrete realization of an IBS scheme.
- In early 2000, bilinear pairing maps are used for concrete realizations of IBE schemes.
- Sakai, Ohgishi and Kasahara (2000) propose an identity-based key-agreement scheme and an IBS scheme.
- Boneh and Franklin (Crypto 2001) propose an IBE scheme. Its security is proved in the random-oracle model.

- Boneh and Boyen (Eurocrypt 2004) propose an IBE scheme whose security can be proved without random oracles.
- Joux (ANTS 2004) proposes a pairing-based three-party key-agreement protocol (not an identity-based protocol).

# Cryptographic Bilinear Map

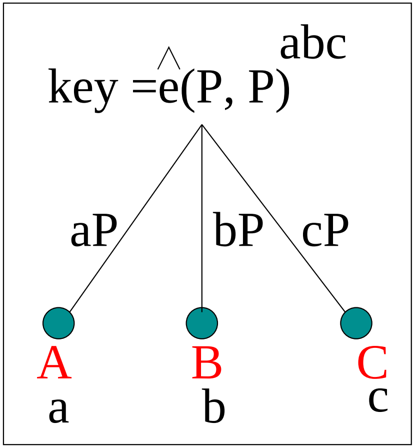
- $G_1, G_2$ two groups of a large prime order $q$.
- $G_1 = \langle P \rangle$ additive, $G_2$ multiplicative, DLP hard.
- Bilinear Map $\quad \hat{e} : G_1 \times G_1 \rightarrow G_2$

  1. Bilinearity : $\hat{e}(aP, bQ) = \hat{e}(P, Q)^{ab}$ for all $P, Q \in G_1$ and $a, b \in Z_q^*$.
  2. Non-degeneracy : $\hat{e}(P, P) \neq 1$.
  3. Computable : $\hat{e}$ can be efficiently computed.

- Examples : Modified Weil, Tate Pairing.

# Three-Party Key Agreement

(Joux, ANST IV 2000, LNCS, Springer)



- $G_1 = \langle P \rangle$ additive, $G_2$ multiplicative group of a large prime order $q$, $\hat{e} : G_1 \times G_1 \rightarrow G_2$ the bilinear map

- security: hardness of BDH problem.
- BDH (Bilinear Diffie-Hellman) Problem in $\langle G_1, G_2, e\rangle$:

  given $\langle P, aP, bP, cP\rangle$ for some $a, b, c \in Z_q^*$, compute $e(P, P)^{abc}$.

## Identity-Based Key Agreement Scheme

- Sakai-Ohgishi-Kasahara (SOK) Key Agreement

## Identity-Based Encryption (IBE) Scheme

- Shamir [1984]
- public key $ID \in \{0,1\}^*$
- An ID-based encryption scheme has four algorithms.

  1. *Setup:* Creates system parameters and *master key*.
  2. *Extract:* Uses master key to generate the private key corresponding to an arbitrary public key string ID.
  3. *Encrypt:* Encrypts messages using the public key ID.
  4. *Decrypt:* Decrypts the message using the corresponding private key of ID.

- Motivation for ID-based encryption : to simplify certificate management in e-mail system, no need to keep a large database for public keys, system derives the public keys by user names.
- We will describe Boneh-Franklin's IBE Scheme

# Boneh-Franklin [2001] IBE Scheme

1. *Setup*: $params = \langle G_1, G_2, \hat{e}, P, q, n, P_{pub}, H_1, H_2 \rangle$,
   $P_{pub} = sP$, $s$ is the master key.
   message space $\mathcal{M} = \{0, 1\}^n$.

2. *Extract*: $Q_{ID} = H_1(ID) \in G_1$, $d_{ID} = sQ_{ID}$.

3. *Encrypt*: $r \in Z_q^*$, $m \in \mathcal{M}$,

$$C = \langle rP, m \oplus H_2(\hat{e}(Q_{ID}, P_{pub})^r) \rangle$$

4. *Decrypt*: $C = \langle U, V \rangle$

$$V \oplus H_2(\hat{e}(d_{ID}, U))$$

# IND-ID-CCA Security Model for IBE Scheme

- The IND-ID-CCA model deals with an adversary who possesses private keys corresponding to identities of its choice $\mathsf{ID}_1, \ldots, \mathsf{ID}_n$ and attacks an identity ID in an ID-based system. Consider the following game between the challenger and an adversary $\mathcal{A}$.
  - **Setup:** The challenger takes a security parameter $k$ and runs the *Setup* algorithm. It gives the adversary the resulting system parameters *params* and keeps the master key secret to itself.

– **Phase 1:** The adversary issues queries $q_1, \ldots, q_m$ where $q_i$ is one of the following two queries.

(a) *Extraction query* $\langle \mathsf{ID}_i \rangle$. The challenger responds by running algorithm *Extract* to generate the private key $d_i$ corresponding to the public key $\mathsf{ID}_i$. It then sends $d_i$ to the adversary.

(b) *Decryption query* $\langle \mathsf{ID}_i, C_i \rangle$. The challenger responds by running algorithm *Extract* to generate the private key $d_i$ corresponding to $\mathsf{ID}_i$. It then runs algorithm *Decrypt* to decrypt the ciphertext $C_i$ using the private key $d_i$. It sends the resulting plaintext to the adversary.

These queries may be asked adaptively, that is, each query $q_i$ may depend on the replies to $q_1, \ldots, q_{i-1}$.

- **Challenge:** Once the adversary decides that Phase 1 is over, it outputs two equal length plaintexts $M_0, M_1$, an identity $\mathsf{ID}$ on which it wishes to be challenged. The only constraint is that $\mathsf{ID}$ did not appear in any private key extraction queries in Phase 1. The challenger picks a random bit $b \in \{0,1\}$ and sets $C = Encrypt(params, \mathsf{ID}, M_b)$. It sends $C$ as the challenge to the adversary.

- **Phase 2:** The adversary issues more queries $q_{m+1}, \ldots, q_n$ where $q_i$ is one of the following two queries.
  (a) *Extraction query* $\langle \mathsf{ID}_i \rangle$ *where* $\mathsf{ID}_i \neq \mathsf{ID}$. Challenger responds as in Phase 1.
  (b) *Decryption query* $\langle \mathsf{ID}_i, C_i \rangle \neq \langle \mathsf{ID}, C \rangle$. Challenger responds as in Phase 1.
- **Guess:** Finally, the adversary outputs a guess $b' \in \{0, 1\}$ and wins the game if $b = b'$.

- We refer such an adversary $\mathcal{A}$ as an IND-ID-CCA adversary.
- The advantage for the adversary $\mathcal{A}$ in attacking the scheme is defined as

$$\mathsf{Adv}(\mathcal{A}) = |Pr[b = b'] - \frac{1}{2}|.$$

- We say that an identity based scheme is secure against an adaptive chosen ciphertext attack (IND-ID-CCA) if no polynomially bounded adversary $\mathcal{A}$ has non-negligible advantage against the challenger.

- The IND-ID-CCA model is the *strongest* acceptable notion of security and the security model for other public key encryption schemes

- Security of Boneh-Franklin's (BF) IBE scheme depends on hardness of BDH problem in $\langle G_1, G_2, \hat{e} \rangle$.

- BDH (Bilinear Diffie-Hellman) Problem in $\langle G_1, G_2, \hat{e} \rangle$:
  - given $\langle P, aP, bP, cP \rangle$ for some $a, b, c \in Z_q^*$, compute $\hat{e}(P, P)^{abc}$.

- This version of BF scheme is IND-ID-CPA secure in the random oracle model.

- **Fujisaki-Okamoto transformation** has been used to extend IND-ID-CCA security.
- Let $\mathcal{E}$ be a probabilistic public key encryption scheme. We denote by $\mathcal{E}_{pk}(M; r)$ the encryption of $M$ using the random bits $r$ under the public key $pk$.
- Fujisaki-Okamoto define the hybrid scheme $\mathcal{E}^{hy}$ as:

$$\mathcal{E}^{hy}_{pk}(M) = \mathcal{E}_{pk}(\sigma; H_3(\sigma, M)), H_4(\sigma) \oplus M$$

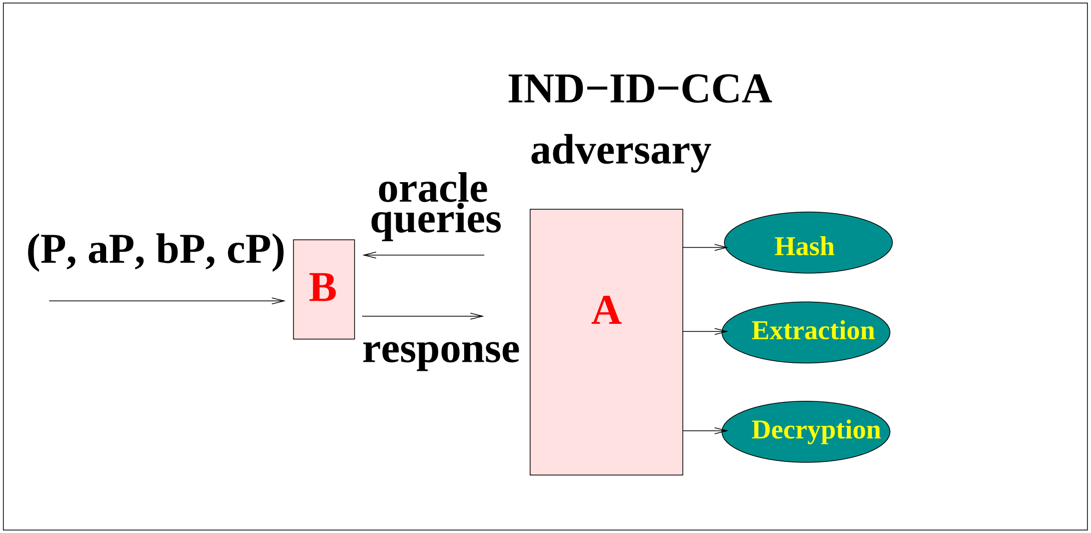
where $\sigma$ is generated at random and $H_3$, $H_4$ are cryptographic hash functions.

- Fujisaki-Okamoto (FO) show that if $\mathcal{E}$ is an IND-CPA secure encryption scheme, then $\mathcal{E}^{hy}$ is an IND-CCA secure encryption scheme in the random oracle model.
- So after applying FO transformation, the resulting ciphertext is

$$C = \langle rP, \sigma \oplus H_2(\hat{e}(Q_{ID}, P_{pub})^r), m \oplus H_4(\sigma) \rangle$$

where $r = H_3(\sigma, m)$.

# Security proof



- $\mathcal{B}$ simulates a real attack environment for $\mathcal{A}$

- $\mathcal{B}$ maintains lists for $H_1, H_2, H_3, H_4$
- $\mathcal{A}$ outputs a guess bit while $\mathcal{B}$ outputs from $H_2$-list
- $\mathcal{A}$ guesses correct bit means $H_2$-query is made on $D = \hat{e}(P, P)^{abc}$ by $\mathcal{A}$ at some point during the simulation
- Thus $D$ appears somewhere in $H_2$-list and $\mathcal{B}$ outputs $D$ with non-negligible probability (solving BDH problem)

# Random Oracle Model

- Assume that all random values are indeed random
- Assume Adversary does not exploit any properties of the hash function
- Assume that hash functions behave idealistically (random public functions)
- Assumptions reduce the strength of a proof

- A random oracle is a function $H : X \to Y$ chosen uniformly at random from the set of all functions $\{h : X \to Y\}$ (we assume Y is a finite set). An algorithm can query the random oracle at any point $x \in X$ and receive the value $H(x)$ in response. Random oracles are used to model cryptographic hash functions such as SHA-1.
- Note that security in the random oracle model does not imply security in the real world. Nevertheless, the random oracle model is a useful tool for validating natural cryptographic constructions.

- Boneh and Boyen gave selective ID model, namely IND-sID-CCA, which is slightly weaker than the model described above. In this model the adversary must commit ahead of the time to the identity that it intends to attack, whereas in the standard model described earlier, the adversary is allowed to choose this identity adaptively.

# Boneh-Boyen's IBE Without Random Oracle

- Protocol Description :
  - *Setup* : The public keys (ID) are assumed to be elements of $Z_q^*$ and messages are elements of $G_2$. Select random elements $x, y \in Z_q^*$ and set $U = xP, V = yP$. The public parameters are $(U, V)$ and the master key is $(x, y)$.
  - *Extract* : Given a public key $\mathsf{ID} \in Z_q^*$, choose a random $r \in Z_q^*$ and compute $K = \frac{1}{\mathsf{ID}+x+ry} P \in G_1$. Output the private key $S_{\mathsf{ID}} = (r, K)$.

- *Encrypt* : To encrypt a message $M \in G_1$ under public key $\mathsf{ID} \in Z_q^*$, pick a random $s \in Z_q^*$ and output the ciphertext $C = \langle s(\mathsf{ID})P + sU, sV, e(P,P)^s\ M\rangle$.
- *Decrypt* : To decrypt a cipher text $C = \langle X, Y, Z\rangle$ using the private key $S_{\mathsf{ID}} = (r, K)$, output $Z/e(X + rY, K)$.

- **Security** : Secure against selective-ID adaptive chosen ciphertext attack (IND-sID-CCA) without random oracles under $q$-DBDHI assumption.

# $k$-DBDHI Problem

- The $k$-Decisional Bilinear Diffie-Hellman Inversion ($k$-DBDHI) problem in $\langle G_1, G_2, \hat{e} \rangle$ :

  $Instance:$ $(P, yP, y^2P, \ldots, y^kP, r)$ for some $y \in Z_q^*$, $r {\in_R} G_2$.

  $Output:$ yes if $r = \hat{e}(P,P)^{\frac{1}{y}} \in G_2$ and output no otherwise.

# IND-sID-CCA Security Model for IBE Scheme

- Consider the following game:
  - **Init:** The adversary outputs an identity $\mathsf{ID}$ where it wishes to be challenged upon.
  - **Setup:** The challenger runs the *Setup* algorithm. It gives the adversary the resulting system parameters em params. It keeps the master-key to itself.

- **Phase 1:** The adversary (adaptively) issues queries $q_1, \ldots, q_m$ where $q_i$ is one of the following two queries.

  (a) *Extraction query* $\langle \mathsf{ID}_i \rangle$ *where* $\mathsf{ID}_i \neq \mathsf{ID}$. The challenger responds by running algorithm *Extract* to generate the private key $d_i$ corresponding to the public key $\mathsf{ID}_i$. It then sends $d_i$ to the adversary.

  (b) *Decryption query* $\langle C_i \rangle$ *for* $\mathsf{ID}$ *or any prefix of* $\mathsf{ID}$. The challenger responds by running algorithm *Extract* to generate the private key $d$ corresponding to $\mathsf{ID}$. It then runs algorithm *Decrypt* to decrypt the ciphertext $C_i$ using the private key $d$. It sends the resulting plaintext to the adversary.

- **Challenge:** The adversary outputs two equal length plaintexts $M_0, M_1$. The challenger sends to the adversary $C = Encrypt(\mathsf{ID}, M_b)$, $b \in_R \{0,1\}$.
- **Phase 2:** The adversary issues more queries $q_{m+1}, \ldots, q_n$ where $q_i$ is one of the following two queries.
  (a) *Extraction query* $\langle \mathsf{ID}_i \rangle$ *where* $\mathsf{ID}_i \neq \mathsf{ID}$. Challenger responds as in Phase 1.
  (b) *Decryption query* $\langle C_i \rangle \neq \langle C \rangle$. Challenger responds as in Phase 1.
- **Guess:** Finally, the adversary outputs a guess $b' \in \{0,1\}$ and wins the game if $b = b'$.

# Sahai-Water's Fuzzy IBE

1. $Setup(U, d)$: $\mathsf{params} = \langle p, G, G_T, e, g\rangle$, $|G| = p$, $G = \langle g\rangle$
   - $\mathcal{U} = \{1, 2, \ldots, n\}$, universe of attributes
   - $d$, system threshold
   - $\alpha \in_R Z_p, Y = e(g, g)^\alpha$
   - $t_i \in_R Z_p, T_i = g^{t_i}$, for each $i \in \mathcal{U}$
   - $\mathsf{PK} = (p, e, G, G_T, T_1, T_2, \ldots, T_n, Y)$
   - $\mathsf{MK} = (\alpha, t_1, t_2, \ldots, t_n)$
2. $KeyGen(\mathsf{PK}, \mathsf{MK}, L)$: $L \subset \mathcal{U}$, $\mathsf{MK} = (\alpha, t_1, t_2, \ldots, t_n)$
   - $q(x)$, a polynomial of degree $d - 1$ with $q(0) = \alpha$

- $d_i = g^{\frac{q(i)}{t_i}}$ for each $i \in L$
- $\mathsf{SK}_L = \{L, \{d_i\}_{i \in L}\}$

3. *Encrypt*$(\mathsf{PK}, W, m)$: Attribute set $W$ of receivers, message $m \in G_T$, $\mathsf{PK} = (p, e, G, G_T, T_1, T_2, \ldots, T_n, Y)$
   - $C = m\, Y^s, C_i = T_i^s$ for each $i \in W$
   - ciphertext $\mathsf{CT}_W = [W, C, \{C_i\}_{i \in W}]$
4. *Decrypt*$(\mathsf{PK}, \mathsf{SK}_L, \mathsf{CT}_W)$: $S = W \cap L$
   - if $|S| < d$, then decryption fails

– if $|S| \geq d$, then

$$m = \frac{C}{\Pi_{i \in S}\, e(d_i, C_i)^{\Delta_{i,S}(0)}}$$

$$\Delta_{i,S}(x) = \prod_{j \in S, j \neq i} \frac{x - j}{i - j}$$

– if $|S \geq d$, then

$$q(x) = \sum_{i \in S} \Delta_{i,S}(x)\; q(i)$$

**Correctness:** $|S| = |W \cap L| \geq d$

$$\frac{C}{\Pi_{i \in S}\, e(d_i, C_i)^{\Delta_{i,S}(0)}} = \frac{m\ e(g,g)^{\alpha s}}{\Pi_{i \in S}\, e(g^{\frac{q(i)}{t_i}}, T_i^s)^{\Delta_{i,S}(0)}}$$

and

$$\prod_{i \in S} e(g^{\frac{q(i)}{t_i}}, T_i^s)^{\Delta_{i,S}(0)} = e(g,g)^{s \sum_{i \in S} \Delta_{i,S}(0)\ q(i)}$$

$$= e(g,g)^{sq(0)} = e(g,g)^{s\alpha}$$

**Security** : Simantically secure assuming DMBDH problem is hard.

# DMBDH Problem

(Decisional Modified Bilinear Diffie-Hellman)

$Instance : \langle g, g^g, g^y, g^z, e(g,g)^\theta \rangle$

$Output :$ yes if $\theta = \frac{xy}{z}$ and output no otherwise.

## Identity-Based Signature (IBS) Scheme

- Shamir's IBS
- Sakai-Ohgishi-Kasahara (SOK) IBS