DISMA Dipartimento di Scienze Matematiche G. L. Lagrange

Gruppo di Crittografia e Teoria dei Numeri

Seminario divulgativo della serie

CRITTOGRAFIA: dalla teoria alle applicazioni

ATTRIBUTE-BASED ENCRYPTION AND SIGNATURES

Prof. Jovan Golic

10 Dicembre 2018 – ore 10:00

Aula Buzano

Dipartimento di Scienze Matematiche

Politecnico di Torino

POLITECNICO DI TORINO

Telsy

Attribute-Based Encryption and Signatures

Jovan Golic

Senior Technical Leader, Security Lab, Telecom Italia

Chairman of NGMN Security Competence Team

NGMN – Next Generation Mobile Networks

Former EIT Digital Action Line Leader for Privacy, Security & Trust

EIT – European Institute of Innovation & Technology

Politecnico di Torino, Torino, December 10, 2018

Data Security

- Data confidentiality – *data intelligible only to desired entities*
- Data integrity – *data received/retrieved in original form*
- Data availability – *data available on request*
- Entity authentication/identification – *persons, organizations, and things creating, sending, and receiving data have authentic identity attributes in real time*

- *Main techniques: encryption for data confidentiality, message authentication codes or digital signatures for data integrity, redundant resources and business continuity agreements for data availability, and user credentials and authentication protocols for entity authentication*

Data Privacy

- **Control:** User should be able to fully control usage of *sensitive data* during the whole life cycle: personal data and metadata, industrial secrets, etc. (e.g., via user consent)
- **Security:** *Sensitive data* should be securely collected, transmitted, stored, processed, shared, and deleted
- **Anonymity:** Untraceability of *identity attributes* + unlinkability
- **Minimality:** *Sensitive data* should be disclosed to a minimum possible extent for a minimum period of time only to entities authorized by the user and to be used only for purposes authorized by the user
- **Important business opportunities**
- Reconciliation with lawful cyber surveillance & investigation
- Mass cyber surveillance and citizen profiling clash with fundamental human rights (no privacy implies no freedom)

Advanced Cryptography for Data Privacy

- ***Secret sharing*** *(no single points of trust and failure)*
- ***Secure multiparty computation*** *(joint computation of functions without disclosing own data)*
- ***Practical homomorphic encryption*** *(processing of encrypted data, e.g., in the cloud)*
- ***Privacy-preserving profiling*** *(without revealing user data, not only pseudonymization and data aggregation)*
- ***Anonymity protocols*** *(e.g., anonymous authentication and signatures, possibly with revocable anonymity)*
- ***Attribute-based encryption*** *(cloud data sharing by applying access policies on encrypted data)*
- ***Searchable encryption*** *(search over encrypted data)*
- ***End2end encryption*** *(possibly, with key escrow – secret sharing & threshold cryptography for lawful interception)*

Attributes

- **Attribute is any verifiable physical or logical property of any entity (e.g., person, organization, thing, service, application, data, multimedia content etc.) which itself is described as a string of characters or bits**
- *For example, attributes can relate to age, role, education, profession, expertise, affiliation, medical data, financial data, individual preferences, orientations etc.*
- Unlike identities, which usually uniquely specify an entity in a group of entities, attributes are more general and are usually considered in various combinations
- *In order to be used electronically, attributes need to be specified and certified by attribute certificates, issued (and digitally signed) by certification authorities upon verification*

Attribute-Based Access Control

- Attribute-based access control relates to access to assets including various data, such as stored files, sensitive data, broadcasted multimedia content, log data, medical records, financial data, data generated by meters/sensors in IoT etc.
- Binary attributes as properties are either satisfied or not; as such, they can be expressed as binary (Boolean) variables
- Non-binary (numerical) attributes can be expressed in terms of multiple binary attributes, e.g., corresponding to integer values or ranges of values; generalizations (?)
- Access policy or access structure is expressed as a predicate (i.e., Boolean function) over binary attributes; access is allowed if the access policy has value 1, and is denied if it has value 0; e.g., monotone structure can be expressed in terms of AND, OR, and threshold gates; non-monotone structure also includes NOT gates; *fine-grained access control*

Access to Sensitive Data

- ***Sensitive data*** *to be accessed or shared may include:*
  - ***Sensitive database*** *files stored on untrusted servers (e.g., personal medical records, IoT data produced by smart meters/sensors or personal wearables, financial data)*
  - ***Pay-per-view*** *multimedia content (e.g., sent out in broadcast)*
  - ***Audit log*** *data to be accessed by authorized entities*
- Attributes can relate to entities accessing data or to data itself (e.g., type of multimedia content and other metadata, type of log data, type of document, etc.)
- Access policy can relate both to entity attributes, typically supported by attribute certificates, and data attributes, specified by entities generating or providing data
- *For data protection and access control, data are segmented and encrypted according to attributes or access policies*

Examples

- An access policy A∨(B∧C) is satisfied if A=1 or (B=1 and C=1)
- An access policy $tr_{(2,4)}$ (A,B,C,D) is satisfied if at least two of four attributes have value 1 (threshold gate)
- **Personal medical records (PMRs):** labeled by types of medical data and expertise and/or affiliation of medical practitioners; access policy defines which practitioners can access given data
- **Pay-per-view broadcast content:** labeled by metadata describing content, such as genre, time (e.g., year, month, date), episodes, seasons, actors, directors; access policy for a given user defines subscription package over metadata
- **Audit log data:** database entries labeled and indexed by metadata as attributes, where access policy (e.g., for forensic analysis) is defined over metadata; e.g., NetFlow logs include IP packet payloads labeled by Netlow information (source/destination IP address, source/destination port number etc.)

Classical Encryption

- Assume that data segments are encrypted by different (symmetric) keys, called **data keys**
- Then these **data keys** need to be distributed to users satisfying the respective access policies, preferably in real time, along with the encrypted data/files
- **If classical public-key or symmetric-key cryptographic infrastructure** is used for real-time transmission of encrypted **data keys** to users, upon verifying the access policy, then:
  - *Communication complexity is linear in number of users (even in broadcast), due to user-specific transmission keys*
  - *Pre-distributed transmission keys from classical cryptographic infrastructure are bound neither to attribute certificates of users and data attributes nor to respective access policies; this renders the system impractical if access policies and/or attributes are dynamic, so that data keys need to change*

Attribute-Based Encryption

- *If Attribute-Based Encryption (ABE) is used for real-time transmission of encrypted data keys to users, then each data key needs to be encrypted only once by using unique (global) ABE public key, and each user satisfying the access policy can decrypt by using its own ABE private decryption key*
- ABE encryption essentially consists in masking the data key with a mask, to be recovered by authorized decryption keys
- *Size of ABE encrypted data key and hence communication complexity are linear in the number of attributes, not users*
- Attributes and access policies are built in combined ABE encryption and decryption functions, so that only the users having the right private decryption keys, which satisfy the access policies for given data, can decrypt and hence obtain the data key and the data; **inherent flexibility** allowing dynamic access policies and/or attributes

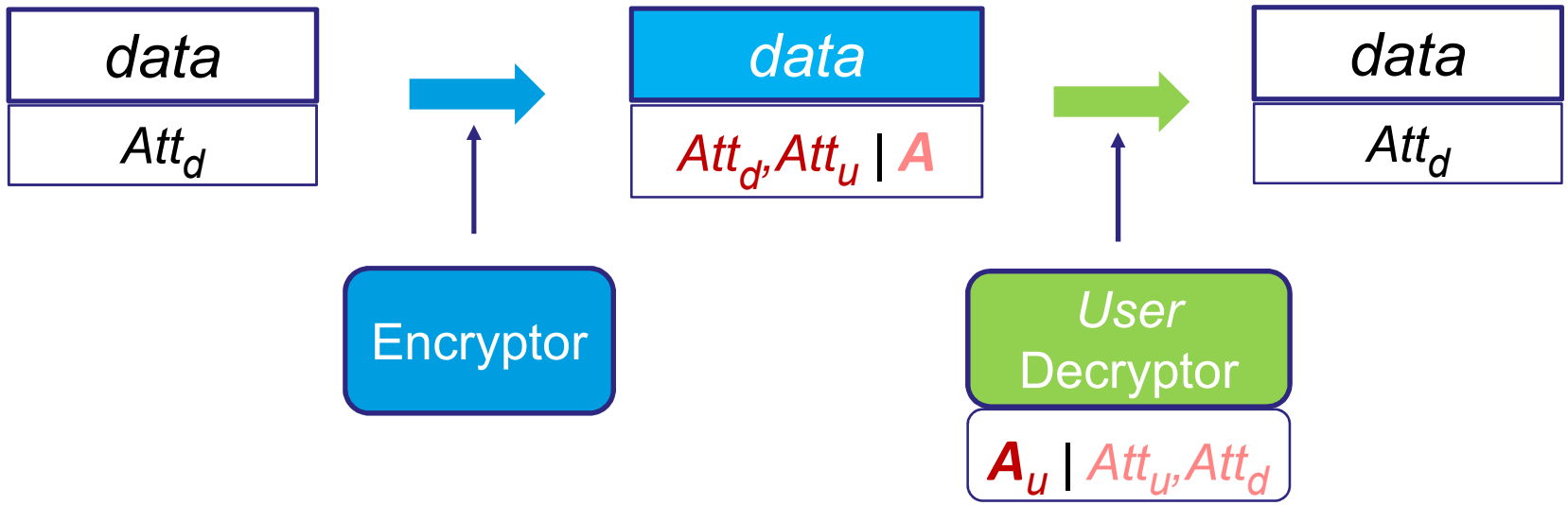
Basic Concepts of ABE (1)

- In ABE system, there is a single public encryption key PK and multiple private decryption keys DK
- **Setup:** Key-management server (KMS) generates PK and its secret master key MK
  - KMS verifies validity of user attributes, either directly or by verifying user attribute certificates
  - By using PK and MK, it generates randomized private decryption keys DK of users, initially or in real time, and distributes them to users over protected channels (e.g., classical PKI)
- In **key-policy ABE (KP-ABE)**, access policies are built in private decryption keys, and an attribute set in ciphertext
- In **ciphertext-policy ABE (CP-ABE)**, attribute sets are built in private decryption keys, and an access policy in ciphertext
- (An attribute set is a set of satisfied attributes)

Basic Concepts of ABE (2)

- In **KP-ABE**, decryption with DK works if and only if an attribute set built in ciphertext satisfies the access policy built in DK *(attribute set can vary with each encryption)*
- In **CP-ABE**, decryption with DK works if and only if an access policy built in ciphertext is satisfied by the attribute set built in DK *(access policy can vary with each encryption)*
- Encryption time & ciphertext size are linear in the number of attributes involved (in both KP-ABE and CP-ABE)
- *Randomized DK are a basis for collusion resistance*
- ***Collusion resistance:*** *Given any number of valid DKs, it is infeasible to decrypt any ciphertext for any new attribute set (in CP-ABE) or new access policy (in KP-ABE)*
- *Randomized encryption implies that one cannot distinguish repeated encryptions of the same message (privacy)*

Basic Concepts of ABE (3)



KP-ABE: attributes $Att_d, Att_u \rightarrow$ ***Enc***

user access policy $A_u = A_u(Att_d, Att_u) = 1 \rightarrow$ ***Dec***

flexible: encryptor can vary attributes, enabling/disabling A_u

CP-ABE: global access policy $A = A(Att_u, Att_d) = 1 \rightarrow$ ***Enc***

attributes $Att_u, Att_d \rightarrow$ ***Dec***

more flexible: encryptor can vary access policy A

Examples

- **CP-ABE:**
  - Encrypt a PMR with embedded access policy: A="doctor with name1" OR B="doctor specialized in area1" AND (C="doctor with at least 10 years of practice" OR D="doctor employed at hospital1")
  - Can decrypt: {A}, {B,C}, {B,D}, {A,B,C}, {A,E}
  - Cannot decrypt: {B}, {C}, {B,F}, {C,E}

- **KP-ABE:**
  - Encrypt a broadcast content with attributes: A="genre1", B="actor1", C="actor2", D="director1", E="production year1"
  - Can decrypt: A, B ∨ C, D ∧ E, A ∨ E, A ∨ F
  - Cannot decrypt: A ∧ F, D ∧ ¬E, B ∧ ¬C, F

Practical Aspects

- *In basic KP-ABE, size of PK is linear in the total number of possible attributes U; in large universe KP-ABE, PK size is linear in the maximum number of attributes effectively used in encryption; in random oracle large universe construction with hash function, PK is fixed in size*
- In CP-ABE, PK size can be fixed or linear in U
- KP-ABE is more suited to data attributes, while CP-ABE is more suited to user attributes
- Expiry or revocation of decryption keys (e.g., due to expiry of embedded attributes or access policies or to key compromise) can be dealt with by including the expiry time/date among the attributes, by periodic refreshing, and by revocation lists
- *In principle, users can cheat by sharing private decryption keys (and attribute certificates if required)*

Attribute-Based Signatures

- Ensure data integrity with non-repudiation, by using private signing keys and public verification keys
- *In Attribute-Based Signature (ABS) scheme, an entity with a private signing key can generate a valid signature with an access policy embedded in the signature if and only if the attribute set embedded in the signing key satisfies this access policy (analogy with CP-ABE)*
  - Unique (global) public key for verification, multiplicity of private signing keys distributed by KMS
  - Unlinkability by randomization of signing, except for the same access policy shared by all valid signatures
  - There exist ABS schemes that can use the same keys as CP-ABE, and such a combined scheme is denoted as ABES; preferable with respect to key management
  - ABS can also be generated from anonymous credentials

Mathematical Hints

- Generalization of Identity-Based Encryption (IBE)
- Made possible by efficiently computable bilinear maps
  - Cyclic multiplicative groups *G* and *H* of prime order *p*, generator *g*

$$e\colon G \times G \to H, \quad e(x^a, y^b) = e(x, y)^{ab}, \quad e(g, g) \neq 1$$

- *In practice, bilinear maps are based on elliptic curve groups – elliptic curve pairings*
  - Additive group: exponentiation – point multiplication on elliptic curve
  - Computations are operations in underlying finite field and pairings
- ***Security reduction to Decisional Bilinear DH Assumption***
- Usage of secret sharing for embedding access structures in private decryption keys (KP-ABE) or ciphertexts (CP-ABE)
- Encryption time of less than 1 ms per attribute can be achieved
- **KP-ABE Goyal, Pandey, Sahai, Waters 2006; Ostrovsky, Sahai, Waters 2007**
- **CP-ABE Bethencourt, Sahai, Waters 2007**

IMSI Encryption in 5G

- Subscriber identity IMSI is the (public) globally unique long-term identity associated with the secret subscriber credential: **IMSI = MCC, MNC, MSIN**
- *Needed for establishing connection, by enabling the mobile operator to find the subscriber credential (pre-shared symmetric key) for authenticating the user to the network*
- Used by passive or active IMSI catchers for location tracking
- In 5G, IMSI will be encrypted for privacy protection
- **3GPP TS 33.501 (Release 15)** mandates the support of the encryption of MSIN only, by using randomized public-key encryption with the public key of the home mobile operator *(for simplicity, only partial privacy)*
- Visited network cannot decrypt, but receives IMSI from home network afterwards, for lawful interception

IMSI Encryption by ABE

- *For full privacy, the whole IMSI needs to be encrypted*
- Classical public-key encryption is an option, but requires a globally trusted PKI for public-key certificates as well as broadcast of public-key certificates to endpoint devices (UEs)
- In **ETSI TS 103.458** (2018-06) and **3GPP TR 33.899** (2017-08), **Telecom Italia** proposed an ABE method for the encryption of the whole IMSI for full privacy:
  - Upon request, each mobile operator receives ABE decryption key from KMS, with embedded MCC, MNC and expiry time
  - UE receives in broadcast MCC, MNC of visited network as well as expiry time and uses them as an attribute for ABE (CP or KP)
  - Large universe construction for fixed-size PK
  - Global KMS can be implemented distributedly, to be more trustworthy, by using secret sharing and threshold cryptography