

FUNCTIONAL ENCRYPTION, AN OVERVIEW



CifrisCloud

24th March 2021

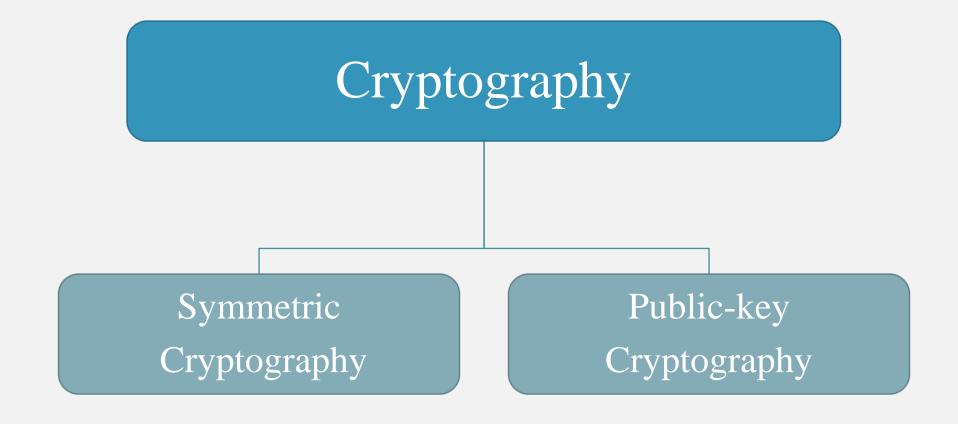
Irene Villa and Carla Mascia,

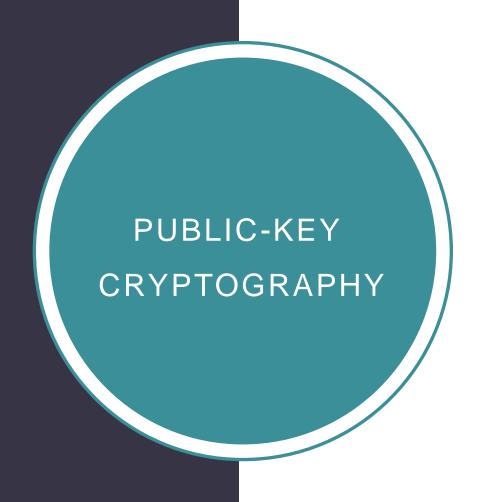
University of Trento



OUTLINE

- 1. Introduction to Functional Encryption (FE)
- 2. Security notions and mathematical assumptions
- 3. Predicate Encryption (PE)
- 4. PE with public index
- 5. PE with private index
- 6. FE beyond predicates
- 7. Other features





Introduced in 1976 by Diffie and Hellman.

Later on, Rivest, Shamir and Adleman presented RSA.

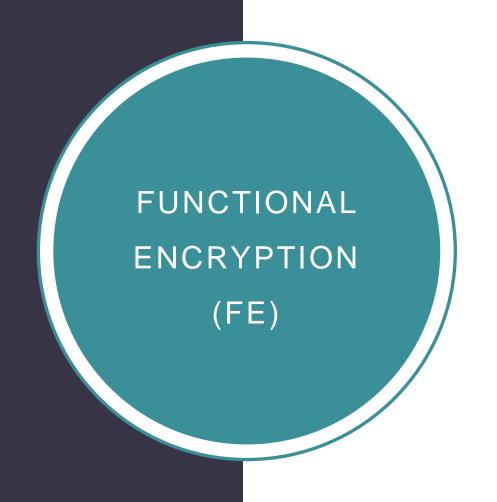
It is now used in secure web communication (e.g. SSH, SSL), Cloud services, ...

- Two parties can securely communicate (over an insecure channel) with each other *without* having an a prior mutual secret.
- A pair of keys (pk, sk) is used:
 - * pk is <u>public</u> and it is used for encryption,
 - sk is <u>private/secret</u> and it is used for decryption.
- Access to encrypted data is **all-or-nothing**.
- A single secret key can decrypt the data.

For many emerging applications such as "cloud" services the notion of PKE is not enough.

For example:

- Spam filtering on encrypted emails
- Expressive access control
- useful in large organisations: corporate, health care, government institutions, ...
- Mining large datasets
- Search over encrypted data
- Perform analysis over the dataset (compute the mean of some transactions)



Introduced in 2005 by Sahai and Waters.

Formalised in 2010 by Boneh, Sahai and Waters.

FE allows:

- a <u>fine-grained access control</u> of encrypted data;
- to learn (only) a specific <u>function</u> of the encrypted data, but nothing else about the data.

FUNCTIONALITY

A **functionality** F defined over (K, X) is a function $F : K \times X \to \Sigma \cup \{\bot\}$ described as a (deterministic) Turing Machine.

The set *K* is called the **key space** and the set *X* is called the **plaintext space**.

The key space contains a special key called the **empty key** denoted by ε .

The functionality F describes the functions of a plaintext that can be learned from the ciphertext. The symbol \bot denotes the failure of the algorithm.

A **functional encryption scheme** for a functionality F defined over (K, X) is a tuple of four PPT algorithms (Setup, KeyGen, Enc, Dec):

- 1. Setup: generate a public key pk and a master secret key mk pair from a security parameter λ ;
- 2. KeyGen: taking in input mk and a key $k \in K$, generate a secret key sk_k ;
- 3. Enc: taking in input pk and a plaintext $x \in X$, generate a ciphertext c;
- 4. Dec: taking in input a ciphertext c and a secret key sk_k return $y \in \Sigma \cup \{\bot\}$.

Correctness

If $sk_k = KeyGen(mk, k)$ and c = Enc(pk, x), then $y = Dec(sk_k, c) = F(k, x)$ with probability 1 (minus a negligible function).



AUTHORITY

 $\mathsf{Setup}(1^\lambda) \to (\mathsf{pk}, \mathsf{mk})$

 $KeyGen(mk,k) = sk_k$

BOB

 sk_k

 $Dec(sk_k, c) = F(k, x)$

ALICE

pk

Cloud

Enc (pk, x) = c

THE SIMPLEST EXAMPLE OF FE

The **standard public-key cryptography** is the simplest example of FE.

Set $K = \{1, \varepsilon\}$ and

$$F(k,x) = \begin{cases} x & k=1\\ |x| & k=\varepsilon \end{cases}$$

where $|\cdot|$ denotes the bit-length function.

The empty key $\varepsilon \in K$ captures all the information about the plaintext that the ciphertext intentionally reveals (e.g., the length of the encrypted plaintext).



Main goal is to prevent **collusion attacks** (attacks that make use of multiple functional secret keys)

Example:

A user can decrypt a ciphertext Enc(pk, x) only if he owns the following attributes:

"over 30" and "Italian citizen".

Mario has the following attributes: "male" and "Italian citizen" Luigi has the following attributes: "over 30" and "San Marino citizen"

Even if Mario and Luigi collude (combine their secret keys), they should not be able to read the encrypted message

INDISTINGUISHABILITY-BASED SECURITY

Game b (random bit) for an adversary A:

- 1. Setup: Run Setup and give pk to A
- 2. Query1: \mathcal{A} adaptively submits queries for $k_i \in K$ (i = 1, ..., n) and is given $sk_i = KeyGen(mk, k_i)$
- 3. Challenge: \mathcal{A} submits two messages $m_0, m_1 \in X$ satisfying

$$F(k_i,m_0) = F(k_i,m_1) \ \, \forall \, i=1,\ldots,n$$
 and is given Enc(pk, m_b)

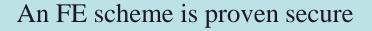
- 4. Query2: \mathcal{A} continues to issue key queries as before (subjected to the same restriction)
- 5. Guess: \mathcal{A} eventually outputs a bit b'

SIMULATION-BASED SECURITY

The view of an adversary can be simulated by a simulator, given access only to secret keys and functions evaluated on the corresponding messages.

Let \mathcal{A} be an adversary that takes as input the public key pk, a set of secret keys $sk_1, ..., sk_\ell$ for keys $k_1, ..., k_\ell$ of its choice ($sk_i = \text{KeyGen}(mk, k_i)$), and a ciphertext Enc(pk, x).

We assume that there exists an algorithm \mathcal{B} , called *simulator*, that given pk and $F(k_1, x), \dots, F(k_\ell, x)$ is able to output the same information about x that \mathcal{A} outputs.





A certain mathematical problem is "hard" to solve (assumption)

Assumptions used to prove security of FE schemes are:

- Different problems based on bilinear groups (pairings);
- Learning with Errors problem;
- Quadratic residuosity, Multivariate Quadratic polynomial problem, ...



BILINEAR GROUPS

A **bilinear group** is a tuple $(p, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, e)$ such that:

- \mathbb{G}_1 , \mathbb{G}_2 , \mathbb{G}_T are (multiplicative) cyclic groups of order a prime p;
- $e: \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_T$ is bilinear, that is, $e(g_1^a, g_2^b) = e(g_1, g_2)^{ab}$ for all $g_1 \in \mathbb{G}_1$, $g_2 \in \mathbb{G}_2$, $a, b \in \mathbb{Z}$,
- *e* is an *admissible* bilinear map:
 - e is efficiently computable,
 - *e* is non-degenerate.

Most used bilinear maps are the <u>(modified) Weil Pairing</u> and the <u>(modified) Tate Pairing</u> defined over the points of an elliptic curve.

SOME PAIRING-BASED PROBLEMS

The **Decisional Diffie-Hellman (DDH) problem** for \mathbb{G}_1 is to distinguish g^{ab} from g^c given (g, g^a, g^b) for random $a, b, c \in \mathbb{Z}_p^*$ and $g \in \mathbb{G}_1^*$.

If the Decisional Bilinear Diffie-Hellman (DBDH) problem for $(q, \mathbb{G}_1, \mathbb{G}_T, e)$ (here $\mathbb{G}_2 = \mathbb{G}_1$) is to

distinguish $e(g,g)^{abc}$ from a random $T \in \mathbb{G}_T$ given (g,g^a,g^b,g^c) for random $a,b,c \in \mathbb{Z}_p^*$ and $g \in \mathbb{G}_1^*$.

The **Decisional Linear (DLIN) problem** for $(q, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, e)$ is to distinguish g^{c+d} from a random $T \in \mathbb{G}_1$, given the tuple $(g, g^a, g^b, g^{ac}, g^{bd}, h, h^a, h^b)$ for random $a, b, c, d \in \mathbb{Z}_p^*, g \in \mathbb{G}_1^*$ and $h \in \mathbb{G}_2^*$.

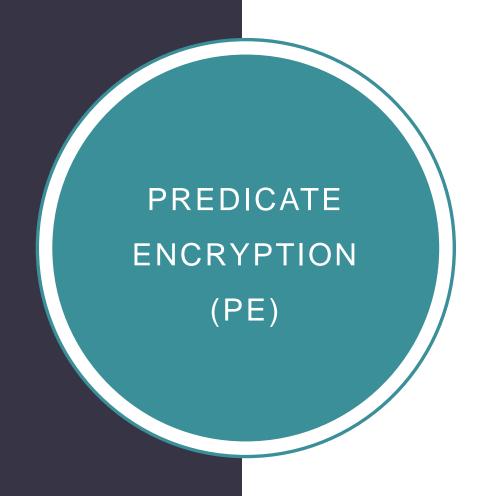
LEARNING WITH ERRORS (LWE)

For an integer $q = q(n) \ge 2$ and an error distribution $\chi = \chi(n)$ over \mathbb{Z}_q ,

the (decisional) LWE problem is to distinguish between the following pairs of distributions:

$${A, As + x}$$
 and ${A, u}$

where $A \in \mathbb{Z}_q^{n \times m}$, $s \in \mathbb{Z}_q^n$, $x \in \chi^m$ and $u \in \mathbb{Z}_q^m$ are all chosen at random.



PE is a subclass of functional encryption that enables fine-grained access to encrypted information such as searching on encrypted data.

In a PE scheme:

A secret key decrypts correctly



information embedded into the ciphertext satisfies a predicate embedded into the secret key.

PREDICATES AND QUERIES

A **predicate** is a Boolean-valued function $P: A \to \{0,1\}$, where A is an arbitrary set, and the elements 0 and 1 are interpreted as the logical values false and true, respectively.

With a predicate P we can perform different queries. Let $a, a' \in A$.

- Equality: $P_h(a) = 1$ if and only if a = b;
- Range query: $P_B(a) = 1$ if and only if $a \in B$;
- Comparison: $P_b(a) = 1$ if and only if $a \le b$;
- **Disjunction Equality**: $P_{b \vee c}(a) = 1$ if and only if (a = b) OR (a = c);
- Conjunction Equality: $P_{b \wedge c}(a, a') = 1$ if and only if (a = b) AND (a' = c).

FORMAL DEFINITION OF PE

A plaintext is a pair (ind, m) $\in I \times M$, where I is called **index set** and M the **payload set**.

A **PE scheme** is defined in terms of a polynomial-time predicate $P: K \times I \rightarrow \{0,1\}$, and

the functionality
$$F$$
 over $K \times (I \times M)$ is defined as
$$F(k, (\text{ind}, m)) = \begin{cases} m & \text{if } P(k, \text{ind}) = 1, \\ \bot & \text{if } P(k, \text{ind}) = 0. \end{cases}$$

Let c = Enc(pk, (ind, m)) and $sk_k = \text{KeyGen}(mk, k)$,

- if P(k, ind) = 1, then $Dec(sk_k, c)$ reveals the payload message m;
- if P(k, ind) = 0, then $Dec(sk_k, c)$ reveals nothing new about m.
- If ind is readable from $c \Rightarrow$ **public-index scheme**
- If ind is <u>not</u> readable from $c \Rightarrow$ **private-index scheme**

Functional Encryption

PE with public index

Identity-Based Encryption (IBE)

Attribute-Based Encryption (ABE)

PE with private index

Anonymous IBE

Hidden Vector Encryption

Inner Product Predicate Encryption

beyond predicate

Inner Product Encryption (IPE)

• • •

IDENTITY-BASED ENCRYPTION (IBE)

Introduced in 1984 by Shamir. First practical IBE systems in 2001 by Boneh-Franklin and Cocks.

- The plaintext, and then the ciphertext, is associated with a string/identity ID $\in I$
- The decryptor, ad then the secret key, is associated with a string/identity $ID' \in K$

The decryption is successful if two identities are equal and the predicate P on $K \times I$ is

$$P(ID', ID) = \begin{cases} 1 & \text{if } ID = ID' \\ 0 & \text{otherwise} \end{cases}$$

A. Shamir, Adi. "Identity-based cryptosystems and signature schemes." Workshop on the theory and application of cryptographic techniques. Springer, 1984.

D. Boneh, and M. Franklin. "Identity-based encryption from the Weil pairing." Annual international cryptology conference. Springer, 2001.

C. Cocks. "An identity-based encryption scheme based on quadratic residues." IMA international conference on cryptography and coding. Springer, 2001.

ATTRIBUTE-BASED ENCRYPTION (ABE)

Introduced by Sahai and Waters in 2005.

- Extension of Identity-based encryption.
- More complex access policies.
- To ciphertexts and keys are associated attributes and access policies.

An access policy represented by a Boolean formula ϕ involves the operations

AND, OR, NOT and THRESHOLD over a set of attributes.

THRESHOLD-POLICY ACCESS CONTROL (FUZZY-IBE)

Scheme proposed by Sahai and Waters in 2005

- The <u>ciphertext</u> is associated with a <u>set of attributes</u> $\omega \in I$
- The secret key is associated with a set of attributes $\omega' \in K$

The decryption is successful if there is enough overlap between ω and ω'

$$P(\omega', \omega) = \begin{cases} 1 & \text{if } |\omega \cap \omega'| \ge d \\ 0 & \text{otherwise} \end{cases}$$

- This threshold provides error-tolerance for coarse-grained access control schemes
- Useful in biometric applications and as base for many other ABE schemes

KEY-POLICY ACCESS CONTROL (KP-ABE)

First scheme proposed by Goyal, Pandey and Sahai in 2006

- The <u>ciphertext</u> is associated with a <u>set of attributes</u> $\omega \in I$
- The secret key is associated with an access policy $\phi \in K$

The access structure specifies which type of ciphertext the key can decrypt

$$P(\phi, \omega) = \begin{cases} 1 & \text{if } \omega \text{ satisfies } \phi \\ 0 & \text{otherwise} \end{cases}$$

- The data owner has limited control over who can decrypt the data

CIPHERTEXT-POLICY ACCESS CONTROL (CP-ABE)

First scheme proposed by Bethencourt, Sahai and Waters in 2007

- The <u>ciphertext</u> is associated with an <u>access policy</u> $\phi \in I$
- The secret key is associated with an set of attributes $\omega \in K$

A ciphertext can be decrypted by a user if and only if "his" attributes satisfy the access policy

$$P(\omega, \phi) = \begin{cases} 1 & \text{if } \omega \text{ satisfies } \phi \\ 0 & \text{otherwise} \end{cases}$$

- The data owner can determine the access policy, i.e. who can decrypt the data
- If needed, the policy can be updated with more flexibility (without distribute new keys)

KP-ABE VS CP-ABE

Example of KP-ABE

In a television channel broadcasting, all its TV series are encrypted with their attributes (e.g. "French", "Horror", "Sitcom", "3 seasons").

A user wants to watch a new show and will look for it using the following access policy:

NOT ("British") AND ("Medical drama" OR "Comedy") AND ("at least 2 seasons")

Example of CP-ABE

The sales department at the Trento branch of a big company is suspected to leak secret information.

The central security group opens an investigation and encrypts the related documents with the following encryption policy:

NOT ("CITY=Trento" AND "DEPT=sales") OR ("SECURITY LEVEL > 5")



IBE and ABE systems allow for expressive forms of access control but

- 1. the index associated with the plaintext is given in the clear,
- 2. no computation on the encrypted data.

Predicate Encryption with private index addresses the first issue:

- Anonymous IBE,
- Hidden Vector Encryption (HVE),
- Inner Product Predicate Encryption (IP-PE).

ANONYMOUS IBE

- Same as IBE but the identity of the plaintext can only be determined with the corresponding private key (the ciphertext does not leak the identity of the recipient).
- Anonymous IBE systems can be used to construct PKE with Keyword Search schemes: PKE that allows an encryptor to make a document searchable by keywords.
- First Anonymous IBE scheme not based on the random oracle model was presented by Boyen and Waters in 2006 (based on the decisional linear assumption).

HIDDEN VECTOR ENCRYPTION (HVE)

Introduced in 2007 by Boneh and Waters.

- Key space: $K = \{(k_1, ..., k_n) \mid k_i \in \{0,1\}^* \cup \{\star\}\} \cup \{\varepsilon\}$
- Index space: $I = (\{0,1\}^*)^n$
- Plaintext space: $I \times M = \{((w_1, ..., w_n), m) \mid (w_1, ..., w_n) \in I, m \in M\}$
- The predicate P on $K \times I$ is

$$P((k_1, ..., k_n), (w_1, ..., w_n)) = \begin{cases} 1 & \text{if } k_i = w_i \text{ whenever } k_i \neq \star, \\ 0 & \text{otherwise} \end{cases}$$

HVE

HVE supports:

- the evaluation of conjunctive equality,
- comparison,
- subset operations

between "attributes" in ciphertexts and "attributes" in secret keys.

Moreover, it supports:

• the conjunctive combination of the abovementioned primitive predicates.

INNER PRODUCT PREDICATE ENCRYPTION (IP-PE)

Introduced in 2008 by Katz, Sahai and Waters.

- Key space: $K = \mathbb{Z}_p^n \cup \{\varepsilon\} = \{(k_1, ..., k_n) \mid k_i \in \mathbb{Z}_p\} \cup \{\varepsilon\}$, where p is a prime
- Index space: $I = \mathbb{Z}_p^n$
- Plaintext space: $I \times M = \{((w_1, ..., w_n), m) \mid (w_1, ..., w_n) \in I, m \in M\}$
- The predicate P on $K \times I$ is

$$P((k_1, \dots, k_n), (w_1, \dots, w_n)) = \begin{cases} 1, & \text{if } \sum_{i=1}^n k_i \cdot w_i \equiv 0 \bmod p, \\ 0, & \text{otherwise} \end{cases}$$

J. Katz, A. Sahai, and B. Waters. "Predicate encryption supporting disjunctions, polynomial equations, and inner products." *annual international conference on the theory and applications of cryptographic techniques*. Springer, Berlin, Heidelberg, 2008.

IP-PE

IP-PE implies:

- anonymous IBE and HVE
- PE schemes supporting
 - Polynomial evaluation of some bounded degree
 - Disjunctions
 - DNF and CNF formulae of some bounded size
 - Threshold queries $(|\omega \cap \omega'| = d)$



So far the functionalities presented allow to recover the entire message, when the predicate is satisfied, P(k, ind) = 1.

Still an all-or-nothing encryption.

What about FE with fine-grained access to the encrypted data where decryption recovers partial information about the encrypted data?

INNER PRODUCT ENCRYPTION (IPE)

Introduced in 2015 by Abdalla, Bourse, De Caro, and Pointcheval.

- Key space: $K = \mathbb{Z}_p^n \cup \{\varepsilon\} = \{k = (k_1, ..., k_n) \mid k_i \in \mathbb{Z}_p\} \cup \{\varepsilon\}$
- Plaintext space: $X = \mathbb{Z}_p^n = \{ \mathbf{x} = (x_1, ..., x_n) \mid x_i \in \mathbb{Z}_p \}$
- Functionality:

$$F(\mathbf{k}, \mathbf{x}) = \langle \mathbf{k}, \mathbf{x} \rangle = \sum_{i=1}^{n} k_i \cdot y_i \in \mathbb{Z}_p$$

• Dec algorithm returns F(k, x).

IPE APPLICATIONS

Different applications in:

- Biometric identification
- Machine learning
- Statistical analysis

Example:

Assume that an hospital makes available patients data to a research institute to perform statical analysis. Since they are sensitive data, instead of providing entire patient records, the hospital encrypts them and associates to each record a vector (containing only relevant data for the analysis).

The institute research is now able to analyse data, without having access to the entire patient record.

OTHER SCHEMES BEYOND PREDICATE

- Some constructions of FE schemes for <u>quadratic functions</u> under standard assumptions
- Some constructions of FE schemes for any <u>arbitrary circuits</u> BUT relying on non-standard assumptions
- Some constructions of FE schemes for arbitrary circuits under standard assumption with bounded collusions (limited number of corrupted secret keys)

OTHER FEATURES

We have presented FE schemes with basic characteristics. Based on the situation, other extra features can be useful.

For example:

- Hierarchical
- Multi-Authority
- Multi-Input
- Multi-Client
- Function Hiding
- Scalability, Revocability, Fast Decryption, Unbounded schemes...

HIERARCHICAL FE

An FE scheme that mirrors an organisational hierarchy, it augments FE with **delegation** capabilities.

A user holding a secret key Sk_f for function f can generate a secret key for a more restricted function f'.

- <u>HIBE</u>: a user holding a secret key Sk_{ID} for identity $ID = (I_1, ..., I_\ell)$ can generate a secret key $Sk_{ID'}$ for a descendent identity $ID' = (I_1, ..., I_{\ell+1})$.
- <u>KP-ABE</u>: a user holding a secret key Sk_{ϕ} for an access policy ϕ can generate a secret key Sk_{ϕ} , for a more restrictive access policy ϕ' .
- <u>CP-ABE</u>: a user holding a secret key sk_k for a set of attributes k can generate sk_k , for $k' \subseteq k$.

MULTI-AUTHORITY FE

FE requires the assumption of a <u>central trusted authority</u> which performs Setup and KeyGen (manage the credentials of every user).

Multi-Authority Functional Encryption (MAFE) addresses this trust issue:

- Multiples authorities in the system, they can:
 - initialize together the Setup, or
 - generate independently their private and public keys, without any interaction.
- Each authority can provide a secret key for a function or a set of attributes of its choice.
- A user has to put together one or more secret keys to decrypt.

MULTI-INPUT FE

Let F be a **an n-input functionality**, that is, defined over $(K, X_1 \times \cdots \times X_n)$.

In a Multi-Input FE scheme, a secret key sk_k decrypts jointly ciphertexts

Enc(pk, x_1), ..., Enc(pk, x_n) coming from the different inputs and obtains $F(k, x_1, ..., x_n)$.

Note that at all the ciphertexts are encrypted using the same pk.

(DECENTRALIZED) MULTI-CLIENT FE

In a **Multi-Client FE** scheme:

- The single input x is divided into n independent components $(x_1, ..., x_n)$.
- Each client C_i does the following encryption: Enc(pk, ek_i, x_i , ℓ), with respect to a label ℓ .
- A user owning a secret key sk_k and ciphertexts $\operatorname{Enc}(\operatorname{pk},\operatorname{ek}_1,x_1,\ell),\ldots,\operatorname{Enc}(\operatorname{pk},\operatorname{ek}_n,x_n,\ell)$ with the label ℓ can compute $F(k,x_1,\ldots,x_n)$ but nothing else about each x_i .

A **Decentralized Multi-Client FE** scheme is a multi-input scheme in which <u>the authority is removed</u> and the clients work together to generate appropriate functional decryption keys.

FUNCTION-HIDING

In many realistic scenarios, we need both

- privacy of the encrypted message,
- **function privacy**, that is secret keys reveal no information about the encoded identities, attributes, access policies, predicates or functions.

Note that no public-key functional encryption scheme is function hiding.

Function privacy can be achieved only in private-key settings (pk is private).

CONCLUSIONS

FE introduced to address some issues that PKE cannot solve, such as:

- fine-grained access control of encrypted data;
- computation of a specific function of the encrypted data.

- Several IBE and ABE schemes, in literature and in real implementations.
- Different private-index PE and IPE schemes in literature, <u>BUT</u> few real implementations (for high computational costs).
- For many dynamic scenarios, it needs more schemes that compute more complex functions.

SOME OTHER REFERENCES

- □ D. Boneh, A. Sahai, and B. Waters. "Functional encryption: Definitions and challenges." *Theory of Cryptography Conference*. Springer, Berlin, Heidelberg, 2011.
- □ Z. Qiao, S. Liang, S. Davis, and H. Jiang. "Survey of attribute based encryption." *15th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*. IEEE, 2014.
- ■S. Gorbunov, V. Vaikuntanathan, and H. Wee. "Predicate encryption for circuits from LWE." *Annual Cryptology Conference*. Springer, Berlin, Heidelberg, 2015.
- □ Z. Brakerski, et al. "Hierarchical functional encryption." 8th Innovations in Theoretical Computer Science Conference (ITCS 2017). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.
- □ C.E.Z. Baltico, D. Catalano, D. Fiore, and R. Gay. "Practical functional encryption for quadratic functions with applications to predicate encryption." *Annual International Cryptology Conference*. Springer, Cham, 2017.
- □ Z. Brakerski, I. Komargodski, and G. Segev. "Multi-input functional encryption in the private-key setting: Stronger security from weaker assumptions." *Journal of Cryptology* 31.2 (2018): 434-520.
- □ J. Chotard, E.D. Sans, R. Gay, D.H. Phan, and D. Pointcheval. "Decentralized multi-client functional encryption for inner product." *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, Cham, 2018.
- □ C. Mascia, M. Sala, and I. Villa. "A survey on Functional Encryption." *In preparation*, 2021.



THANK YOU FOR YOUR ATTENTION ANY QUESTION?