ISSN 2188-286X

マス・フォア・インダストリ研究　No.4

Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques

Institute of Mathematics for Industry
Kyushu University

編　集　穴田　啓晃
安田　貴徳
櫻井　幸一
寺西　勇

九州大学マス・フォア・インダストリ研究所

About the Mathematics for Industry Research

The Mathematics for Industry Research was founded on the occasion of the certification of the Institute of Mathematics for Industry (IMI), established in April 2011, as a MEXT Joint Usage/Research Center – the Joint Research Center for Advanced and Fundamental Mathematics for Industry – by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in April 2013. This series publishes mainly proceedings of workshops and conferences on Mathematics for Industry (MfI). Each volume includes surveys and reviews of MfI from new viewpoints as well as up-to-date research studies to support the development of MfI.

October 2014
Yasuhide Fukumoto
Director
Institute of Mathematics for Industry

Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques

Mathematics for Industry Research No.4, Institute of Mathematics for Industry, Kyushu University
ISSN 2188-286X
Editors: Hiroaki Anada, Takanori Yasuda, Kouichi Sakurai and Isamu Teranishi
Date of issue: 29 January 2016
Publisher:
Institute of Mathematics for Industry, Kyushu University
Motooka 744, Nishi-ku, Fukuoka, 819-0395, JAPAN
Tel +81-(0)92-802-4402, Fax +81-(0)92-802-4405
URL http://www.imi.kyushu-u.ac.jp/
Printed by
Social Welfare Service Corporation Fukuoka Colony
1-11-1, Midorigahama, Shingu-machi Kasuya-gun, Fukuoka, 811-0119, Japan
TEL +81-(0)92-962-0764 FAX +81-(0)92-962-0768

Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques

IMI Workshop of the Joint Research Projects

Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques

September 1th – 3th, 2015

Industry-University-Government Collaboration Innovation Plaza

3-8-34 Momochihama Sawara-ku Fukuoka 814-0001, Japan

Sponsored by

Institute of Mathematics for Industry (IMI),

Kyushu University

Organized by

Hiroaki Anada, Takanori Yasuda and Kouichi Sakurai

Institute of Systems, Information Technologies and Nanotechnologies (ISIT)

and

Isamu Teranishi

NEC Corporation

Acknowledgements

The research presented by Sushmita Ruj was a part of JSPS-DST Japan-India Bilateral Collaborative Research Program.

Concerning her visit and participation to this workshop, Sushmita Ruj was partially supported by a *kakenhi* Grant-in-Aid for Scientific Research (C) 15H02711 from Japan Society for the Promotion of Science.

Concerning his visit and participation to this workshop, Masashi Une was partially supported by a *kakenhi* Grant-in-Aid for Scientific Research (C) 15H02711 from Japan Society for the Promotion of Science.

Preface

Privacy Protection and Decentralized Control are two hot areas of research due to rapid growing of Cloud Computing and Internet of Things for these several years. Especially topics such as attribute-based access control, homomorphic encryption, multi-authority cryptography as well as proof-of-work consensus in Bitcoin are exciting. Their evolution is greatly due to the use of fruitful mathematical structures like pairings on elliptic curves, ideal lattices of polynomial rings.

The purpose of this workshop was to discuss the functions, mathematical structures, meaningful applications and how to answer to realistic requirements. For those purposes, 10 distinguished lectures with the titles in the program and one panel discussion are held with more than 30 attendees.

Hence, my sincere thanks are to those lecturers and attendees of this workshop. I hope this lecture note will be read among more researchers and developers who are interested in the above areas.

Hiroaki Anada, Representative of Organizers

Table 1 List of attendees

Sushmita RUJ	Chen-Mou CHENG	Anirban BASU
Masayuki YOSHINO	Katsuyuki TAKASHIMA	Hitoshi OKADA
Sherman CHOW	Goichiro HANAOKA	Le Trieu PHONG
Yoshihiro SHIKATA	Masashi UNE	Tomoko ADACHI
Keiichi IWAMURA	Yasuyuki MURAKAMI	Makoto ISHIKAWA
Yuntao WANG	Ye YUAN	Keishi MABUCHI
Chi CHENG	Rui XU	Samiran BAG
Hui ZHAO	Yaokai FENG	Junpei KAWAMOTO
Yun Ju HUANG	Shinichi MATSUMOTO	Kirill MOROZOV
Masaya YASUDA	Tsuyoshi TAKAGI	Isamu TERANISHI
Takanori YASUDA	Kouich SAKURAI	Hiroaki ANADA

Photograph 1. Part of attendees in front of the venue.

Photograph 2. Sceneries in the workshop lectures

Photograph 3. Snaps in Banquet at Momochi-hama

Acknowledgements. We appreciate to the photographer Dr. Anirban Basu (the right).

Program

Tuesday, September 1, 2015

14:00 - 14:10 Opening Remark

14:10 - 15:00 Invited Lecture

"Social Implications of the Decentralized Virtual Currency: A Public Policy Standardization Perspective"

Hitoshi OKADA (National Institute of Informatics (NII))

15:20 - 16:00 International Invited Lecture

"Attribute-Based Access Control in Mobile Clouds"

Sushmita RUJ (Indian Statistical Institute, India)

16:00 - 16:40 Invited Lecture

"Order-Preserving Encryption Secure Beyond One-Wayness"

Isamu TERANISHI (NEC Corporation)

16:40 - 17:00 Photo Session

Wednesday (Morning), September 2, 2015

09:50 - 10:00 Opening Remark of the Second Day

10:00 - 10:40 Lecture

"Fast and Secure Linear Regression and Biometric Authentication with Security Update"

Le Trieu PHONG (Nat. Ins. Info. and Comm. Tech. (NICT))

11:00 - 11:40 Invited Lecture

"Homomorphic Encryption - are we there yet?"

Anirban BASU (KDDI R&D Lab.)

11:40 - 12:20 Invited Lecture

"Cryptography for Cloud Service"

Masayuki YOSHINO (Hitachi, Ltd.)

Wednesday (Afternoon), September 2, 2015

14:00 - 14:40 Invited Lecture

"Cryptography for Availability: The Case of Secure Cloud Storage"

Sherman Chow (The Chinese University of Hong Kong)

15:00 - 15:40 Invited Lecture

"Decentralized Attribute-Based Cryptosystems"

Katsuyuki TAKASHIMA (Mitsubishi Electric Corp.)

15:40 - 16:20 Invited Lecture

"Dynamic Threshold Public-key Encryption with Decryption Consistency from Static Assumptions"

Goichiro HANAOKA (Nat. Inst. of Adv. Ind. Sci. and Tech. (AIST))

Thursday, September 3, 2015

09:50 - 10:00 Opening of the Third Day

10:00 - 10:40 Lecture

"On the Application of Clique Problem for Proof-of-Work in Cryptocurrencies"

Samiran Bag (Kyushu University / ISI)

11:00 – 12:00 Panel Discussion

Core Panelist: Masashi UNE (Institute for Monetary and Economic Studies, Bank of Japan)

Moderator: Kirill MOROZOV (IMI, Kyushu University)

(a): Mathematical structure behind decentralized cryptocurrencies: Formal? Robust? Secure?

Sushmita RUJ, Hitoshi OKADA

(b): Multi-authority cryptographic primitives: Usability and practical impact

Katsuyuki TAKASHIMA, Goichiro HANAOKA

12:00 - 12:10 Ending Comments

Hiroaki ANADA (ISIT)

Presentation

Social Implications of the Decentralized Virtual Currency: A Public Policy Standardization Perspective

Hitoshi Okada
National Institute of Informatics, Japan
okada@nii.ac.jp

Abstract

Bitcoin is a decentralized which enables the unique distribution of electronic value from one person to another without the existence of a centralized issuer. Virtual currency circulates in an open-looped system as if it were real money, whereas existing electronic money circulates in a closed-looped system. The decentralization issue of virtual currency raises a question concerning the seigniorage profit, which ought to be under state monopoly. This paper explains why the seigniorage profit is attributable to the state for what reason currency issuance should be decentralized. The author discusses the ideal public policy and international standardization for virtual currency in order for decentralized currency to achieve justice as fairness.

References

[1] Satoshi Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, November 2008.

[2] Friedrich August von Hayek: Decentralization of Money, 1976.

[3] Financial Crimes Enforcement Network (FinCEN): Application of FinCEN's Regulations to Persons Administering, Exchanging, or Using Virtual Currencies, FIN-2013-G001, March 18, 2013.

[4] Financial Crimes Enforcement Network (FinCEN): Application of the Definition of Money Transmitter to Brokers and Dealers in Currency and other Commodities, FIN-2008-G008, September 10, 2008.

[5] European Central Bank (ECB): Virtual Currency Schemes, October 2012.

[6] Burggraeve R.: Emmanuel Levinas et la socialité de l'argent Un philosophe en quête de la réalité journalière. La genèse de Socialité et argent ou l'ambiguïté de l'argent, PEETERS Publisher, 1997.

[7] U.S. House Subcommittee on Domestic and International Monetary Policy: The Future of Money—Part 2 Hearing, October 11, 1995, Washington: Government Printing Office, 1996. (Y4.B 22/l: 104-27 /PT).

[8] United States of America: Code of Federal Regulations, Title 31, Article 1010, Item 100 (m).

[9] Nick Szabo: Bit Gold. http://unenumerated.blogspot.jp/2005/12/bit-gold.html

[10] Wei Dai: b-Money. http://www.weidai.com/bmoney.txt

[11] Adam Back: Hashcash: A Denial of Service Counter-Measure, Technical Report, August 2002.
[12] Group of 10, Report, 1997. http://www.bis.org/publ/gten01.htm
[13] Directive 2000/46/EC of the European Parliament and of the Council of 18 September 2000 on the taking up, pursuit of and prudential supervision of the business of electronic money institutions, O.J. L 275, 27 October 2000.
[14] Directive 2007/64/EC of the European Parliament and of the Council of 13 November 2007 on payment services in the internal market amending Directives 97/7/EC, 2002/65/EC, 2005/60/EC and 2006/48/EC and repealing Directive 97/5/EC, OJ L 319, 5.12.2007.
[15] European Commission: Legal Analysis of a Single Market for the Information Society, SMART 2007/0037.
[16] European Central Bank: Virtual Currency Schemes, ECB (Frankfurt am Main, Germany), October 2012.
[17] European Banking Authority: EBA Opinion on 'virtual currencies'.
http://www.eba.europa.eu/documents/10180/657547/EBA-Op-2014-08+Opinion+on+Virtual+Currencies.pdf
[18] National Crime Agency: UK's National Crime Agency moves against Silk Road, NCA arrests Silk Road suspects, NCA News Release, National Crime Agency UK, 8th October 2013.
[19] Ben Regnard-Weinrabe, Mark Taylor and Rachel Savary of Hogan Lovells, London: Virtual Currencies, the Risks and the Regulatory Radar.
[20] Q&As 164 and 255 at "Your Questions on PSD", Payment Services Directive 2007/64/EC
http://ec.europa.eu/internal_market/payments/docs/framework/transposition/faq_en.pdf
[21] HMRC: UK Bitcoin Exchanges Don't Have to Register under Money Laundering Regulations
http://www.coindesk.com/hmrc-uk-bitcoin-exchanges-dont-have-to-register-under-moneylaundering-regulations/
[22] Bank of England: The economics of digital currencies, September 2014.
http://www.bankofengland.co.uk/publications/Documents/quarterlybulletin/2014/qb14q3digitalcurrenciesbitcoin2.pdf

Reference (in Japanese)

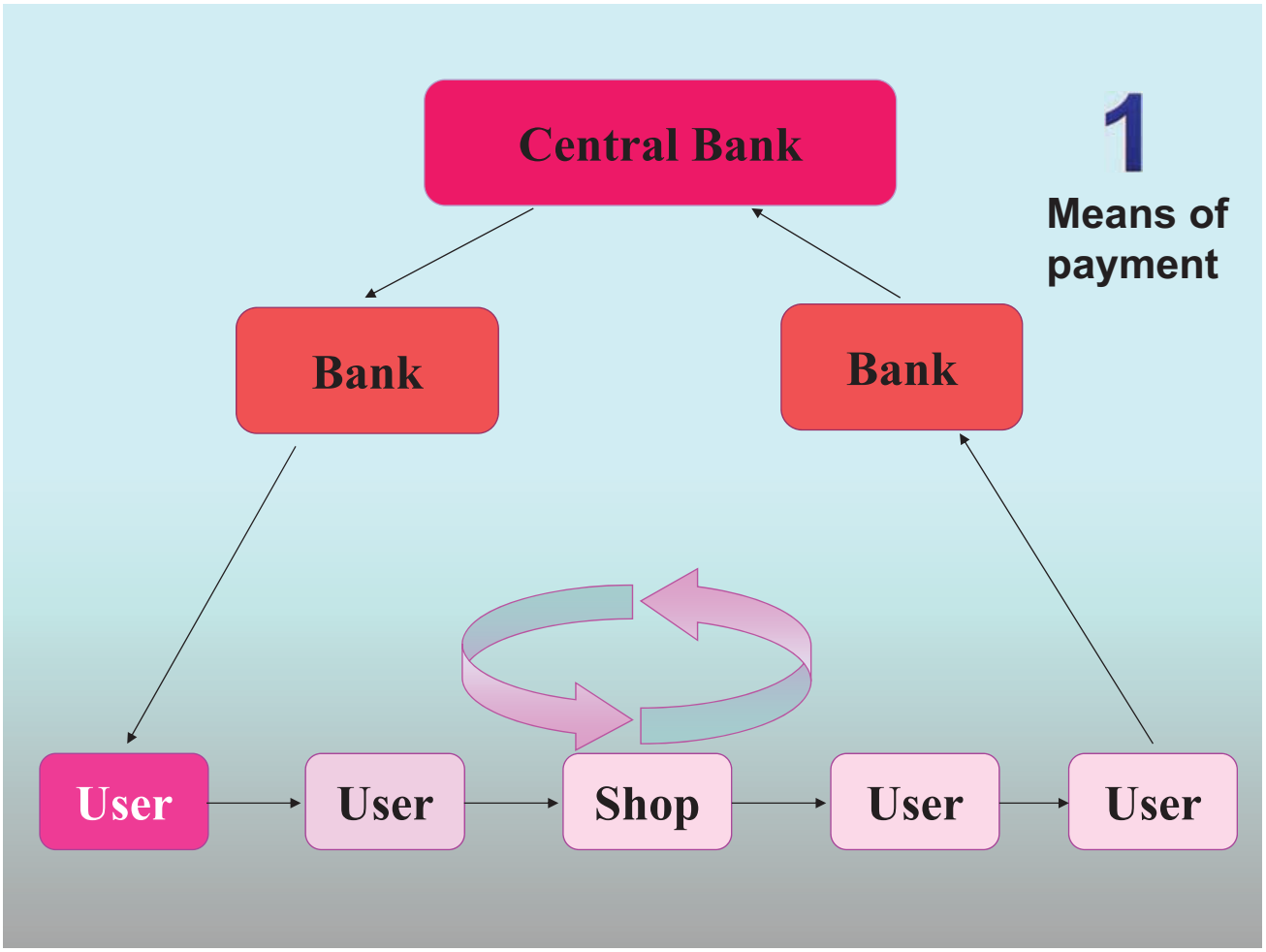
[23] Hitoshi Okada, Ikuo Takahashi, and Shigeichiro Yamasaki: "Kasou Tsu-ka: Bitcoin Wo Kai-Bou Suru', ("Virtual Currency: Anatomy of Bitcoin"), Toyo Keizai Inc., Tokyo, Japan, May 29th, 2015. (ISBN：9784492681381)

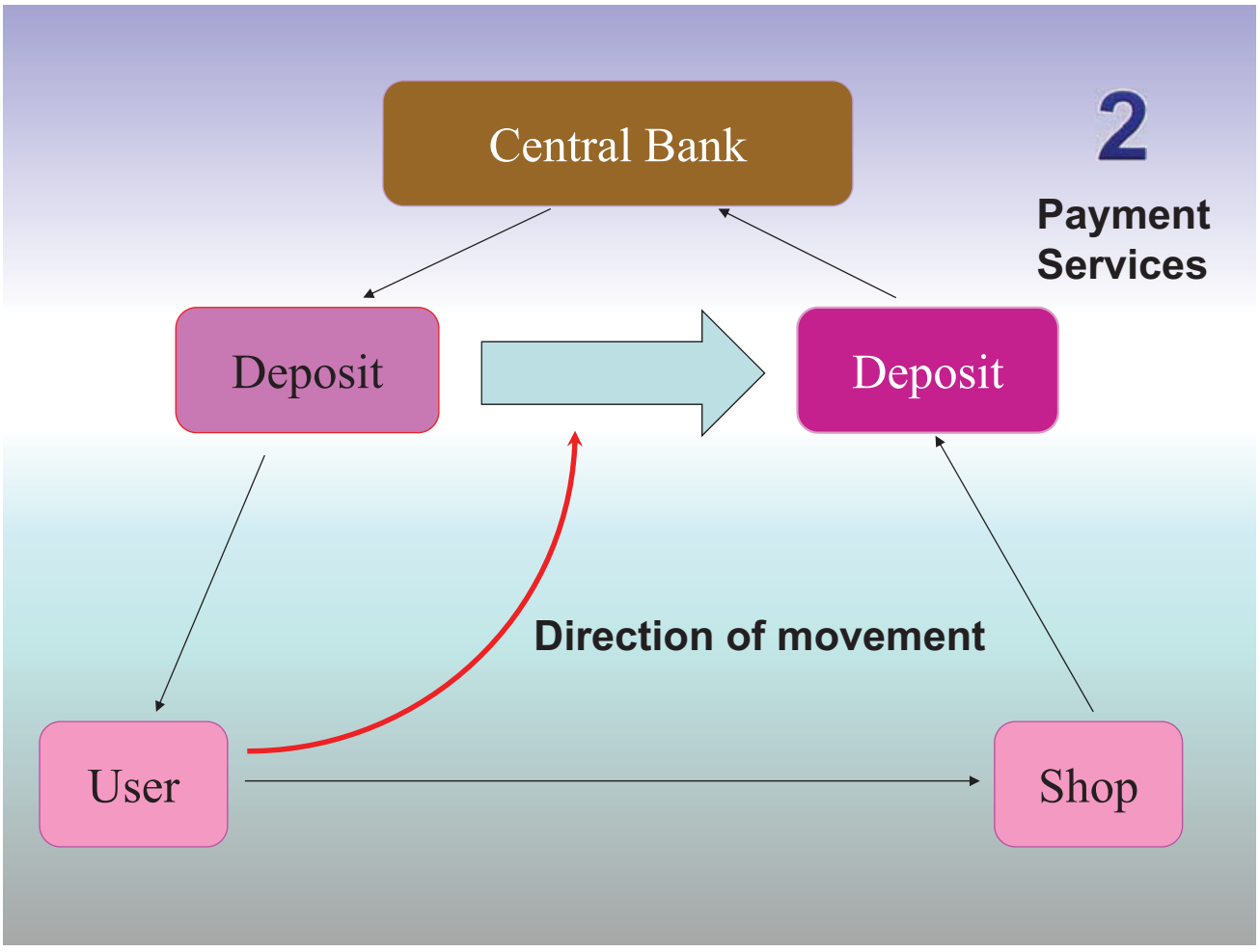
References [1] to [22] are quoted from the reference list of the book [23].

Social Implications of the Decentralized Virtual Currency: A Public Policy Standardization Perspective

National Institute of Informatics
Hitoshi OKADA

Means or Services

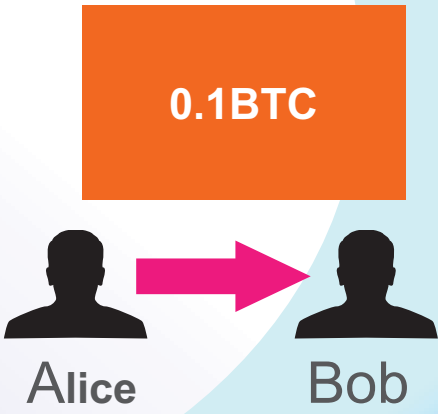

Central Bank
1
Means of payment
Bank
Bank
User
User
Shop
User
User

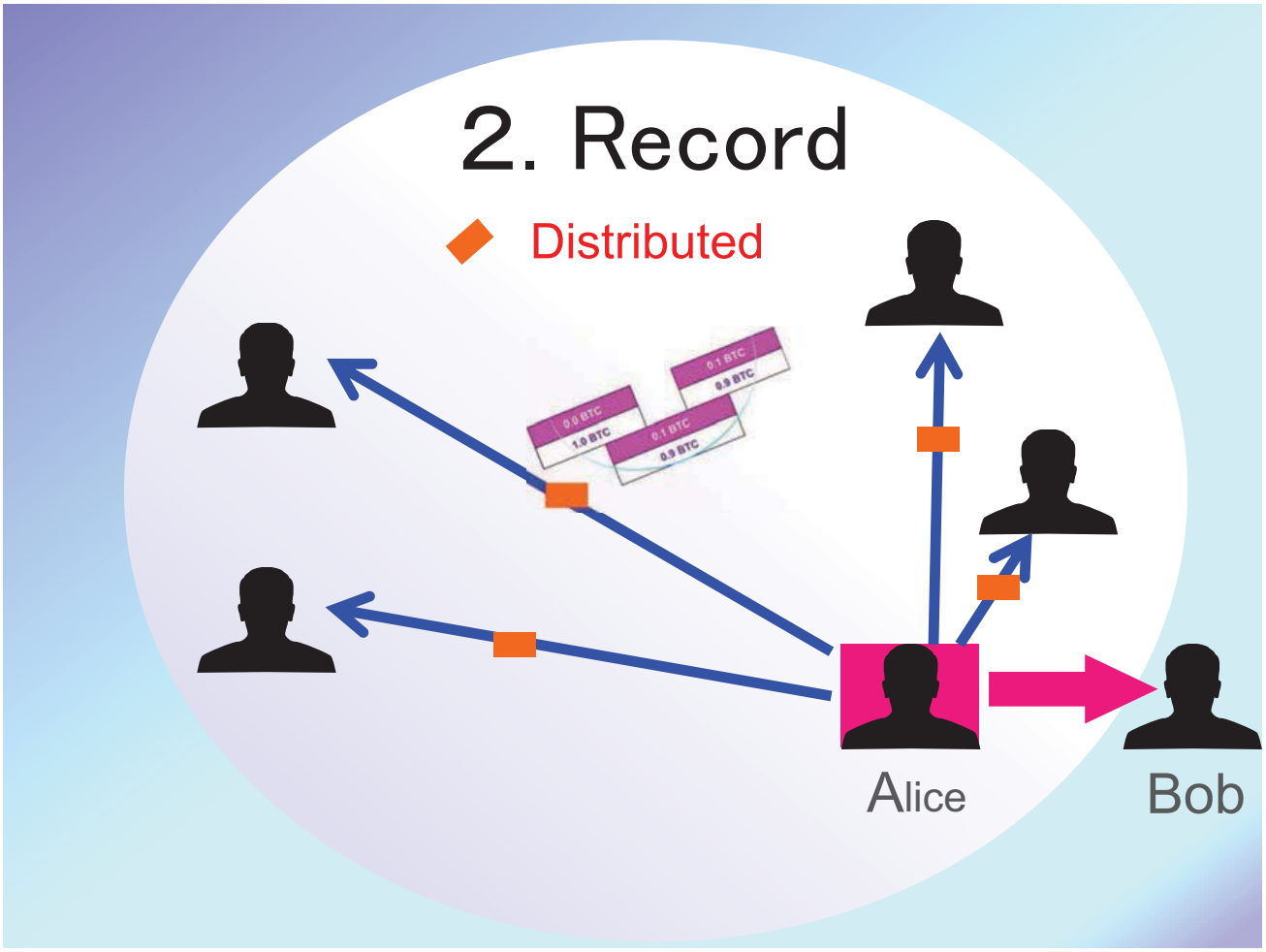

Central Bank
2
Payment Services
Deposit
Deposit
Direction of movement
User
Shop

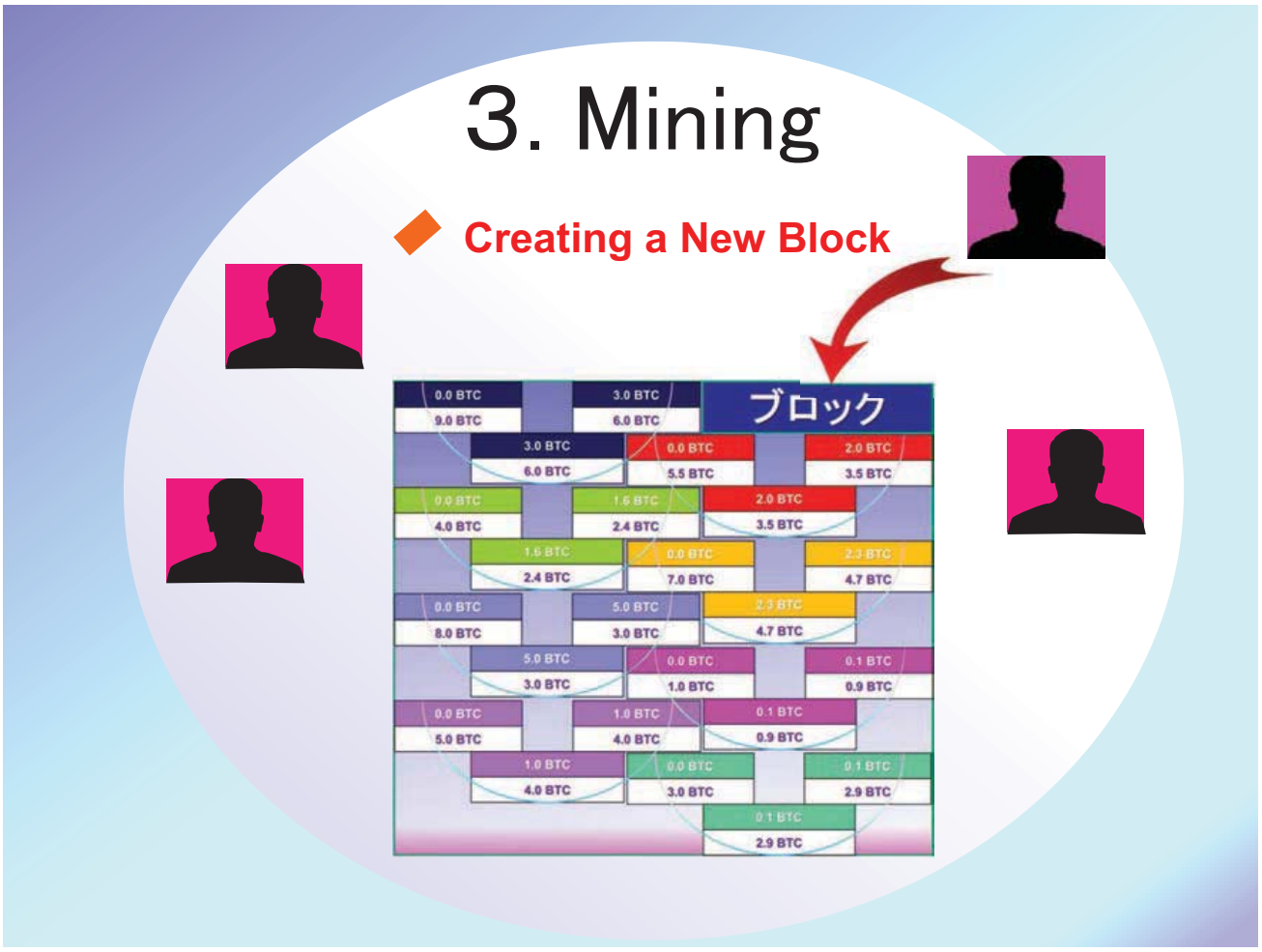
Bitcoin Eco System

1. Payment

- Peer to Peer

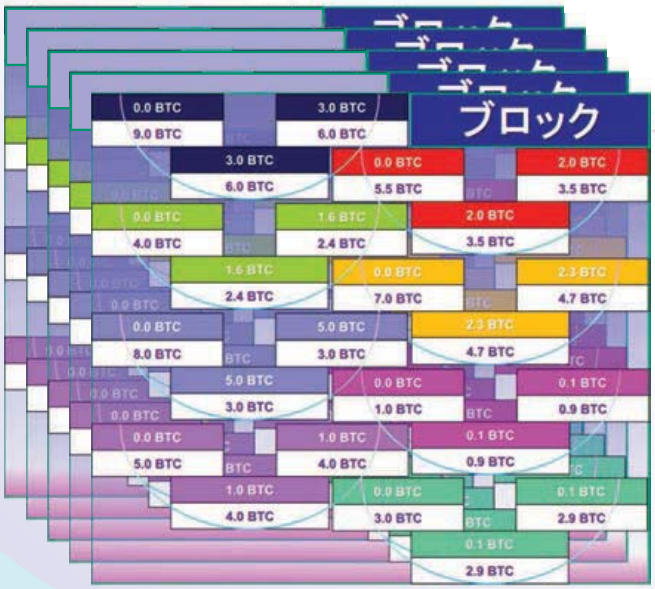



2. Record
Distributed
Alice
Bob


3. Mining
Creating a New Block
ブロック

4. Reward

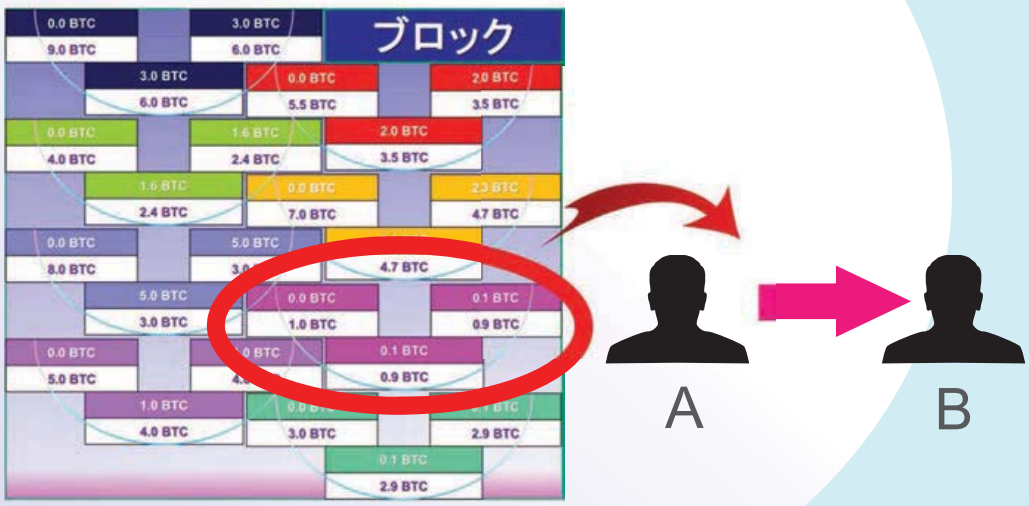
- Chained to the Blockchain



25BTC

5. Authorize

- Confirm the transactions



6. Issuing

- Issuing new currency

Definition of Currency

(j) *Commodity.* Any good, article, service, right, or interest described in section 1a(4) of the Commodity Exchange Act ("CEA"), 7 U.S.C. 1a(4).

(k) *Common carrier.* Any person engaged in the business of transporting individuals or goods for a fee who holds himself out as ready to engage in such transportation for hire and who undertakes to do so indiscriminately for all persons who are prepared to pay the fee for the particular service offered.

(l) *Contract of sale.* Any sale, agreement of sale, or agreement to sell as described in section 1a(7) of the CEA, 7 U.S.C. 1a(7).

(m) *Currency.* The coin and paper money of the United States or of any other country that is designated as legal tender and that circulates and is customarily used and accepted as a medium of exchange in the country of issuance. Currency includes U.S. silver certificates, U.S. notes and Federal Reserve notes. Currency also includes official foreign bank notes that are customarily used and accepted as a medium of exchange in a foreign country.

(n) *Deposit account.* Deposit accounts include transaction accounts described in paragraph (ccc) of this section, savings accounts, and other time deposits.

Guidance

FIN-2013-G001

Issued: March 18, 2013

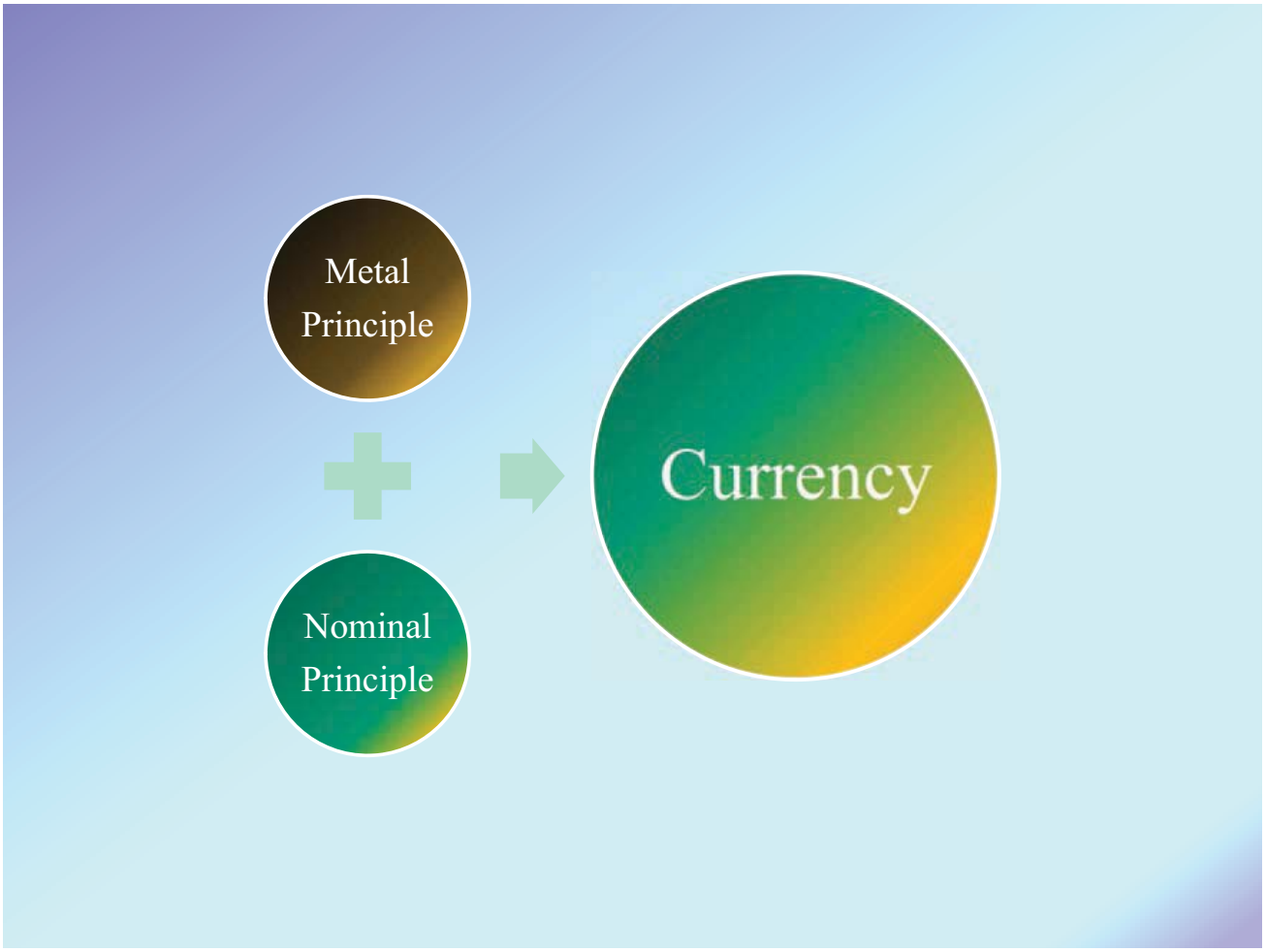
Subject: Application of FinCEN's Regulations to Persons Administering, Exchanging, or Using Virtual Currencies

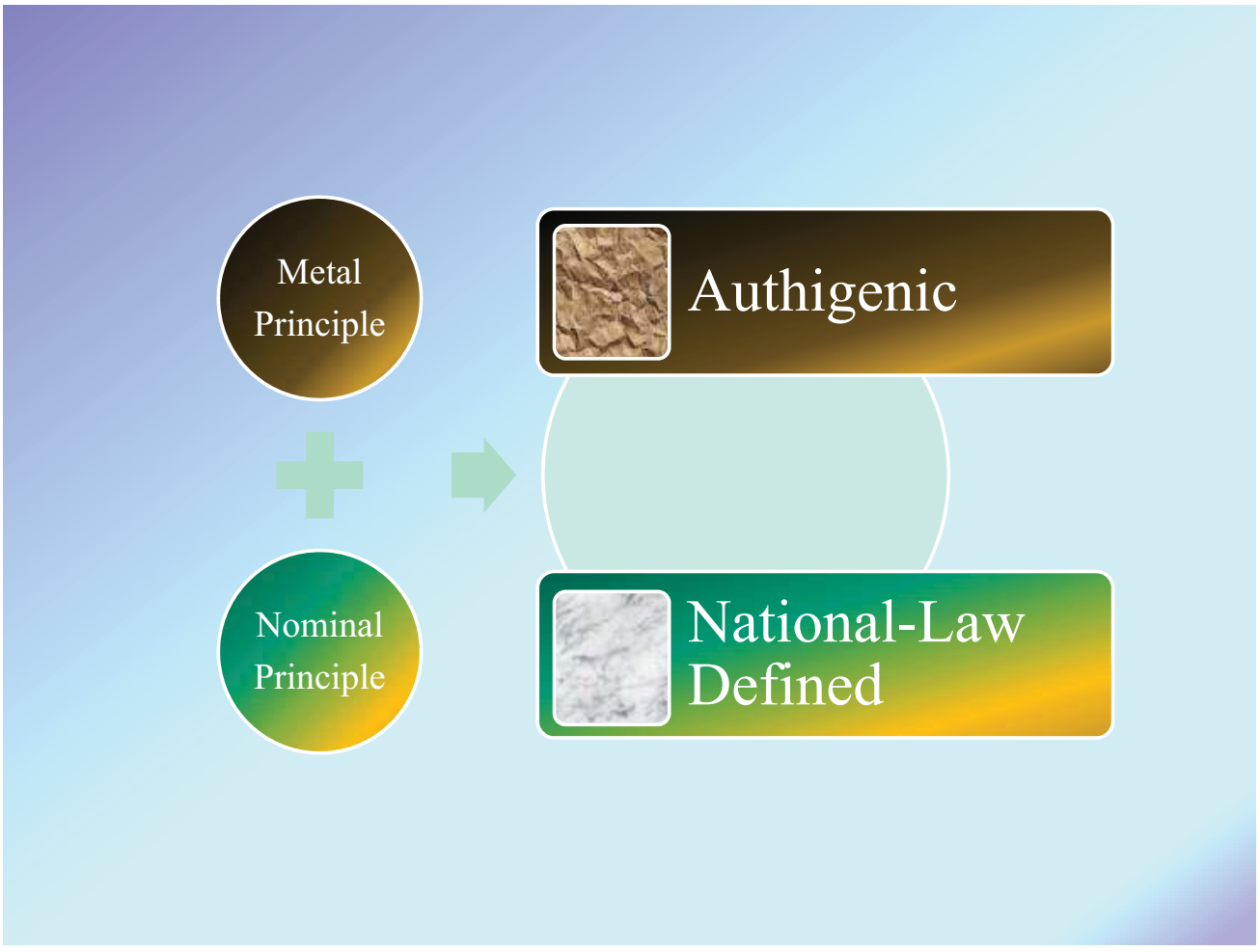
---

The Financial Crimes Enforcement Network ("FinCEN") is issuing this interpretive guidance to clarify the applicability of the regulations implementing the Bank Secrecy Act ("BSA") to persons creating, obtaining, distributing, exchanging, accepting, or transmitting virtual currencies.[1] Such persons are referred to in this guidance as "users," "administrators," and "exchangers," all as defined below.[2] A user of virtual currency is **not** an MSB under FinCEN's regulations and therefore is not subject to MSB registration, reporting, and recordkeeping regulations. However, an administrator or exchanger is an MSB under FinCEN's regulations, specifically, a money transmitter, unless a limitation to or exemption from the definition applies to the person. An administrator or exchanger is not a provider or seller of prepaid access, or a dealer in foreign exchange, under FinCEN's regulations.

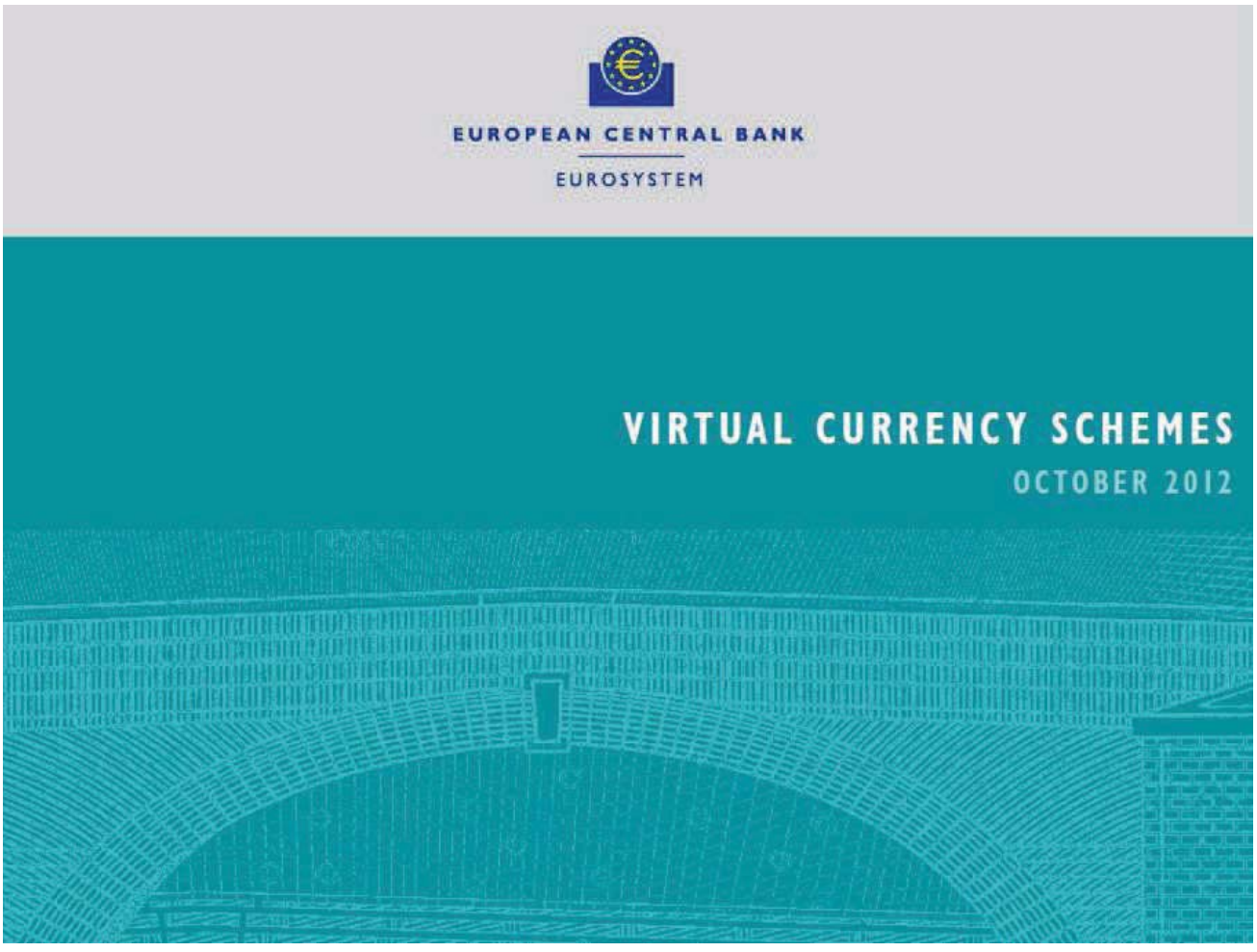
Currency vs. Virtual Currency

FinCEN's regulations define currency (also referred to as "real" currency) as "the coin and paper money of the United States or of any other country that [i] is designated as legal tender and that [ii] circulates and [iii] is customarily used and accepted as a medium of exchange in the country of issuance."[3] In contrast to real currency, "virtual" currency is a medium of exchange that operates like a currency in some environments, but does not have all the attributes of real currency. In particular, virtual currency does not have legal tender status in any jurisdiction. This guidance addresses "convertible" virtual currency. This type of virtual currency either has an equivalent value in real currency, or acts as a substitute for real currency.


Metal
Principle
Nominal
Principle
Currency


Metal
Principle
Nominal
Principle
Authigenic
National-Law
Defined

Function of Currency



EU study on the

Legal analysis of a Single Market for the Information Society

New rules for a new age?

7. Electronic payments
8. Electronic contracting

Public Policy

http://fukuroh.air-nifty.com/katsudou/files/kachikirokuenglish.pdf

Self or Co-Regulation

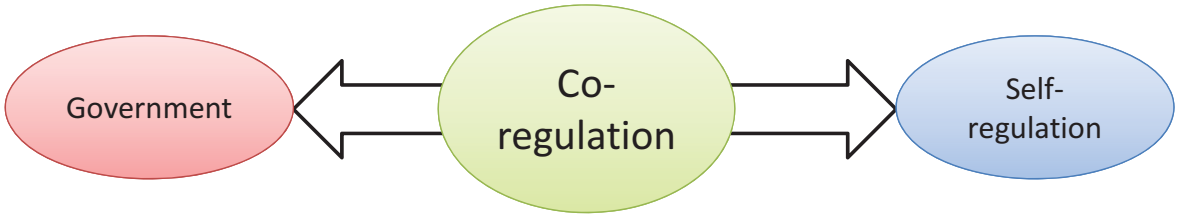
What is co-regulation?

Co-regulation means

Self-regulation (market) ＋

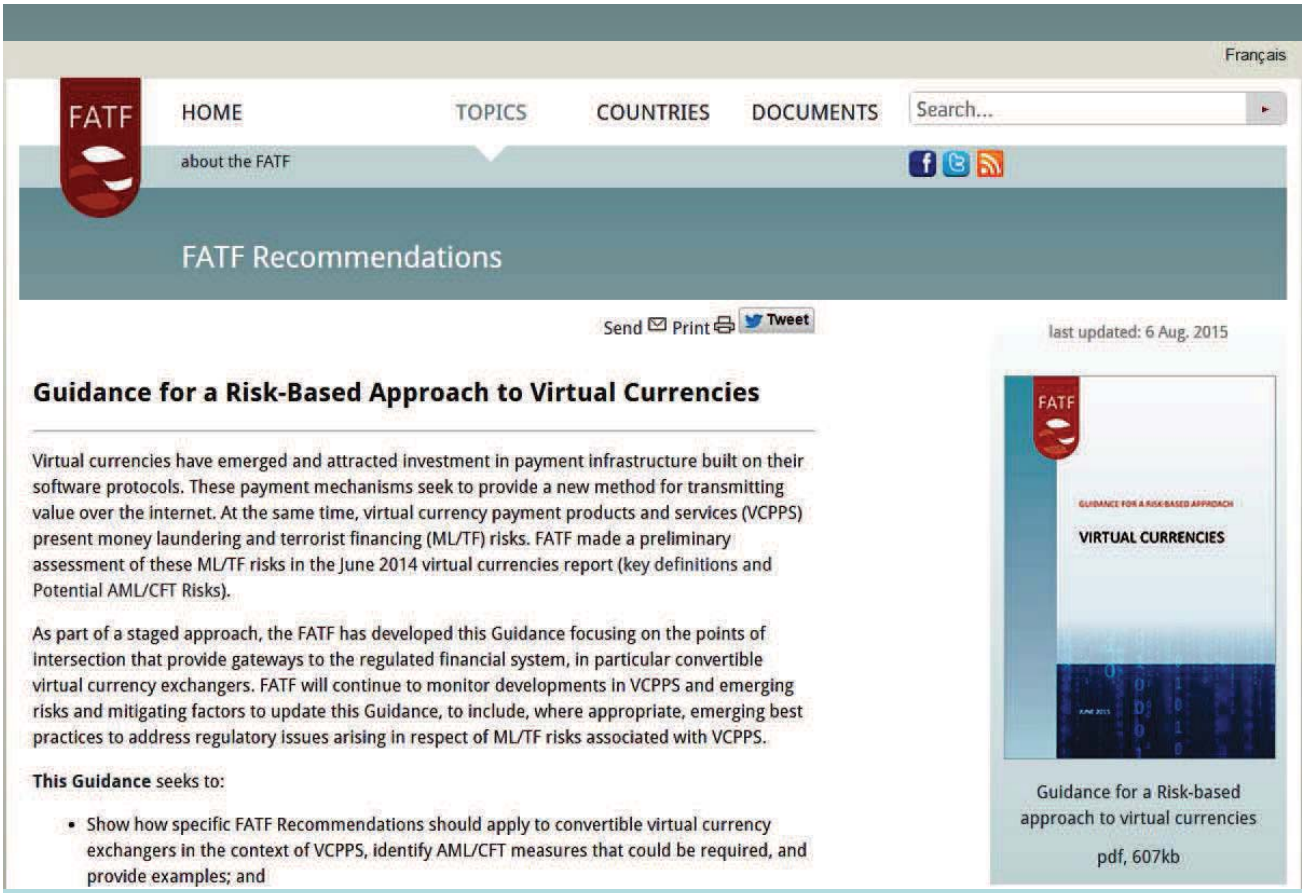
Government-regulation

Establishing the complementary relationship between market and government is the main object of co-regulation

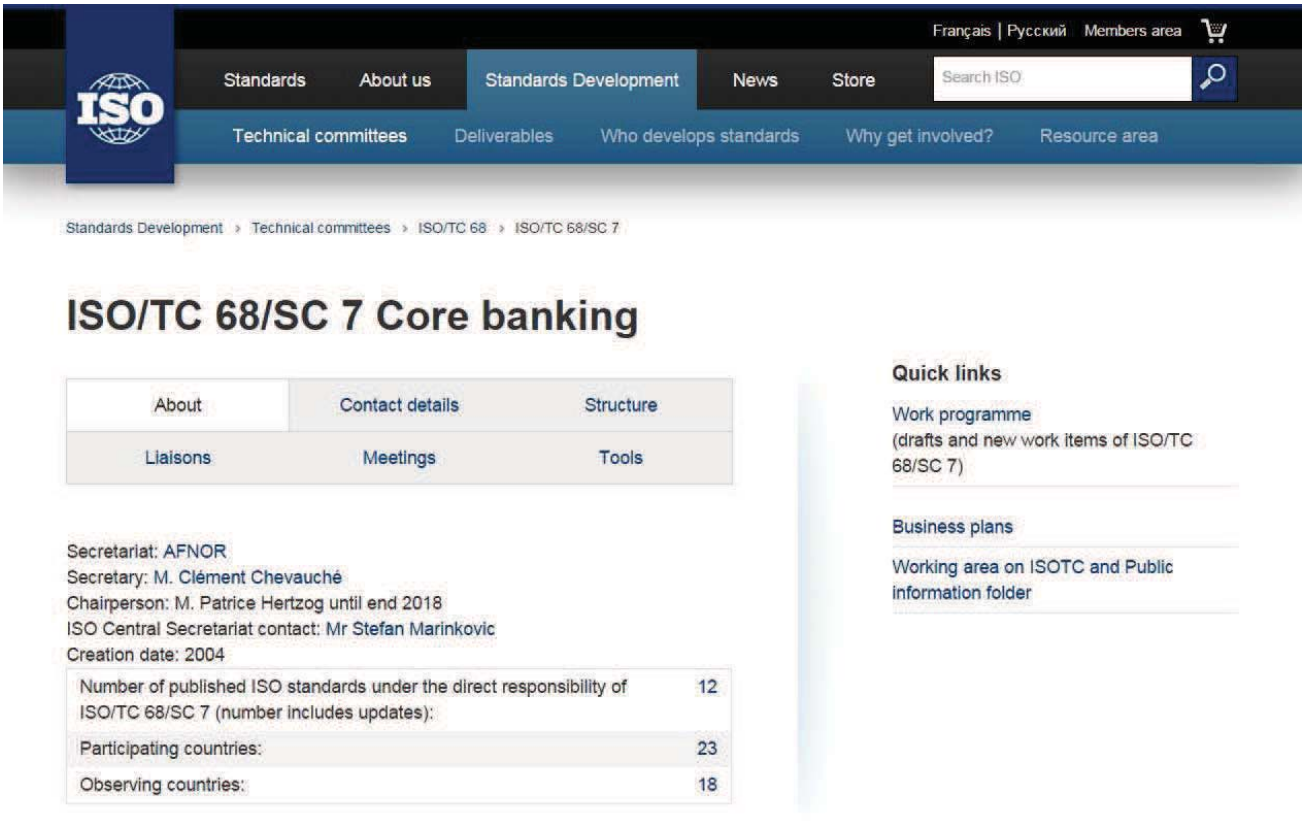


Dr. Naoto IKEGAI

International Public Policy



International Standardization



IMI / September 1th, 2015

National Institute of Informatics

JAPAN

Dr. Hitoshi OKADA

http://www.nii.ac.jp/en/faculty/society/okada_hitoshi/

Attribute-Based Access Control in Mobile Clouds

Sushmita Ruj

Indian Statistical Institute
sush@isical.ac.in

Fine grained access control is a requirement for data stored in untrusted servers like clouds. Owing to the large volume of data, decentralized key management schemes are preferred over centralized ones. Often encryption and decryption are quite expensive and not practical when users access data from resource constrained devices. We propose a decentralized attribute based encryption (ABE) scheme with fast encryption and outsourced decryption. The main idea is to divide the encryption into two phases, offline preprocessing phase which is done when the device is otherwise not in use and an online phase when the data is actually encrypted with the policy. This makes encryption faster and more efficient than existing decentralized ABE schemes. For decryption outsourcing, data users need to generate a transformed version of the decryption key allowing an untrusted proxy server to partially decrypt the ciphertext without gaining any information about the plaintext.

Consider the common scenario where data owners want to upload their data for long-term storage to untrusted servers such as the cloud. The data may initially reside in low-power devices such as mobile phones, smartcards or wireless sensors. These low-power devices may have been used to collect the data. The aim is to store the data over a long time and allow multiple users to access the data. Cloud Service Providers (CSPs) today provide such seemingly unlimited storage facilities and is thus rapidly gaining popularity among individual data owners as well as small and medium-sized enterprises with limited budgets.

In spite of the benefits provided by CSPs, they are assumed to be malicious and data owners generally do not trust them with their sensitive data. So, any data stored in the cloud must be encrypted. Moreover, data owners may wish to impose access control measures on data so that only users who have certain credentials can access it. For example, a hospital may wish to upload to the cloud the results of a clinical trial recording the response of cancer patients to a new drug. This data is extremely sensitive and the hospital may want only the doctor attending a patient or a researcher involved in the drug discovery to have access to the data. Encryption schemes such as attribute-based encryption (ABE) provide great flexibility in terms of access control on encrypted data and are ideally suited for this scenario.

In practical settings, decentralized or multi-authority [4] ABE schemes are very useful, because they do not need any central authority for generation and distribution of decryption keys related to different attributes. For example, the doctor who wants to access a patient's health record for diagnosis may be provided the relevant key by the concerned hospital, whereas a medical researcher may be provided access to the same data by a medical research organization.

The use of these sophisticated encryption schemes pose one severe problem. The encryption and decryption phases are usually very costly, involving several bilinear

pairing operations, and resource constrained devices are not suitable for performing such operations fast enough. One solution to the costly encryption problem is to divide the encryption phase into an offline phase and an online phase, such that, most of the costly operations are performed offline when the user does not immediately expect the encryption to be completed, the device is charging or otherwise not in use. The online phase has little or no computations so that users can get on with their work without the device's performance being affected in any respect.

As observed by Hohenberger and Waters [3], this method allows one to fit almost all of the encryption work into the pre-processing or offline phase, before gaining any knowledge about the message to be encrypted or the attributes related to successful decryption. This was done in the centralized setting. Data users may be relieved from performing costly decryption operations by outsourcing such operations to a proxy server. The proxy server, using a transformed decryption key, partially decrypts the ciphertext. However, the partial decryption process should not reveal any information to the malicious proxy server. Then, the data user needs to perform only a few simple operations to derive the final plaintext from the transformed/partially decrypted ciper-text. An alternative to this method is to make the decryption process fast such that it involves only a constant number of bilinear operations. Earlier work on decryption outsourcing [2]was done in the centralized setting. Using our proposed scheme, the burden on mobile devices is greatly reduced and time for encryption and decryption is much faster than existing scheme.

This [1] is a joint work with Sourya Joyee De and appears in IEEE Globecom 2015.

References

[1] S. J. De and S. Ruj. Decentralized access control on data in the cloud with fast encryption and outsourced decryption. In *IEEE Globecom*, 2015.

[2] M. Green, S. Hohenberger, and B. Waters. Outsourcing the decryption of ABE ciphertexts. In *20th USENIX Security Symposium, San Francisco, CA, USA, August 8-12, 2011, Proceedings*, 2011.

[3] S. Hohenberger and B. Waters. Online/offline attribute-based encryption. In *Proceedings of Public-Key Cryptography - PKC 2014 - 17th International Conference on Practice and Theory in Public-Key Cryptography, Buenos Aires, Argentina, March 26-28, 2014*, pages 293–310, 2014.

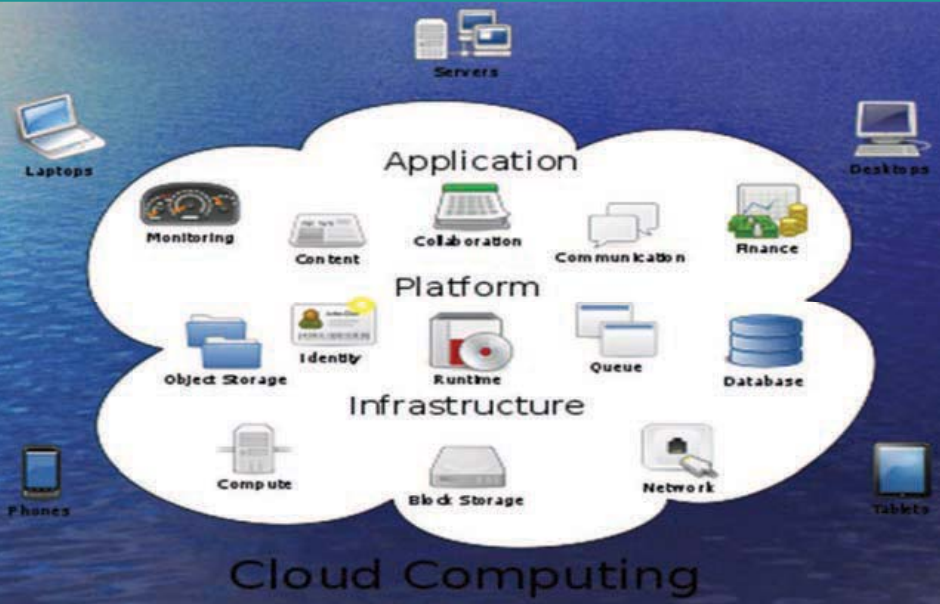
[4] A. B. Lewko and B. Waters. Decentralizing attribute-based encryption. In *Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011.*, pages 568–588, 2011.

Attribute Based Access Control in Mobile Clouds

Sushmita Ruj
Cryptology Research Group
R.C. Bose Center for Cryptology & Security
Indian Statistical Institute, Kolkata
http://www.isical.ac.in/~sush
Email: sush@isical.ac.in

Clouds

Why buy when we can rent?



Ref: Wikipedia

2

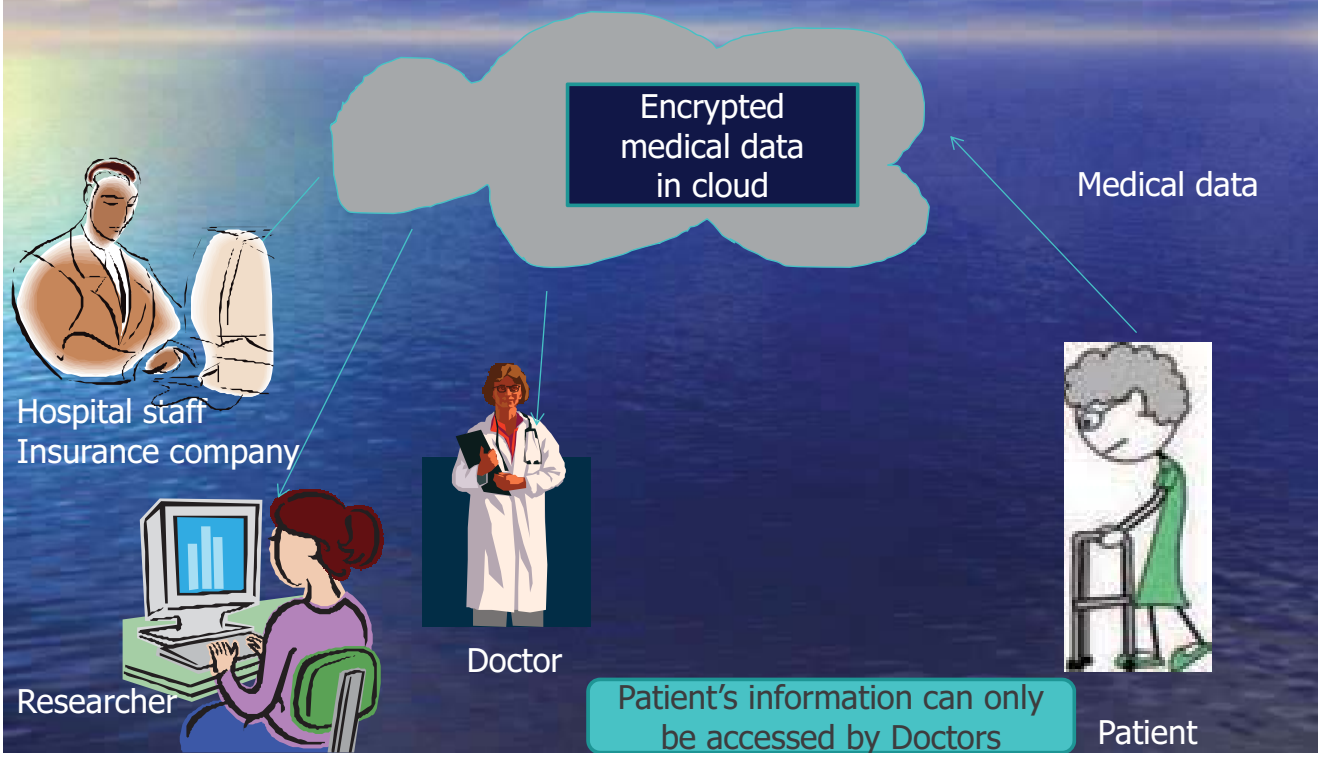
Security issues in Cloud Computing

- A user's data should be protected against adversaries or other users
- Cloud should be oblivious to the data stored
- Cloud should be oblivious to data it is computing
- Cloud should be accountable for its services

Cloud service provider as adversary

- Read/change data
- CSP might not provide the desired amount of redundancy
- Might not provide the amount of storage as specified in the SLA
- Might not provide enough computational resources as specified in the SLA

Storage of medical records



Social Networks

- facebook
- flickr
- LinkedIn

- Share personal information with certain members
- Hide information from others

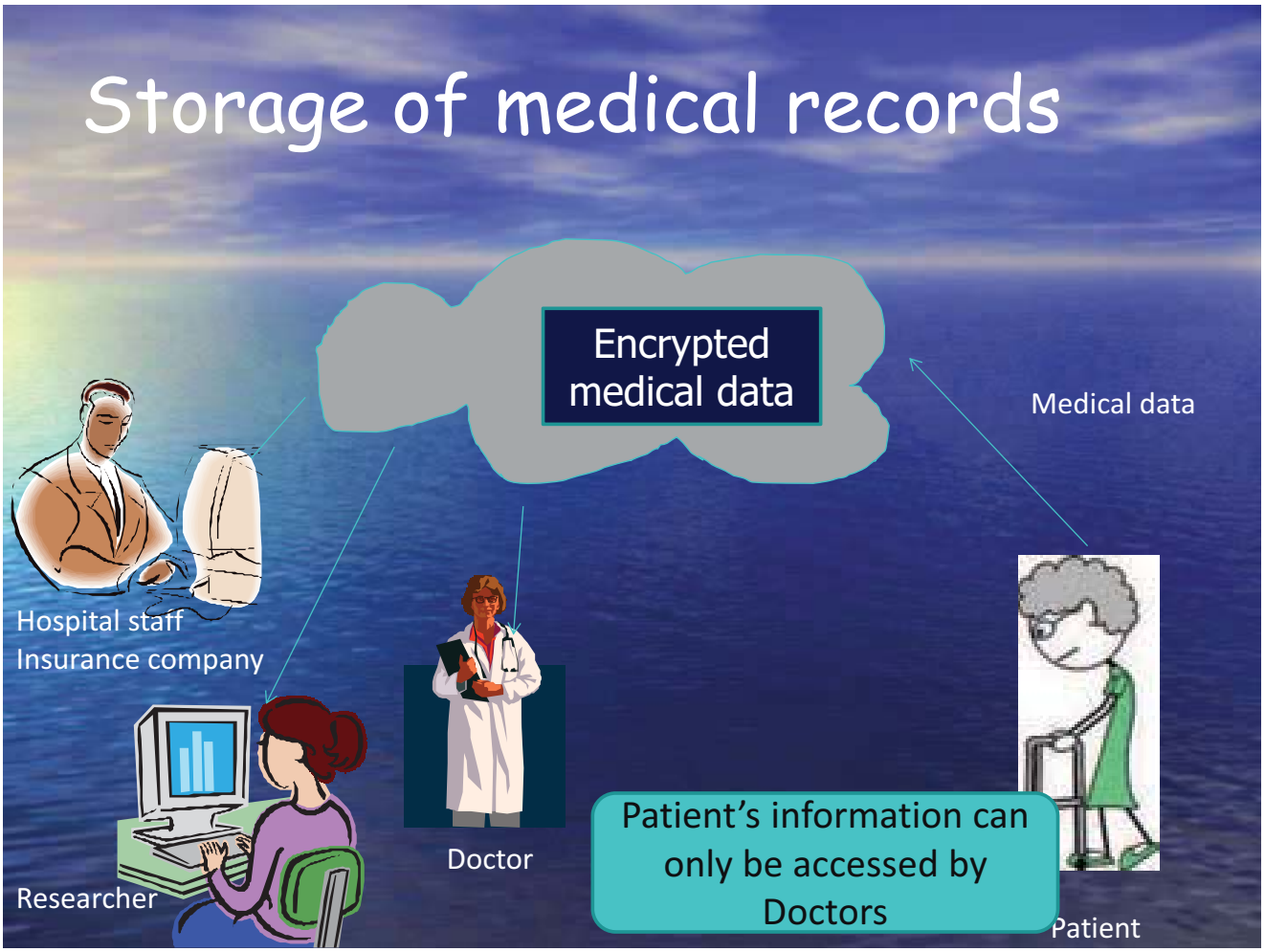
Ways to achieve Access Control

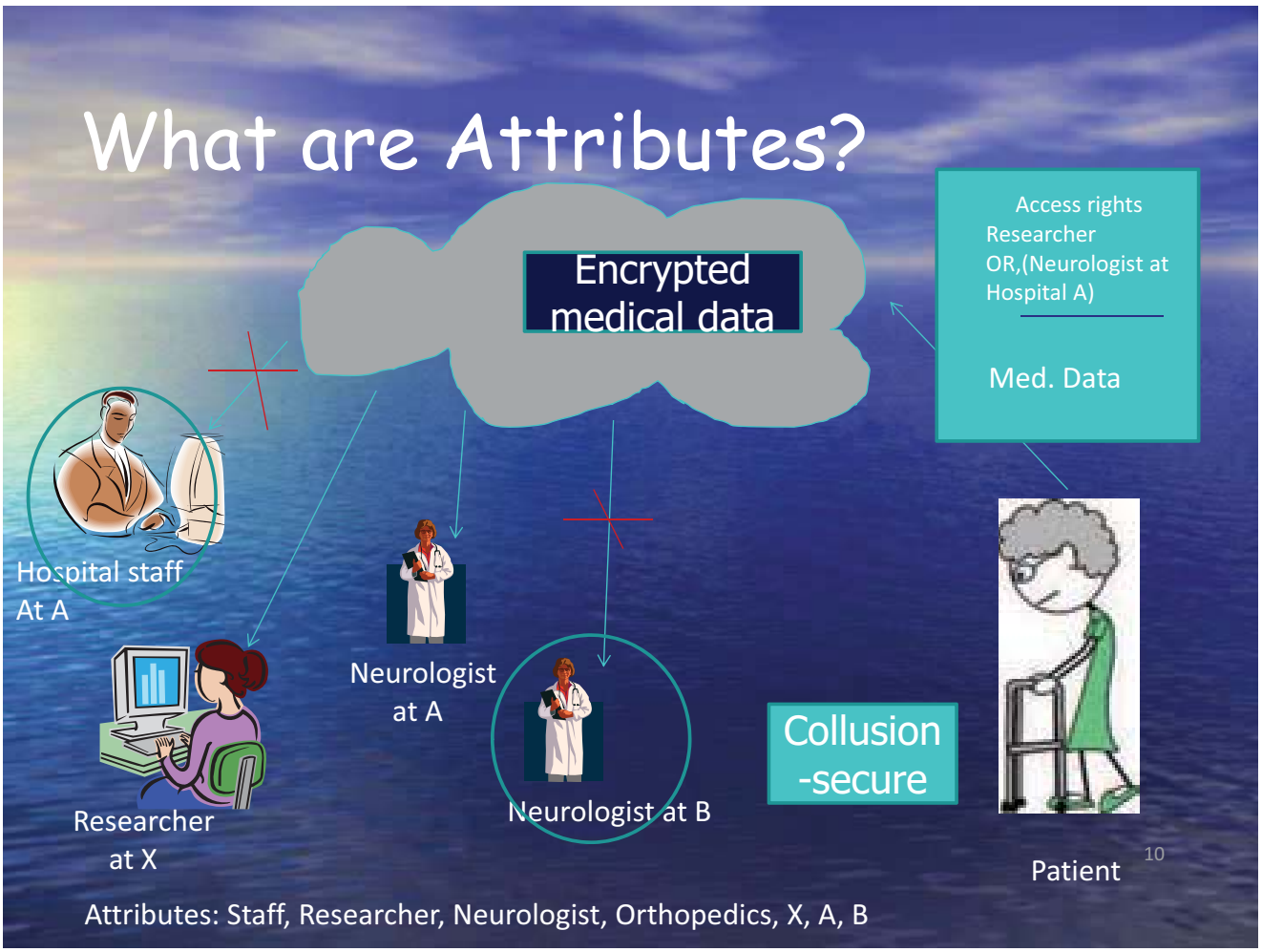
- User based Access Control
  - Access control list attached to data
  - Not a feasible solution when there are many users, example clouds
- Role-Based Access Control
  - Access based on specific role
  - Does not support fine grained access control
- Give each user a public/secret key pair
- Encrypt each message with public key of authorized user
  - Same data has to be encrypted multiple times.
- Attribute based access control
  - Provides fine-grained access control

What is Attribute Based Access Control?

- Users have attributes rather than roles
- Doctor working between 9 am to 5 pm
- Doctor specializes at cardiology
- Has 10 years experience
- Works in Hospital A and research lab R

8


Storage of medical records
Encrypted medical data
Medical data
Hospital staff
Insurance company
Doctor
Researcher
Patient's information can only be accessed by Doctors
Patient


What are Attributes?
Encrypted medical data
Access rights
Researcher
OR,(Neurologist at Hospital A)
Med. Data
Hospital staff
At A
Neurologist at A
Researcher at X
Neurologist at B
Collusion -secure
Patient
10
Attributes: Staff, Researcher, Neurologist, Orthopedics, X, A, B

Access control in Clouds

Giving access to authorized users

- Preventing unauthorized user to access
- Data is encrypted
- Technical challenge: Making the access control mechanism collusion secure =
- Two or more users cannot collude and access records, which each cannot access individually

Attributes

- Attributes are tags that a user can have
- Only users with certain tags have access
- Key distribution centers distribute keys and attributes to users
- Sender encrypts message under certain attributes/access policies
- Receiver can decrypt if it has matching structure/attributes This is the basis of

ATTRIBUTE BASED ENCRYPTION

or

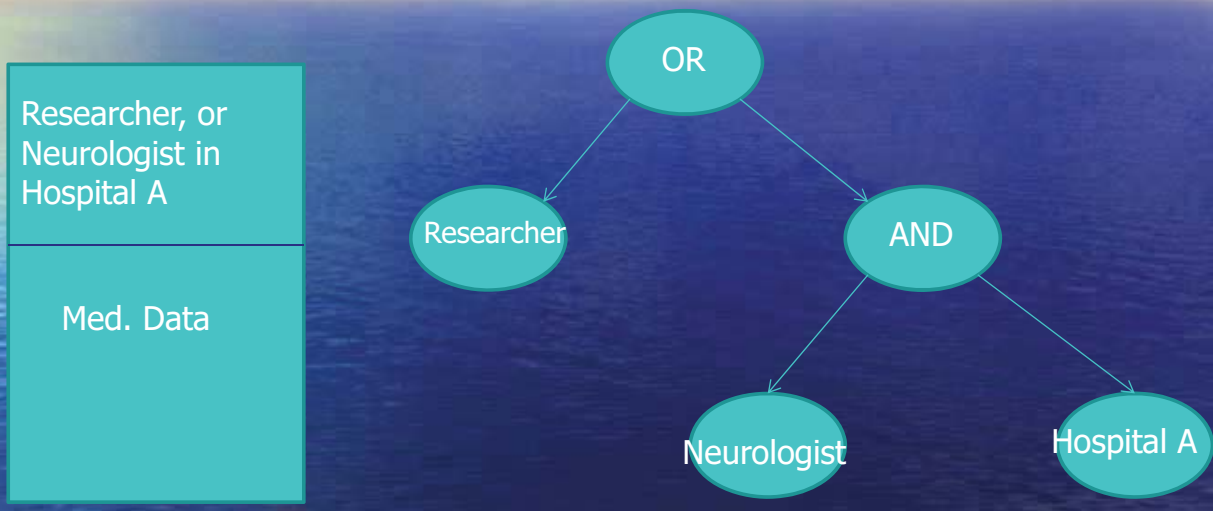
ABE

12

ATTRIBUTE BASED ENCRYPTION ABE

- Select set of attributes
- Assign some attributes to each user and deliver corresponding keys
- Decide the structure of access policy
- Give access to users whose assigned attributes satisfy the access policy
- No need to address individual users unless particular user needs to be revoked by changing some attribute keys
- Sahai and Waters, Eurocrypt 2005
- Goyal el al, CCS'06 …

Attributes access tree

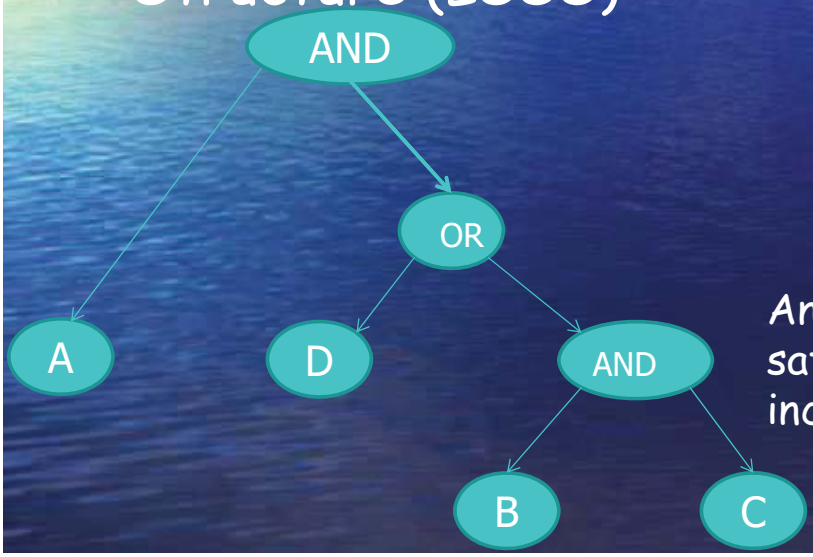


Secret keys and access tree delivered by KDC (key distribution center) to senders (data providers) and receivers (consumers)

14

LSSS

- Access structures can also be represented using Linear Secret Sharing Structure (LSSS)

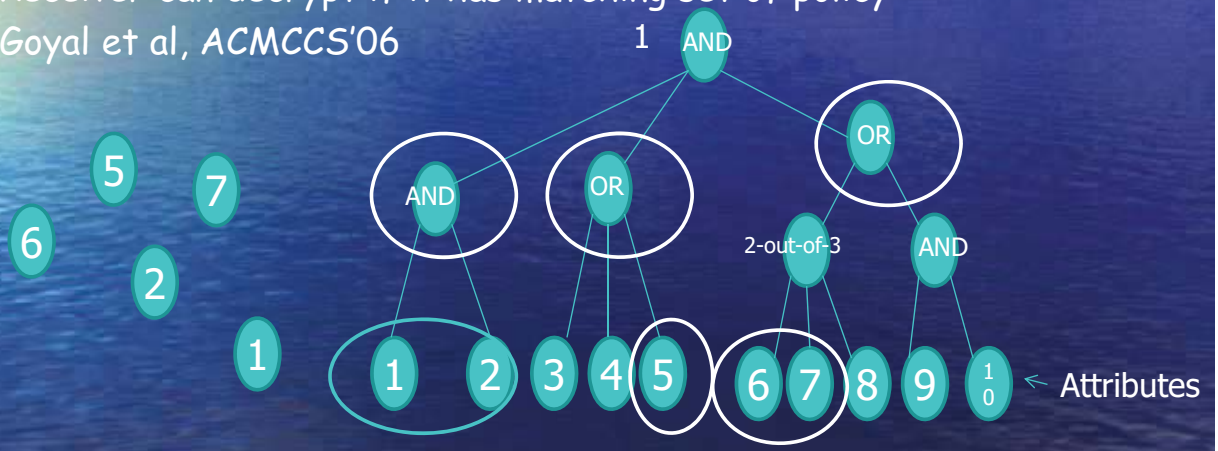


$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Any set of attributes that satisfy this policy will be include (1,0,0) in its span.

Key-policy ABE

- Sender encrypts the data encrypted with certain attributes.
- Receivers have keys with inbuilt access policy
- Receiver can decrypt if it has matching set of policy
- Goyal et al, *ACMCCS'06*



Ciphertext encrypted with attributes

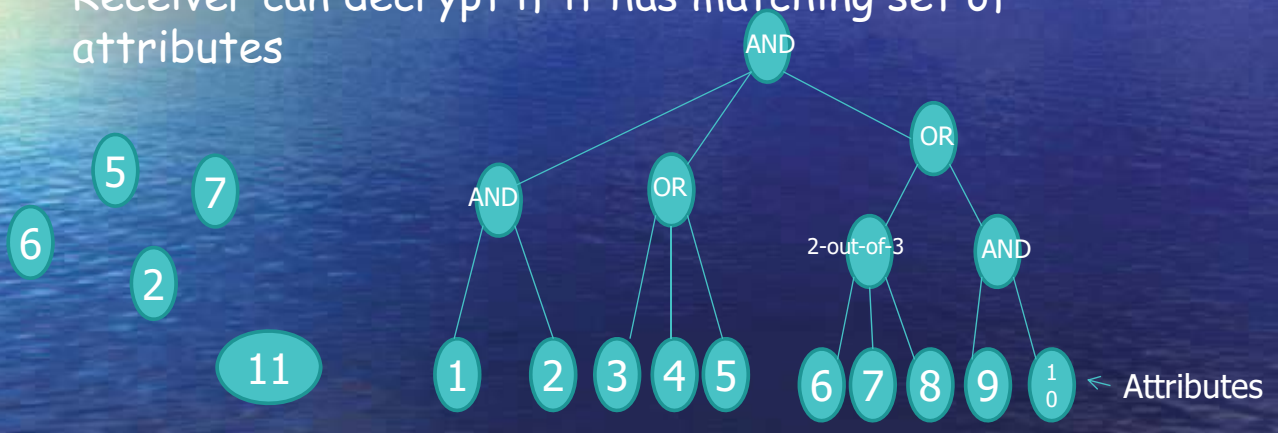
Can decrypt

Receiver

16

Key-policy ABE

- Key-policy ABE: Sender has specific access policy
- Sender encrypts message using key corresponding to the access policy
- Receiver can decrypt if it has matching set of attributes



Ciphertext encrypted with attributes

Cannot decrypt

Receiver

17

Ciphertext-policy ABE

- Reverse of KP-ABE
- Receiver has set of attributes, sender encrypts under an access policy
- Receiver can decrypt if it has matching attributes
- Bethencourt-Sahai-Waters 2007

Access policy

Can decrypt

Receiver

18

Bilinear Pairings

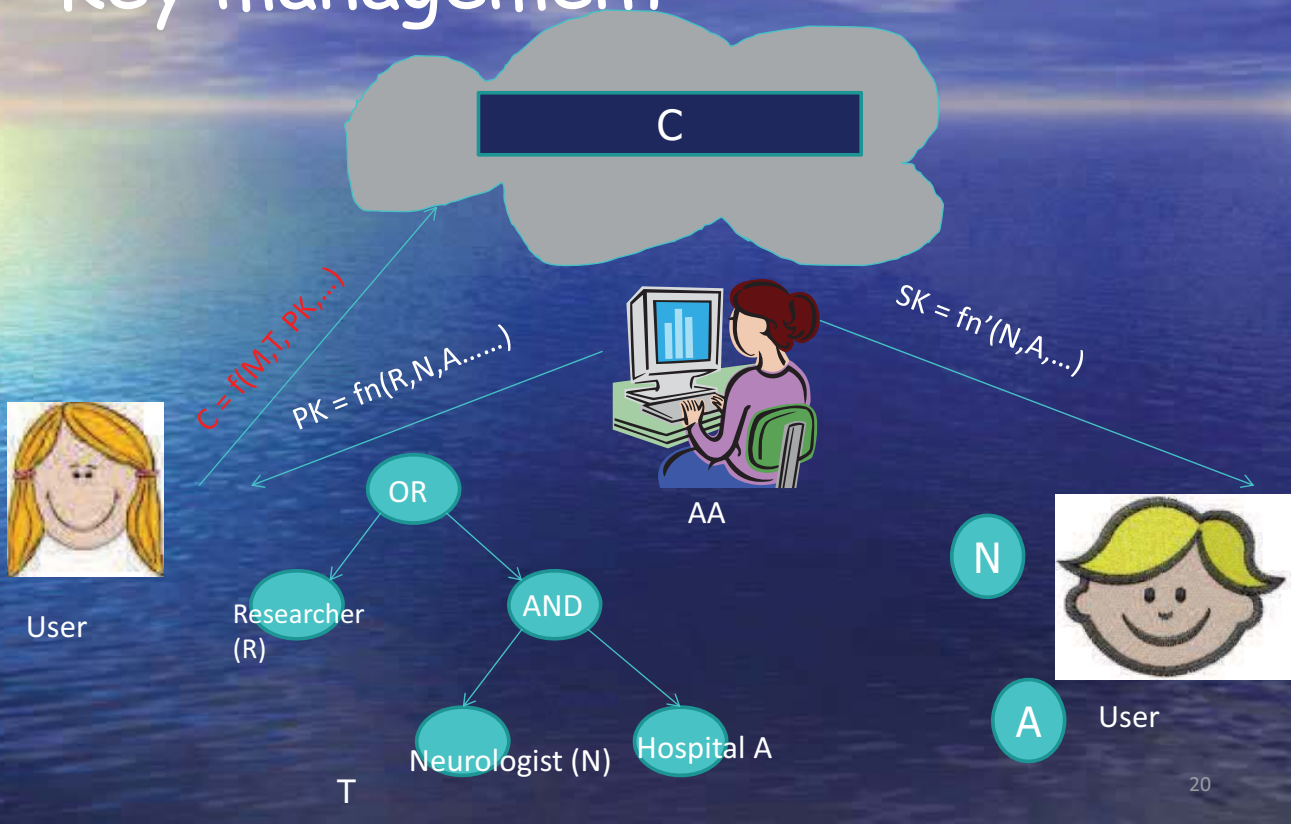
- G, G_T are groups of order p (prime)
- $e : G \times G \rightarrow G_T$ is an a bilinear map if:

-Non degenerate
$e(g,g) \neq 1$

-Bilinear: $e(g^a,g^b) = e(g,g)^{ab}$, $a,b \in Z_p^*$, $g \in G$

-e can be computed efficiently

Key management



20

Requirements, specific to clouds

- Too big for one administrative authority to handle key management
- Large number of users, large access policies, efficient encryption and decryption
- Privacy protection
- Often data is accessed using mobile devices, so what can be done to make access easier?

21

Different flavors of ABE

- KP-ABE, CP-ABE
- Type of access rights: Threshold, monotonic, non-monotonic
- One key distribution center or many : Chase, Chase-Chow, Lewko-Waters, Constant-size ciphertexts/keys
- Hidden policy

22

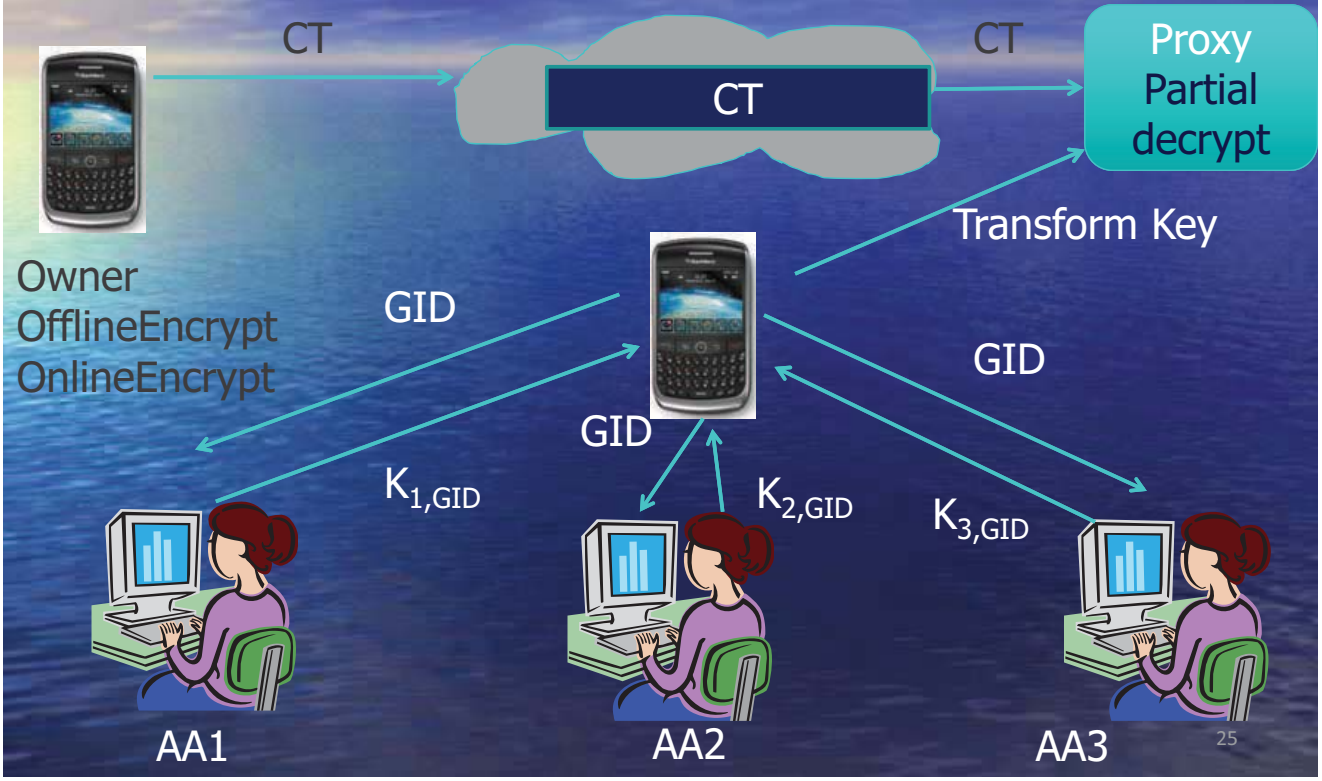
Related work

- Lewko-Waters, "Decentralized ABE", 2011
- Green-Hohenberger-Waters, "Outsourcing decryption of ABE ciphertexts", USENIX, 2011
- Hohenberger-Waters, "Online-Offline ABE", PKC, 2014
- Last two problems were done in a single KDC setting
- We address them in the decentralized setting

Features of our scheme

- Decentralized access control
- Efficient encryption by diving encryption into two phases, online and offline phases
- Outsourced decryption such that the proxy does partial decryption.
- This mobile device at the user end performs few computations

Overview of our scheme



Lewko-Water's scheme

- GlobalSetup: Sets up the system parameters
- AuthoritySetup: For each attribute belonging to an authority generates public and secret keys.
- Encrypt: Uses the intermediate ciphertext to generate ciphertext and sends to the storage server.
- KeyGen: Generates the secret keys corresponding to the attribute policies that user has
- Decrypt: User uses the secret keys to decrypt the ciphertext

Extra steps in our scheme

- GlobalSetup: Sets up the system parameters
- AuthoritySetup: For each attribute belonging to an authority generates public and secret keys.
- Encrypt: Offline and online stages
- KeyGen: Generates the secret keys corresponding to the attribute policies that user has
- Key Transform
- Partial Decrypt
- FullDecrypt

Lewko Waters Scheme in Brief

- Global setup: Choose bilinear group of prime order p. g is a generator of G, H maps global identities GID to elements of G
- Authority setup: For each attribute i belonging to the authority,

-PK = $\{e(g,g)^{\alpha_i}, g^{y_i}\}$

-SK = $\{\alpha_i, y_i\}_{i \in Z_p}$

Lewko Waters Scheme in Brief

- Encrypt: Choose secret $s \varepsilon Z_p$, random vector $v \varepsilon Z_p^n$ with first entry s, $w \varepsilon Z_p^n$, with first entry 0 and calculate:
- $\lambda_i = A_i.v$, $\omega_i = A_i.w$ (A_i is the ith row of A)
- $C_0 = m\, e(g,g)^s$,
- $C_{1,i} = e(g,g)^{\lambda i} e(g,g)^{\alpha i r i}$
- $C_{2,i} = g^{ri}$
- $C_{3,i} = g^{y' i r i} g^{\omega' I}$
- Ciphertext = $\{C_0, \{C_{1,i}, C_{1,i}, C_{1,i}\}_{i \varepsilon Zp}\}$
- Number of pairings is proportional to the number of attributes in the policy.
- We will reduce this in online phase of our scheme

Lewko Waters Scheme in Brief

- KeyGen: The attribute authority gives the key $K_{i,GID}$ corresponding to attribute i to user with GID,

$$K_{i,GID} = g^{\alpha i} H(GID)^{yi}$$

Lewko Waters Scheme in Brief

- Decrypt:
- The user chooses $c_x \varepsilon Z_p$, such that

  $\Sigma c_x A_x = (1,0,0..0)$,

  If such c_x values are not found, then decrption fails. For each room x satisfying the policy,

  $\Pi(C_{1,x} \cdot e(H(GID), C_{3,x})/e(K_{x,GID}, C_{2,x}))^{cx} = \Pi(e(g,g)^{\lambda x} e(H(GID),g)^{wx})^{cx}$

  $=e(g,g)^{s}$ ($v(1,0,0..) =s$, $w(1,0,0,..)=0$)

  Number of pairings proportional to number of attributes needed to satisfy the policy. We will reduce this in our scheme

Our Algorithm

- GlobalSetup: Sets up the system parameters
- AuthoritySetup: For each attribute belonging to an authority generates public and secret keys.
- OfflineEncrypt: Performed at data owner's device. Computes intermediate ciphertext
- OnlineEncrypt: Uses the intermediate ciphertext to generate ciphertext and sends to the storage server.
- KeyGen: Generates the secret keys corresponding to the attribute policies that user has
- KeyTransform: Generates a transformation key to enable the proxy to partially decrypt the ciphertext
- PartialDecrypt: Performed by the proxy. Generates partial ciphertext using the transformation key
- FullDecrypt: User uses the partially decrypted ciphretext and secret keys to fully decrypt the ciphertext

Setup Stage

- Generating global Parameters: Bilinear group of prime order p, g, H (which maps identities GID to elements in G)
- Authority setup: For each attribute i, choose secret keys SK =a_i ,y_i
- PK = $e(g,g)^{ai}$, g^{yi} is published

Offline Encrypt

- For each policy i, choose $\lambda'_i, a'_i\ \omega'_i,\ y_i',\ r_i\ \varepsilon Z_p$
- $C'_{1,i} = e(g,g)^{\lambda' i} e(g,g)^{\ a'iri}$
- $C'_{2,i} = g^{ri}$
- $C'_{3,i} = g^{y'iri} g^{\omega' I}$
- $CT_{1,i} = PK^{ri}_{1,i}\ e(g,g)^{\ -a'iri} = e(g,g)^{\ (ai-ai')\ ri}$
- $CT_{2,i} = PK^{ri}_{2,i}\ e(g,g)^{\ -riyi'} = e(g,g)^{\ (yi-yi')\ ri}$
- $IC = \{C'_{1,i}, C'_{2,i}, C'_{3,i}\}_{i=\{1,2,..P\}}$
- $IS = \{CT_{1,i}, CT_{1,i}\}_{i=\{1,2,..P\}}$
- Preprocessing occurs without of message

Online Encrypt

- The phase is performed at the owner's device.
- Input: Message m, IS, IC, LSSS $A_{l \times n}$
- Choose $s \varepsilon Z_p$, $v \varepsilon Z_p^n$ with first entry s, $w \varepsilon Z_p^n$, with first entry 0: $\lambda_i = A_i.v$, $\omega_i = A_i.w$
- $C_0 = m\, e(g,g)^s$,
- $C'_{4,i} = \lambda_I - \lambda'_I$
- $C'_{5,i} = \omega_i - \omega'_I$
- $CT = \{C_0, \{C'_{1,I}, C'_{2,i}, C'_{3,i}, C'_{4,i}, C'_{5,i}, PK, PK\}_{i \varepsilon P}\}$
- $C_{1,i} = C_{1,i}' \, CT_{1,i}$ $e(g,g)$ $C'_{4,i} = e(g,g)^{\lambda i} e(g,g)^{\alpha i r i}$
- $C_{2,i} = g^{ri}$
- $C_{3,I} = C_{3,i}' \, CT_{2,i}$ $e(g,g)^{C'_{5,I}} = g^{y'iri} g^{\omega' Ii}$
- Reduction in number of pairing operations.

KeyGen

- Performed at the attribute authorities
- Input: GID (User id), Attribute (i), SK, GP
- On presenting attribute i to the suitable authority, authority calculates
- $K_{i,GID} = g^{\alpha i} H(GID)^{yi}$
- Sends $K_{i,GID}$ to user

KeyTransform

- Client calculates transformation key
- Input: $K_{i,GID}$
- Client chooses random $z \varepsilon Z_p$
- $T_{i,GID} = (K_{i,GID}^{1/z}, H(GID)^{1/z})$
- $T_{i,GID}$ is the transformation key to enable proxy to do partial decryption

PartialDecrypt

- This phase is performed at the proxy
- Input: CT, $K_{i,GID}$, GP
- Proxy chooses $c_x \varepsilon Z_p$, such that

  $\Sigma c_x A_x = (1,0,0..0)$,

If such c_x values are not found, then decrption fails.

- It calculates $C_{1,i} = C_{1,i}' \, CT_{1,i} \, e(g,g)^{C'_{4,i}}$

$C_{2,i} = C_{2,i}'$, $C_{3,i} = C_{3,i}' \, CT_{2,i} \, e(g,g)^{C'_{5,i}}$

PartialDecrypt (contd..)

- $CT_1 = \Pi(e(H(GID)^{1/z}), C_{3,i})/e(K_{i,GID}^{1/z}, C_{3,i})^{ci})$
- $CT_2 = \Pi(C_{1,i})^{ci}$
- Proxy sends (CT_1, CT_2) to data user

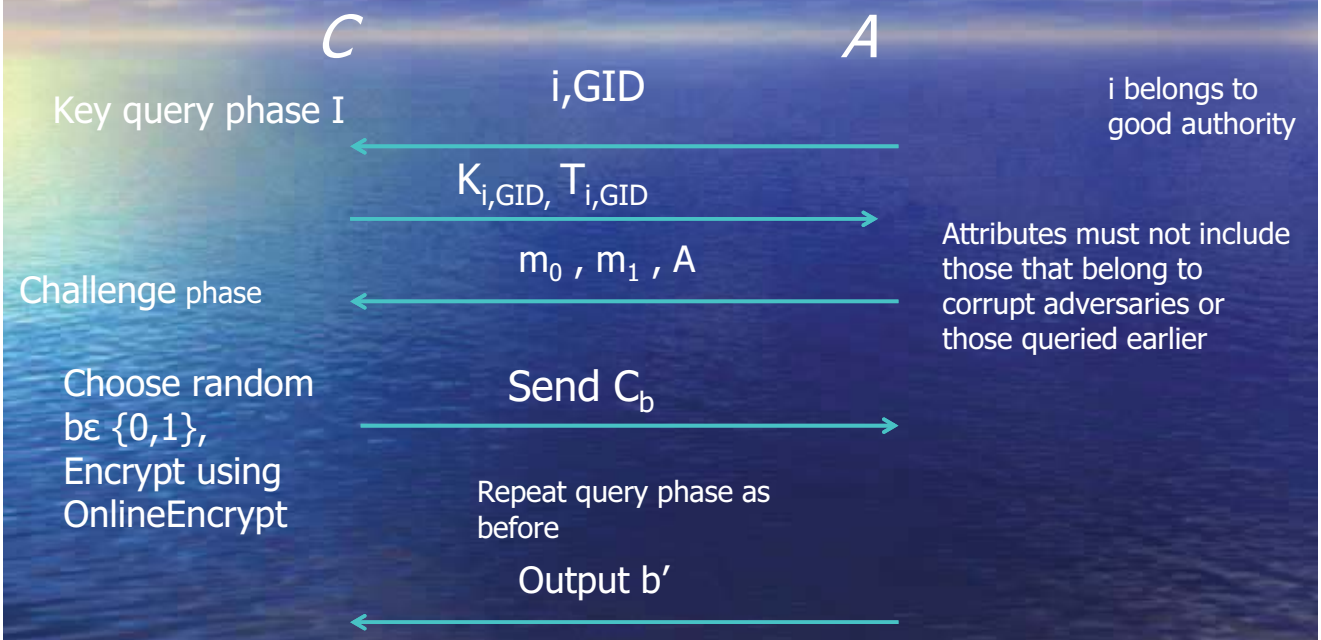
FullDecrypt

- This step is performed at the user
- $CT_2' = (CT_1)^{1/z}$
- $CT = CT_2'.CT_1$
- $CT^z = e(g,g)^s$
- $m = C_0 / CT^z$

No pairing needed, only exponentiations.

Security of the scheme

OO-OD-MA-CPABE Game $_{A,\Pi}(\lambda,U)$



If b = b', the attacker wins and the output of the security game is 1.

Security of the scheme (contd..)

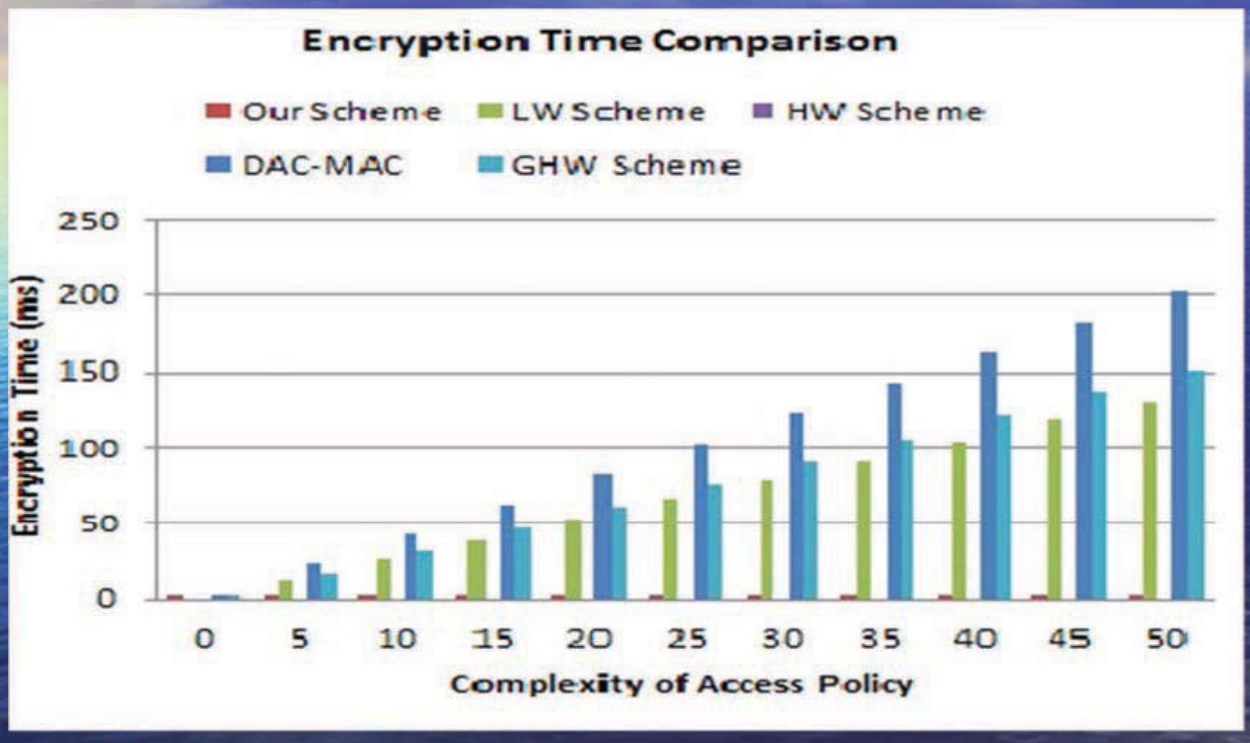
- An online/offline multi-authority CPABE scheme is secure (for attribute universe U) against static corruption of authorities if for all probabilistic polynomial time adversaries A , there exists a negligible function negl such that:

$$\Pr[\text{OO-OD-MA-CPABE Game}_{A,\Pi}(\lambda,U) = 1] \leq 1/2 + \text{negl}(\lambda)$$

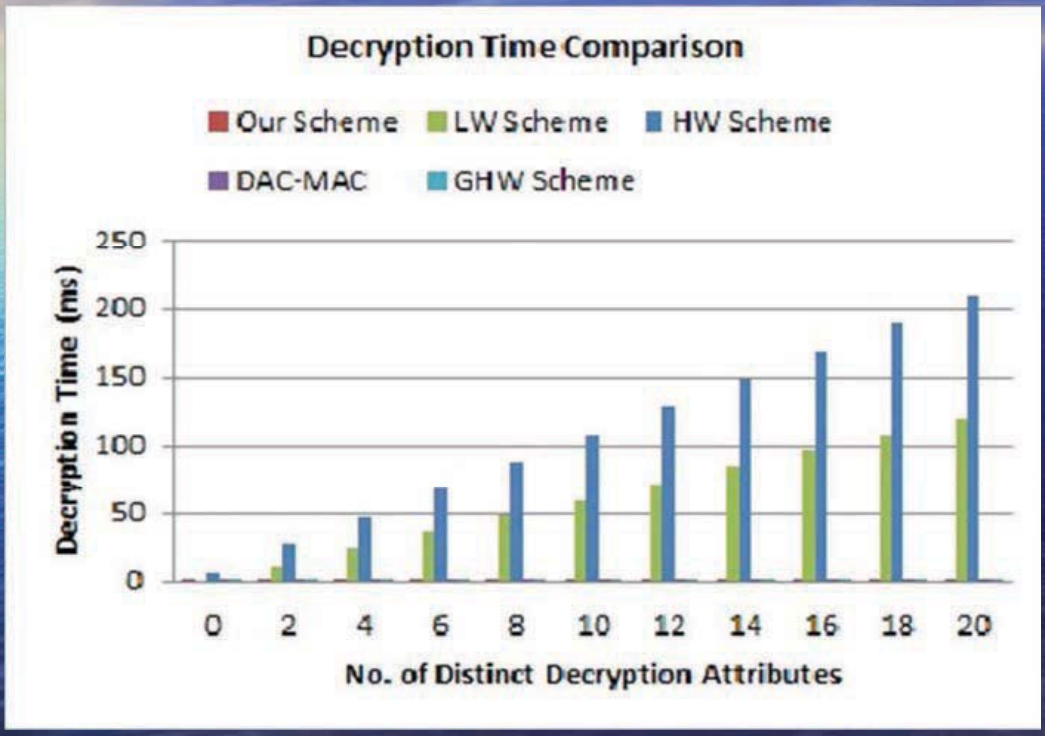
Security of the scheme (contd..)

- The above online-offline multi-authority CPABE scheme with outsourced decryption is secure with respect to the above Definition under the assumption that the scheme of Lewko and Water's scheme is secure.

Comparison with existing schemes



Comparison with existing schemes



Comparison with existing schemes

Encryption Algorithm	Operations	Ciphertext Size
RW Scheme [18]	$1E_T + (5P_{enc}+1)E_1 + 2P_{enc}M_1 + 1M_{Z_p}$	$(3P_{enc}+1)\|G_1\| + \|G_T\| + \|m\| + ln$
HW Scheme (Offline) [9]	$1E_T + (5P_{enc}+1)E_1 + 2P_{enc}M_1 + 1M_{Z_p}$	$(3P_{enc}+1)\|G_1\| + 2P_{enc}\|Z_p\| + ln$
HW Scheme (Online) [9]	$2P_{enc}S_{Z_p} + P_{enc}M_{Z_p}$	
LW Scheme [15]	$3P_{enc}E_T + 2P_{enc}E_1 + (P_{enc}+1)M_T + P_{enc}M_1 + P_{enc}S_{Z_p} + 2P_{enc}M_{Z_p} + H$	$2P_{enc}G_1 + (P_{enc}+1)G_T + \|m\| + ln$
RSN Scheme (without signature) [20]	$2P_{enc}E_T + (3P_{enc}+1)E_1 + P + (P_{enc}+1)M_T + P_{enc}M_1$	$2P_{enc}\|G_1\| + (P_{enc}+1)\|G_T\| + \|m\| + ln$
DAC-MAC [25]	$(4P_{enc}+2)E_1 + E_T + P_{enc}M_T + P_{enc}M_1 + P_{enc}M_{Z_p}$	$(4P_{enc}+2)\|G_1\| + \|G_T\| + \|m\| + ln$
GHW Scheme [6]	$(3P_{enc}+1)E_1 + P_{enc}M_1 + M_T + E_T + P_{enc}H$	$(2P_{enc}+1)\|G_1\| + \|G_T\| + \|m\| + ln$
Our scheme, offline (Section IV)	$4P_{enc}E_T + 4P_{enc}E_1 + 2P_{enc}M_T + P_{enc}M_1, P_{enc}S_{Z_p} + 4P_{enc}M_{Z_p}$	$4P_{enc}\|G_1\| + (3P_{enc}+1)\|G_T\| + 2P_{enc}Z_p + \|m\| + P_{enc}^2$
Our scheme, online (Section IV)	$1E_T + 1M_T + 2P_{enc}S_{Z_p}$	

Comparison with existing schemes

Decryption Algorithm	Operations
RW Scheme [18]	$(3P_{dec}+1)\mathbb{P}+(2P_{dec}+1)M_T+P_{dec}E_T$
HW Scheme [9]	$(3P_{dec}+2)\mathbb{P}+2P_{dec}E_T+(P_{dec}+1)\mathbb{E}_1+(3P_{dec}+1)M_T, P_{dec}M_{Z_p}+(P_{dec}-1)\mathbb{S}_{Z_p}$
LW Scheme [15]	$P_{dec}E_T+(3P_{dec}-1)M_T+2P_{dec}\mathbb{P}$
RSN Scheme [20]	$2P_{dec}\mathbb{P}+3P_{dec}M_T+H$
DAC-MAC, data user's end [25]	E_T+M_T
GHW Scheme [6]	E_T+M_T
Our scheme, data user's end (Section IV)	$2E_T+2M_T+M_{Z_p}$
Our scheme, proxy server's end (Section IV)	$3P_{dec}E_T+(5P_{dec}-3)M_T+P_{dec}\mathbb{E}_1+2P_{dec}M_1+2P_{dec}\mathbb{P}+H$

Open Problems

- Revocation of attributes and users to be considered
- Verifiable outsourcing to be performed
- To reduce communication overheads the ciphertext size is to be reduced
- How to account for hidden attribute set

Related work in Clouds

- Hierarchical storage : Wang et al (ACM CCS 2010)
- Access control for digital health records:
  -Yu et al. (ASIACCS 2010)
  -Li et al. (SecureComm 2009)
- Privacy-preserving : Zhao et al (2011)
- Distributed Access control:
  -Ruj et al (IEEE TrustCom 2011)

Privacy preserving access control with authentication.
-Ruj et al (ACM/IEEE CCGrid 2012, IEEE TPDS)

Distributed Access control

- Use multiple KDCs, prevent single point failure
- Ruj et al (IEEE TrustCom 2011)
- Privacy preserving access control with authentication:
  Ruj et al (ACM/IEEE CCGrid 2012, IEEE TPDS' 2014)
- For digital health records : Li et al. (SecureComm 2009)

Temporal & Comparison based access control

- Handles the case of time-varying/range attributes
- Eg: Valid during 9 pm-5 pm or weight between 60kg-90 kg etc
- Comparison based encryption: Zhu et al. (ACM Codaspy 2011)
- Temporal Attribute based encryption: Zhu et al. (IEEE INFOCOM 2011)
- Temporal access control with user revocation: Balani-Ruj (IEEE TrustCom 2014)

- This paper is a variation of our paper

Sourya Joyee De and Sushmita Ruj, "Decentralized Access Control on Data in the Cloud with Fast Encryption and Outsourced Decryption", Globecom'15. (accepted).

Students and Sponsors









Department of Science and Technology (DST)

DEFENCE R&D ORGANISATION · MINISTRY OF DEFENCE

JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE · FOUNDED 1932 ·

SAMSUNG

http://www.isical.ac.in/~sush

Curiouser and curiouser!

Software Implementation of Bilinear Pairing

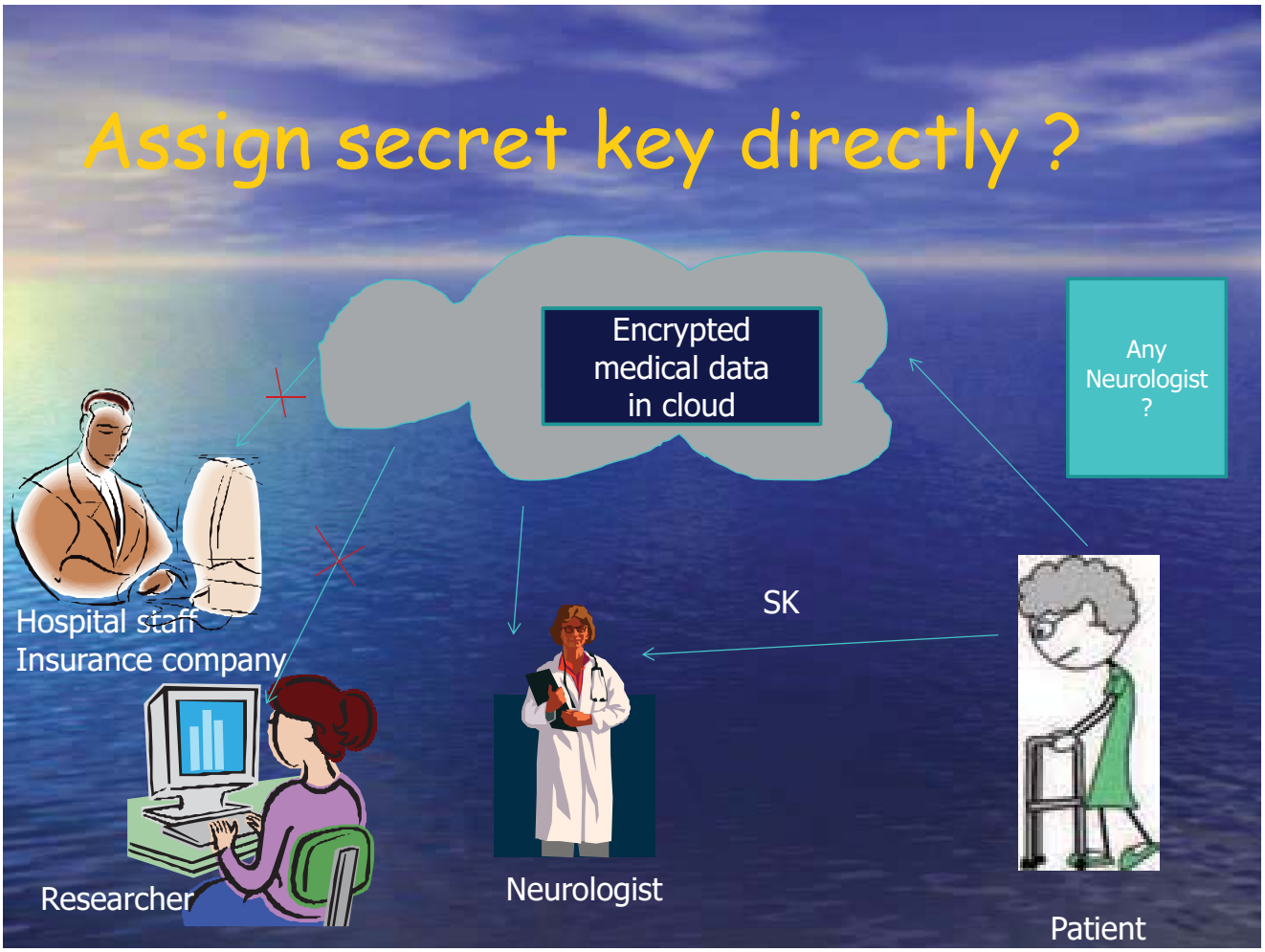
- Choosing pairing friendly curves
- Weil and Tate pairings on Elliptic curves
- Computed using Miller's algorithm
- Pairing Based Cryptography (PBC)

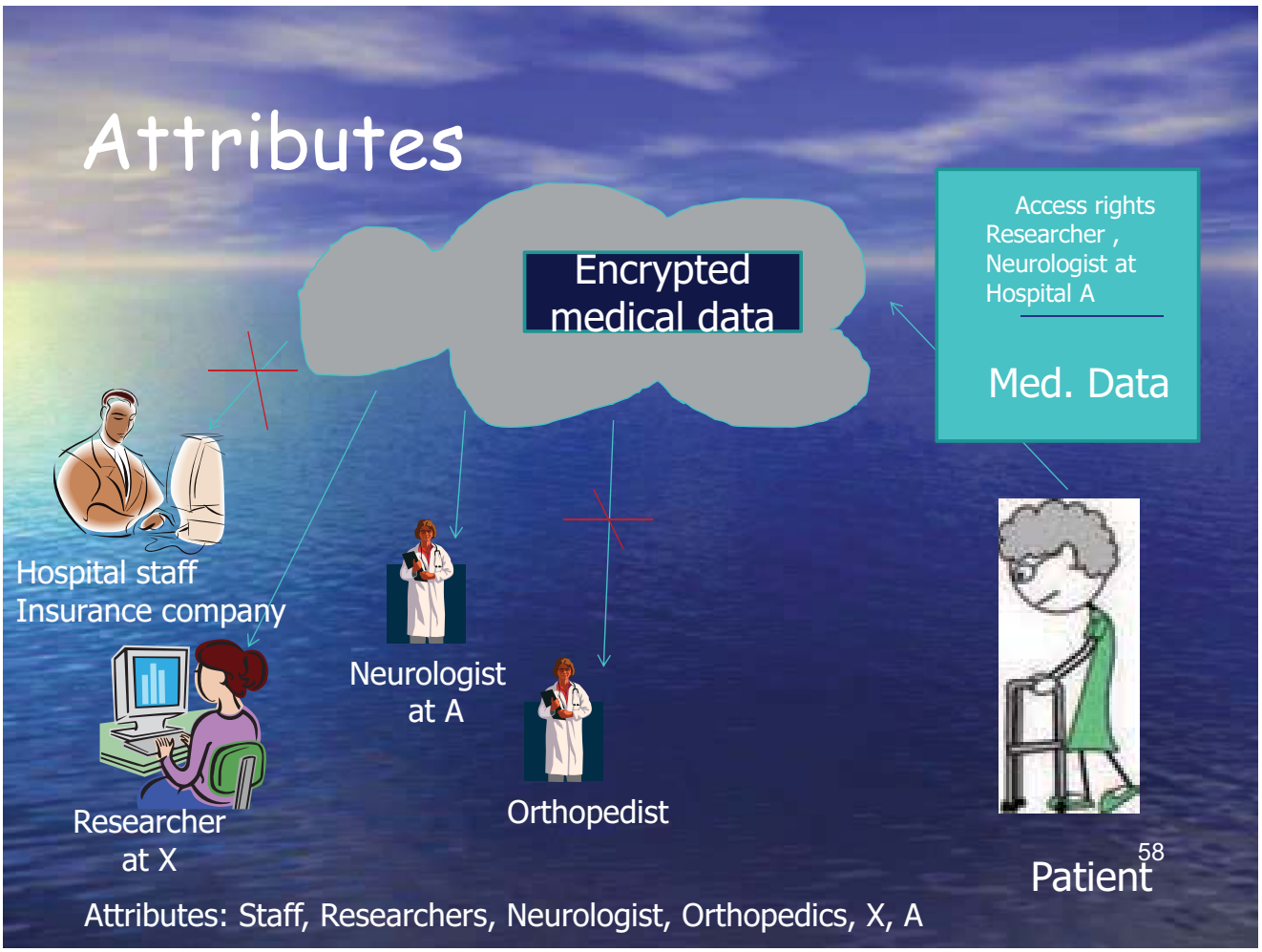
- C library built on GMP (GNU Math Precision) library
- In built algorithms for pairings
- choose the group size (e.g. elliptic curve, group size 159)
- Type of curve (Type d)

Related work in Clouds

Li et al (2010) – Storing health records

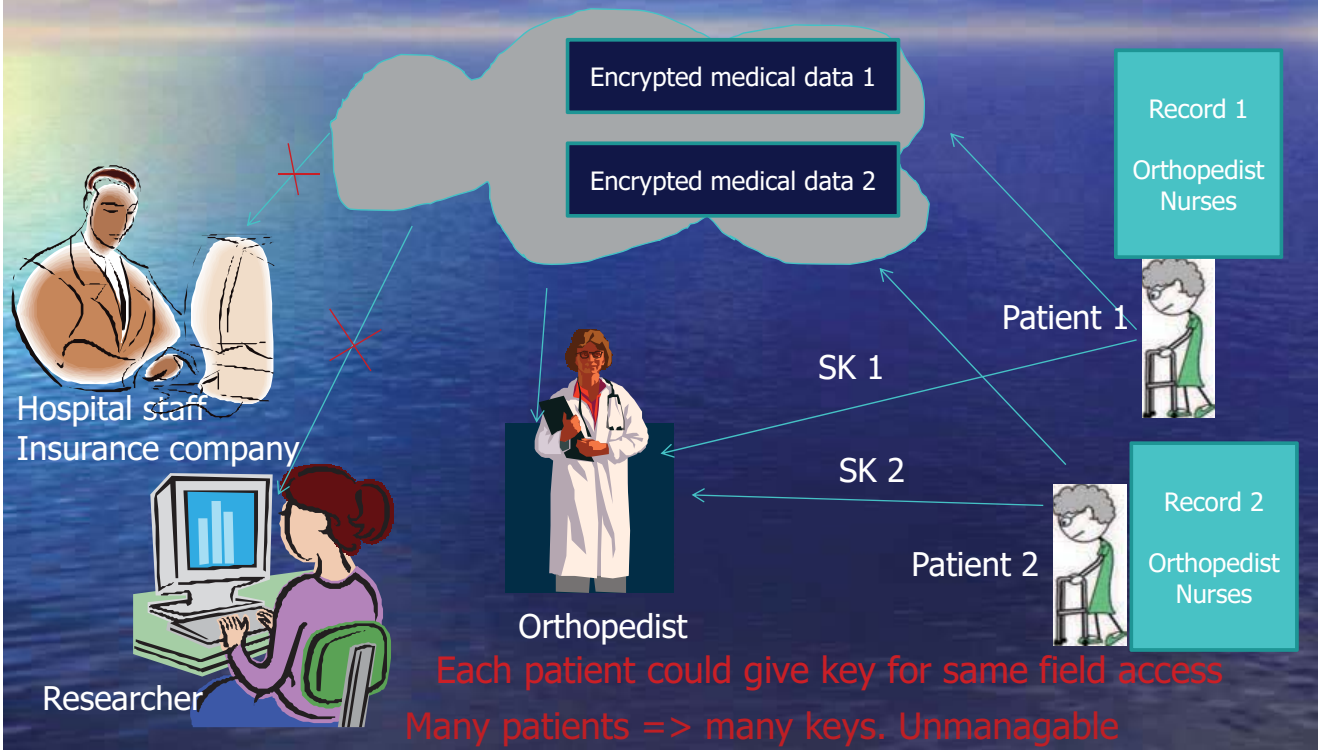
- Wang et al (2010) – Hierarchical storage
- Zhao et al (2011) – Privacy-preserving
- Ruj et al (IEEE TrustCom 2011) – Distributed Access control
- Ruj et al (ACM/IEEE CCGrid 2012)- Privacy preserving access control with authentication.
- Ruj and Nayak (IEEE Trans. Smart Grids, 2012)-A decentralized security framework for data ag-
- gregation and access control in smart grids.
- Green et al (2011) Access control with computation outsourcing


Assign secret key directly ?
Encrypted medical data in cloud
Any Neurologist ?
Hospital staff
Insurance company
SK
Researcher
Neurologist
Patient


Attributes
Encrypted medical data
Access rights
Researcher , Neurologist at Hospital A
Med. Data
Hospital staff
Insurance company
Neurologist at A
Orthopedist
Researcher at X
Patient
Attributes: Staff, Researchers, Neurologist, Orthopedics, X, A

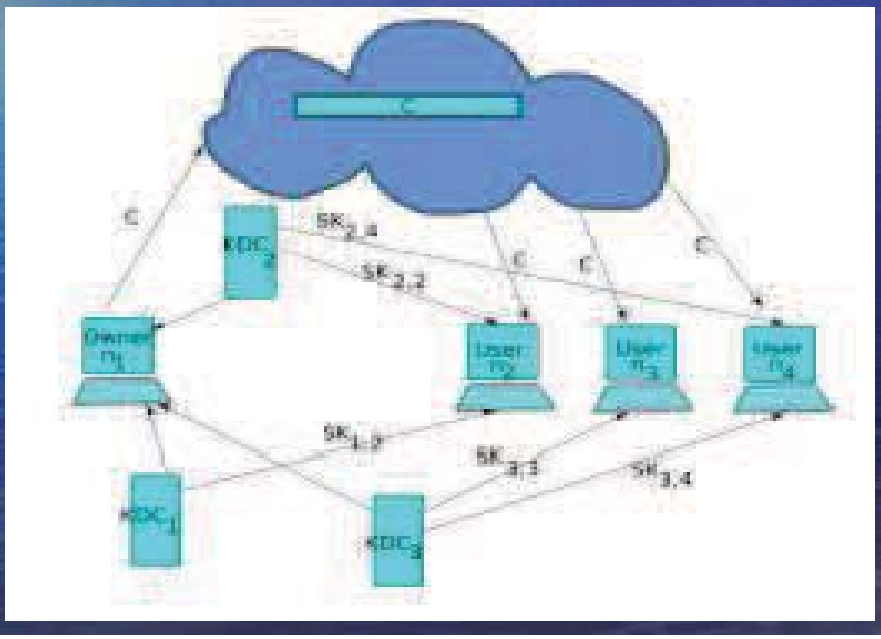
58

Storage of medical records: no KDC

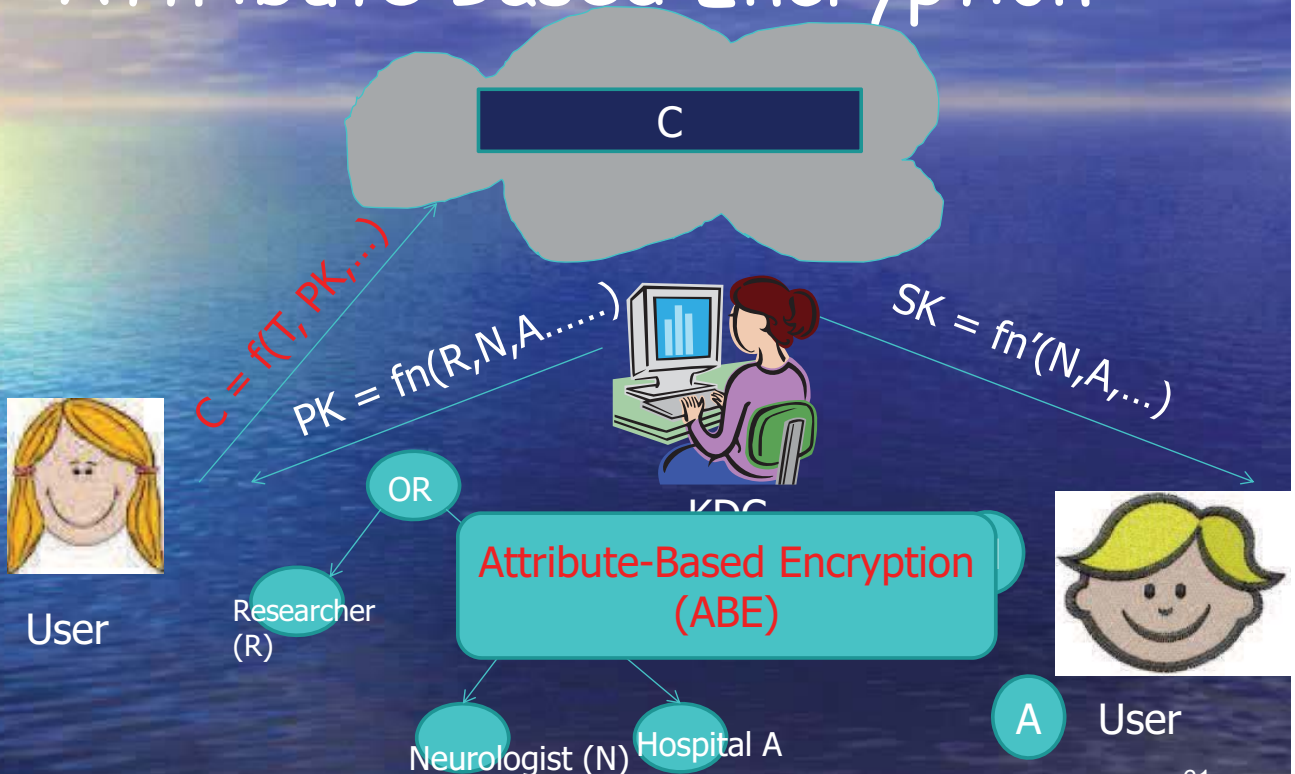


Cloud Model

- Stores data without decrypting it/modifying it



Attribute Based Encryption



61

KeyGen and KeyTransform

- Performed by the device
- KeyGen: For user with id GID, for each attribute i, compute $K_{i,GID} = g^{\alpha i} H(GID)^{yi}$
- KeyTransform: User chooses a random $Z \varepsilon Z_p$ and calculates Transform key $T_{i,GID} = (K_{i,GID}{}^{i/z}, H(GID)^{i/z})$
- $T_{i,GID}$ sent to the proxy

Encryption

- OfflineEncrypt: Generates intermediate ciphertext CT_{1j}, CT_{2j} where j is an attribute in the policy and intermediate state IS
- OnlineEncrypt: Takes the message M, access matrix, IC, IS and generates the final ciphertext
-

KeyGen and KeyTransform

- Performed by the device
- KeyGen: For user with id GID, for each attribute i, compute $K_{i,GID} = g^{\alpha i} H(GID)^{yi}$
- KeyTransform: User chooses a random $Z \varepsilon Z_p$ and calculates Transform key $T_{i,GID} = (K_{i,GID}^{\ i/z}, H(GID)^{\ i/z})$
  - $T_{i,GID}$ sent to the proxy

Decrypt

- PartialDecrypt: Done at the proxy, using the transformation key
- Converts ciphertext CT to CT' = (CT_1, CT_2)
- Requires 2*|A_{owner}| pairing operations where A_{owner} is the number of attributes required for decryption
- FullDecrypt $CT_2' = CT_2^{1/z}$, $C' = CT_2' . CT_1$
- $M = C_0/(C')^z$ No pairing operation is needed, one exponentiation needed.

ABE: Sahai and Water's Idea

- c-out-of-n access structure
- Secret key: $t_1, t_2, \ldots t_W$, $u \varepsilon Z_p$
- Public parameters: $pk_1 = T_1 = g^{t_1}$ $pk_2 = T_2 = g^{t_2}$, …., $pk_w = g^{t_w}, Y = e(g,g)^u$
- Degree c polynomial p(x), s.t. p(0) = u
- Secret keys: $sk_i = g^{\wedge}(p(i)/t_i)$, i is an attribute of user
- Encryption: Choose s
- Ciphertext $C = MY^s$, $\langle T_i^s, \rangle$ i is an attribute

ABE: Sahai and Water's Idea

- Matching set of attributes, calculate
- $e(sk_i, pk_i) = e(g^{p(i)/t_i}, g^{st_i}) = e(g,g)^{p(i)s}$
- If atleast c attributes matching, then calculate $e(g,g)^{p(0)s}$ Lagrange interpolation
- So, $Y^s = e(g,g)^{us}$
- M can be obtained

Key policy ABE (Goyal): sender

Total w Attributes:

Researcher, Neurologist, Orthopedist, …, Hospital A, Hospital B

Secret keys:

$t_1, t_2, \ldots t_W$, u chosen at random from {0,1,…,p-1}

Public parameters (keys):

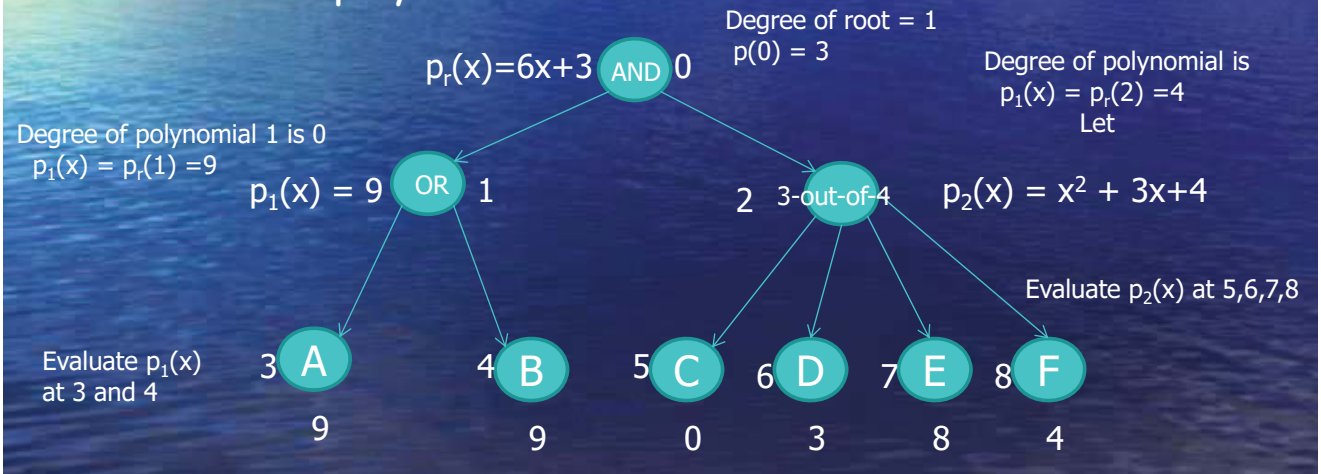
g^{t_1} g^{t_2} g^{t_3} $g^{t_{w-1}}$ g^{t_w}

Key for receiver

- Each receiver receives access tree/structure
- Each node has an unique arbitrary index in {0,1,…q-1}
- One polynomial $p_i(x)$ for each node i in the tree
- If node is c-out-of-d threshold gate,
  then, each polynomial $p_i(x)$ has degree c-1
- For root, $p_r(0) = u$
- For any other node i, $p_i(0) = p_{parent}(i)$
- For each leaf node, calculate value v of the polynomial of parent at the leaves
- Secret key of user = $\{g^v\}$ for each attribute v of receiver

KDC to user

- Consider following tree
- q=11, u =3
- Generate polynomials



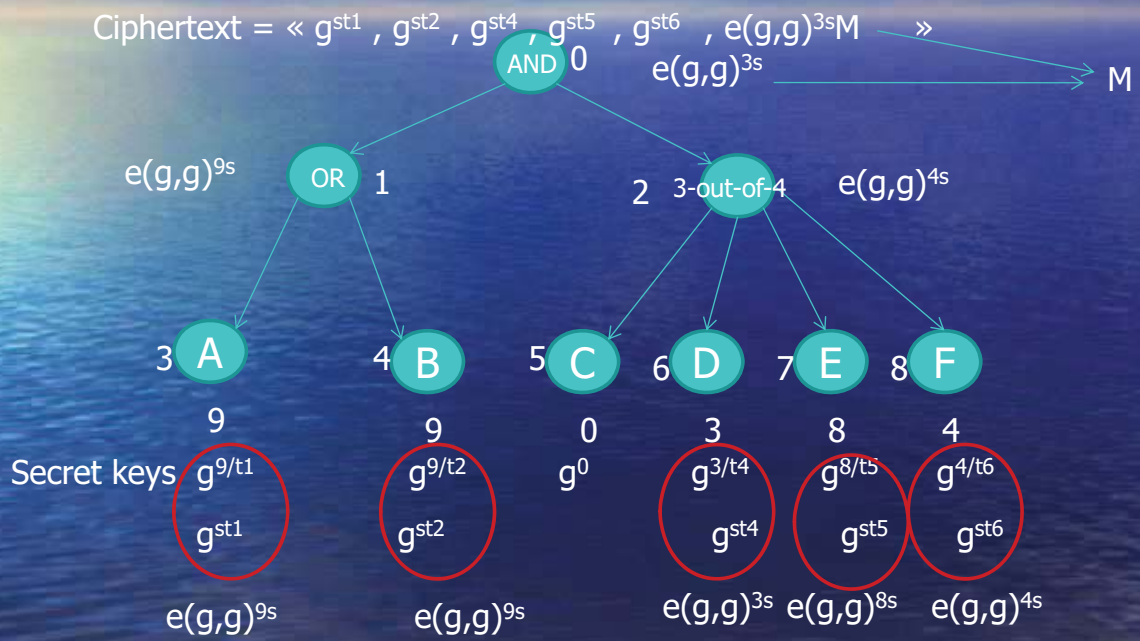
Give secret keys $g^{9/t1}$, $g^{9/t2}$, g^0 , $g^{3/t4}$, $g^{8/t5}$, $g^{4/t6}$ to user

ABE Encryption

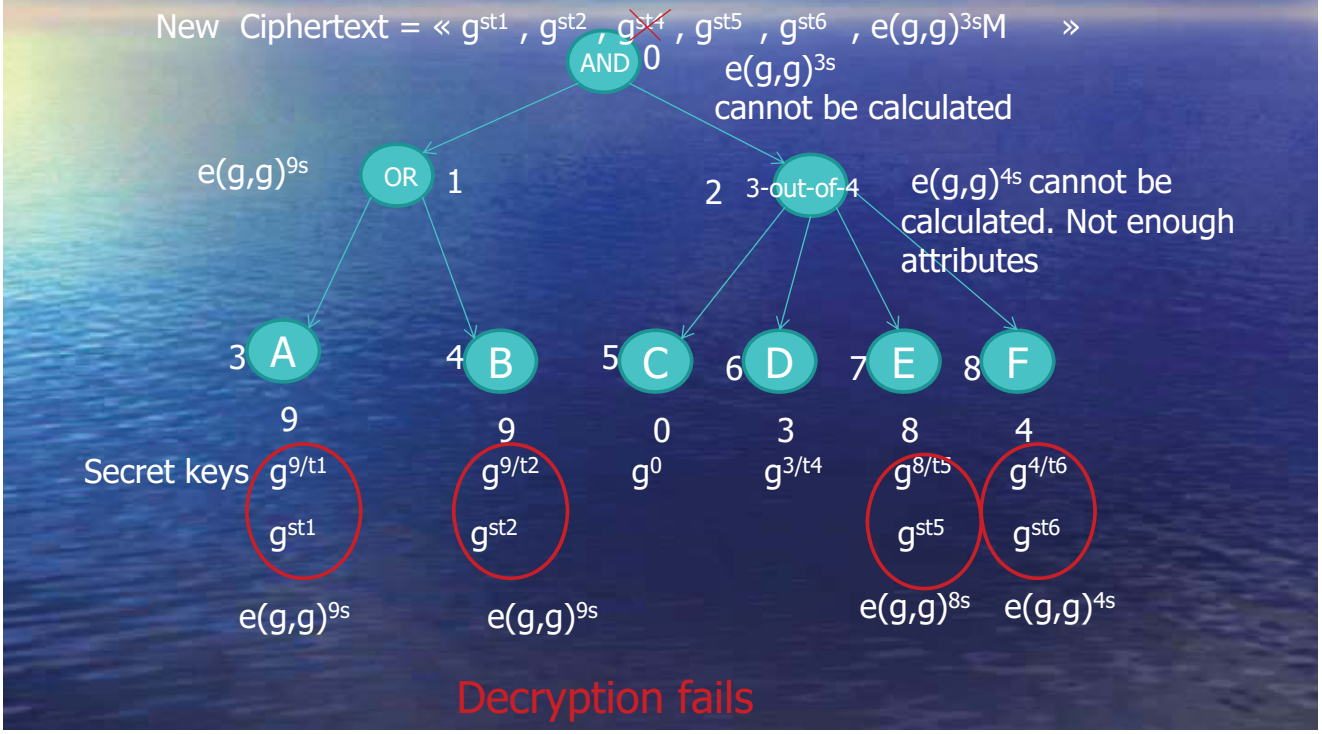
- Choose a random number s in {0,1,…,q-1}
- Raise the public parameters of sender to s
- Send ciphertext

g^{st_1} g^{st_2} g^{st_3} $g^{st_{w-1}}$ g^{st_w} $e(g,g)^{us}$ M

ABE: encryption and decryption



ABE: encryption and decryption



Collusion secure

- Polynomial different for different user
- Must have at least c attributes, to calculate p(0)
- Two users cannot combine attributes and calculate p(0)

Order-Preserving Encryption Secure Beyond One-Wayness

Isamu Teranishi (Joint work with Moti Yung and Tal Malkin)

NEC Corporation

teranisi@ah.jp.nec.com

An *Order-Preserving Encryption* (OPE) is a symmetric encryption over the integers such that ciphertexts preserve the numerical orders of the corresponding plaintexts. That is, if plaintexts m and m' satisfy $m < m'$, then their ciphertexts satisfy $\mathsf{Enc}_K(m) < \mathsf{Enc}_K(m')$.

OPE is attractive since it allows one to simultaneously perform very efficiently over encrypted data numerous fundamental database operation, range queries (i.e., finding all messages m within a given range $\{i, , j\}$).

Furthermore, OPE is more efficient than these other primitives. For instance, the simple matching operation realized by OPE only requires logarithmic time in the database size, while the same operation realized by, say, searchable encryption, needs linear time in the size, which is too costly for a database containing a few millions data items.

Despite its importance, security of OPE is far from being understood at this time. It is known that no OPE scheme can satisfy a "naturally defined" indistinguishability notion, which means that, even the most fundamental problem: "what plaintext information can be semantically hidden" is open.

In this talk, we define a new and weaker indistinguishability notion and propose a new OPE scheme satisfying this notion. We then show that an OPE scheme satisfying our indistinguishability notion can hide lower order bits of a plaintext and satisfies a known one-wayness notion about OPE.

Order-Preserving Encryption Secure Beyond One-Wayness

Isamu Teranishi（NEC）
Moti Yung (Google, Columbia University)
Tal Malkin (Columbia University)

Order Preserving Encryption (OPE)

Secret Key Encryption Scheme s.t.

- Plaintext and Ciphertext Spaces are intervals of the set of integers.
- It satisfies the order-preserving property:

$$m < m' \Leftrightarrow Enc_K(m) < Enc_K(m')$$

Application

- OPE can be used in encrypted outsourced database
- (**Range Query**) Because OPE enables one to find documents m satisfying

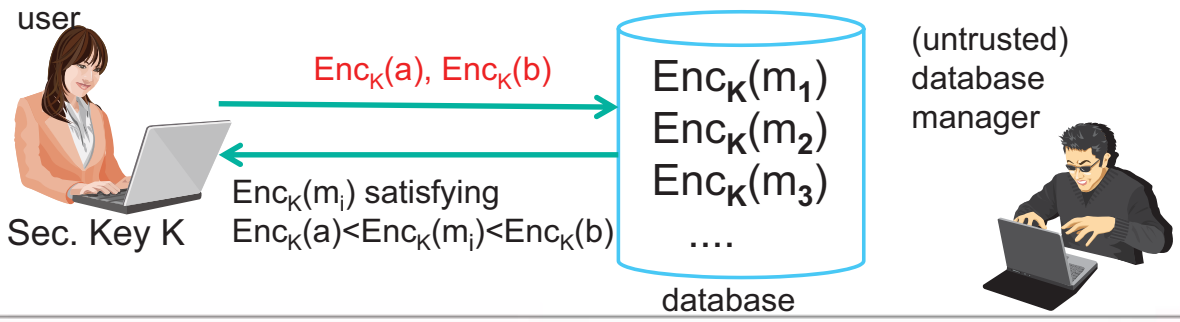
$$a<m<b$$

without decrypting ciphertexts.

- In fact, due to the order-pres. property, one can find such m by checking whether

$$Enc_K(a) < Enc_K(m) < Enc_K(b)$$

holds or not.



Subject and Results of This Paper

- However, security of OPE is far from being understood at this time.
  - In fact, a naturally defined indistinguishability notion (IND-O-CPA) cannot be achievable (under some natural condition) [1].

- In this paper we tackle the following fundamental problem for OPE:

what exactly must OPE leak?,

and what can it hide?

- And we show a positive results for it:
  - Define a weaker indistinguishability notion, (X,T,q)-IND, for OPE than the known (unachievable) one while the known result[2]is about one-wayness
    - the notion is natural in the database setting mentioned before.
    - the notion can ensure that secrecy of lower bits of plaintext.
  - Propose a new OPE scheme satisfying our indistinguishability notion.

[1] Boldyreva, Chenette, Lee, O'Neill: Order-Preserving Symmetric Encryption. EUROCRYPT 2009: 224-241

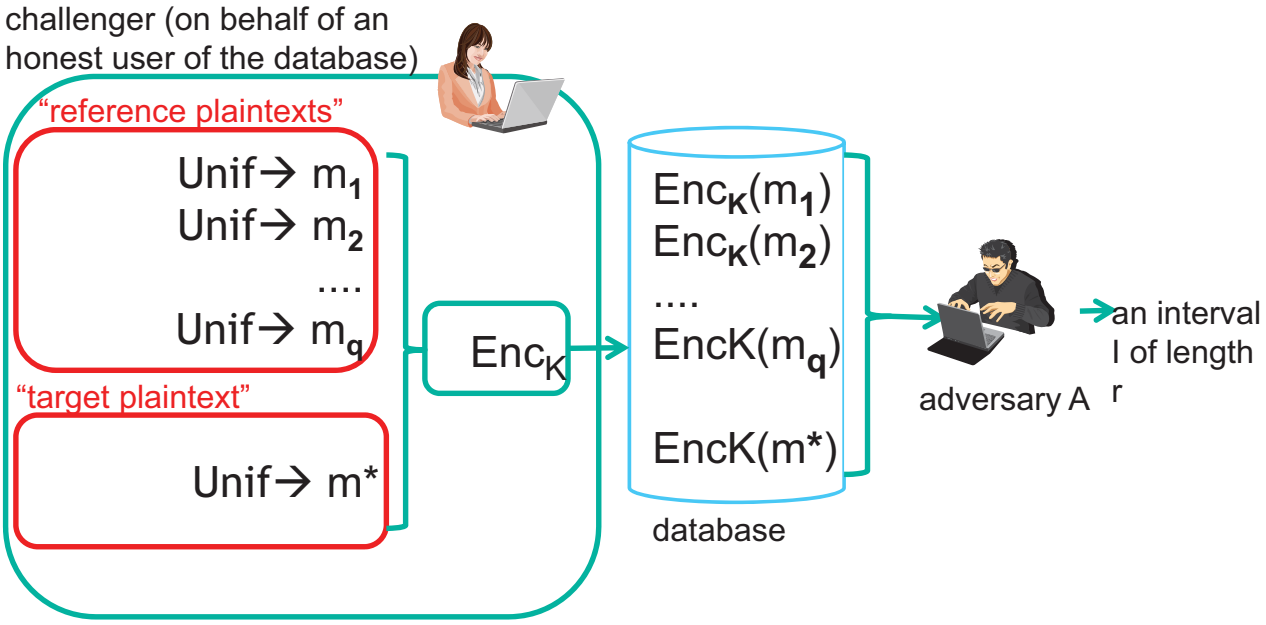
Rest of This Talk

- Our Definition of Indistiguishability Notion
- Our Results
- Construction of Our scheme
- Security Proof

- Our Definition of Indistiguishability Notion
- Our Results
- Construction of Our scheme
- Security Proof

Review of (r,q+1)-WOW (Window-OneWay)

Our security notion is obtained by modifying the following known one-way based notion, (r,q+1)-WOW [2]

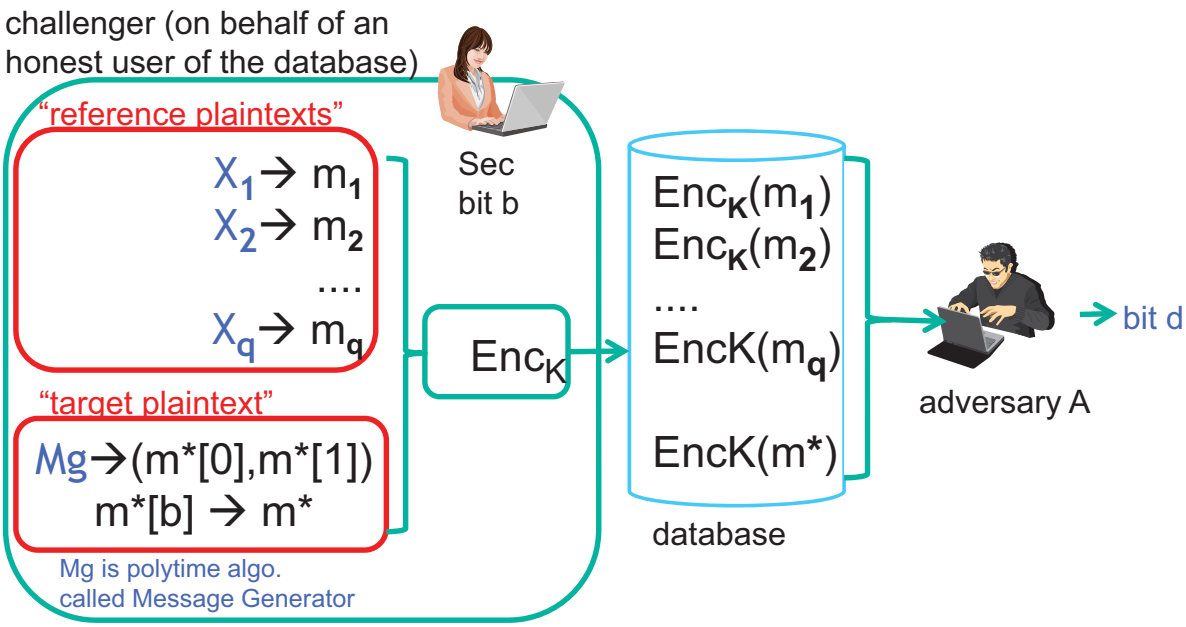


$\forall$ A (polytime) Pr[m* $\in$ I] $\leq$ neg(Mess. Sp. Size)

[2] Boldyreva, Chenette, O'Neill: Order-Preserving Encryption Revisited: Improved Security Analysis and Alternative Solutions. CRYPTO 2011: 578-595

Our notion (X,T,q)-IND

Here X =(X_1,...,X_q) be a tuple of (indep.) distributions on the Mess. Sp.



$\forall$Mg (polytime) whose output satisfies |m*[0]-m*[1]| <T
$\forall$ A (polytime) |Pr[d=b]-1/2| $\leq$ neg(Mess. Sp. Size)

Why |m*[0]-m*[1]| <T ?

- In our def., we require a message generator Mg to output (m*[0],m*[1]) satisfying

$$|m^*[0]-m^*[1]| < T$$

- This is because otherwise, an OPE is broken easily using the following idea [1]:
  - The order-pres. property

$$m < m' \Longrightarrow Enc_K(m) < Enc_K(m')$$

    means that Enc_K is monotone increasing.
  - Hence, if we allow an adversary to select (m*[0],m*[1]) such that

$$m^*[1] - m^*[0]$$

    is large, the difference

$$Enc_K(m^*[1]) - Enc_K(m^*[0])$$

    has to become noticeably large.
  - Therefore, the adversary can distinguish $Enc_K(m^*[0])$ and $Enc_K(m^*[1])$ easily.

Property of (X,T,q)-IND

Our (X,T,q)-IND implies that the least significant log T bits of a plaintext are hidden from the adversary in our database setting.

Proof (rough idea)

- Consider the following two messages:
  - m*[0] : any message
  - m*[1] : lower log T bits are selected randomly and the other bits are the same as those of m*[0]
- Then, it holds that

$$|m^*[0]-m^*[1]| < T,$$

  which is our condition for (X,R,q)-IND.
- Hence, $Enc_K(m^*[0])$ is indis. from $Enc_K(m^*[1])$.
- Recall that the lower log T bits of m*[1] is random.
- This means that an adversary given $Enc_K(m^*[0])$ cannot know the lower log T bits of m*[0].

- Our Definition of Indistiguishability Notion
- Our Results
- Construction of Our scheme
- Security Proof

Our Result (Informal)

Very roughly, we construct an OPE scheme such that

Main Thm.(informal) if min-entropies of $X_1,\ldots,X_q$ are large, our scheme is (X,T,q)-IND for a large T. (Here $X=(X_1,\ldots,X_q)$.)

To formalize the above statement, we give some def.

- The **min-entropy** of random variable X_i on a Mess. Sp. is

$$H_\infty(X_i) := \min \{ - \log \Pr[X_i = m] \mid m \in \text{Mess. Sp.}\}$$

- It is known that the min-entropy of X_i has to less than that of Unif on Mess. Sp:

$$H_\infty(X_i) \leq H_\infty(\text{Unif}) \ (= \log \#(\text{Mess. Sp.}))$$

- So we define "***normalized***" **min-entropy** of X as follows:

$$H^*_\infty(X_i) := H_\infty(X_i) / H_\infty(\text{Unif}) \leq 1$$

- for a *tuple* $X=(X_1,\ldots,X_q)$ of random variables, we also define

$$H^*_\infty(X) := \min_i H^*_\infty(X_i)$$

Our Result (Formal)

We construct an OPE scheme E[α, β] satisfying the following property:

Main Thm (Formal):
For a tuple of (indep) rand. variable X =(X_1,...,X_q) satisfying

$$H^*_\infty(X) > \beta,$$

our scheme E[α, β] satisfies

$$(X, M^\alpha, q)\text{-IND}$$

for any $0 < \alpha < \beta$.

Here M is Mess. Sp.Size.

Our scheme is based on a PRF and the above result holds under security of PRF.

Corollaries

- Recall that our (X,M^α,q)-IND can hide lower bits of a plaintext
- Hence, the following corollaries hold (under the same assumption as above).

Corollary: Our scheme E[α, β] can hide fraction α of lower bits of plaintexts for any $\alpha < \beta$ satisfying $\beta < H^*_\infty(X)$.

In particular, if X is a tuple of the Unif distributions, it follows that

Corollary: Our scheme E[$\alpha, 1$] can hide any fraction of lower bits of plaintexts.

(r,q+1)-WOW of Our Scheme.

- We can show the following fact as well:

Theorem: (Unif^{q},T,q)-IND implies (r,q+1)-WOW for suitable r.

- In particular, we can conclude the following corollary:

Corollary: Our scheme satisfies (M^{s},q+1)-WOW for any $0<s<1$

- In the case of the known scheme [1], it is shown that
  - the known scheme is (1,q+1)-WOW
  - but it is *not* (M^{s},q+1)-WOW for $s > 1/2$.

Hence, our scheme achieve (r,q+1)-WOW for better parameter r than the known scheme [1].

- Our Definition of Indistiguishability Notion
- Our Results
- Construction of Our scheme
- Security Proof

Construction (1/4)

We construct our scheme in the following two steps:

- First, we construct a scheme
  - which satisfies our (X,M^{α},q)-IND without assuming any computational assumption.
  - But the enc. and dec. of this scheme requires super-polytime

→ Today we talk about this scheme

- Second, we improve the above scheme
  - Here we use the "lazy sampling" technique [2],
  - So we use a PRF
  - and the security of this scheme is based on PRF.
  - The scheme achieves poly-time enc. and dec. costs.

→ See our paper for this scheme

Construction (2/4)

For an encryption function Enc_K, we let

$$R := Enc_K(0)$$

$$D[i] := Enc_K(i) – Enc_K(i\text{-}1)$$

Then we can write $Enc_K(m)$ as follows:

$$Enc_K(m) = R + \Sigma_{i=1}^{m} D[i].$$

Therefore, a design of Enc_K can be reduced to the selections of R and D[i].

Construction (3/4)

How to select D[i]:

we set D[i] ← small value with high probability,
but set it to a "large random value" with low probability.

Specifically,

- Let p be a "small" fixed value.
- Take a coin r[i] which becomes 1 with high prob 1-p.
- if (r[i] = 1)
  - D[i] ← small value (say, 1).
- Otherwise,
  - $D[i] \xleftarrow{\$} \{1,...,L\}$,

    where L = large value (say, $2^{poly(SecParam)}$)

We take a value R in a similar manner

Construction (4/4)

Then we set

Key K← (R,D[1],...,D[M]), (Here Mess.Sp={0,...,M})

$$Enc_K(m) \leftarrow R + \Sigma_{i=1}^{m} D[i].$$

But the problems are that,

when the Mess. Sp. size M is super-polynomial of SecParam,

- the above key K is ***not*** polysize
- the above Enc_K is ***not*** polytime

So, finally, we improve the above scheme using "lazy sampling" technique [1].

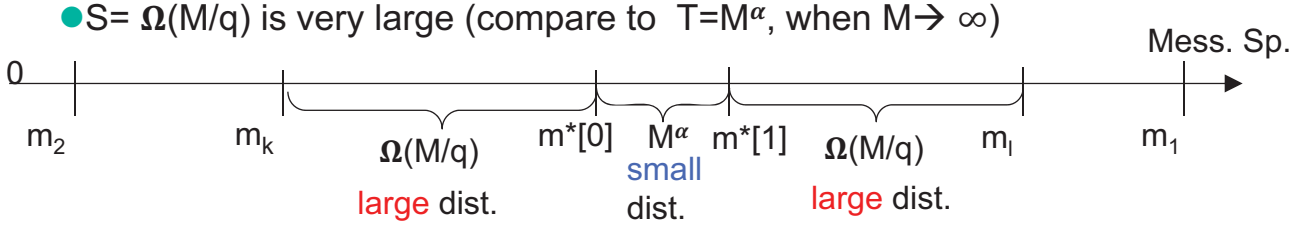
- We omit the explanation of this final part. See our paper.

- Our Definition of Indistiguishability Notion
- Our Results
- Construction of Our scheme
- Security Proof

(X,M$^{\alpha}$,q)-IND of Our Scheme

(Proof)

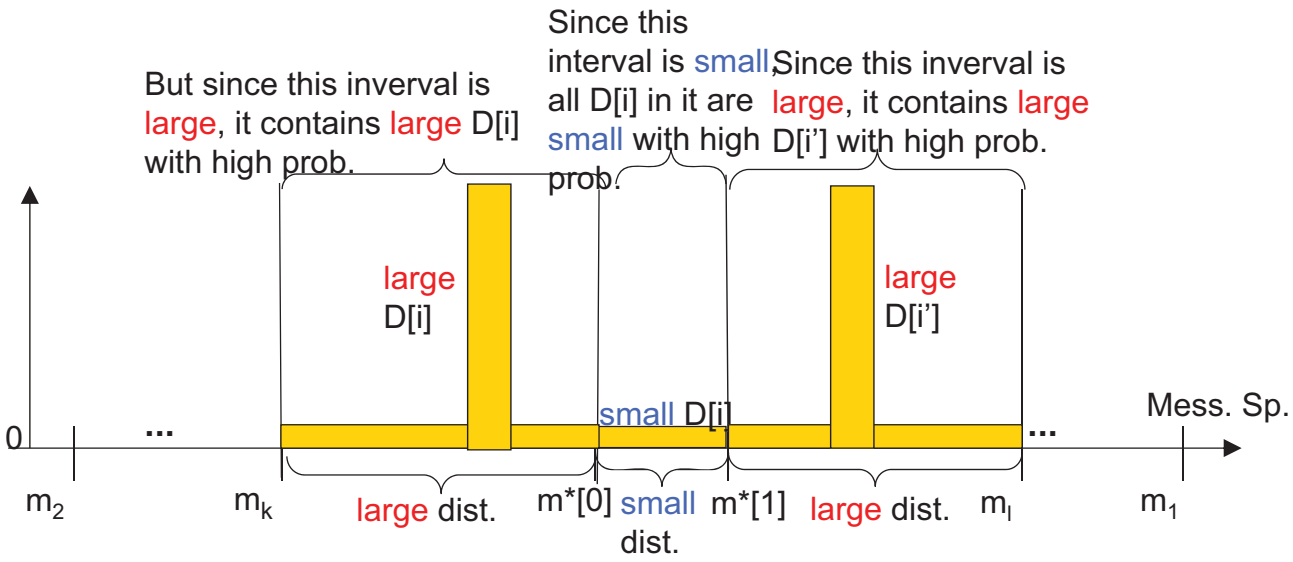
- Consider the Mess. Sp. ={1…M}
  - Due to the def. of (X,M$^{\alpha}$,q)-IND , messages m*[0] and m*[1] of the challenge have to be within the distance T=M$^{\alpha}$.
  - Since α<1, the distance T=M$^{\alpha}$ is small compare to M (when M→ ∞)
  - Recall that we consider the case where components of X has high min-entropy.
  - Hence, the reference messages distributes “almost uniformly at random”.
  - Hence, with high probability, m_1,...m_q are without distance S= Ω(M/q) from m*[0] and m*[1]
  - S= Ω(M/q) is very large (compare to T=M$^{\alpha}$, when M→ ∞)



(X,M$^{\alpha}$,q)-IND of Our Scheme

Recall that we take D[i] as follows:

- with high probability D[i] ← small value.
- with small probability D[i] becomes large random value.

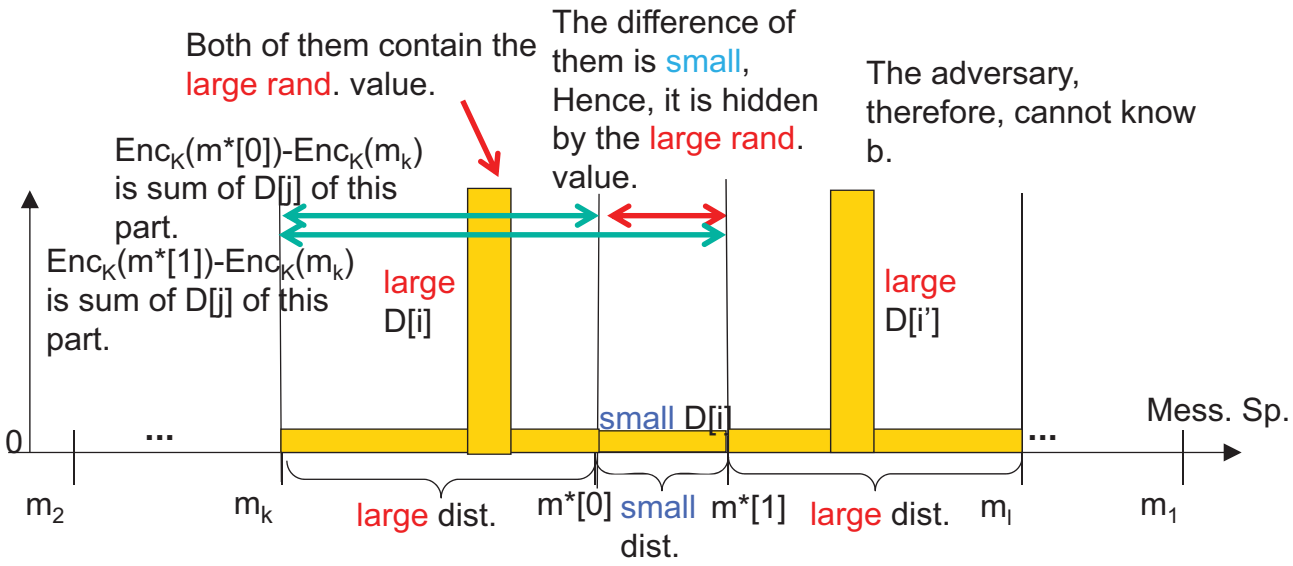


(X,M$^{\alpha}$,q)-IND of Our Scheme

Consider an adversary who want to know b from

$$\mathrm{Enc}_K(m^*[b]) - \mathrm{Enc}_K(m_k) \quad (\text{for } m_k < m^*[0])$$

$$= \Sigma_{j=m_k}^{m^*[b]} D[j] \quad (\text{by definition.})$$

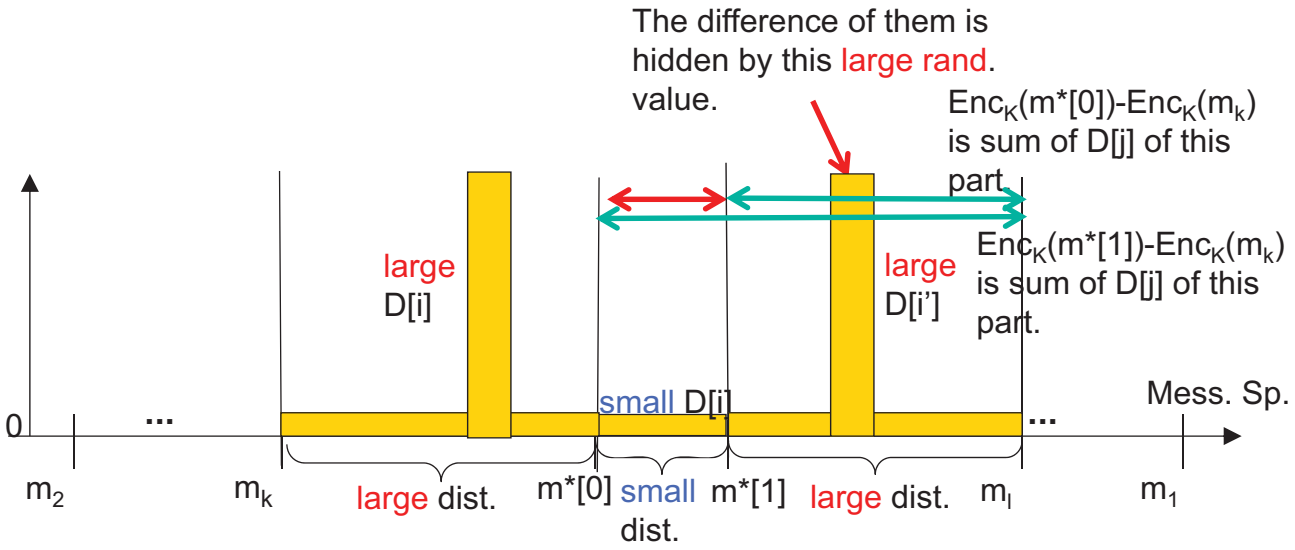


(X,M$^{\alpha}$,q)-IND of Our Scheme

- Similarly, even if an adversary tries to know b from

  $Enc_K(m_l) - Enc_K(m^*[b])$ (for $m_l > m^*[1]$),

  he cannot know it due to a similar reason.



Conclusion

- OPE is very powerful for encrypted database
- but so far, security for it is poorly understood beyond just onewayness the encryption
- We proposed a new indistinguishability notion for OPE.
- This notion can ensure secrecy of lower bits of a plaintext.
- We construct a new OPE scheme which satisfies our new ind. notion.
- In some application hidden lower bits is significant security property like physical measurement, may be trade secret.
- Many question are remaining open.

Thank you

IMI Workshop: Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques

September 1–3, 2015, Kyushu University

Fast and Secure Linear Regression and Biometric Authentication with Security Update*

Le Trieu Phong

phong@nict.go.jp

National Institute of Information and Communications Technology (NICT), Japan

Imaginatively, the storage and computation on the cloud can be seen as storing data and performing computations on a huge and globally available "machine". Formally, a definition of cloud computing is given by NIST in [4] saying that it is a "model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction."

The benefits of cloud computing come with threats, typically one of which is that plain data stored in the cloud may be accessed unwillingly. Promisingly, homomorphic encryption can balance the situation, as it enables the computations over the data even in encrypted form, on which the discussion can go back far to Rivest et al. [8] in 1978. Specifically, a client can store encrypted data on the cloud to enjoy the service, but at the same time can ensure that no useful information is leaked to the storage.

We explicitly explore features of an LWE-based homomorphic encryption scheme, and exploit those features in designing extremely efficient and secure systems in cloud computing. Details are given below.

1. A homomorphic encryption scheme: we make explicit and analyze a homomorphic public key encryption scheme, which is a variant of Regev's scheme [7]. By making the scheme explicit, we newly discover that it has a very *flexible* encoding of plaintexts. By "*flexible*", we mean two things: (a) the scheme natively handles binary strings, real numbers, and the computations (additions, multiplication) over them; and (b) in the scheme, the message length can properly vary with applications. The novel encoding of plaintexts is exploited in depth in following secure systems.

2. Fast and secure linear regression: we show that secure linear regression can be accomplished extremely fast in time and modest in communication. In particular, in the scenario of outsourced computation with an honest-but-curious server and a client, our system processes a simulated dataset of 10^8 records each with 20 features in 10 minutes at the server (Xeon E5-2660 v3, 2.60GHz, 20 threads) and 0.38 second (1 thread) at the client, with only 280 kilobytes of communication from the server to the client. These are extremely fast compared with the best previous work using Paillier encryption equipped with garbled circuits (8.75 hours at the server, [6]).

3. Fast and secure biometric system: we show that binary strings are processed extremely efficiently by our scheme in biometric authentication, in which two binary templates are compared by XORing. The computation can be done in a secure way

*Based on the joint work [3] with Yoshinori Aono, Takuya Hayashi, and Lihua Wang.

where the templates are encrypted, so that it can be performed by an honest-but-curious server. Our ciphertext size is at least a half smaller than previous results by [10], and the computation at the server is dramatically (more than 1000x) improved.

4. Key rotation and security update: this is very unique to our homomorphic encryption scheme, where we show that encrypted data (stored in the server in above systems) can be key-rotated, and generally security-updated in a non-naive way. The task of key rotation is recommended by NIST [5], PCI DSS [1], and OWASP [2]. We design algorithms allowing key rotation and security update over encrypted data without any plain data recovery, and hence help to prevent any data breach (on the cloud servers) specifically in above systems. Relating to but different from our work, [9] also consider key rotation via all-or-nothing transform.

References

[1] `https://www.pcisecuritystandards.org/documents/Prioritized_Approach_V2.0.pdf`.

[2] `https://www.owasp.org/index.php/Cryptographic_Storage_Cheat_Sheet`.

[3] Y. Aono, T. Hayashi, L. T. Phong, and L. Wang. Fast and secure linear regression and biometric authentication with security update. Cryptology ePrint Archive, Report 2015/692, 2015. `http://eprint.iacr.org/`.

[4] P. Mell and T. Grance. The NIST definition of cloud computing. `http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf`.

[5] National Institute of Standards and Technology (NIST). Recommendation for Key Management: Part 1: General (Revision 3). `http://csrc.nist.gov/publications/nistpubs/800-57/sp800-57_part1_rev3_general.pdf`. Accessed: 2014, January 16.

[6] V. Nikolaenko, U. Weinsberg, S. Ioannidis, M. Joye, D. Boneh, and N. Taft. Privacy-preserving ridge regression on hundreds of millions of records. In *2013 IEEE Symposium on Security and Privacy, SP 2013, Berkeley, CA, USA, May 19-22, 2013*, pages 334–348. IEEE Computer Society, 2013.

[7] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In H. N. Gabow and R. Fagin, editors, *STOC*, pages 84–93. ACM, 2005.

[8] R. L. Rivest, L. Adleman, and M. L. Dertouzos. On data banks and privacy homomorphisms. *Foundations of secure computation*, 4(11):169–180, 1978.

[9] D. Watanabe and M. Yoshino. Key update mechanism using all-or-nothing transform for network storage of encrypted data. *IEICE Transactions*, 98-A(1):162–170, 2015.

[10] M. Yasuda, T. Shimoyama, J. Kogure, K. Yokoyama, and T. Koshiba. Practical packing method in somewhat homomorphic encryption. In DPM/SETOP 2013, volume 8247 of *Lecture Notes in Computer Science*, pages 34–50. Springer, 2013.

Fast and Secure Linear Regression and Biometric Authentication with Security Update

Presenter: **Le Trieu PHONG** (NICT, Japan)

Joint work with
***Y. Aono, T. Hayashi, L. Wang** (NICT, Japan)*

Workshop of IMI, Kyushu University:
Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques
1-3 September 2015

Agenda

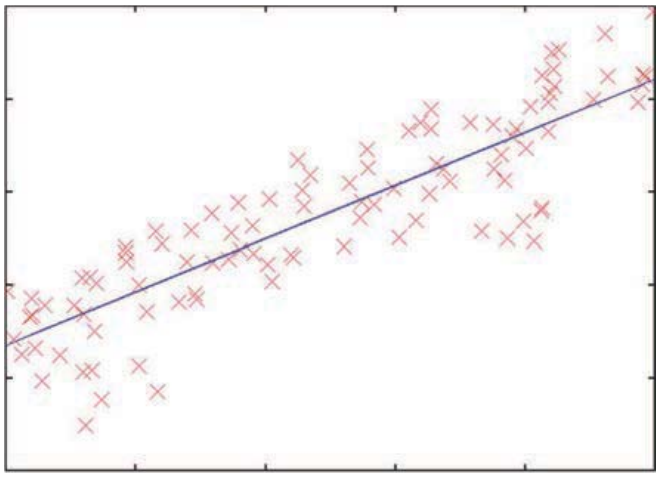
① Secure Linear Regression

② A LWE-based homomorphic encryption with flexible encodings

③ Security Update

④ Secure Biometric Authentication

Secure linear regression



Goal

- Compute the best fit line
- Even if the data items (in red) are encrypted

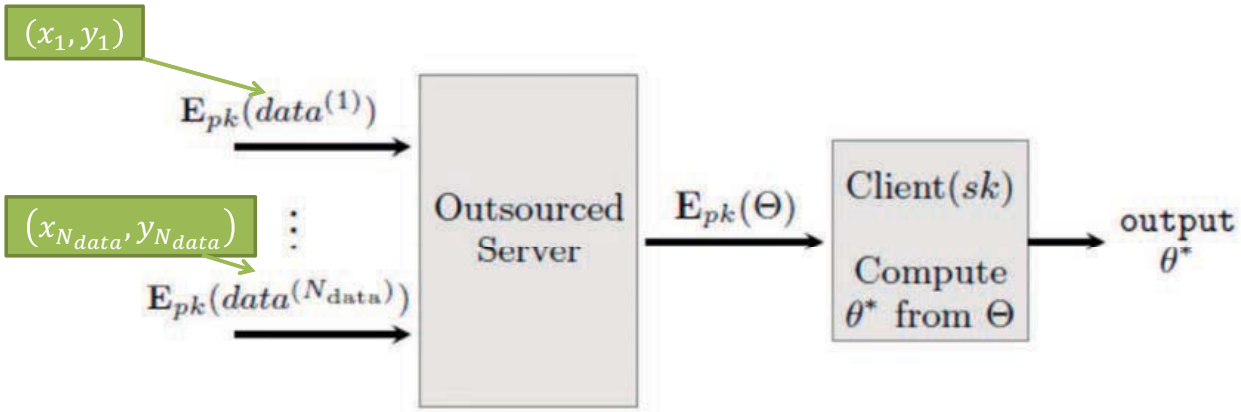
Why secure linear regression?

- The data may be sensitive.
- The data sources may be geographically-distributed, so that cloud computing makes sense.

Warfarin pharmacogenetic dosing algorithm			
		5.6044	
-		0.2614 x	Age in decades
+		0.0087 x	Height in cm
+		0.0128 x	Weight in kg
-		0.8677 x	*VKORC1* A/G
-		1.6974 x	*VKORC1* A/A
-		0.4854 x	*VKORC1* genotype unknown
-		0.5211 x	*CYP2C9* *1/*2
-		0.9357 x	*CYP2C9* *1/*3
-		1.0616 x	*CYP2C9* *2/*2
-		1.9206 x	*CYP2C9* *2/*3
-		2.3312 x	*CYP2C9* *3/*3
-		0.2188 x	*CYP2C9* genotype unknown
-		0.1092 x	Asian race
-		0.2760 x	Black or African American
-		0.1032 x	Missing or Mixed race
+		1.1816 x	Enzyme inducer status
-		0.5503 x	Amiodarone status
=	Square root of weekly warfarin dose**		

The International Warfarin Pharmacogenetics Consortium, ***Estimation of the Warfarin Dose with Clinical and Pharmacogenetic Data***, N Engl J Med 2009; 360:753-764

Outsourced computation model



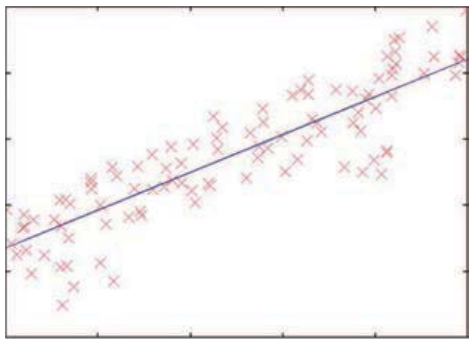
- Numbers of data can be big, e.g., $N_{data} = 126{,}890{,}000$.
- Server is semi-honest (having pk).
- Client is honest (having sk).
- Naively, $\Theta = \theta^*$ but it is not a must.
  - E.g., $\Theta = (\theta_1, \theta_2)$ and $\theta^* = \theta_1/\theta_2$ as division is not supported by homomorphic encryption.

Mathematics of linear regression

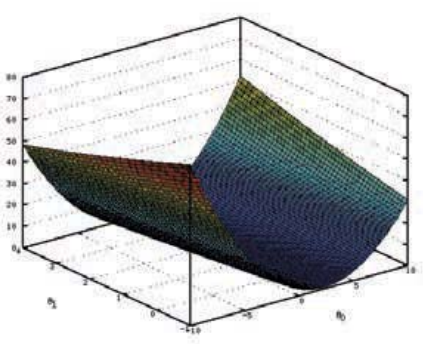
Input: $(x_i, y_i) \in \boldsymbol{R^2}, 1 \le i \le N$

Find: $\boldsymbol{Y = a^* X + b^*}$ minimizing

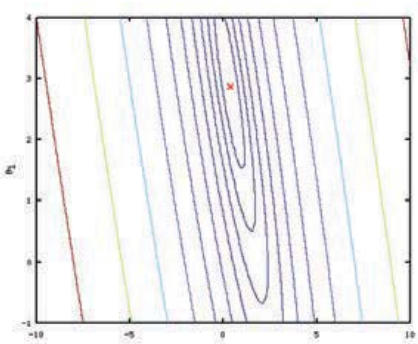
$$F(a,b) = \sum_{i=1}^{N} (y_i - (ax_i + b))^2$$



■ $F(a,b)$ is smooth-convex



$F(a,b)$



The contour of $F(a,b)$

Trick 1: expanding $F(a,b)$

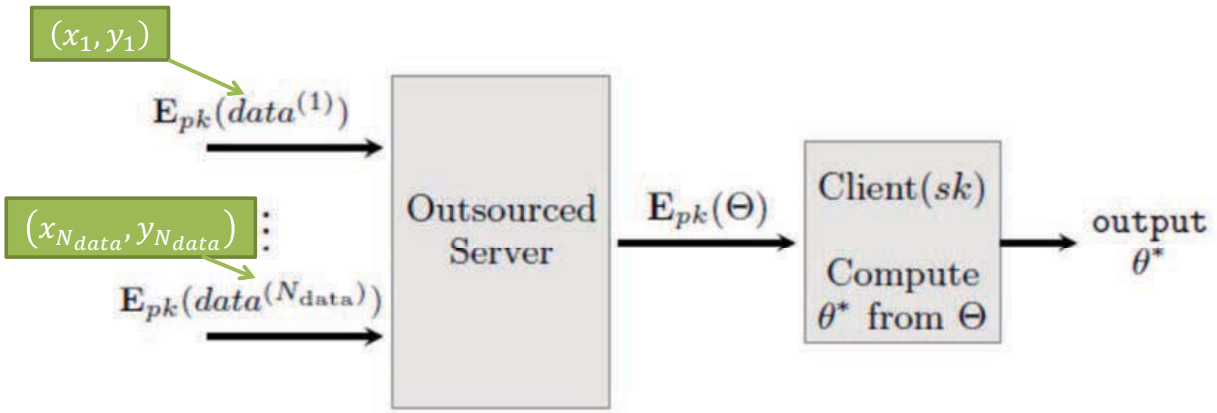
$$F(a,b) = \sum_{i=1}^{N} (y_i - (ax_i + b))^2$$

$$= a^2 \cdot \theta_{a^2} + b^2 \cdot \theta_{b^2} + ab \cdot \theta_{ab} + a \cdot \theta_a + b \cdot \theta_b + \theta_0$$

$$\theta_{a^2} = \sum_{i=1}^{N} x_i^2 \qquad \theta_{b^2} = N \qquad \theta_{ab} = \sum_{i=1}^{N} 2x_i \qquad \theta_a = \sum_{i=1}^{N} (-2x_i y_i) \qquad \theta_b = \sum_{i=1}^{N} (-2y_i) \qquad \theta_0 = \sum_{i=1}^{N} y_i^2$$

These 6 coefficients only depend on the data

Trick 2: exploiting $\Theta \neq \theta^*$

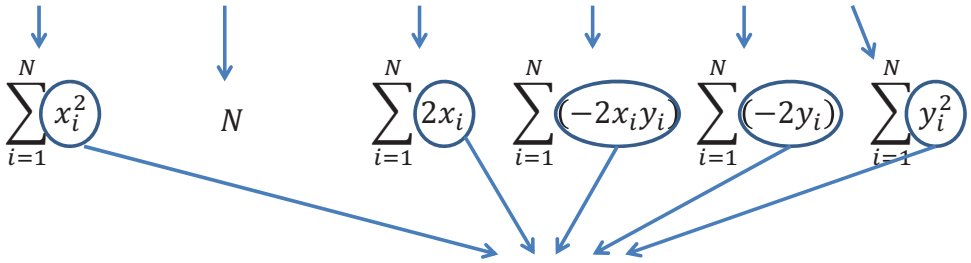


- Naively, $\Theta = \theta^*$ but it is not a must.
- We use, $\Theta = (\theta_{a^2}, \theta_{b^2}, \theta_{ab}, \theta_a, \theta_b, \theta_0)$ and $\theta^* = (a^*, b^*) = \mathbf{argmin}\, F(a,b)$.
- **argmin** is efficient since $F(a,b)$ is smooth-convex.

Trick 3: data pre-processing

$$F(a,b) = \sum_{i=1}^{N} (y_i - (ax_i + b))^2$$

$$= a^2 \cdot \theta_{a^2} + b^2 \cdot \theta_{b^2} + ab \cdot \theta_{ab} + a \cdot \theta_a + b \cdot \theta_b + \theta_0$$

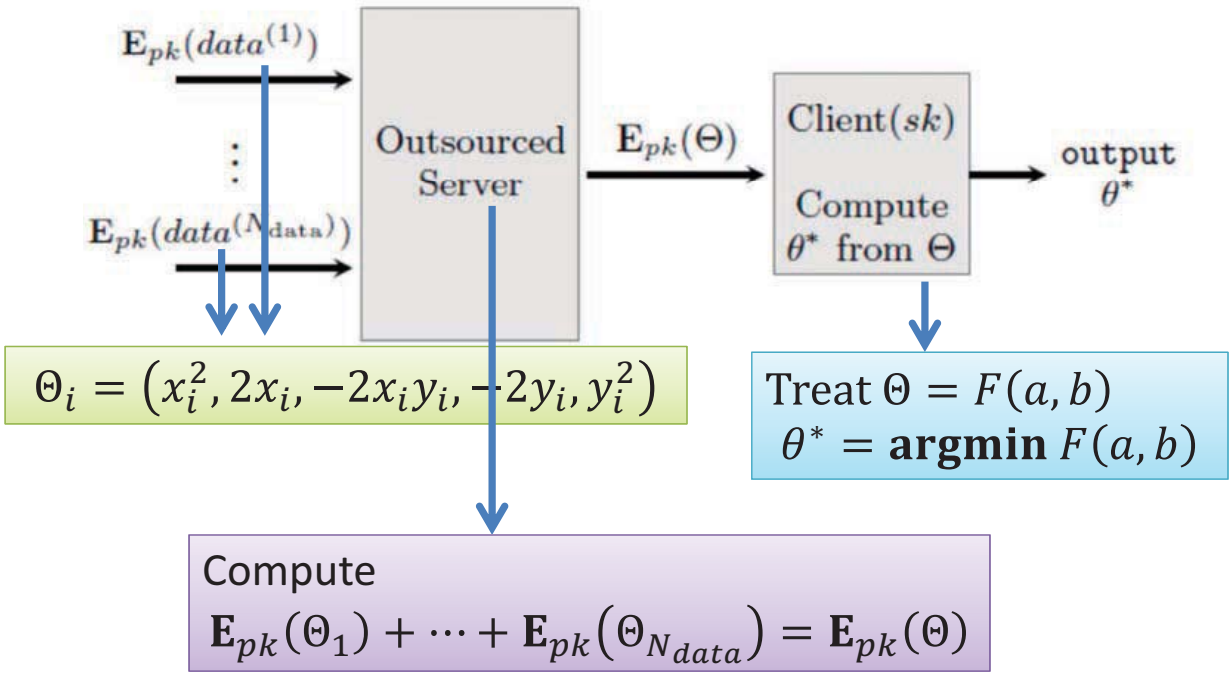


These numbers are from (x_i, y_i)

■ Data source i with (x_i, y_i) does the following:

1. Let $dat_i = (x_i^2, 2x_i, -2x_i y_i, -2y_i, y_i^2)$ of 5 real numbers
2. Encrypt $CT_i = \mathbf{E}_{pk}(dat_i)$ and send to the server.

Our secure linear regression



Costs of parties in our system

Table 4. Costs of parties in dataset size N_{data} and dimension d.

	Storage	Computation
Server	$O(N_{\text{data}})$	$O(N_{\text{data}})$
Client	N/A	$O(d^2)$
Data sources	N/A	$O(d^2)$

	Communication
Each data source → server	$O(d^2)$
Server → client	$O(d^2)$

So what homomorphic encryption is needed?

- Only additive homomorphism (so far).
- Dealing with real numbers.
- Paillier, Ring-LWE-based schemes can be used.

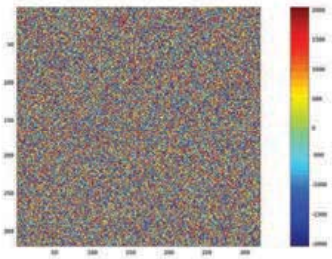
- We will present a LWE-based scheme.
- Having smaller ciphertext sizes (than known ring-LWE schemes).
- Having properties unknown to Paillier (key rotation, security update, post-quantum).

Encryption in our LWE-based scheme

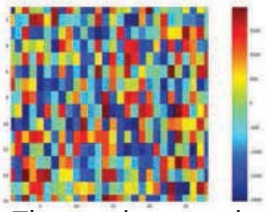
- $pk = (A, P)$ of two matrices
- $m \in Z_p^{1\times l}$: plaintext
  (p can be $2^{30}+1$, $l = 64$)
- Modulus q, big, can be 2^{114}
- **Encryption:**

$$CT = \underbrace{e_1[A|P] + p[e_2|e_3]}_{\text{computationally random mask}} + [0|m] \in Z_q^{n+l}$$

- Regev-based, secure under LWE.



Matrix A



The random mask

13

Flexible message space

- $p = 2, l = 1$: essential Regev's encryption.
- **Our observation**:
  ℓ can vary 'almost' independent of (n, q, s, p)

Theorem 1. *The scheme in Figure 1 is CPA-secure under the LWE assumption. Specifically, for any poly-time adversary $\mathcal{A}$, there is an algorithm $\mathcal{D}$ of essentially the same running time such that*

$$\mathbf{Adv}^{\mathrm{cpa}}_{\mathcal{A}}(\lambda) \le (l+1) \cdot \mathbf{Adv}^{\mathrm{LWE}(n,s,q)}_{\mathcal{D}}(\lambda).$$

- $l \approx 16128$: secure linear regression.
- $l = 2048$: secure biometric authentication.

14

Homomorphism in our scheme

- $Enc_{pk}(m) = e_1[A|P] + p[e_2|e_3] + [0|m]$
- $Enc_{pk}(m') = e_1'[A|P] + p[e_2'|e_3'] + [0|m']$
- **Addition** is vector addition over Z_q^{n+l}
- **Multiplication** is the tensor product

$$CT \otimes CT' \in Z_q^{(n+l)\times(n+l)}$$

- **Combination** (mult-then-add)

$$\sum_{i=1}^{N_{add}} CT_i \otimes CT_i' \in Z_q^{(n+l)\times(n+l)}$$

15

Linking *ciphertext addition* to *real number addition*

- Real numbers

$$a = a_{-L}2^{-L} + \cdots + a_0 + a_1 2^1 + \cdots + a_\ell 2^\ell \in \mathbb{R}$$
$$b = b_{-L}2^{-L} + \cdots + b_0 + b_1 2^1 + \cdots + b_\ell 2^\ell \in \mathbb{R}$$

- Encryption of a, b:

$$CT_a = \mathbf{E}_{pk}([a_{-L}, \dots, a_0, \dots, a_\ell])$$
$$CT_b = \mathbf{E}_{pk}([\underbrace{b_{-L}, \dots, b_0, \dots, b_\ell}_{\in Z_p^{L+\ell+1}}])$$

$$CT_{a+b} = \mathbf{E}_{pk}([a_{-L} + b_{-L}, \dots, a_0 + b_0, \dots, a_\ell + b_\ell])$$

16

Adding N_{data} ciphertexts

- To prevent rounding

$$N_{data} < \frac{p}{2}$$

- In secure linear regression, we can take

$$p = 2^{30} + 1 = 1{,}073{,}741{,}825$$

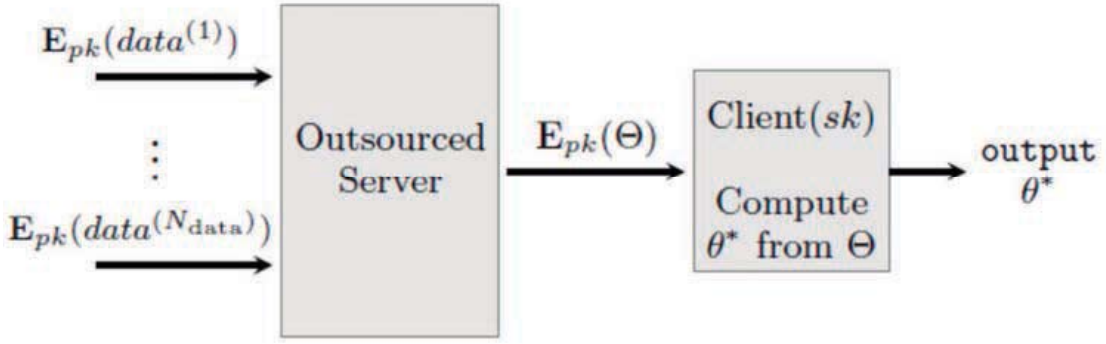
- To tolerate, for example,

$$N_{data} = 126{,}890{,}000$$

Secure training and prediction

■**Secure training**: from given dataset, producing θ^*

As seen, we use only additive homomorphism of Enc



■**Secure prediction**: Having θ^*, use it with new coming data

We will use multiplicative homomorphism of Enc

Example of secure prediction

- $\theta^* = (\theta_0^*, \ldots, \theta_d^*)$ is public.
- Patient data is secret.
- All patient's data can be stored encryptedly.
- Yet the result can be obtained as follows:

$$\theta_0^* + \theta_1^* \cdot \mathbf{E}_{pk}(age) + \theta_2^* \cdot \mathbf{E}_{pk}(height) + \theta_3^* \cdot \mathbf{E}_{pk}(weight) + \theta_4^* \cdot \mathbf{E}_{pk}\ (genome) + \cdots$$

θ^* ↓ Patient's data ↓

Warfarin pharmacogenetic dosing algorithm			
		5.6044	
-		0.2614 x	Age in decades
+		0.0087 x	Height in cm
+		0.0128 x	Weight in kg
-		0.8677 x	*VKORC1* A/G
-		1.6974 x	*VKORC1* A/A
-		0.4854 x	*VKORC1* genotype unknown
-		0.5211 x	*CYP2C9* *1/*2
-		0.9357 x	*CYP2C9* *1/*3
-		1.0616 x	*CYP2C9* *2/*2
-		1.9206 x	*CYP2C9* *2/*3
-		2.3312 x	*CYP2C9* *3/*3
-		0.2188 x	*CYP2C9* genotype unknown
-		0.1092 x	Asian race
-		0.2760 x	Black or African American
-		0.1032 x	Missing or Mixed race
+		1.1816 x	Enzyme inducer status
-		0.5503 x	Amiodarone status
=	Square root of weekly warfarin dose**		

From *tensor product* to *real number multiplication*

- Real numbers

$$a = a_{-L}2^{-L} + \cdots + a_0 + a_1 2^1 + \cdots + a_\ell 2^\ell \in \mathbb{R}$$
$$b = b_{-L}2^{-L} + \cdots + b_0 + b_1 2^1 + \cdots + b_\ell 2^\ell \in \mathbb{R}$$

- Important equation

$$ab = \left([2^{-L} \cdots 2^\ell]\begin{bmatrix} a_{-L} \\ \vdots \\ a_\ell \end{bmatrix}\right)\left([b_{-L} \cdots b_\ell]\begin{bmatrix} 2^{-L} \\ \vdots \\ 2^\ell \end{bmatrix}\right) = [2^{-L} \cdots 2^\ell]\left(\begin{bmatrix} a_{-L} \\ \vdots \\ a_\ell \end{bmatrix}[b_{-L} \cdots b_\ell]\right)\begin{bmatrix} 2^{-L} \\ \vdots \\ 2^\ell \end{bmatrix}$$

Real number multiplication

Tensor product of binary vectors

Secure prediction in our system

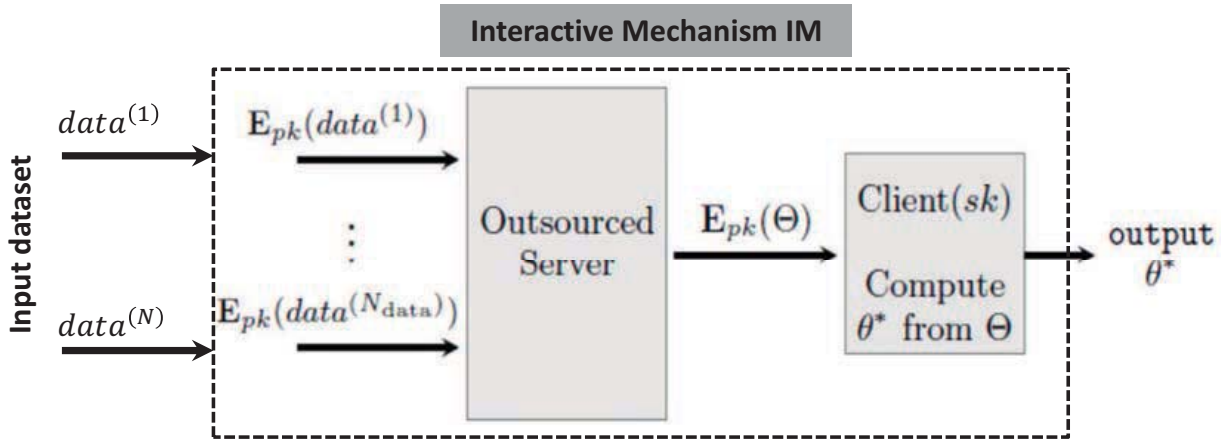
Server computes the following:

$$\mathbf{E}_{pk}(\theta_0^*) \otimes \mathbf{E}_{pk}(1) + \mathbf{E}_{pk}(\theta_1^*) \otimes \mathbf{E}_{pk}(x_1) + \cdots + \mathbf{E}_{pk}(\theta_d^*) \otimes \mathbf{E}_{pk}(x_d)$$

Decryption by patient's sk will yield:

$$\theta_0^* + \theta_1^* x_1 + \cdots + \theta_d^* x_d$$

Differential privacy of interactive mechanism

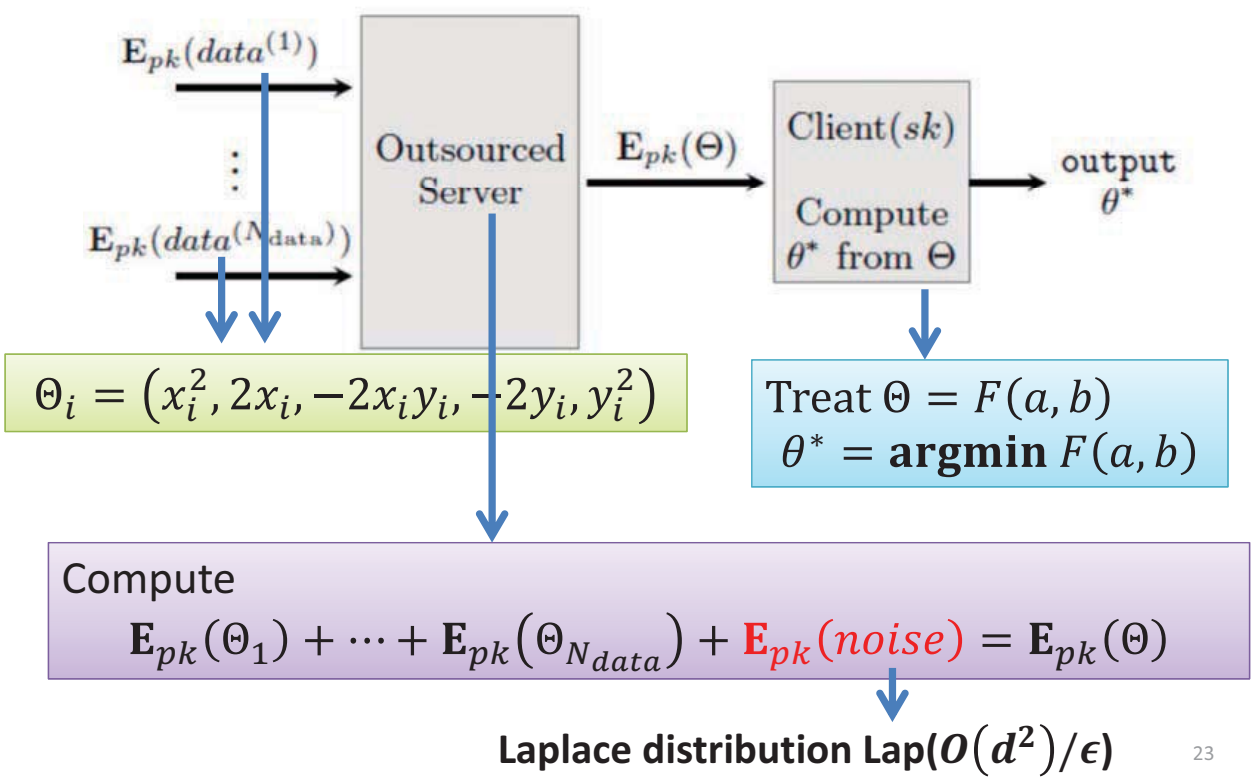


IM satisfies ϵ-differential privacy iff

$$\Pr[\mathbf{IM}(I) \to \theta^*] \leq \exp(\epsilon) \cdot \Pr[\mathbf{IM}(I') \to \theta^*]$$

where input datasets I and I' differ at only one item.

Adding differential privacy to our system



Achieving differential privacy

Theorem

If $noise$ is of $\mathrm{Lap}(O(d^2)/\epsilon)$ distributions, our system viewed as an interactive mechanism satisfies ϵ-differential privacy.

※d is the data dimension.

※Proof uses the following paper:

Jun Zhang, Zhenjie Zhang, Xiaokui Xiao, Yin Yang, Marianne Winslett:
Functional Mechanism: Regression Analysis under Differential Privacy. PVLDB 5(11): 1364-1375 (2012)

Agenda

① Secure Linear Regression

② A LWE-based homomorphic encryption with flexible encodings.

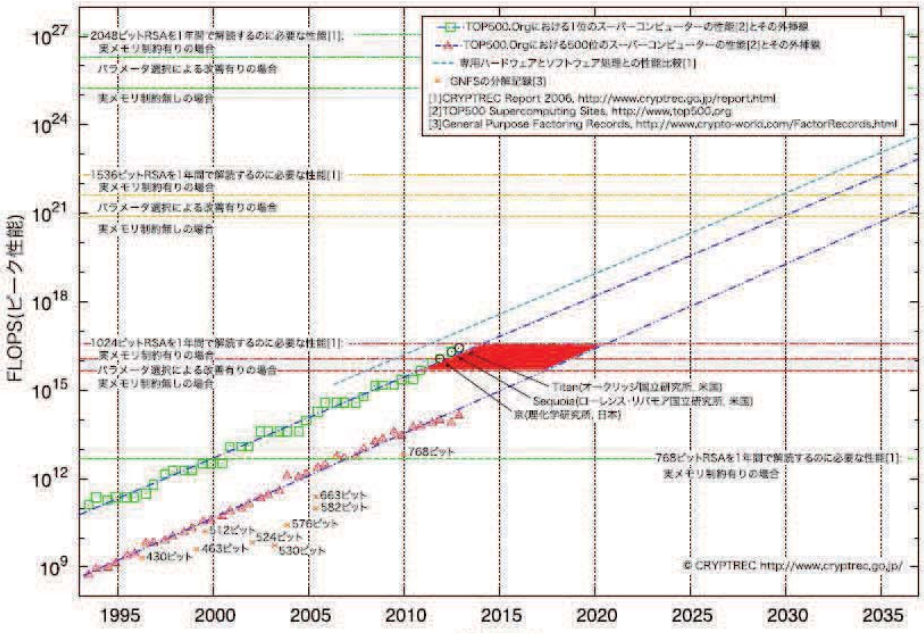
③ Security Update

④ Secure Biometric Authentication

25

Encryption is weakened by years

http://www.cryptrec.go.jp/report/c12_sch_eb.pdf



26

Encryption is weakened by years

- Use a very long encryption key at the beginning?

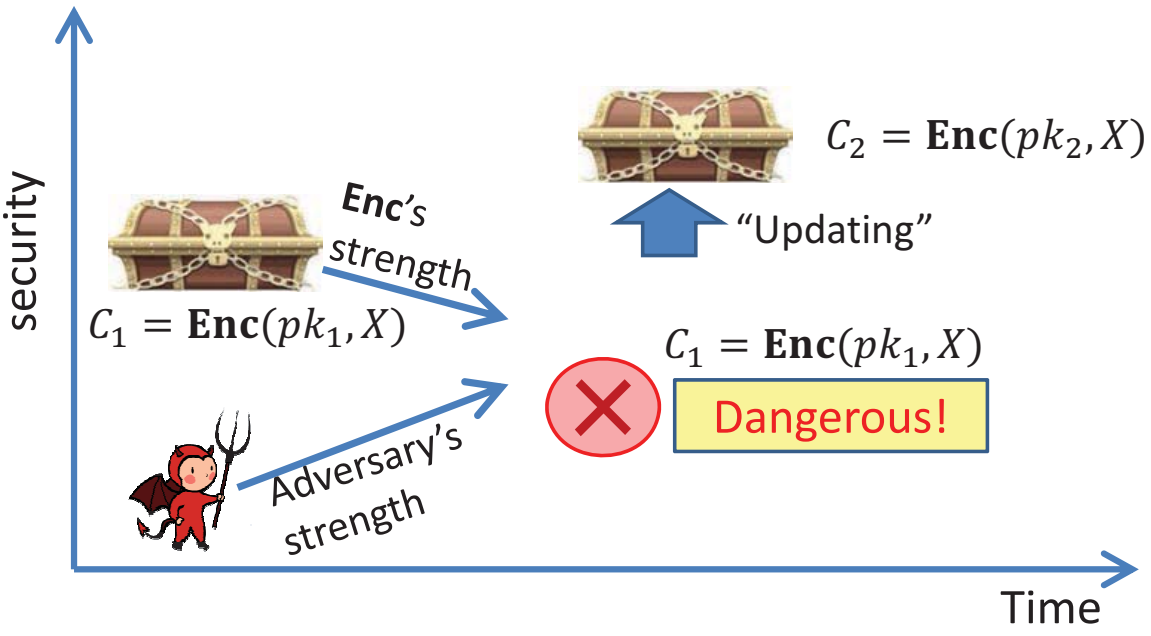
RSA's paper(1978)

We recommend that n be about 200 digits long. Longer or shorter lengths can be used depending on the relative importance of encryption speed and security in the application at hand. An 80-digit n provides moderate security against an attack using [illegible] nology; using 200 digits provides a margin of safety against future developments.

- Suggested key length of 200 digits is broken in 2005
- < 30 years

27

"Updating" encryption



28

Security update (bird's-eye view)

Core techniques:

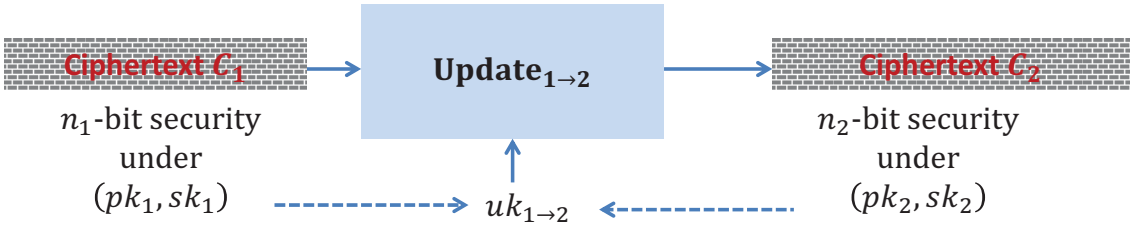
- **Dimension switching** of vectors

$$CT_1 \in Z_q^{400} \longrightarrow CT_2 \in Z_q^{500}$$

Multiply by a dedicated matrix of size 400×500

- **Re-randomization** by adding $\mathbf{Enc}(0)$

Security update (more details)



- **[Dimension switching]**

$$C_2' = \mathbf{DimSwitch}(uk_{1\to2}, C_1)$$

so that sk_2 can decrypt

- **[Re-randomization]**

$$C_2 = C_2' + \mathbf{Enc}(pk_2, 0)$$

so that C_2 has n_2 bit security and homomorphisms are unaffected

Lattice problems in dimension $\sqrt{n}$ (STOC 2013)

- We also use the fact that LWE(q, n, s) gets harder with larger dimension n (number of variables) for fixed q, s

Uses of dimension switching

Table 1. Usages of the dimension switching technique.

Dimension switching	Exploited in	Main purpose
(high → low) $n_2 < n_1$	BGV12, BV11	efficiency improvement in FHE
(equal) $n_2 = n_1$	ABPW13, CCL+14	PRE, obfuscation
(equal, or low → high) $n_1 \leq n_2$	Here	key rotation and security update

[BGV12] Zvika Brakerski, Craig Gentry, Vinod Vaikuntanathan:
(Leveled) fully homomorphic encryption without bootstrapping. ITCS 2012: 309-325

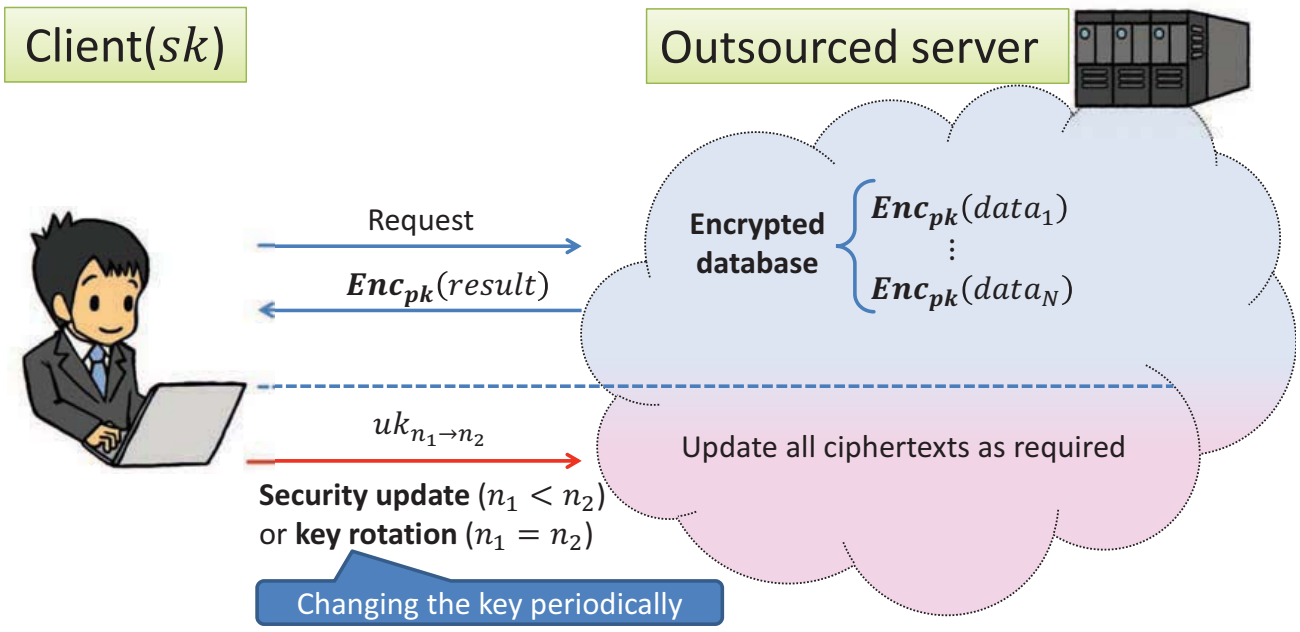
[BV11] Zvika Brakerski, Vinod Vaikuntanathan:
Efficient Fully Homomorphic Encryption from (Standard) LWE. FOCS 2011: 97-106

[ABPW13] Yoshinori Aono, Xavier Boyen, Le Trieu Phong, Lihua Wang:
Key-Private Proxy Re-encryption under LWE. INDOCRYPT 2013: 1-18

[CCL+13] Nishanth Chandran, Melissa Chase, Feng-Hao Liu, Ryo Nishimaki, Keita Xagawa:
Re-encryption, Functional Re-encryption, and Multi-hop Re-encryption: A Framework for Achieving Obfuscation-Based Security and Instantiations from Lattices. Public Key Cryptography 2014: 95-112

Vision of application

- Outsourced computation with security update



Agenda

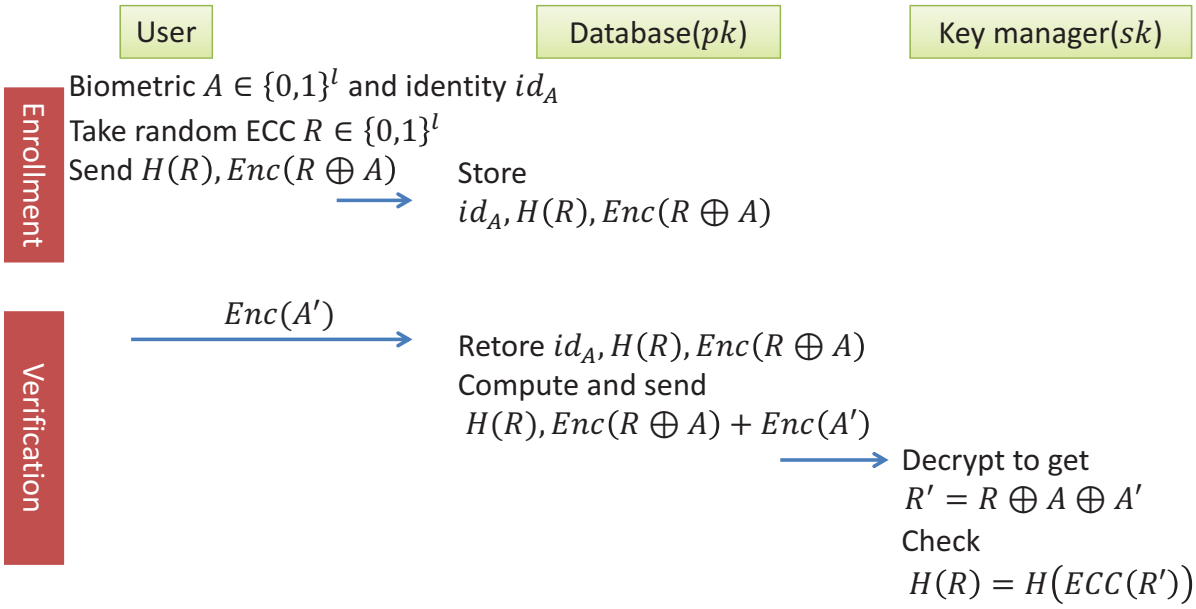
① Secure Linear Regression

② A LWE-based homomorphic encryption with flexible encodings.

③ Security Update

④ Secure Biometric Authentication

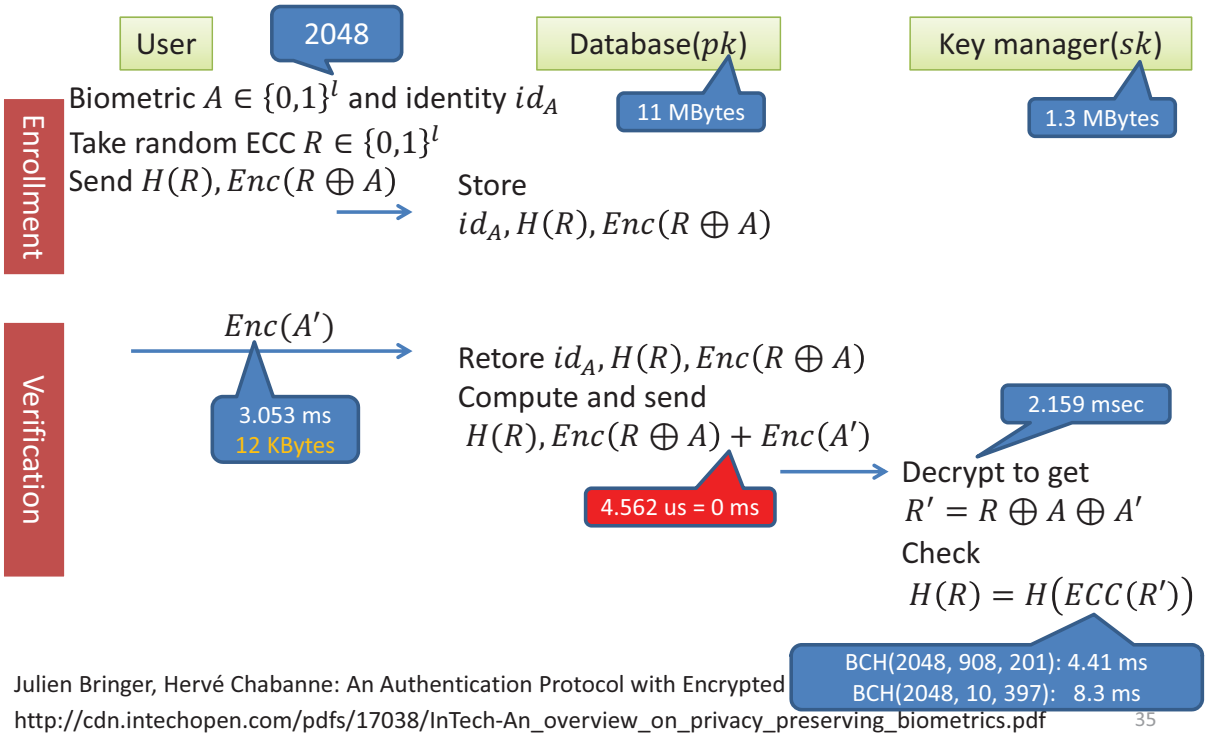
Secure Biometric Authentication
(use any additively homomorphic Enc)



Julien Bringer, Hervé Chabanne: An Authentication Protocol with Encrypted Biometric Data. AFRICACRYPT 2008
http://cdn.intechopen.com/pdfs/17038/InTech-An_overview_on_privacy_preserving_biometrics.pdf

Using our scheme with $l = 2048$

(128-bit security)



Julien Bringer, Hervé Chabanne: An Authentication Protocol with Encrypted
http://cdn.intechopen.com/pdfs/17038/InTech-An_overview_on_privacy_preserving_biometrics.pdf

Conclusion

① Secure Linear Regression

② A LWE-based homomorphic encryption with flexible encodings.

- Only additive homomorphism for training
- Multiplicative homomorphism for predicting
- Processing real numbers of fixed precision in ①

③ Security Update

- New use of a known technique (dimension switching)

④ Secure Biometric Authentication

- Our scheme can also process bit strings

■Paper is at https://eprint.iacr.org/2015/692.pdf containing experiments, discussions, extended models for several clients..

Homomorphic Encryption – are we there yet?

Anirban Basu

KDDI R&D Laboratories

basu@kddilabs.jp

With the proliferation of services offered on the Internet, large volumes of personal data have been collected and stored in cloud computing environments by different organisations to better understand and tailor goods and services for potential individuals. The recent advances in the Internet of Things usher yet another era of rapid growth of data, a lot of which is of a very personal nature, e.g., detailed physiological characteristics collected by wearable sensors or the trajectory data of the movement of individuals. While developing sophisticated intelligence to make sense of this data is a high priority, the privacy of individuals and even organisations is at stake. Even more so, when such rich datasets need to be shared across organisational boundaries.

Privacy-preserving analysis of data has become an increasingly important field of research in recent times. Various types of anonymisation and random perturbation of the data can help with guaranteeing certain privacy levels but these come at a cost of data utility. The other alternative is homomorphic encryption, which enables computing over encrypted data. In this talk, we explore how far is homomorphic encryption from being a reality.

To do so, we present two of our recent works that utilise homomorphic encryption and have been tested on real world cloud computing platforms. In [1], we discuss how an additively homomorphic encryption can be used to query a public cloud based classifier for collaborative filtering. We show evaluations of this scheme using datasets with prototypes built atop real world cloud computing platforms. In [2], we demostrate how a lightweight and practical (sender-)anonymous message routing network can be built and deployed on the cloud utilising additive homomorphic encryption.

We conclude with the outlook that a mixture of partially homomorphic schemes along with other techniques are practical in some real world application scenarios.

References

[1] A. Basu, J. Vaidya, H. Kikuchi, and T. Dimitrakos. Privacy-preserving collaborative filtering on the cloud – practical implementation experiences, In proc: IEEE Cloud, Santa Clara, CA, USA. 2013.

[2] A. Basu, J. C. Corena, J. Vaidya, J. Crowcroft, S. Kiyomoto, S. Marsh, Y. S. Van Der Sype, T. Nakamura, Lightweight practical private one-way anonymous messaging, In proc: IFIP WG 11.11 International Conference on Trust Management (IFIPTM), Hamburg, Germany, 2015.

Homomorphic Encryption – are we there yet?

Anirban Basu

KDDI R&D Laboratories, Japan

02 September 2015
Kyushu University

KDDI R&D LABS

At a glance

1. Who am I?
2. Why are we listening to this?
   - The privacy problem and homomorphic encryption
3. Collaborative filtering and privacy
   - SlopeOne predictors for CF
   - Privacy-preserving CF
   - Deployment on SaaS engines (PaaS clouds)
   - PPCF experimental results and inferences
4. Lightweight and practical anonymous message routing
   - Why?
   - The state of the art
   - Sender anonymity through message unlinkability
   - Anonymous messaging experimental validation
5. Epilogue
   - What can we conclude?

KDDI R&D LABS

Profile[1]

- Researcher within the Information Security Group at KDDI R&D Laboratories.
- Post-doc (Tokai University, Japan), 2013: privacy preserving collaborative filtering.
- PhD (University of Sussex, UK), 2010: a reputation framework for computer networks.
- BEng (University of Sussex, UK), 2004: augmented reality visualisation.
- Hobbies: programming, photography, travelling, cycling.
- Home town: Chandannagar, West Bengal, India.

KDDI R&D LABS

[1] See: `http://www.linkedin.com/in/anirbanbasu`.

Bigger the data, worse the privacy

- Big data is getting bigger – the Internet of Things!
- Good news: more data to analyse and build intelligence.
- Bad news: privacy of data is a growing concern.
- One cloud adoption barrier: privacy of data, both individual and organisational.

KDDI R&D LABS

Privacy – what to do?

- Data release through anonymisation, perturbation: privacy and utility do not agree.
- Why do we not encrypt all the data?
  - But computing something meaningful (i.e., data mining) over that encrypted data?
- Homomorphic encryption: compute blindly over encrypted data.
- The elephant in the room: is homomorphic encryption magical or mythical? How far is it from reality?

Homomorphic encryption – brief background

- Generally speaking: $f(m_1, m_2) \equiv g(c_1, c_2)$:
  - function f on plaintext messages m_i is equivalent to a function g over ciphertexts of these messages c_i.
- Different classes of homomorphic encryption:
  - additive,
  - multiplicative,
  - somewhat homomorphic, and
  - fully-homomorphic.

The magic of additive homomorphic encryption

- $\mathcal{E}(m_1 + m_2) = \mathcal{E}(m_1) \cdot \mathcal{E}(m_2)$
  - The encryption of the sum of two plaintexts is the modular multiplication of their individual ciphertexts.
- $\mathcal{E}(m_1 \cdot m_2) = \mathcal{E}(m_1)^{m_2}$
  - The encryption of the multiplication of two plaintexts is the modular exponentiation of the ciphertext of one by the other plaintext.
- Examples of such cryptosystems: Paillier, Elliptic Curve ElGamal, Damgård-Jurik.

KDDI R&D LABS

Is this practical?

- Fully-homomorphic encryption is still somewhat far from realistic applicability.
- Practice: partially homomorphic encryption and a mixture of various other encryption techniques for specific application scenarios.
- Two application scenarios: (privacy preserving) collaborative filtering (2013) and anonymous message routing (2015).

KDDI R&D LABS

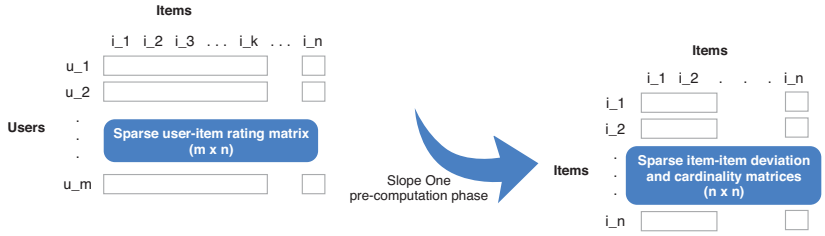
Recommendation and collaborative filtering

- Collaborative filtering (CF) employs opinions of the community.
- Problem is with the privacy of rating data.

KDDI R&D LABS

SlopeOne CF

- Precomputation: Δ: item-item deviation matrix; ϕ: item-item cardinality matrix.



- Standard SlopeOne based prediction: $r_{u,x}$, the rating from user u on item x

$$r_{u,x} = \frac{\sum_{a|a\neq x}(\delta_{x,a} + r_{u,a})\phi_{x,a}}{\sum_{a|a\neq x}\phi_{x,a}} = \frac{\sum_{a|a\neq x}(\Delta_{x,a} + r_{u,a}\phi_{x,a})}{\sum_{a|a\neq x}\phi_{x,a}}$$

KDDI R&D LABS

Privacy-preserving collaborative filtering (PPCF)

- The unencrypted SlopeOne based prediction:

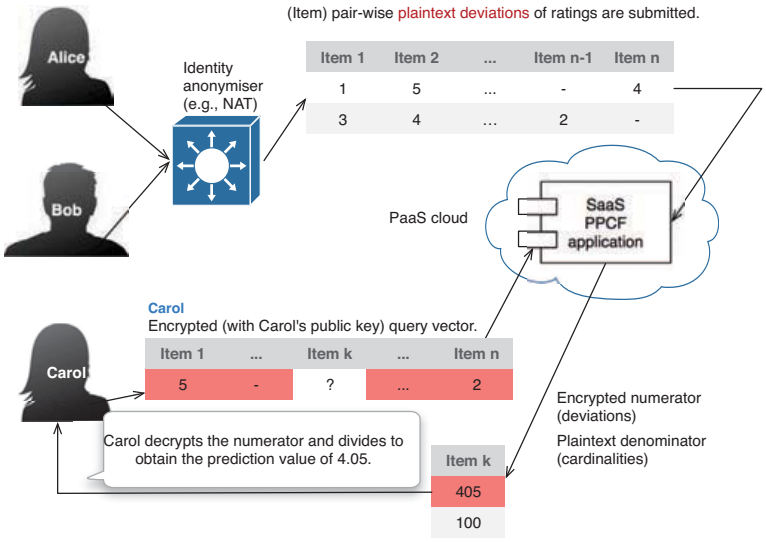
$$r_{u,x} = \frac{\sum_{a|a\neq x}(\Delta_{x,a} + r_{u,a}\phi_{x,a})}{\sum_{a|a\neq x}\phi_{x,a}}$$

- Over an additively homomorphic encrypted domain:

$$r_{u,x} = \frac{\mathcal{D}(\mathcal{E}(\sum_{a|a\neq x}\Delta_{x,a})\prod_{a|a\neq x}(\mathcal{E}(r_{u,a})^{\phi_{x,a}}))}{\sum_{a|a\neq x}\phi_{x,a}}$$

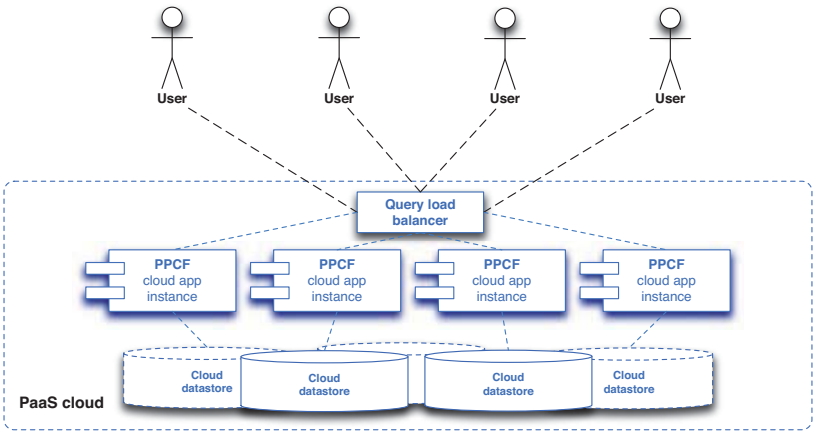
KDDI R&D LABS

PPCF on the cloud



KDDI R&D LABS

PPCF deployment scenario as a SaaS



KDDI R&D LABS

App Engine (GAE/J) versus Elastic Beanstalk (EBS)

	GAE/J	EBS
Java software stack	Limited	Full
Scalability	Very high	High but pricey
Unit performance	Average	Configurable
Data storage	Distributed BigTable, SQL	Distributed SimpleDB, SQL
Vendor lock-in	Yes, partially	No
Free quota	Daily	One-off, first year
Frontend access	HTTP, SPDY (SSL)	HTTP, HTTPS

KDDI R&D LABS

Performance test

- Speed of query processing.
- Varying length of the query vector.
- Varying concurrent user requests.
- Single-threaded or multi-threaded query vector processing.

Google App Engine setting

- Instance class: F4 (2400MHz, 512MB RAM).
- Maximum idle instances: automatic.
- Maximum pending latency: 10ms.
- Datastore: master-slave, not high-replication.

Amazon Elastic Beanstalk setting

- Instance class: t1.micro EC2 instances (min: 1, max: 8).
- Load balancer increase (by one instance) trigger: over 70% CPU utilization in 1 minute.
- Load balancer decrease (by one instance) trigger: below 40% CPU utilization in 1 minute.
- Datastore: MySQL RDBMS on t1.micro EC2 instance.

KDDI R&D LABS

The datasets used

	Jester	MovieLens 100K
Users	73,421	943
Items	100	1,682
Range	$[-10.00 \quad 10.00]$	$\{1, 2, 3, 4, 5\}$
Ratings	4,100,000	100,000
Rating density	55.8%	6.3%
Min. rating	-9.95	1
Max. rating	10.0	5
Rating mean	0.744	3.539
Data points[2]	4,950	983,206
Density	100%	69.5%

KDDI R&D LABS

[2] "Data points" and "Density" refer to Slope One data points and their density.

Graph legend

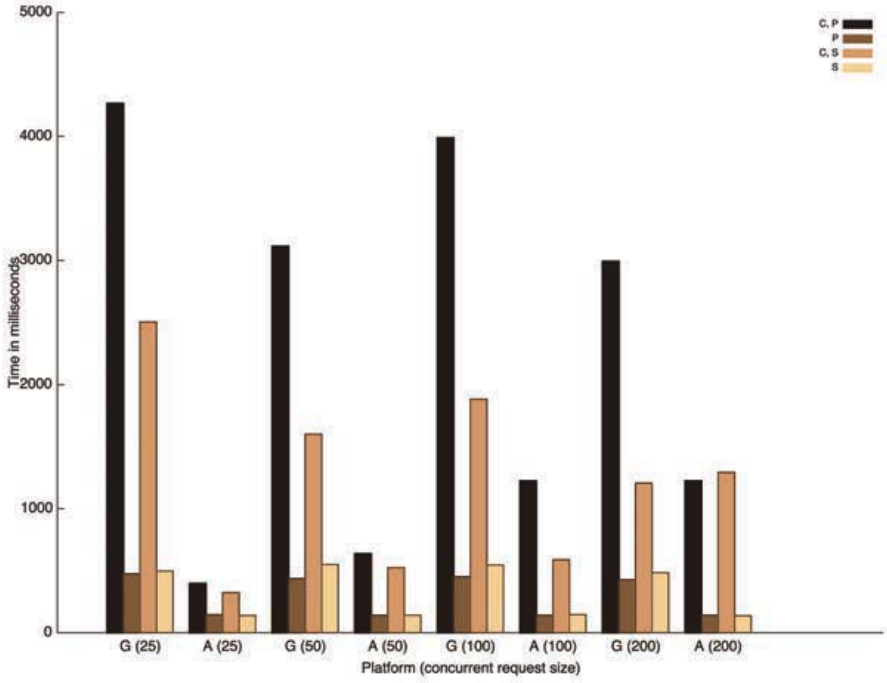
Heads up for the experiment categories

S: single-threaded, single query, P: multi-threaded, single query; C, S: single-threaded, concurrent query and C, P: multi-threaded, concurrent query.

KDDI R&D LABS

Jester dataset (1024-bits)

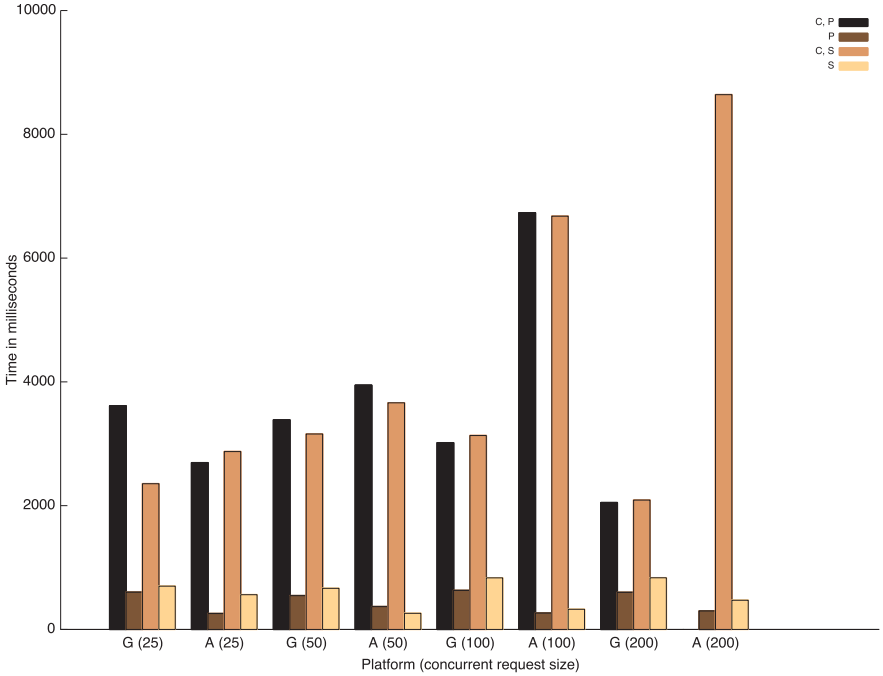
- Example query vector size: 20.



KDDI R&D LABS

MovieLens dataset (1024-bits)

- Example query vector size: 50. EBS partial failure: datastore scalability.



KDDI R&D LABS

General observations

	GAE/J[3]	EBS[4]
Response to short bursty load	Very fast	Slow
Response to steady load	Steady	Steady
Parallel query vector processing	Good	Not necessarily good
Configurability	Limited	High
Ease of deployment	High	Moderately difficult
Running cost	Low, can be capped	High

- Google App Engine is better suited for the type of application and deployment setup we had.
- GAE/J is better with applications that receive high user requests but take relatively short time to process each request.

[3] Google App Engine.
[4] Amazon Elastic Beanstalk.

KDDI R&D LABS

Query processing time estimation

- Estimated query size for a 30s turn-around time, single-threaded processing and with only one query at a time.
- Google App Engine is generally slower but performs better with concurrent loads.

	GAE/J	EBS	Theoretical sizes[5]
Encrypted query vector size	1376 items	3274 items	–
HTTP POST size (numeric IDs[6])	698KB	1.624MB	520n bytes
HTTP POST size (string IDs)	731KB	1.698MB	544n bytes

[5] For Paillier 2048-bits, n query items; ignoring other POST overheads.
[6] Excludes overhead of JSON packaging.

Intermission

Anonymous communication

- Encourages free speech: no fear of reprisal.
- End-to-end encrypted messaging is *not* anonymous communication in the context of this talk.

Motivating use cases

- Anonymous opinions.
  - Present: insufficient 'anonymity' guarantees in existing survey systems, e.g., Survey Monkey.
  - Future: anonymise survey participants.
- Anonymous micro-blogging.
  - Present: micro-blogging platforms, e.g., Twitter, identify bloggers, or re-posters.
  - Future: anonymise micro-bloggers.
- Limit participation to specific groups with private information retrieval, blind signatures.
- Another use case: anonymously posting data to a public cloud-based classifier.

But there are tools that already do this

- Tor – the well-known anonymous network.
- Generalising: high-latency systems (mix networks) and low-latency systems (e.g., onion routing).
- Specialised configurations/permissions, e.g., opening ports through the firewall.
- Pre-existing paths in Tor, potentially breakable[7].
- Recall the adjectives for the title of this work: *lightweight* and *practical*?

KDDI R&D LABS

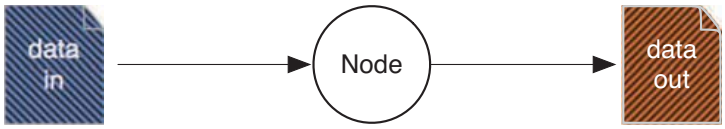
[7] See: `https://blog.torproject.org/blog/one-cell-enough`.

Why are we different?

- We are proposing an anonymous messaging scheme that:
  - provides sender anonymity (*not* recipient anonymity);
  - works without any specialised network configurations – pure HTTP(S), HTML and Javascript;
  - works with a public untrusted cloud – our router (!);
  - preserves secrecy of the message; and
  - works even with some dishonest participants.

KDDI R&D LABS

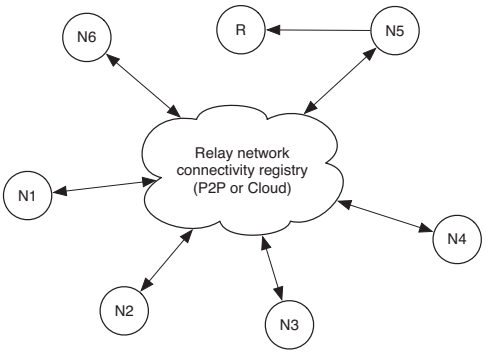
The crux of message unlinkability



- If the ingress and egress messages look indistinguishable then it is hard to tell (traffic analysis aside!) if a message going into a node is the one coming out.
- Have nodes to forward messages around before sending it to the final recipient.

KDDI R&D LABS

Message forwarding network



- Public recipient, R. Public untrusted router – the cloud or a P2P network.
- Any node n_i can either forward a message or send it to R.
- Example: R thinks that the message is from n_5 but it could be from any other node.

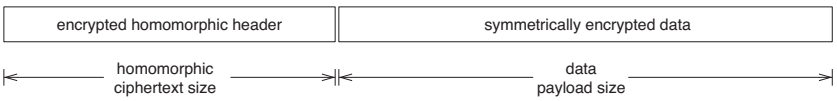
KDDI R&D LABS

How is it done?

- Alter every ingress encrypted (with recipient's public key) message at node n_i to generate the egress message for node n_{i+1} as $\mathcal{E}(m)_{n_{i+1}} = \mathcal{E}(m)_{n_i} \cdot \mathcal{E}(0)$.
- Forward the egress message with probability p_f or send it to the final recipient with probability $1 - p_f$.
- Recipient: $m = \mathcal{D}(\mathcal{E}(m)_{n_k})$.
- Ensure that the messages are of the same size, e.g., $|m| = 2048$ bits.

KDDI R&D LABS

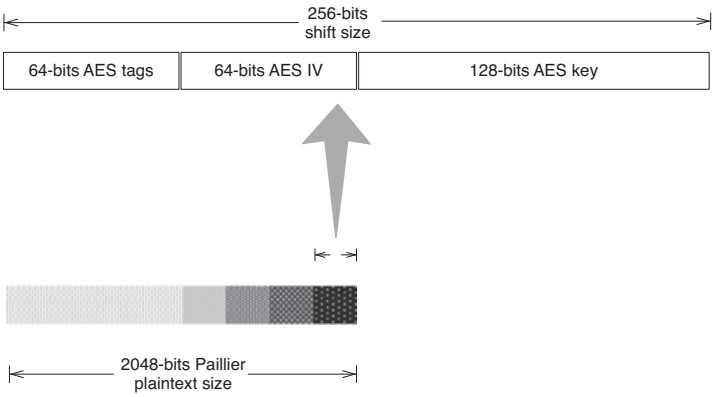
Large size of m – hybrid encryption?



- Apply symmetric key encryption on the message.
- Store random symmetric keys in a homomorphic header.
- Recipient decrypts message in multiple rounds of symmetric key decryption with keys obtained from the header.
- Will need to break messages apart and pad to maintain fixed sizes.
- Limitation: forwarding hop count.

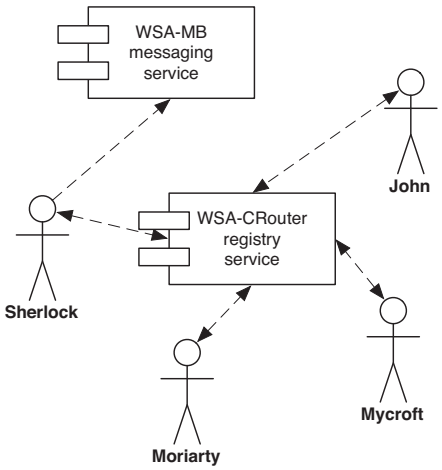
KDDI R&D LABS

Homomorphic header: an example with AES



- At any node n_i, egress message header:
  $\mathcal{E}(h_{n_{i+1}}) = \mathcal{E}(h_{n_i})^{2^{|k|+|p|}} \cdot \mathcal{E}(k_{n_i}||p_{n_i})$.
- Keys added with left shifts by $|k| + |p|$, but not many shifts before information is lost.

Real world: Sherlock's secret message



- Moriarty and his network: who was the actual sender?

Experimental validation

- HTML5 and Javascript client.
- Google App Engine (Java) cloud-based apps (F2 instance class: 1200MHz CPU, 256MB RAM).
- Demo: who wants to be Sherlock?

The public cloud-based router.

A public message board.

KDDI R&D LABS

Client side sender and forwarder nodes

- (mobile) Chrome 35.0.1916.38/iOS 7.1.1, iPhone 5S, 4G network;
- (desktop) IE 10.0.9200.16899S/Windows 8.0, 3GHz Intel Core i7 processor, 16GB RAM, 1Gbps wired network;
- (desktop) Firefox 29.0/Ubuntu Linux 14.10, 3GHz Intel Core i7 Extreme processor, 16GB RAM, 1Gbps wired network; and
- (laptop) Chrome 35.0.1916.114/Mac OS X 10.9, Macbook Air, 1.8GHz Intel Core i7 processor, 8GB RAM, 54Mbps IEEE 802.11g network.

KDDI R&D LABS

Client side performance

Platform	Mean forwarding time
(mobile) Chrome/iOS	70.9s
(desktop) IE/Windows	17.2s
(laptop) Chrome/OS X	4.06s
(desktop) Firefox/Linux	1.89s

- Bottleneck is the performance of Paillier in Javascript.
- A randomised time delay may actually help against traffic analysis attacks.
- Good news: high-performance lattice crypto in Javascript.

KDDI R&D LABS

Concluding remarks – are we there yet?

- Practical applications of partial homomorphic encryption.
- Cloud-based classifiers: collaborative filtering (this talk), support vector machines, decision trees.
- Anonymous messaging routing.
- Short-term future: partial homomorphic encryption and various encryption techniques.

KDDI R&D LABS

Thank you for your time!

Any questions?

KDDI R&D LABS

Cryptography for Cloud Service

Masayuki Yoshino (Joint work with Hisayoshi Sato)

Hitachi, Ltd.

masayuki.yoshino.aa@hitachi.com

Progress in networking technology and an increase in the demand for computing resources have prompted many organizations to outsource their computer environments. This situation has resulted in a new computing model, often called cloud computing, which gives users ability to operate software and hardware functions virtually in a moment. This ability is supplied as service with users via network, which we call cloud service.

On the one hand, cloud service allows a large amount of data accumulation on low-cost cloud storage and bigdata analysis there with rich cloud applications to general users. The cloud applications are often designed using open source code, which result in providing high-quality software with low-cost development. Even start-up companies are able to access the cloud service and gain benefits in their business.

On the other hand, this trend may invite privacy violation issues. The cloud service promotes accumulating a variety of data including private information of users and sensitive information of organizations. Although access to the data is protected from intruders, the privilege administrators are freely able to manipulate the data, which results in serious security incidents in our real life. E.g. some administrators stole and resoled the customer information to a mail listing broker.

In this talk, we introduce our research concept related to privacy protection technology and some instances for achieving secure cloud service. One of the instances uses the Apache Solr [1], which has been developed as an open source software and widely used nowadays as a full-text search engine. It is infeasible for the Apache Solr to implement most similar search techniques such as the morphological analysis and the n-gram models. The cloud storage service applying it is able to store, share and search data on cloud storage with high-speed performance thanks to their search index techniques. Most requirements may be satisfied with the cloud storage service except for the privacy violation risk against privilege administrators. Therefore, we have applied our searchable encryption scheme [2] to the Apache Solr and have developed a full-text search engine to perform the functions for encrypted documents [3]. Thanks to the encryption search function, users can enjoy secure cryptographic cloud storage, which is protected by even powerful privilege administrators.

References

[1] Apache Solr official homepage. http://lucene.apache.org/solr/

[2] Masayuki Yoshino, Ken Naganuma, Hisayoshi Sato. Symmetric Searchable Encrytion for Database Applications. IEEE NBis 2011, pp 657–662(2011).

[3] News release. Hitachi Solutions. Hitachi Solutions Announces Availability of Credeon Secure Full-text Search http://www.hitachi-solutions.com/news/release/2015/0428.html

HITACHI
Inspire the Next

Cryptography for Cloud Service

Hitachi, Ltd., Research & Development Group

Masayuki Yoshino

HITACHI
Inspire the Next

Contents

HITACHI
Inspire the Next

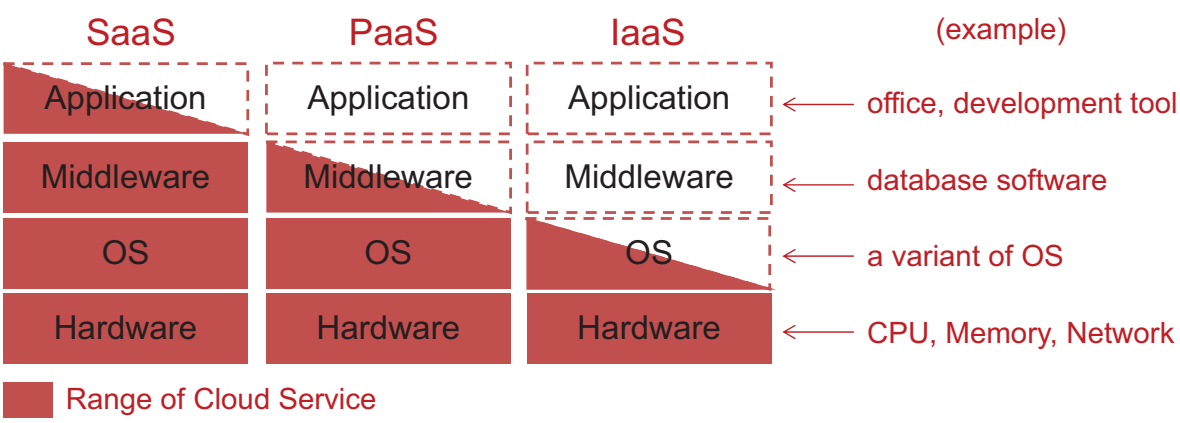
1. Information Circulation / Big Data Analysis Age

Classification of Cloud Service

HITACHI
Inspire the Next

Virtual functions of computer resources supplied via network

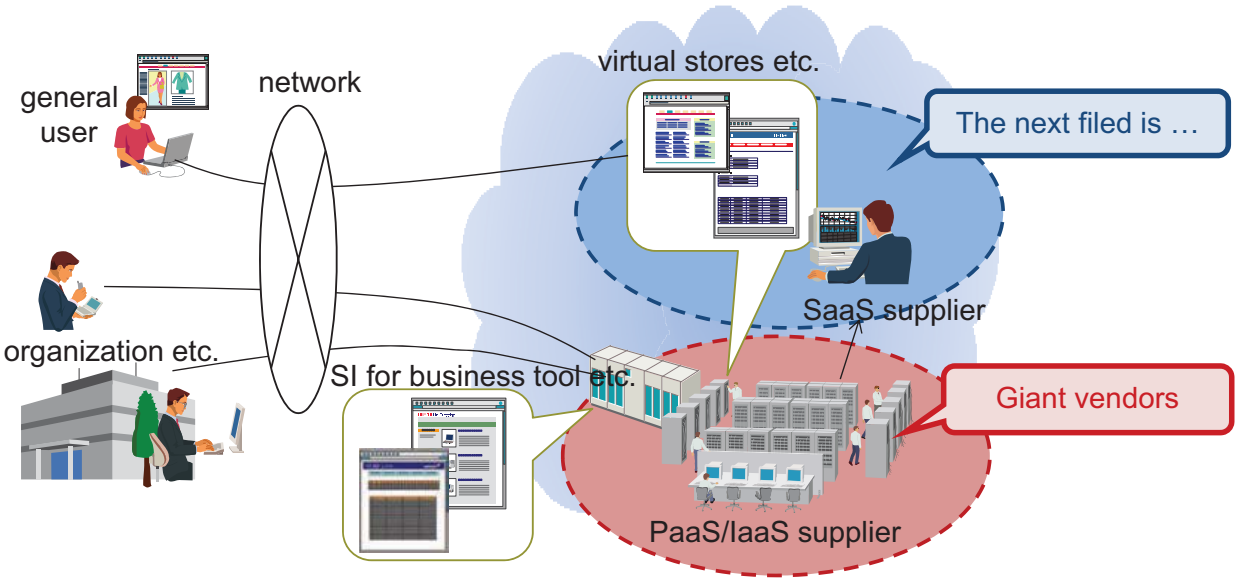
abbrev.	official name	explanation
SaaS	Software as a Service	service of application function
PaaS	Platform as a Service	service of platform for applications
IaaS	Infrastructure as a Service	service of physical device



Progress of Cloud Service

HITACHI
Inspire the Next

Virtual functions of computer resources supplied via network



Point

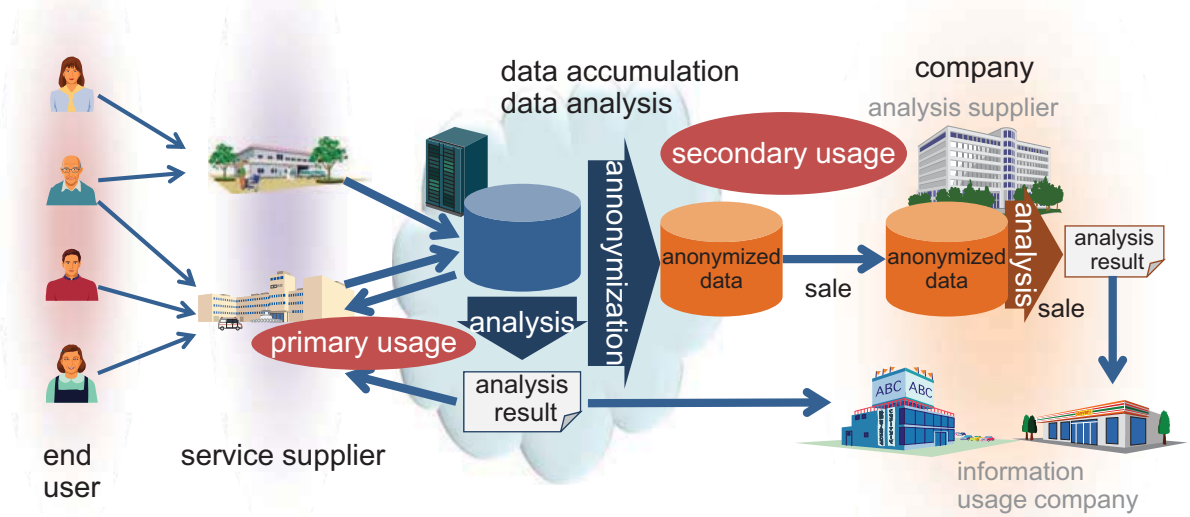
- Low-cost system integration & maintenance
- Easy data analysis: even start-up company can perform big data analysis

 4

Information Circulation and Big Data Analysis

HITACHI
Inspire the Next

This model is formulated at several industry segments such as advertisement, retail sales, health care, railway business etc.

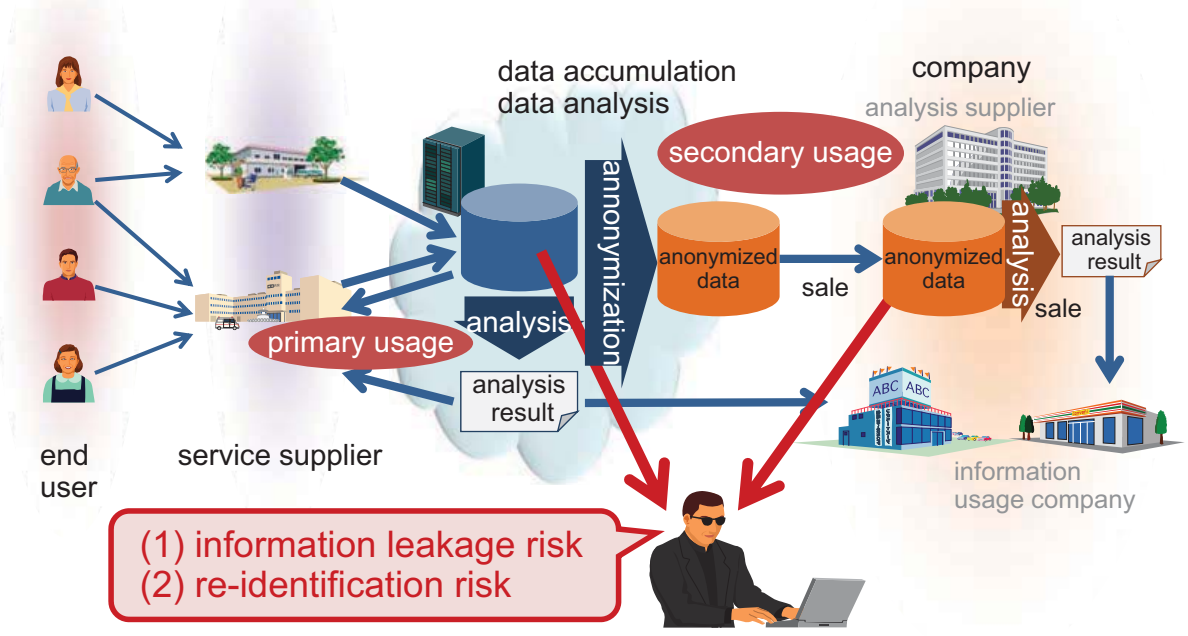


Point

- Data may include personal & sensitive information of end users
- Data is utilized by several organizations

 5

This model is formulated at several industry segments such as advertisement, retail sales, health care, railway business etc.



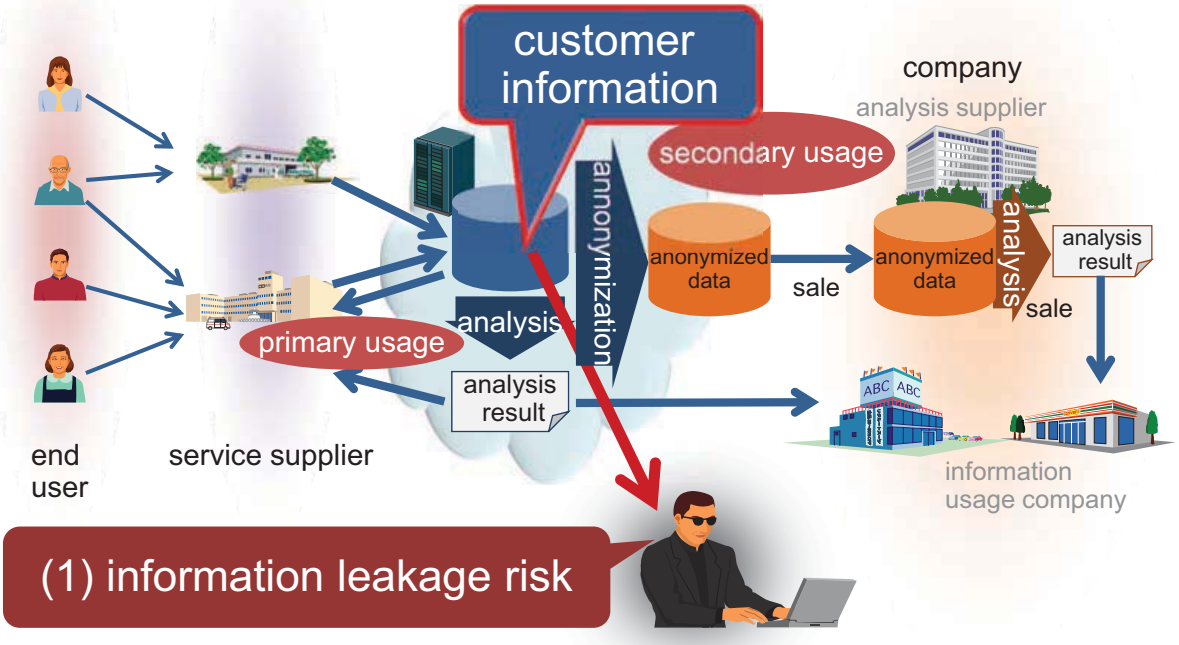
HITACHI
Inspire the Next

2. Technology for Secure Information Circulation

Information Circulation and Big Data Analysis

HITACHI
Inspire the Next

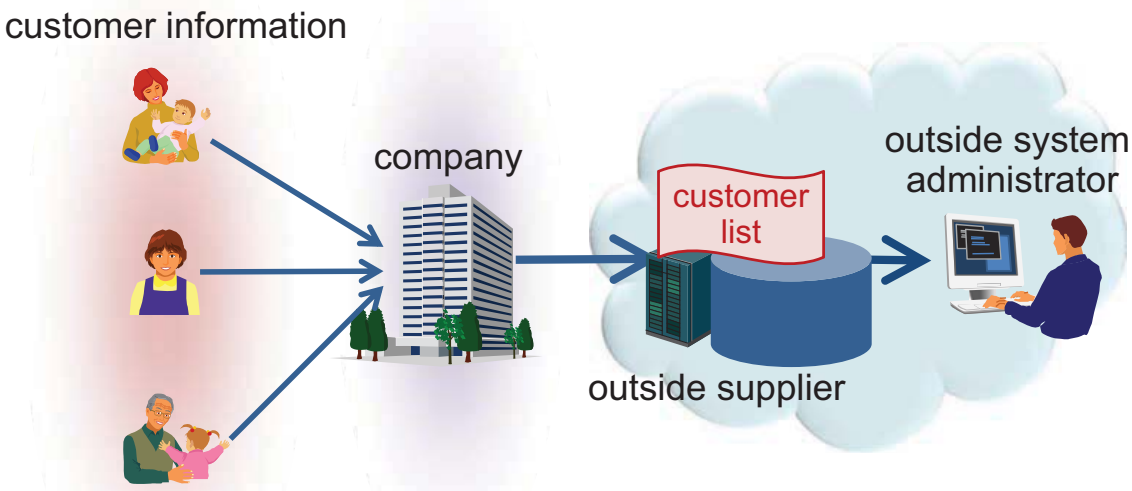
This model is formulated at several industry segments such as advertisement, retail sales, health care, railway business etc.



Recent Trend of Personal Information Leakage

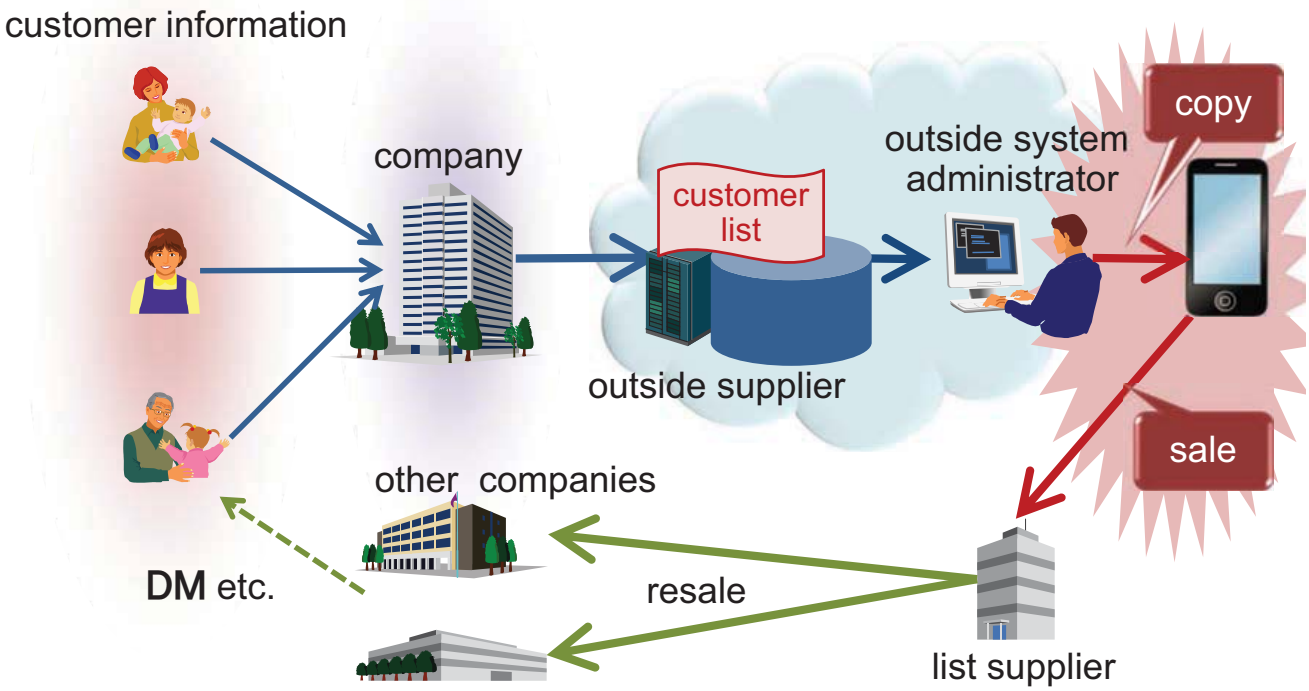
HITACHI
Inspire the Next

Case: system administrators stole and sold customer information



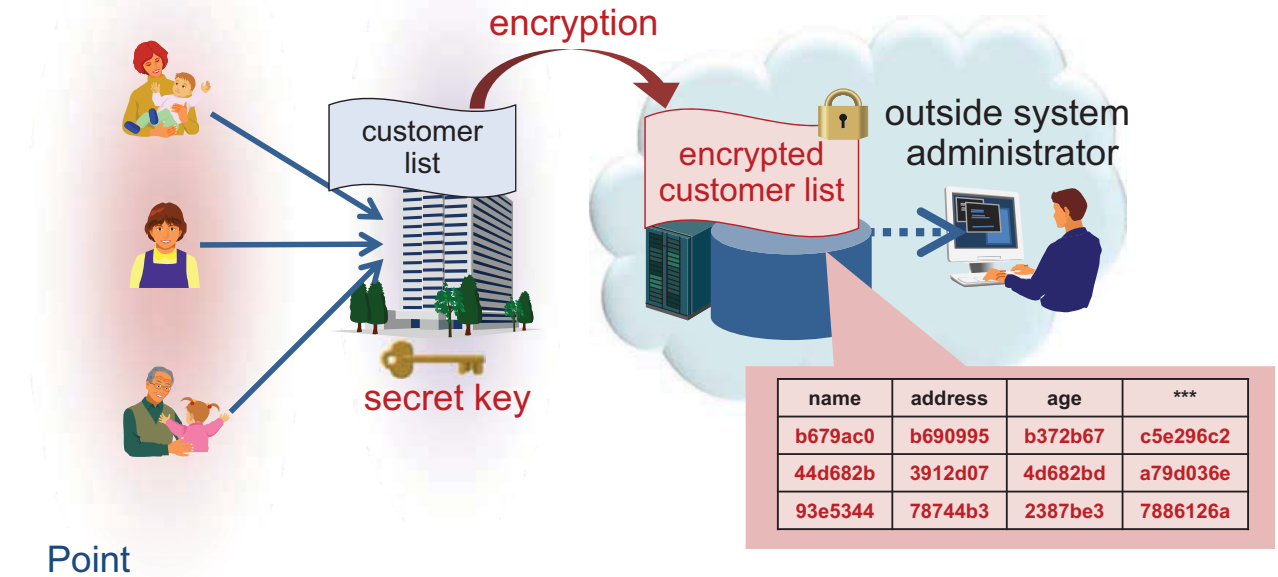
HITACHI
Inspire the Next

Case: system administrators stole and sold customer information



HITACHI
Inspire the Next

Secret key should not stored closed to ciphertext
Secret key should be managed in isolation



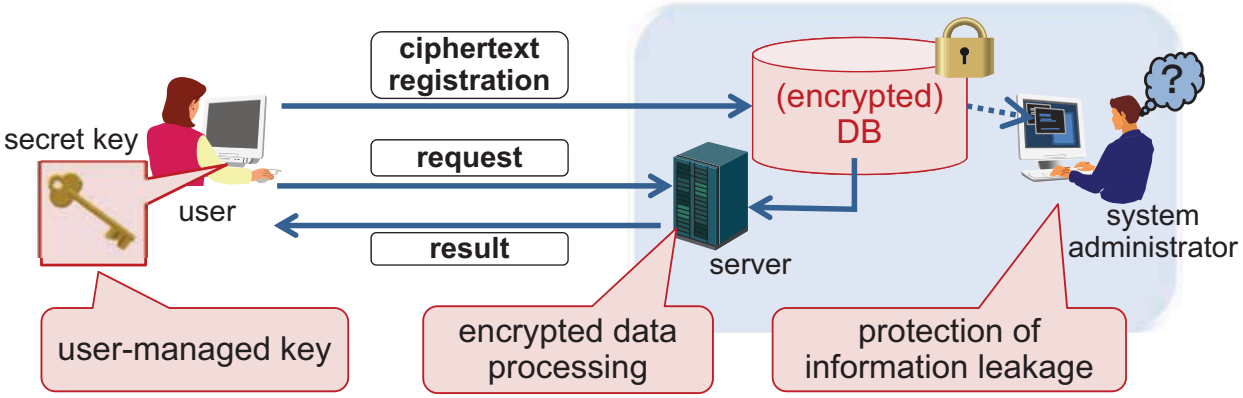
name	address	age	***
b679ac0	b690995	b372b67	c5e296c2
44d682b	3912d07	4d682bd	a79d036e
93e5344	78744b3	2387be3	7886126a

Point

- Secret key should not be given to system administrators
- Classical encryption scheme does not work in this scenario

Achieving information utilization and information security

(e.g. encrypted big data analysis)



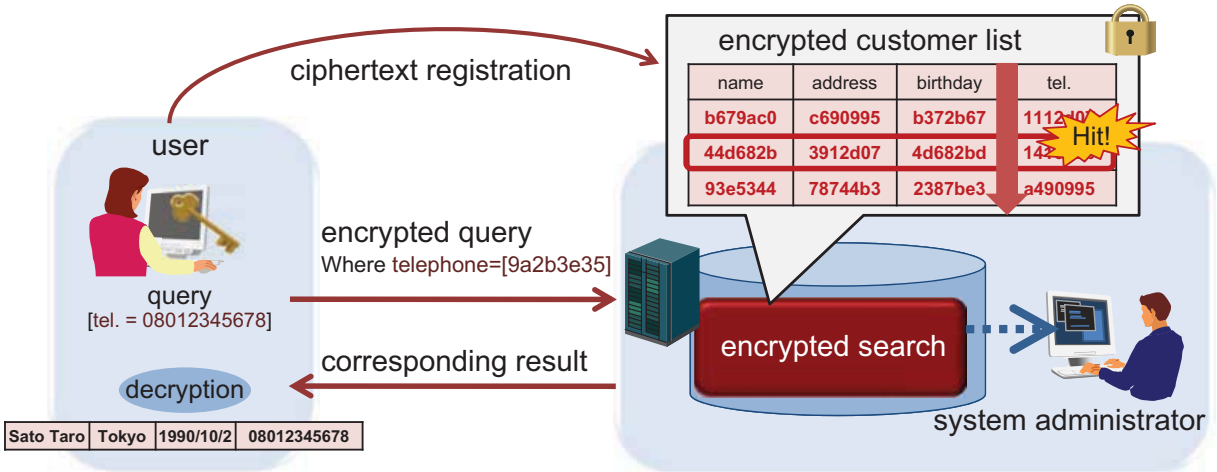
Requirements creating frontier of new business
- practical processing speed comparable to plaintext performance
- explicit function benefit and simple composition

■ Function

(1) encryption, (2) decryption, (3) search encrypted data using encrypted query

■ Hitachi scheme

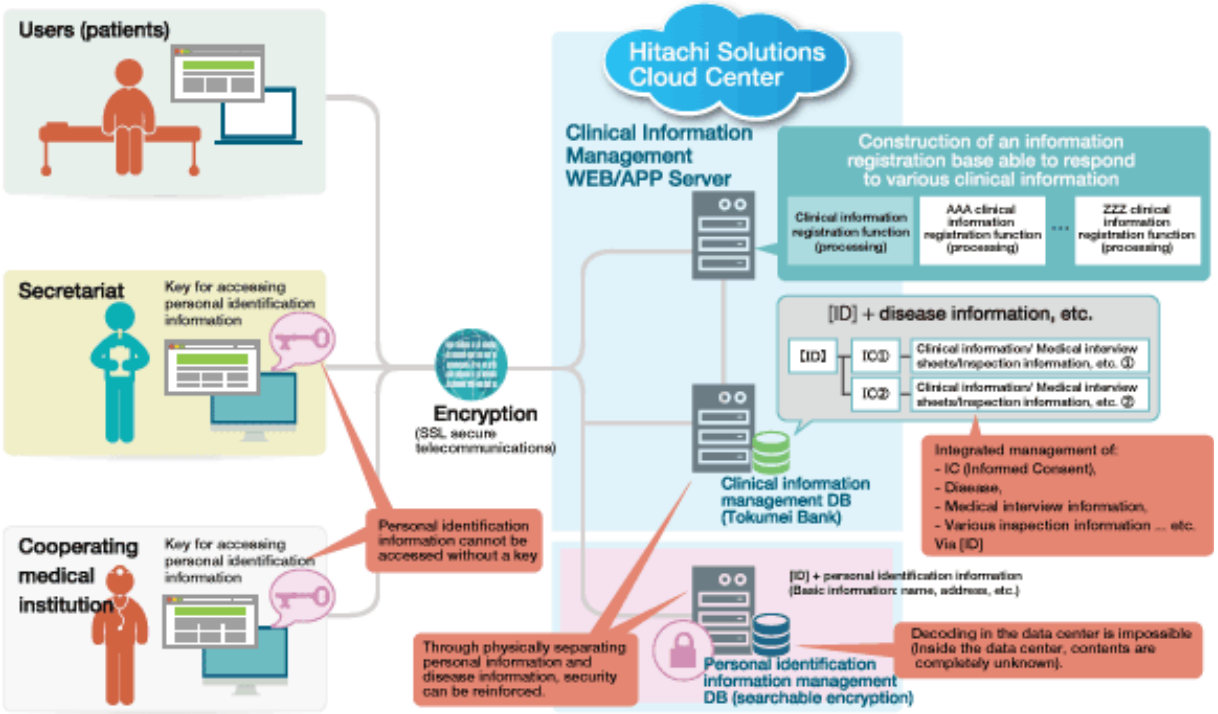
- Benefit
  - encryption prevents information leakage from cloud administrators
  - search function makes cloud service available
- Speedy
  - comparable to plaintext performance: 1 million searches / 1 second



Example of Our Business

HITACHI
Inspire the Next

"Remudy WEB Patient Information Registration System"
- launched with National Center of Neurology and Psychiatry on 2015 April
- in accordance with ethical guidelines for clinical research

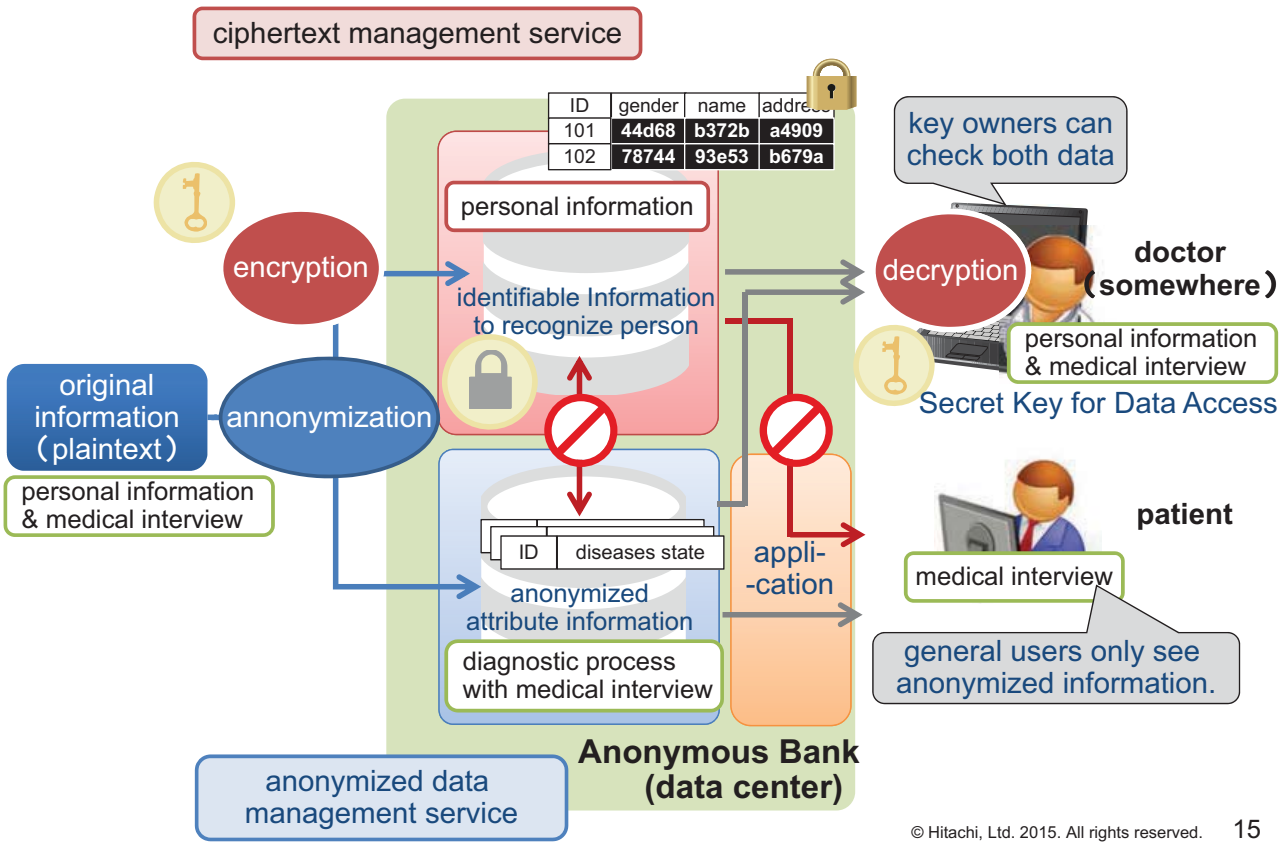


・Official homepage: http://www.remudy.jp/

 14

Sketch of the Example implementing Our Technology

HITACHI
Inspire the Next



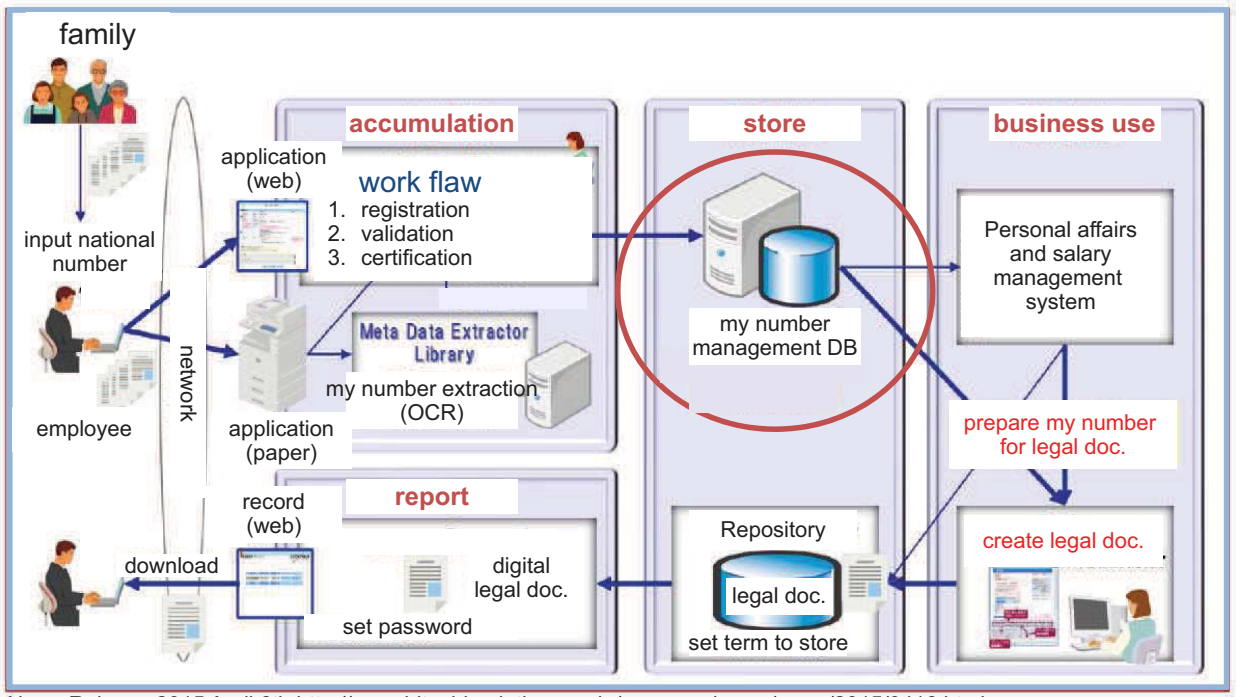
 15

F.Y.R. Management of National Identification Number

HITACHI
Inspire the Next

“My Number Solution”

- news release: cloud service to store “my number” and put it on legal documents



News Release 2015 April 6th http://www.hitachi-solutions.co.jp/company/press/news/2015/0416.html

Example of Our Product

HITACHI
Inspire the Next

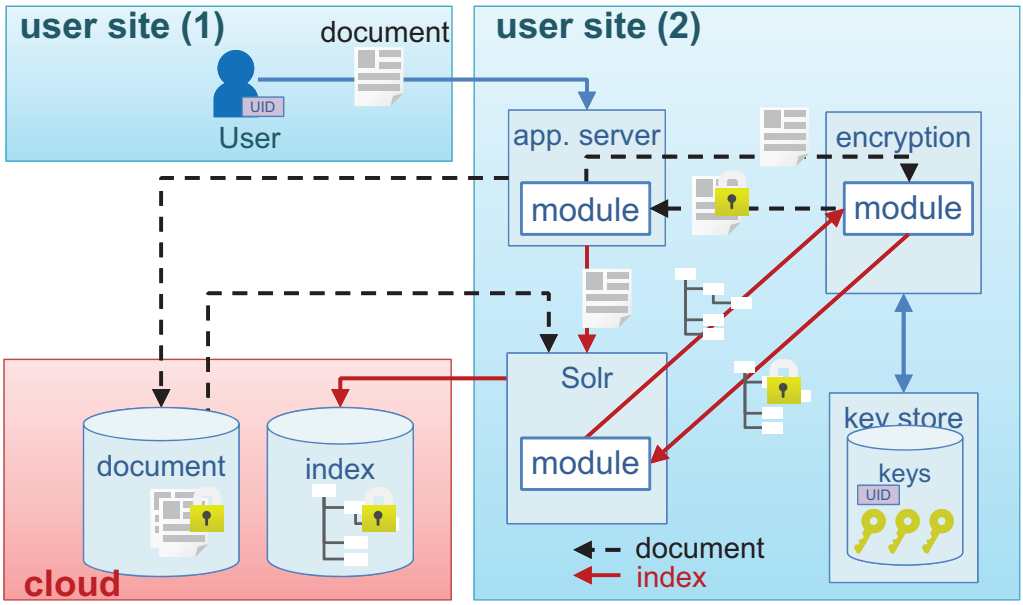
The 1st product of Hitachi searchable encryption supports the infrastructure

Apache Solr

Secure Full-Text Search using Apache Solr

HITACHI
Inspire the Next

"Credeon Secure Full-text Search"
- released on 2015 April
- providing modules to run on a full-text search engine

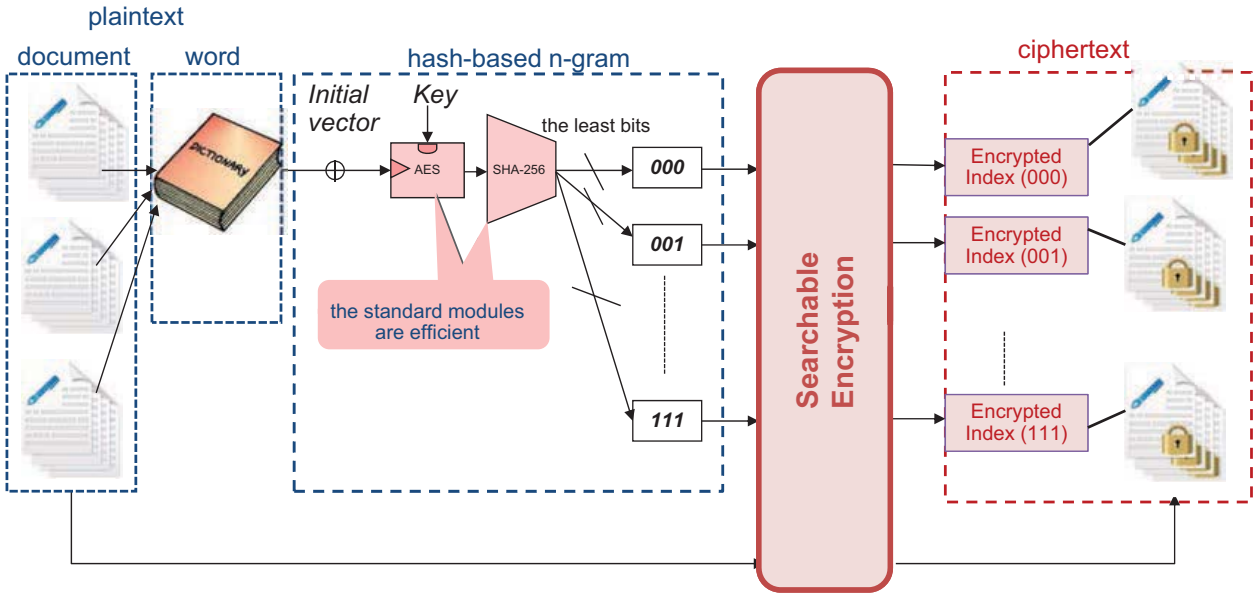


http://www.hitachi-solutions.com/securesearch/
news release 2015/4/28
http://web.hitachi-solutions.com/news/release/2015/0428.html

Design for Efficient Full-Text Search

HITACHI
Inspire the Next

- Accelerating speed to handle big data
  - morphological analysis
  - n-gram models

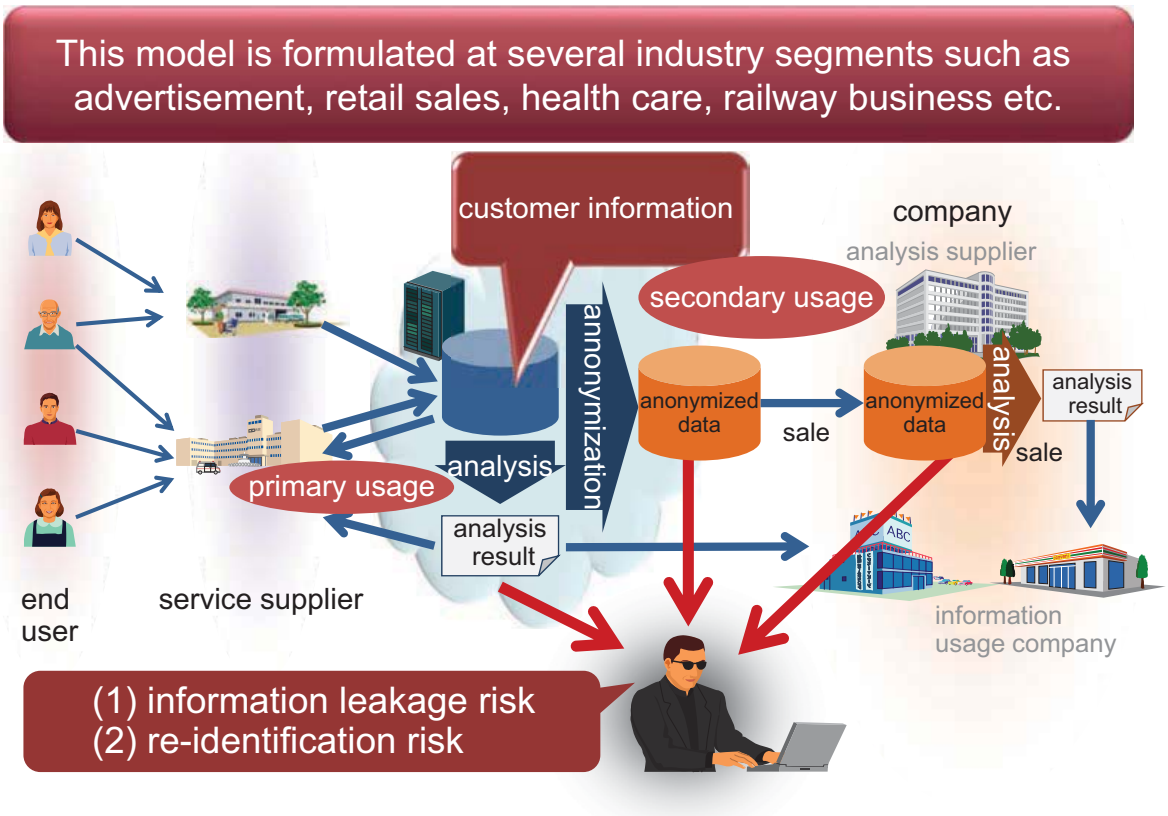


HITACHI
Inspire the Next

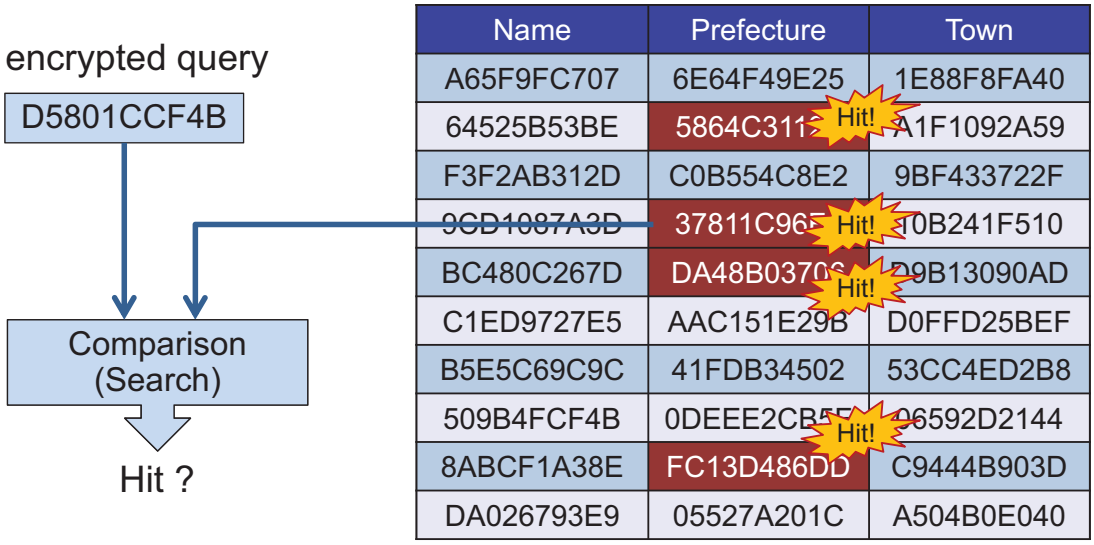
3. Technology for Privacy-Protected Big Data Analysis

Information Circulation and Big Data Analysis

HITACHI
Inspire the Next



Count on frequency of specific items, which are encrypted

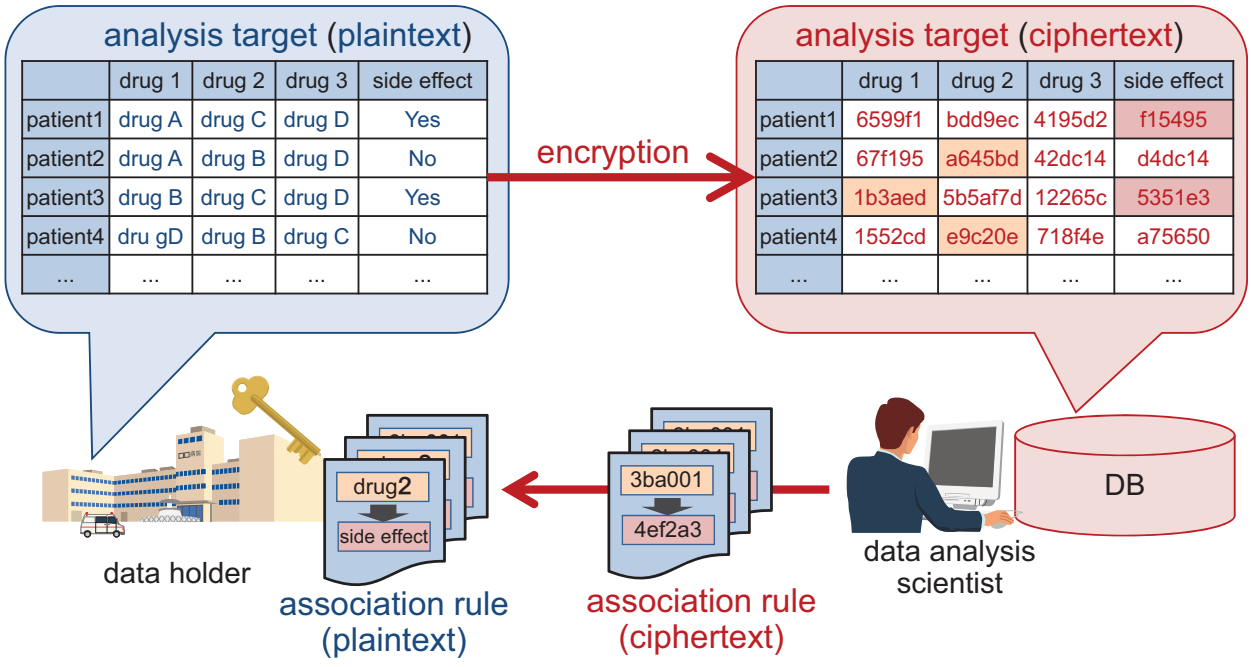


It is feasible to learn frequency based on search results

※Note that security model against cloud administrator is not consistent with the case of searchable encryption.

Privacy-Protection Big Data Analysis(2)

- Associations rule analysis extracting dependency relation of all items
- Analyze 100,000 records in 10 minutes on a normal computer



	drug 1	drug 2	drug 3	side effect
patient1	drug A	drug C	drug D	Yes
patient2	drug A	drug B	drug D	No
patient3	drug B	drug C	drug D	Yes
patient4	dru gD	drug B	drug C	No
...	...	...	...	...

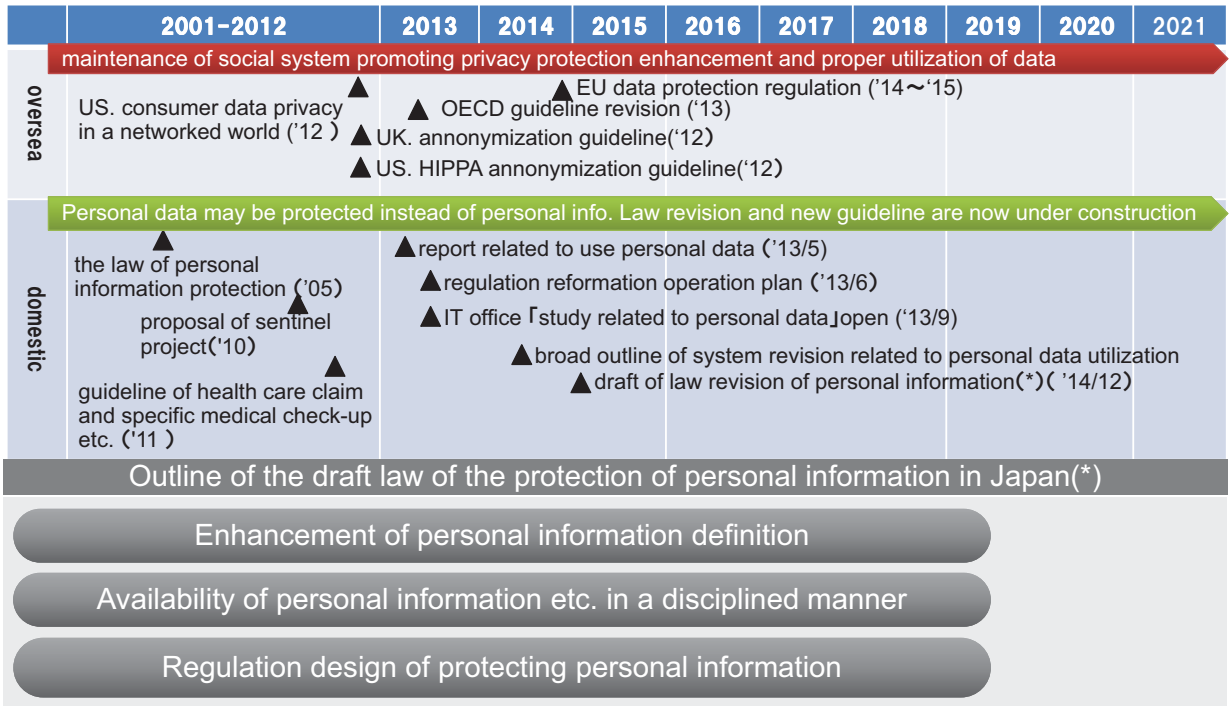
	drug 1	drug 2	drug 3	side effect
patient1	6599f1	bdd9ec	4195d2	f15495
patient2	67f195	a645bd	42dc14	d4dc14
patient3	1b3aed	5b5af7d	12265c	5351e3
patient4	1552cd	e9c20e	718f4e	a75650
...	...	...	...	...

News Release 2014 Jan 21th, http://www.hitachi.co.jp/New/cnews/month/2014/01/0121b.html

Situation related to Personal Data Utilization

HITACHI
Inspire the Next

- Act of revising the protection of personal information in US and EU
- In Japan, the draft law is being under Diet deliberation

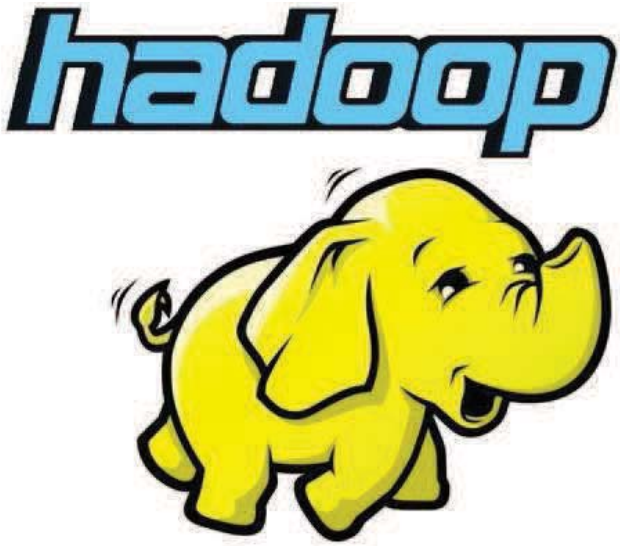


HIPPA: Health Insurance Portability and Accountability Act (*) Cabinet Secretariat IT office, 2014 December 19th

Example of Our Product

HITACHI
Inspire the Next

The 1st version of Hitachi k-annonymization supports the infrastructure



Simple Anonymization

HITACHI
Inspire the Next

original personal data

name	tel.	address	age	gender	height	weight	disease
Taro Yamada	0332581111	Yoyogi, Shibuya-ku, Tokyo	35	M	175	65	diabetes
Hanako Hitachi	0332581111	Shinjuku, Shinjuku-ku, Tokyo	42	F	160	59	heart disease
Taichi Yoshida	0423231111	Miyako-shi, Okinawa	100	M	190	60	heart disease
Jiro Sato	03293888	Hongo, Bunkyo-ku, Tokyo	33	F	155	45	Cancer
.....	.....	Naha-shi, Okinawa	81	F	145	55	heart disease

simple anonymization

Re-identification by Combination of attributes

swapping

name	tel.	address	age	gender	height	weight	disease
~~Taro Yamada~~	~~0332581111~~	Yoyogi, Shibuya-ku, Tokyo	35	M	175-189	65	heart disease
~~Hanako Hitachi~~	~~0332581111~~	~~Shinjuku, Shinjuku-ku, Tokyo~~	~~42~~	~~F~~	~~160-164~~	59	diabetes
Taichi Yoshida	0423231111	Miyako-shi, Okinawa	100	M	181-	60	heart disease
~~Jiro Sato~~	~~03293888~~	**	**	**	**	**	**
.....	.....	Naha-shi, Okinawa	81	F	141-145	55	heart disease

cutting

top coding

grouping

・linkable anonymization
・unlinkable anonymization

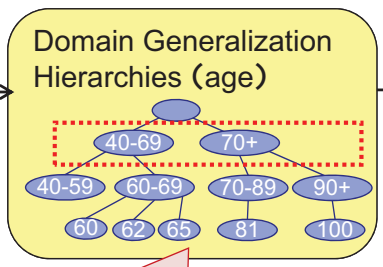
k-Anonymization (Conventional method)

HITACHI
Inspire the Next

Transform the data such that there always exist at least k same records in the data

address	age	sex
Yoyogi, Shibuya-ku, Tokyo	68	M
Shinjuku, Shinjuku-ku, Tokyo	63	F
Yoyogi, Shibuya-ku, Tokyo	68	M
Miyakojima-shi, Okinawa	100	M
Hongo, Bunkyo-ku, Tokyo	65	F
Ikebukuro, Toshima-ku, Tokyo	63	F
Naha-shi, Okinawa	81	F

Before anonymization



k-Anonymization

address	age	sex
Tokyo	40-69	M
Tokyo	40-69	F
Tokyo	40-69	M
Okinawa	70-	-
Tokyo	40-69	F
Tokyo	40-69	F
Okinawa	70-	-

k-anonymized data (k=2)

Problem 1
k-Annonymization needs domain generalization hierarchies, which must be prepared (usually by human hand).
Furthermore, quality of anonymized data depends on quality of generalization hierarchies.

Problem 2
It takes a long time to anonymize a large amount of data.

Hitachi's k-Anonymization

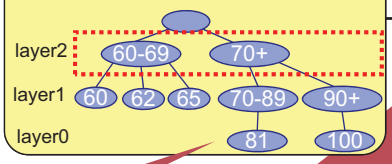
HITACHI
Inspire the Next

- Automated generation of domain generalization hierarchies based on the frequency
  ➔ Information loss is suppressed (25% down)
- By distributed computing, 2.5 million data can be anonymized in 12 minutes (10 times faster)

address	age	sex
Yoyogi, Shibuya-ku, Tokyo	68	M
Shinjuku, Shinjuku-ku, Tokyo	63	F
Yoyogi, Shibuya-ku, Tokyo	68	M
Miyakojima-shi, Okinawa	100	M
Hongo, Bunkyo-ku, Tokyo	65	F
Ikebukuro, Toshima-ku, Tokyo	63	F
Naha-shi, Okinawa	81	F

Before Anonymization

Automated generation of domain generalization hierarchies (age)



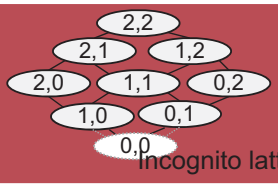
Distributed k-anonymization

address	age	sex
Tokyo	40-69	M
Tokyo	40-69	F
Tokyo	40-69	M
Okinawa	70-	-
Tokyo	40-69	F
Tokyo	40-69	F
Okinawa	70-	-

k-Anonymized data (k=2)

Data with less frequency are placed on lower layer.

The dominant search process (70% of all) is distributed into several nodes, which are performed by Hadoop in parallel



Incognito lattice

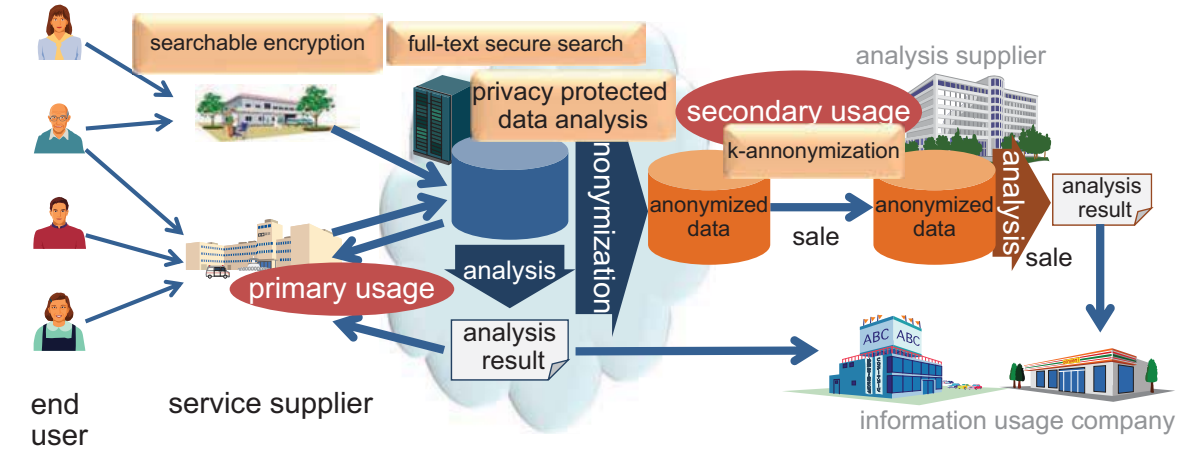
HITACHI
Inspire the Next

4. Conclusion

Conclusion

HITACHI
Inspire the Next

Technology for information circulation and big data analysis is going to apply the following model



Future issues for creating new market

① Technology for manipulating encryption data	② Transparency of Security policy and implementation	③ Development of secure cloud with secure application	④ Promotion including maintenance such as PIA and rules

HITACHI
Inspire the Next

Cryptography for Availability The Case of Secure Cloud Storage

Sherman S. M. Chow

Department of Information Engineering, The Chinese University of Hong Kong
sherman@ie.cuh.edu.hk

Confidentiality, integrity, and availability, are three important aspects of cyber security. Cryptography can help to ensure the former two [1, 3, 4, 5, 6, 7]. This talk explores the remaining one, availability, in the context of cloud storage.

Nowadays, many organizations outsource data storage to the cloud such that a data owner of an organization can easily share data with other members, whom in turn may operate on the data [9]. To verify that the data remains intact on the cloud, Proof of Retrievability (PoR) and Provable Data Possession (PDP), are proposed.

(I): To outsource as much as possible, the users may delegate the auditing to a third party auditor (TPA). We discuss how to enable privacy-preserving public auditing, such that the TPA cannot learn the cloud data [10]. We also discuss how the TPA can perform audits for multiple users simultaneously and efficiently.

(II): We note that public verifiability enables a untrusted cloud to infer the identity of a data owner, which may have security implications in an enterprise setting. We propose a simple and efficient approach to ensure cloud data integrity without sacrificing the anonymity of data owners nor requiring significant verification metadata with the help of a security-mediator (SEM) [8]. The SEM does not learn anything about the data to be uploaded. We also extend to the multi-SEM model for higher availability.

Both ideas can be applied to a wide range of secure cloud storage systems [2].

References

[1] Melissa Chase, Sherman S. M. Chow. Improving Privacy and Security in Multi-Authority Attribute-Based Encryption. ACM CCS, p. 121-130, 2009.

[2] Fei Chen, Tao Xiang, Yuanyuan Yang, and Sherman. S.M. Chow. Secure Cloud Storage meets with Secure Network Coding. IEEE Trans. Computers, to appear.

[3] Sherman S. M. Chow. Removing Escrow from Identity-Based Encryption. PKC, p. 256-276, 2009.

[4] Sherman S.M. Chow, Cheng-Kang Chu, Xinyi Huang, Jianying Zhou, Robert H. Deng. Dynamic Secure Cloud Storage with Provenance. Festschrift J. Quisquater (Crypt.&Sec.). p. 442-464, 2012

[5] Sherman S.M. Chow, Yi Jun He, Lucas Chi Kwong Hui, Siu-Ming Yiu. SPICE - Simple Privacy-Preserving Identity-Management for Cloud Environment. ACNS, p. 526-543, 2012

[6] Cheng-Kang Chu, Sherman S.M. Chow, Wen-Guey Tzeng, Jianying Zhou, and Robert H. Deng. Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage. TPDS, 25(2):468-477, 2014.

[7] Yue Tong, Jinyuan Sun, Sherman S.M. Chow, Pan Li. Cloud-Assisted Mobile-Access of Health Data With Privacy and Auditability. IEEE J. Biomedical & Health Info. 18(2), p. 419-429, 2014.

[8] Boyang Wang, Sherman S.M. Chow, Ming Li, Hui Li. Storing Shared Data on the Cloud via Security-Mediator. IEEE Intl. Conf. on Distributed Computing Sys., (ICDCS), p. 124-133, 2013.

[9] Boyang Wang, Ming Li, Sherman S.M. Chow, Hui Li. A Tale of Two Clouds: Computing on Data Encrypted under Multiple Keys. CNS, p. 337-345, 2014.

[10] Cong Wang, Sherman S.M. Chow, Qian Wang, Kui Ren, Wenjing Lou. Privacy-Preserving Public Auditing for Secure Cloud Storage. IEEE Trans. Computers 62(2), p. 362-375, 2013.

Cryptography for Availability

The Case of Secure Cloud Storage

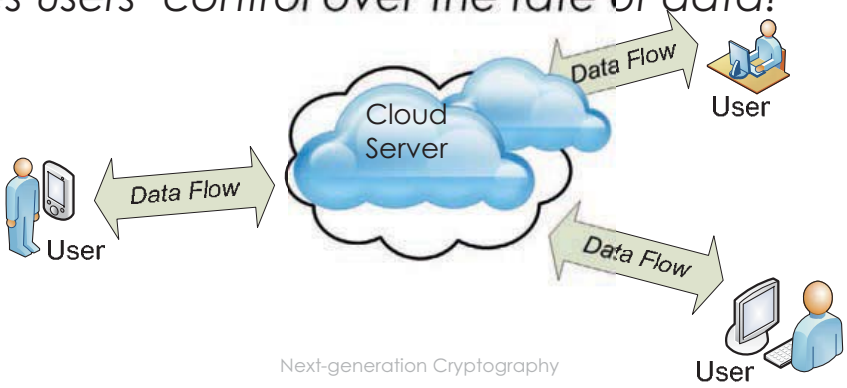
Sherman S. M. Chow
Department of Information Engineering
Chinese University of Hong Kong

Introduction

- Secure Cloud Storage
- Auditing Protocols
  - Simple examples
  - Desirable properties
- Building Blocks
- Concrete Constructions

Basic Settings of Cloud Storage

- Long-term reliable storage is expensive
- Client stores (long) file with server
  - Online backup, Software as a service (SaaS), etc.
- *Eliminates users' control over the fate of data!*



What do you trust?

- Broad range of threats for data integrity
  - Int.: management errors, software bugs, …
  - Ext.: malware, economically motivated attacks, …
    - Amazon S3-Feb., Jul. 08; Gmail-Dec. 06; Apple MobileMe-Jul. 08…
- Cloud servers might behave *unfaithfully*
  - Discard rarely accessed data for monetary reason
  - Hide data loss incident for reputation
- *If cloud deleted your copy, nothing you can do?*
- *Crypto. "for" Availability, at least complain with evidence!*

POR / PDP

- *Audit on data integrity and availability?*
- A proof of retrievability (POR) is a compact proof
  - by a file system (*prover*) to a client (*verifier*) that
  - a target file *F* is intact,
  - in the sense that the client can fully recover it.
- PORs: Proofs of Retrievability for Large Files
  - Ari Juels, Burton S. Kaliski Jr., CCS '07, p. 584-597
- Provable Data Possession at Untrusted Stores.
  - Giuseppe Ateniese, Randal C. Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary N. J. Peterson, Dawn Song, CCS '07, p. 598-609

Papers covered in this talk

- Privacy-Preserving Public Auditing for Secure Cloud Storage
  - Cong Wang, **C**, Qian Wang, Kui Ren, Wenjing Lou
  - IEEE Trans. Computers 62(2) p. 362-375, 2013
- Storing Shared Data on the Cloud via Security-Mediator
  - Boyang Wang, **C**, Ming Li, Hui Li
  - IEEE Intl. Conf. on Distributed Comp. Sys. (ICDCS) 2013, p. 124-133

Outline

- Complete delegation
  - Do you want to do the verification yourself?
  - If you are not the verifier, can you trust the verifier?
- Privacy issue
  - Is it a problem to reveal the signer / (meta) data generator?
  - Can you trust the cloud?

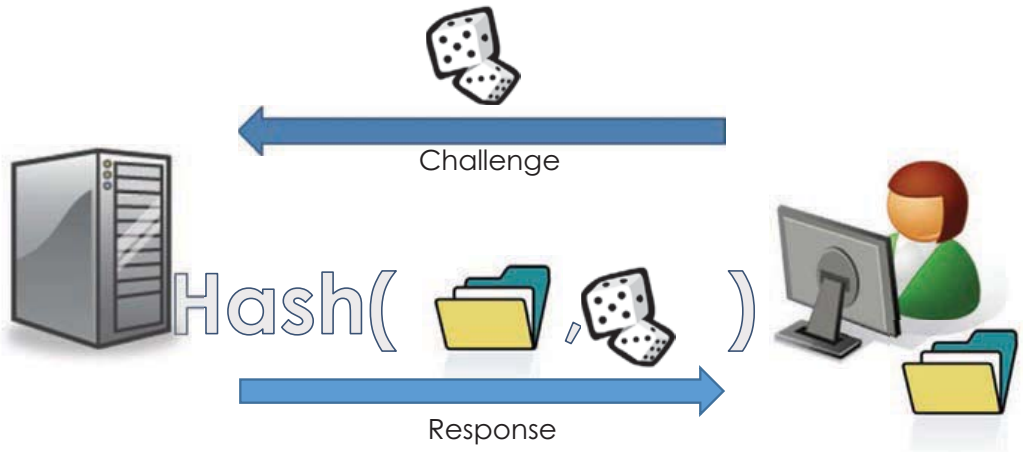
Remind me, why cloud storage?

- *Relief of the burden for storage management*
  - Why the user still needs to *worry* about auditing?
  - We want a *complete* solution of data outsourcing
- A third party auditor (TPA) can take care of it
  - E.g. IT department of an enterprise
  - Single party to handle many clients/storages
  - Continuous correctness assurance
  - Should not introduce additional *vulnerabilities*

Confidentiality of Data

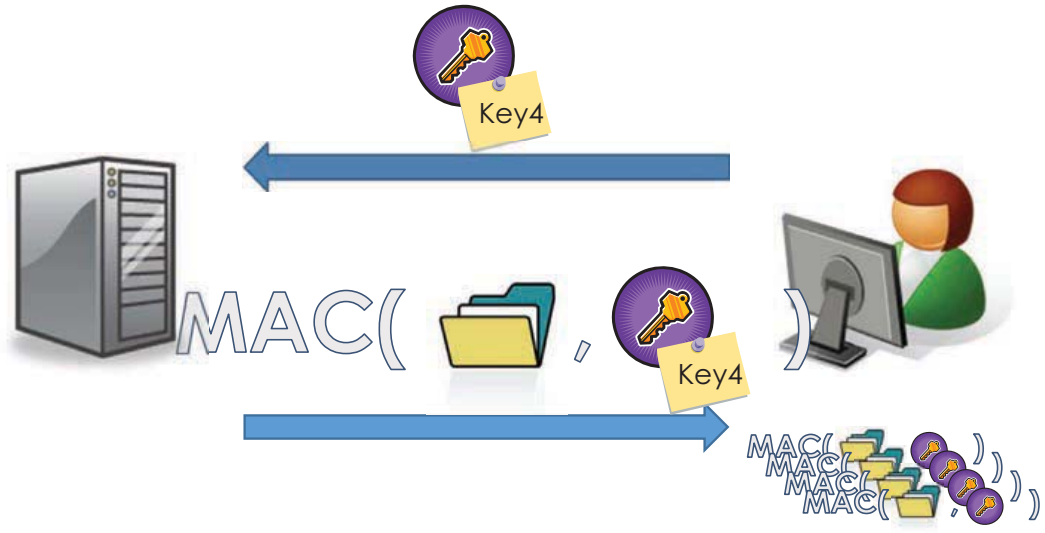
- The TPA should not learn any user data
  - Privacy, financial interests
  - Law: US Health Insurance Portability & Accountability Act
- Encryption
  - reduces the problem to (complex) key management
  - may be an overkill in some cases
- *Publicly-verifiable* Proof of Retrievability?
  - "Doesn't need to know the data" doesn't necessarily mean "Cannot recover the data"

E.g. 1: Challenge + Message Digest



[Kotla, Alvisi, Dahlin, Usenix ATC 2007]

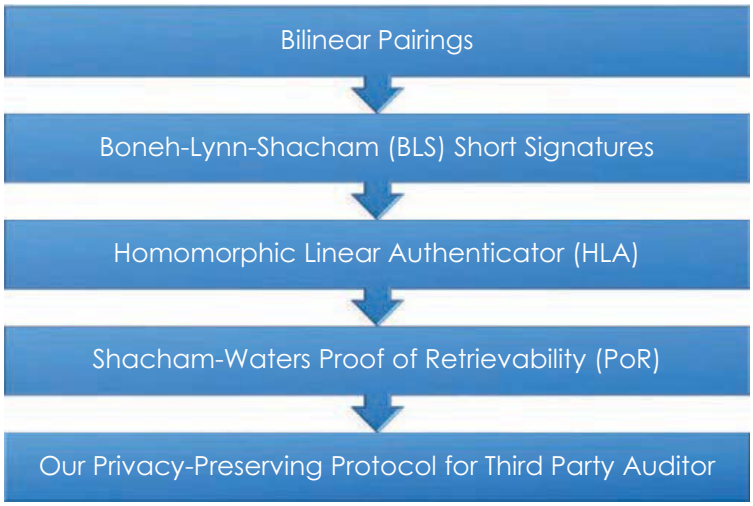
E.g. 2: Message Authentication Code



System Criteria

- Storage overhead
- Computation
  - server side and client side
  - including number of block reads
  - initial setup by the client (must process all parts)
  - server must "touch" parts of the file being audited
- Communication
- Unlimited use
- Stateless verification

Roadmap



Bilinear Map / Pairings

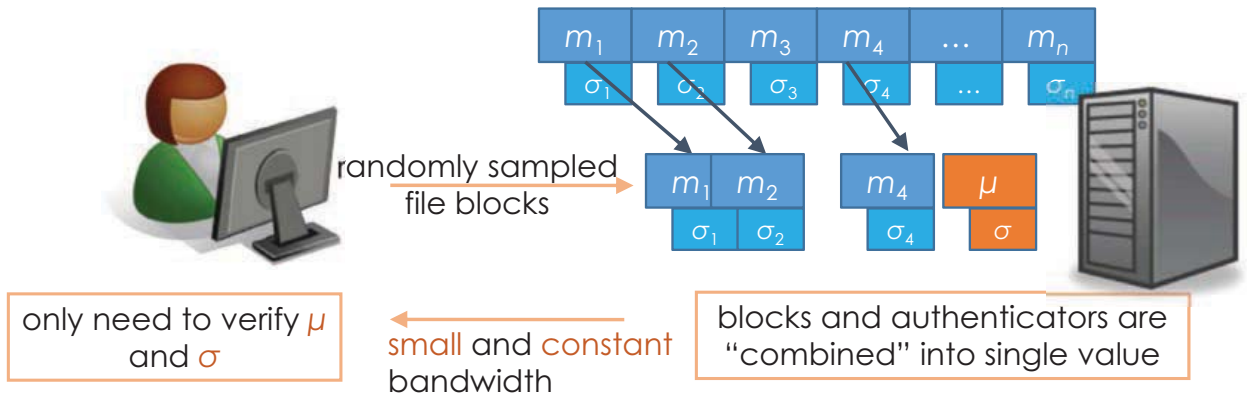
- Let **G** and $\mathbf{G}_T$ be 2 multiplicative groups
  - of prime order p
- **Computational Diffie-Hellman (CDH)** problem:
  - given $(g, g^a, g^b) \in \mathbf{G}^3$, compute g^{ab}
- e: **G** x **G**→$\mathbf{G}_T$ is bilinear:
  - $\forall u, v \in \mathbf{G}, a, b \in \mathbf{Z}_p, e(u^a, v^b) = e(u, v)^{ab} = e(u^{ab}, v)$
- So it is easy to decide if a tuple is CDH

Boneh-Lynn-Shacham Signature

- H: {0, 1}* → **G**
- Private signing key $sk = x \in \mathbf{Z}_p$
- Public key $pk = g^x$
- Sign(sk, m): $\sigma = [H(m)]^x$
- Ver(pk, σ, m): `Valid' iff $e(\sigma, g) = e(H(m), pk)$

- Correctness: $e(\sigma, g) = e(H(m)^x, g) = e(H(m), g^x)$

Proof of Retrievability



Homomorphic Linear Authenticator (HLA)

- Let σ_1, σ_2 be 2 authenticators on m_1, m_2 resp.
- $(\sigma_1)^a(\sigma_2)^b$ is an "authenticator" on $(m_1)^a(m_2)^b$
- Easily forgeable?
- "Linear combination"
- BLS signature: $[H(m)]^x$
- H's (pseudo-)randomness gives unforgeabiliity
- Easier if the message is an exponent

Shacham-Waters HLA (Asiacrypt'08, JoC'13)

- $sk = x \in \mathbf{Z}_p$, $pk = g^x$, $u \in \mathbf{G}$, H: {0, 1}* → **G**

- Auth(sk, m_i, i): $\sigma_i = [H(name \mid\mid i)\ u^{m_i}]^x$
  - *name* is randomly chosen from a large domain

- Ver(pk, σ_i, m_i, i, *name*): (let $W_i = name \mid\mid i$)
- Output `1' iff $e(\sigma, g) = e(H(W_i), pk)\ e(u^{m_i}, pk)$

- Intuition: cannot find y such that $H(W)^y = H(W')$

Homomorphic Property

- $\sigma_i = [H(W_i)\ u^{m_i}]^x$, $e(\sigma_i, g) = e(H(W_i), pk)\ e(u^{m_i}, pk)$

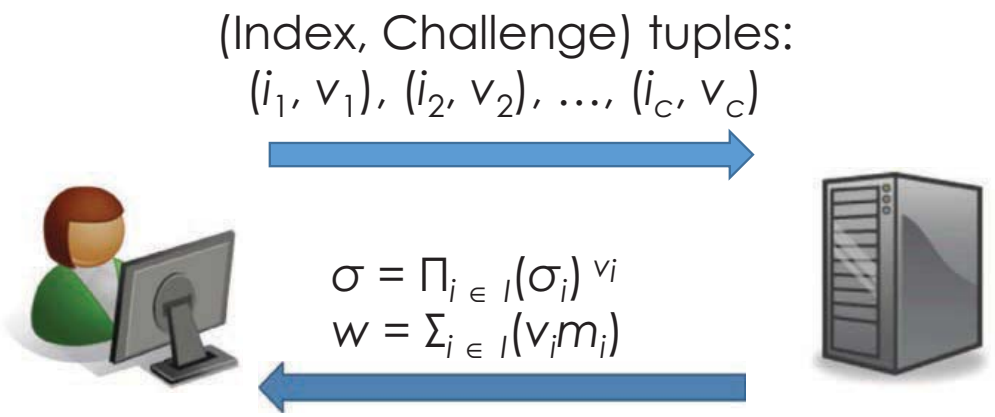
- $\sigma_i = [H(W_i)\ u^{m_i}]^x$, $\sigma_j = [H(W_j)\ u^{m_j}]^x$

- Suppose $\sigma = (\sigma_i)^a(\sigma_j)^b$
- $e(\sigma, g) = (e(\sigma_i, g))^a(e(\sigma_j, g))^b$

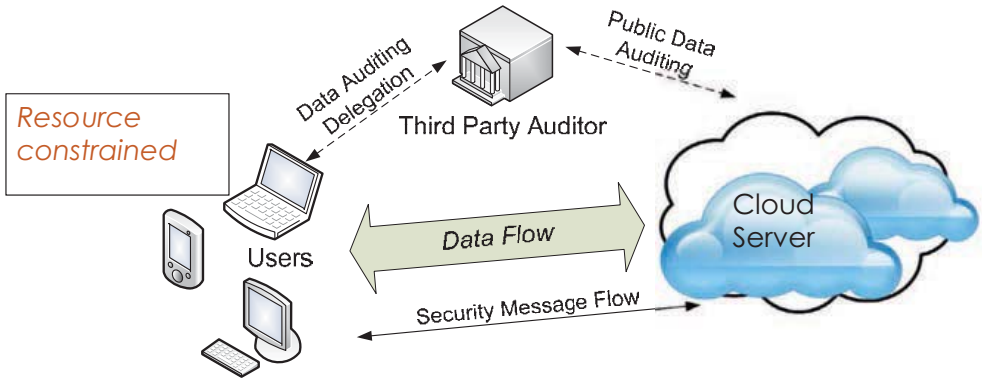
$= e(H(W_i)\ {}^aH(W_j)^b, pk)\ e(u^{a(m_i) + b(m_j)}, pk)$

Linear combination in the exponent: $a(m_i) + b(m_j)$

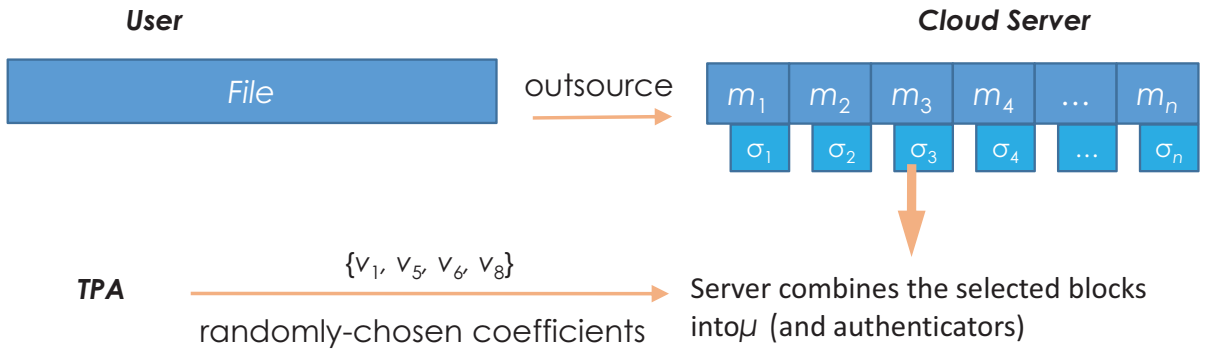
(Public-Verifiable) PoR from HLA



Our Settings

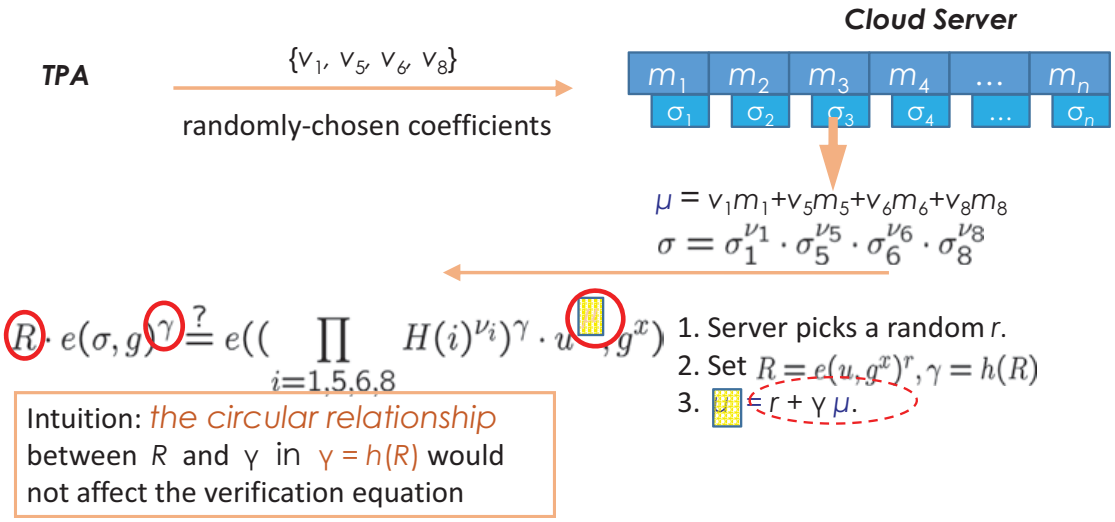


Data Privacy / Confidentiality Problem



- $\mu = v_1m_1+v_5m_5+v_6m_6+v_8m_8$ leaks partial information of data to TPA.
- Direct adoption of the technique is unsuitable for public auditing.

Privacy-Preserving 3rd Party Auditing



Probability of Detection

$$(i_1, v_1), (i_2, v_2), \ldots, (i_c, v_c)$$

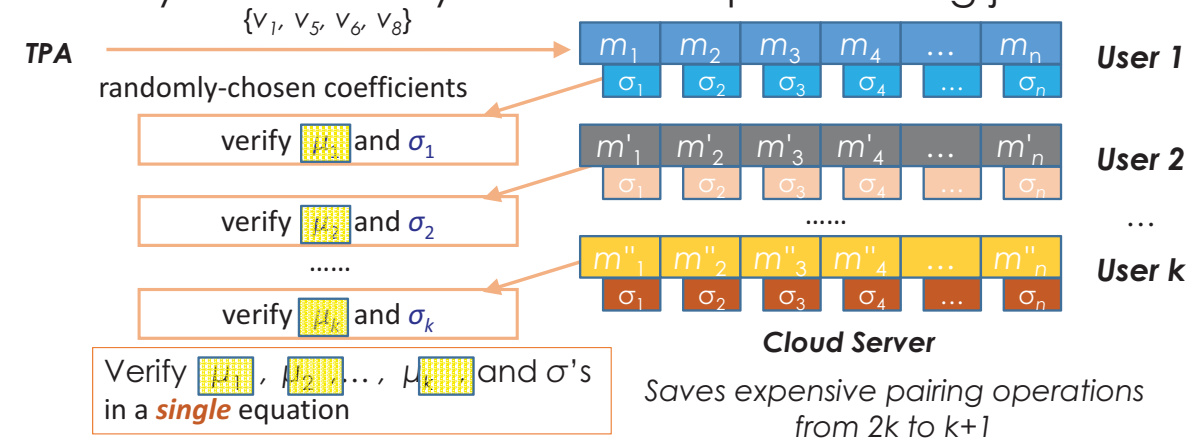
- How to choose c?
- $P = 1 - (1 - t)^c$ for t fraction of corrupted data
- When t = 1%, c = 300 (460) for P = 95% (99%)
- [Ateniese *et al.* CCS 2007]

Storage/Communication Trade-off

- $\sigma_i = [H(W_i)\ u^{mi}]^x$
- 1 block needs 1 authenticator => 2x storage
- Aggregating more in an authenticator!
- Extend the public key to include many u's
- $\sigma_i = [H(W_i)\ \Pi u_j^{m(i,j)}]^x$
- $(1 + 1/s)x$ storage (where $s = |\{u_j\}|$)
- Communication bandwidth becomes O(s)
  - Both the underlying and our hiding technique

Batch Auditing

- Individually auditing can be tedious/inefficient.
- TPA may concurrently handle multiple auditing jobs.

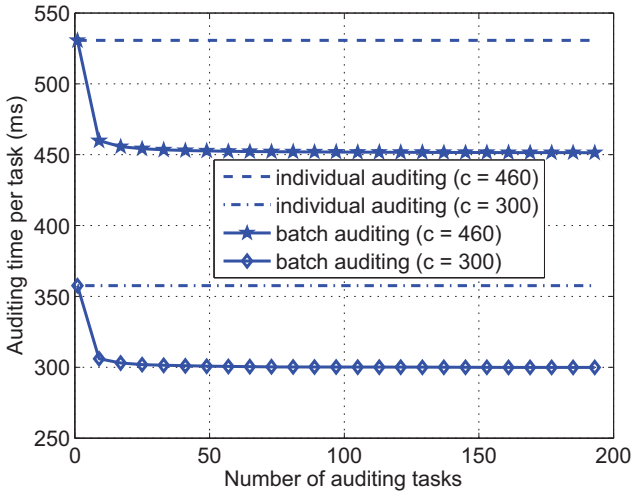


Cost of Privacy

- Intel Core 2 processor 1.86 GHz, 2048M RAM
- Amazon 4-EC2 , 7.5G Ram, 850G storage

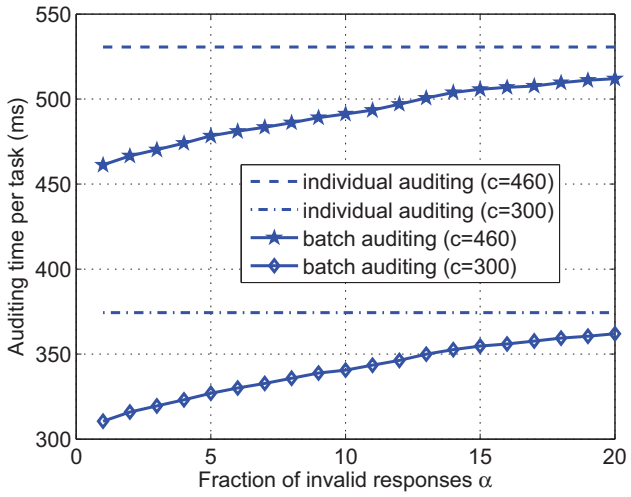
Tradeoff s = 1/10	Our Scheme @ IEEE-TC		Shacham-Waters @ Asiacrypt08	
Sampled blocks c	460	300	460	300
Server time (ms)	335.17 / 361.56	219.27 / 242.29	333.23 / 342.91	217.33 / 223.64
TPA time (ms)	530.60 / 547.39	357.53 / 374.32	526.77 / 543.35	353.70 / 370.29
Comm. (byte)	160 / 1420	160 / 1420	40 / 220	40 / 220

Batch vs. Individual Auditing



Saved > 11% (14%) of per-task auditing time

Sorting out Invalid Response

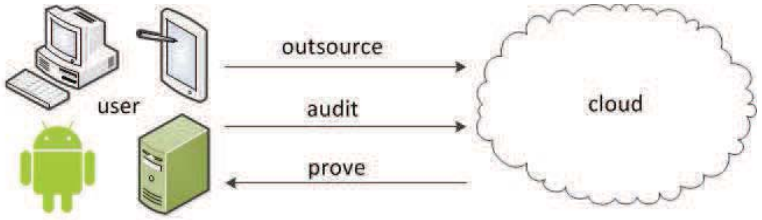


Zero-Knowledge TPA-PoR

- Our protocol only hides (linear combination of) *m*
- The aggregated authenticator σ also leaks *m*
  - Extract an individual authenticator from σ
  - Solve discrete logarithm to recover *m* in exponent
  - Or Brute-force (but offline) attack
- Apply the same technique on σ
- Same asymptotic performance
- "Worry-free" Data outsourcing

More Privacy Issues

- So far we talked about how to let a third-party to audit cloud storage with data privacy
  - can be generalized to other protocols
  - can be extended to handle dynamic data
- Now we will talk about User Privacy for Multi-User Setting



Multi-User Collaborative Environment

- Most existing solutions only focus on personal data
- Not readily extensible to shared data

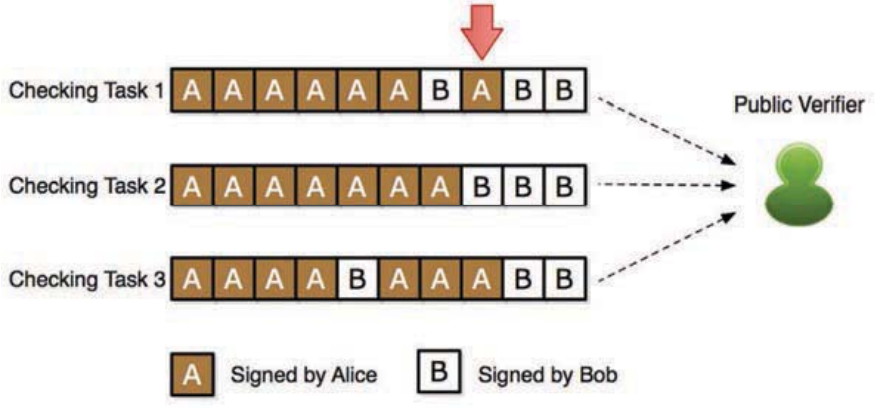
- Collaboratively edit
- Each block is signed by one user
- Once a block is modified, compute a signature for it
- Different blocks are signed by different users

New Privacy Issues

- Reveal private identities to public verifiers
- Need to know which public key should be used for verifying each block
- The binding between a public key and an identity is unique under PKI

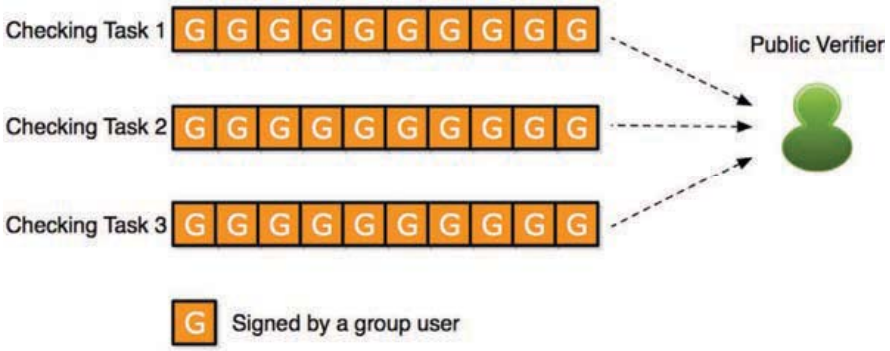
Privacy Leak

- Alice is a key user (always signs more blocks).
- The 8th block is more important (modified by diff. users).



Users want Privacy

- Public verifiers are not able to learn which *group user* or *particular block* is more important than others.
- (Also need techniques like private information retrieval.)



Existing Work

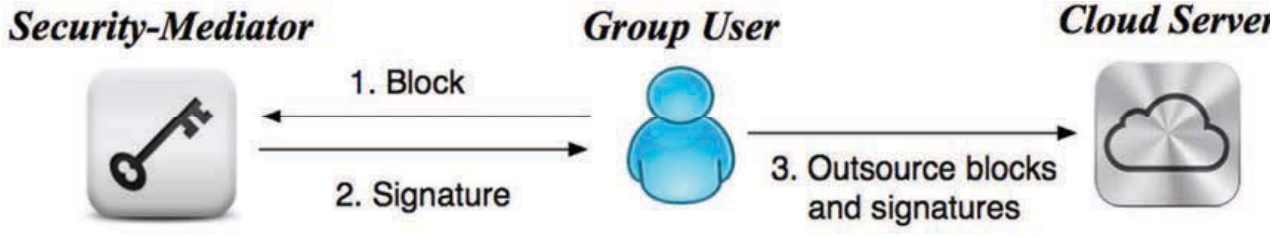
- Significant overhead on storing verification metadata
  - IEEE Cloud 2012 (ring signature based)
  - ACNS 2012 (group signature based)
- Significant overhead on key management and distribution
  - IEEE ICC 2013 (sharing a common key)

Design Goals

- Check the integrity of shared data
- Verify without downloading the entire data
- Preserve identity privacy from public verifiers
- Without introducing significant overhead

Security-Mediator based Approach

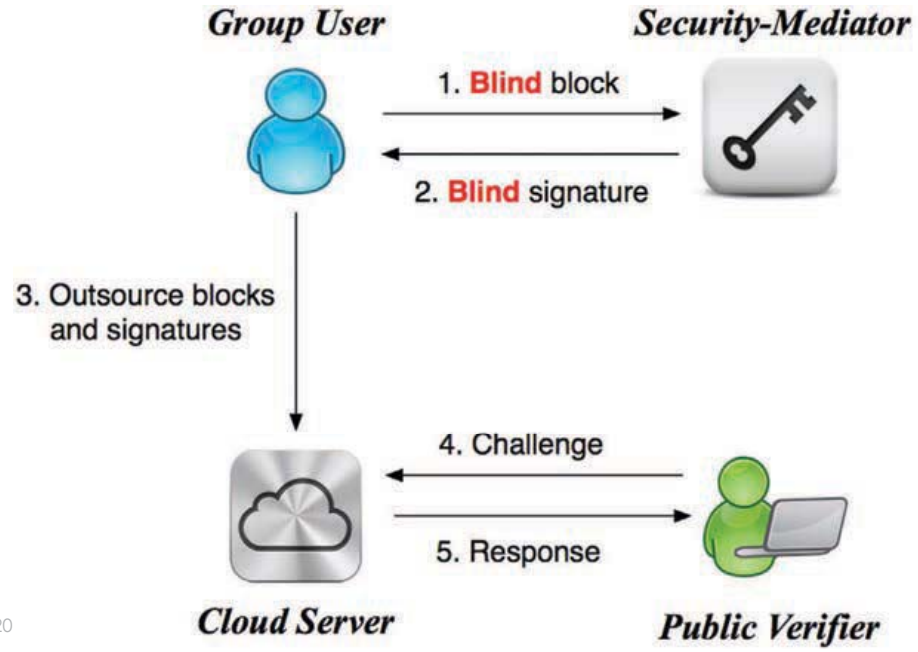
- SEM provides signing services
- Sign blocks for each user before data outsourcing
- The entire shared data are signed by the SEM
- Check data integrity without revealing identities



Challenges and Solutions

- Minimize the trust on this SEM
  - Should not reveal data
  - Should not reveal identities from signatures on cloud
- Utilize Blind Signatures +PDP
  - All the blocks sent to the SEM are blind.
  - Signatures generated by the SEM are blind.
  - Each user recovers "real" signatures from blind ones.
  - A "real" signature and its blind version are unlinkable.

System Model



Comparison with Previous Work

- Signatures can be easily computed
  - Compared to ring/group signatures
- Checking integrity is as the same as PDP
  - No extra storage overhead for achieving anonymity
- Extra communication cost on signature generation
  - Can be reduced by aggregating each block before sending it to the SEM -->1/*k*
  - (Notations in this paper are not the same as the previous one)

Improve Efficiency

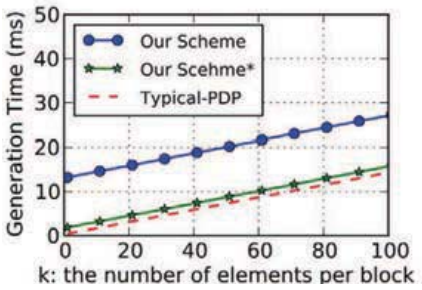
- *k*: the number of elements in a block
  - A trade-off between computation and comm. cost
- Batch verification
  - Verify multiple blind signatures simultaneously before recovering the "real" ones
- Sampling Strategies
  - Randomly selecting a small number of c blocks instead of all the n blocks, where c << *n*.

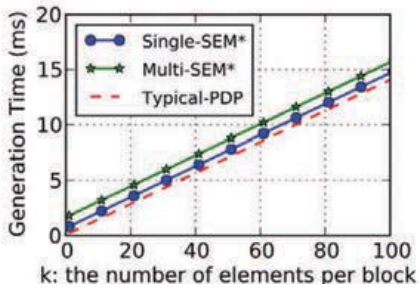
Multi-SEM Model

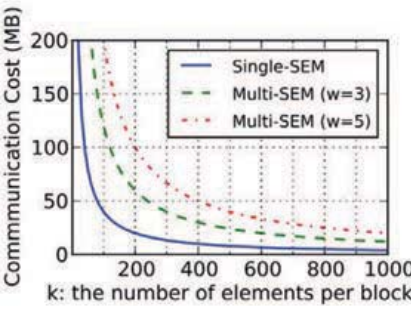
- Single SEM model may cause single-point failures
- Extend our scheme to multi-SEM model
  - Shamir secret sharing
  - As long as the majority of multiple SEMs are secure
  - Get multiple blind signatures on one block from multiple SEMs
  - Recover a "real" signature from multiple blind ones

Performance (Signature Generation)

- Shared Data 2GB, each element 20 Bytes
- *k*: the number of elements per block

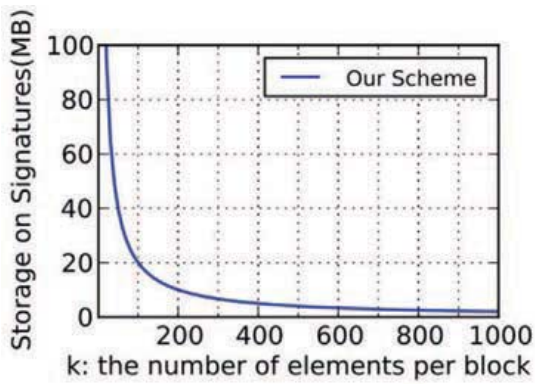






Performance (Storage Cost)

- Shared Data 2GB, each element 20 Bytes
- *k*: the number of elements per block



Trade-off between efficiency and security

- Integrity checking when *k* = 1,000

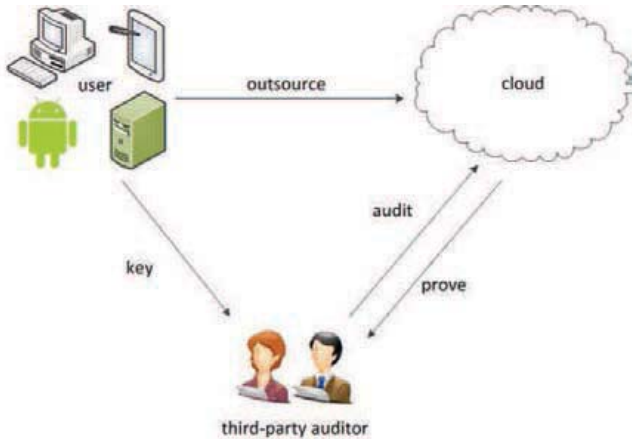
- Shared Data 2GB, each element 20 Bytes

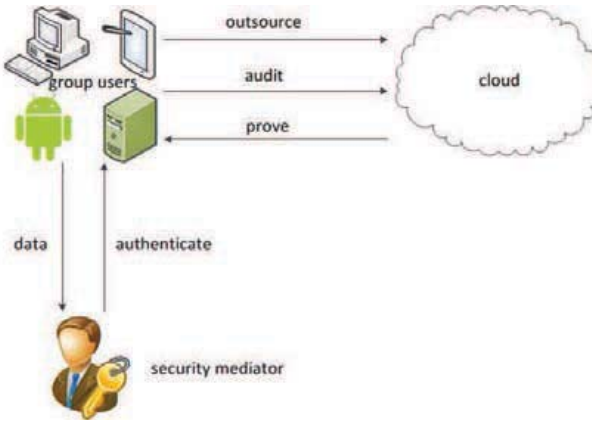
Cost	*n* = 100,000	*c* = 460
Computation	14.15 seconds	0.21 seconds
Communication	2.27MB	30.37KB

Concluding Remarks

- Introduce a Security-Mediator for anonymity
- Minimize the trust on the Security-Mediator
- Check shared data integrity efficiently
- No extra storage overhead compared to PDP

Third Party Auditing & Anonymity from SEM





sherman@ie.cuhk.edu.hk

Decentralized Attribute-Based Cryptosystems

Katsuyuki Takashima

Mitsubishi Electric

Takashima.Katsuyuki@aj.MitsubishiElectric.co.jp

Attribute-based encryption (ABE) [10, 3, 7] gives a new way of sharing encrypted data, in which receiver sets can be specified by an expressive relation (of keys and ciphertexts). We mainly deal with ciphertext-policy ABE (CP-ABE) schemes, where a secret key is associated with a set of attributes and a ciphertext is associated with a (non-monotone span program) access structure. Recently, a versatile and privacy-enhanced variant of digital signatures has been studied, which is called attribute-based signatures (ABS) [6, 9]. Several ABS schemes are converted from CP-ABE schemes, however, the conversions are not straightforward, since no counterpart of a signature in the ABS exists in the CP-ABE, and the privacy property for signature is specific in ABS.

The basic concept of ABE/ABS, however, has a serious problem that only a single authority exists in a system. Therefore, the single authority should issue to all users their secret keys (certificates/credentials) associated with all attributes in the system. If the party is corrupted, the system will be totally broken. To overcome the drawback, the concept of multi-authority ABE [1] and ABS [6] (MA-ABE, MA-ABS), was introduced, in which there are multiple authorities and each authority is responsible for issuing a secret key associated with a category or sub-universe of attributes. The existing MA-ABE (MA-ABS) schemes, however, still have a problem that a special *central* authority is required in addition to multiple authorities regarding attributes, and if the central authority is corrupted, the security of the system will be totally broken. Any MA-ABE (resp. MA-ABS) scheme with no central authority is called *decentralized* MA-ABE (DMA-ABE, resp. DMA-ABS) scheme. Recently, Lewko and Waters [5] presented the first DMA-ABE (based on a composite-order pairing group), and Okamoto and Takashima [8] proposed the first DMA-ABS, in which no central authority and *no trusted setup* are required. We [8] also presented an adaptively secure DMA-ABE scheme based on prime order pairing groups without no trapdoor for initial setup.

Security of both the schemes was proved in the random oracle model, which has been considered as a drawback so far. In the final part of our presentation, we [11] give the DMA-ABE (DMA-ABS) from indistinguishability obfuscation (iO) [2] by applying the techniques in [4] *without random oracles*.

References

[1] M. Chase. Multi-authority attribute based encryption. In *TCC 2007*, pages 515–534, 2007.

[2] S. Garg, C. Gentry, S. Halevi, M. Raykova, A. Sahai, and B. Waters. Candidate indistinguishability obfuscation and functional encryption for all circuits. In *FOCS 2013*, pages 40–49, 2013.

[3] V. Goyal, O. Pandey, A. Sahai, and B. Waters. Attribute-based encryption for fine-grained access control of encrypted data. In *ACM CCS*, pages 89–98, 2006.

[4] S. Hohenberger, A. Sahai, and B. Waters. Replacing a random oracle: Full domain hash from indistinguishability obfuscation. In *EUROCRYPT 2014*, pages 201–220, 2014.

[5] A. B. Lewko and B. Waters. Decentralizing attribute-based encryption. In *EUROCRYPT 2011*, pages 568–588, 2011.
[6] H. K. Maji, M. Prabhakaran, and M. Rosulek. Attribute-based signatures. In *CT-RSA 2011*, pages 376–392, 2011.
[7] T. Okamoto and K. Takashima. Fully secure functional encryption with general relations from the decisional linear assumption. In *CRYPTO 2010*, pages 191–208, 2010. Full version: `http://eprint.iacr.org/2010/563`.
[8] T. Okamoto and K. Takashima. Decentralized attribute-based signatures. In *PKC 2013*, pages 125–142, 2013. Full version: `http://eprint.iacr.org/2011/701`.
[9] T. Okamoto and K. Takashima. Efficient attribute-based signatures for non-monotone predicates in the standard model. *IEEE T. Cloud Computing*, 2(4):409–421, 2014. The preliminary version appeared in the proceedings of PKC 2011. Full version: `http://eprint.iacr.org/2011/700`.
[10] A. Sahai and B. Waters. Fuzzy identity-based encryption. In *EUROCRYPT 2005*, pages 457–473, 2005.
[11] K. Takashima. Decentralized attribute-based cryptosystems from indistinguishability obfuscation. In *SCIS 2014*, pages 3B3–4, 2014. (In Japanese).

Decentralized Attribute-Based Cryptosystems

Kyushu Univ. IMI Workshop

2015 / 9 / 2

Katsuyuki TAKASHIMA (Mitsubishi Electric)

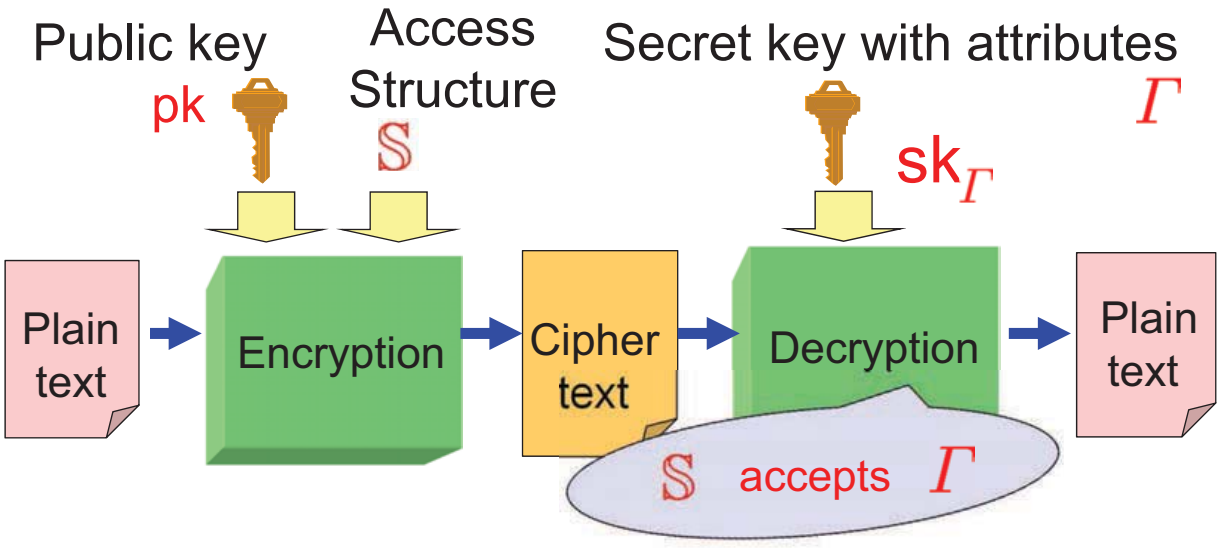
Agenda

- Attribute-Based Encryption (ABE) and Attribute-Based Signatures (ABS)
- Decentralized Multi-Authority (DMA) Ciphertext-Policy (CP-) ABE and ABS
- DMA ABE / ABS without Random Oracle (RO) from indistinguishability Obfuscation (iO)

ABE and ABS

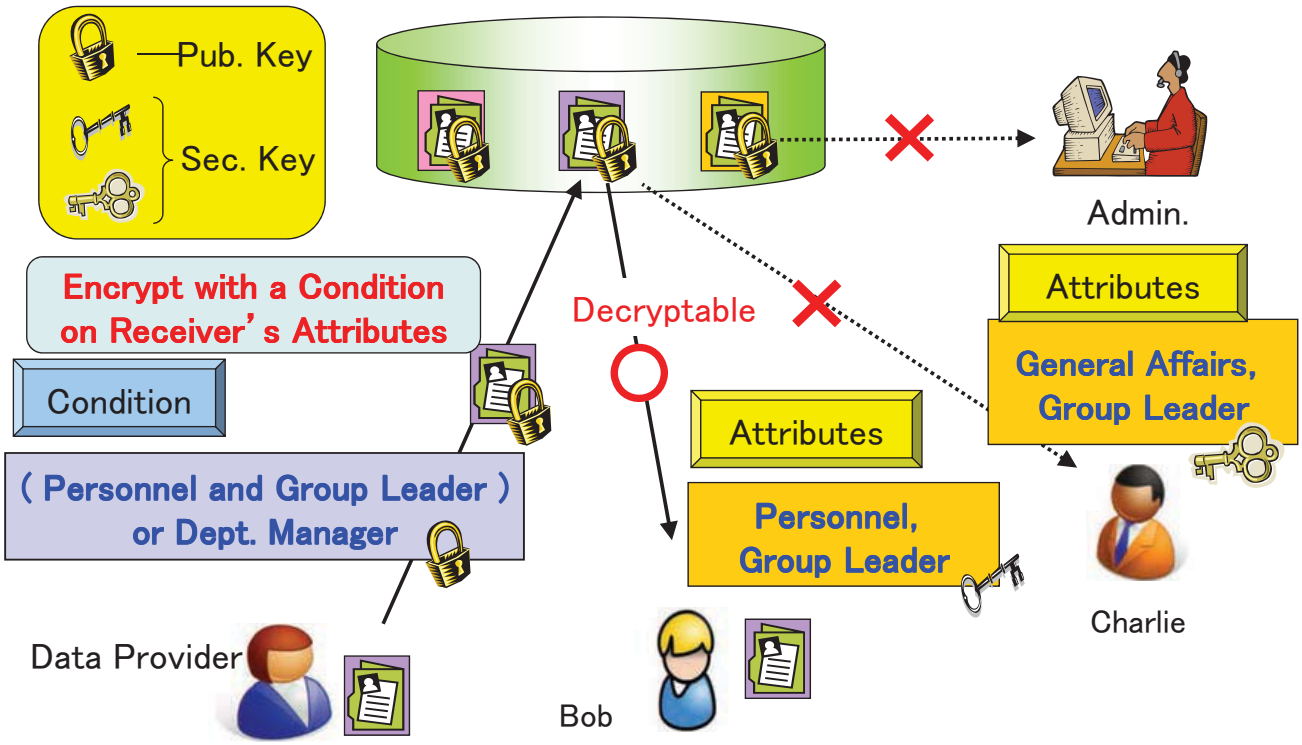
3

Ciphertext-Policy Attribute-Based Encryption (CP-ABE)



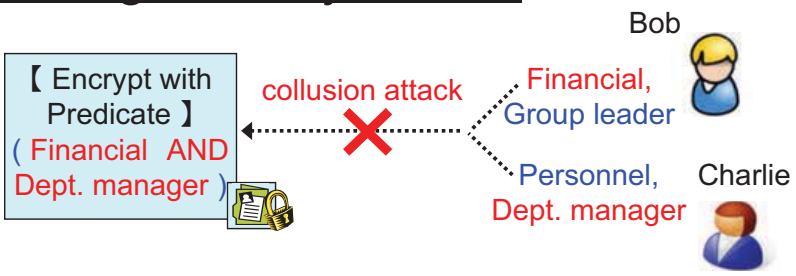
4

Application 1 of ABE: Information-Sharing Platform

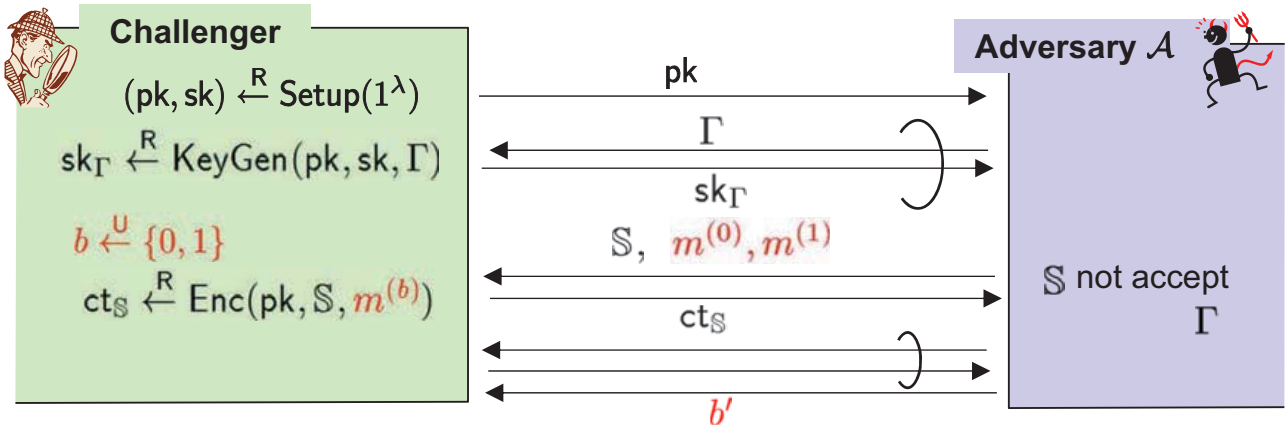


Payload-Hiding Security of ABE

- Key point:
  Collusion-resistance
  = To prevent collusion attack



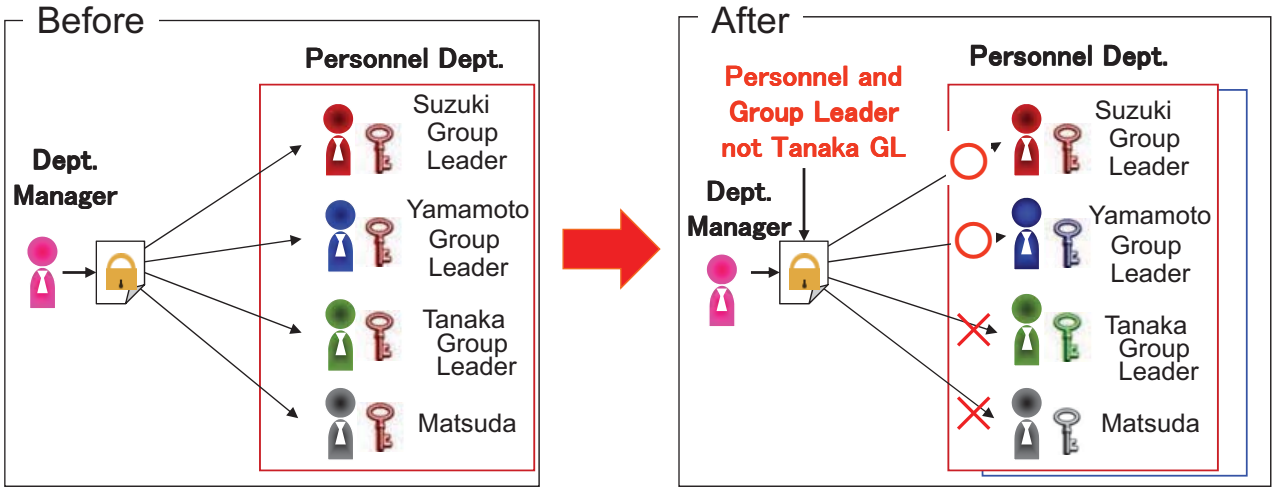
- Game-based Definition



Adversary $\mathcal{A}$ wins if $b = b'$. $\mathrm{Adv}_{\mathcal{A}}^{\mathrm{ABE,PH}} := \Pr[\mathcal{A} \text{ wins}]$.
The ABE scheme is payload-hiding $\overset{\text{def}}{\Leftrightarrow}$ $\mathrm{Adv}_{\mathcal{A}}^{\mathrm{ABE,PH}}$ is negligible for any $\mathcal{A}$

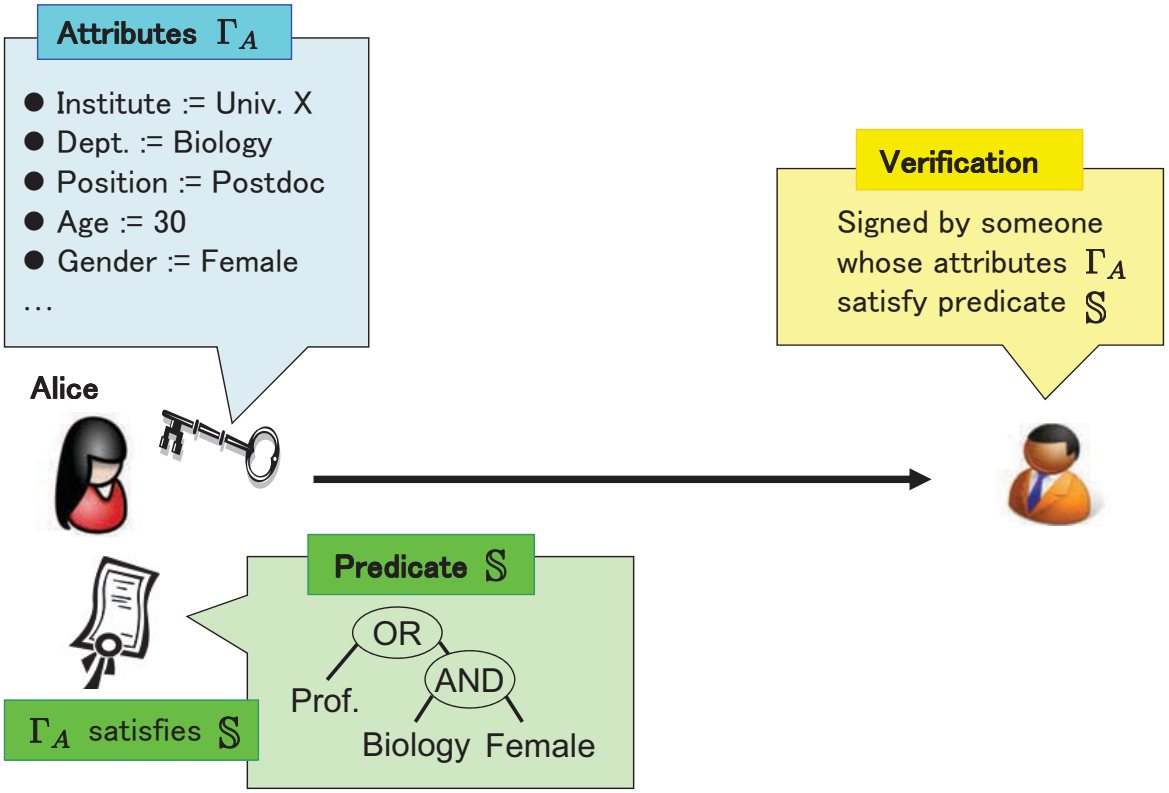
Application 2 of ABE: Fine-Grained Broadcast

Broadcast an encrypted message with fine-grained access control

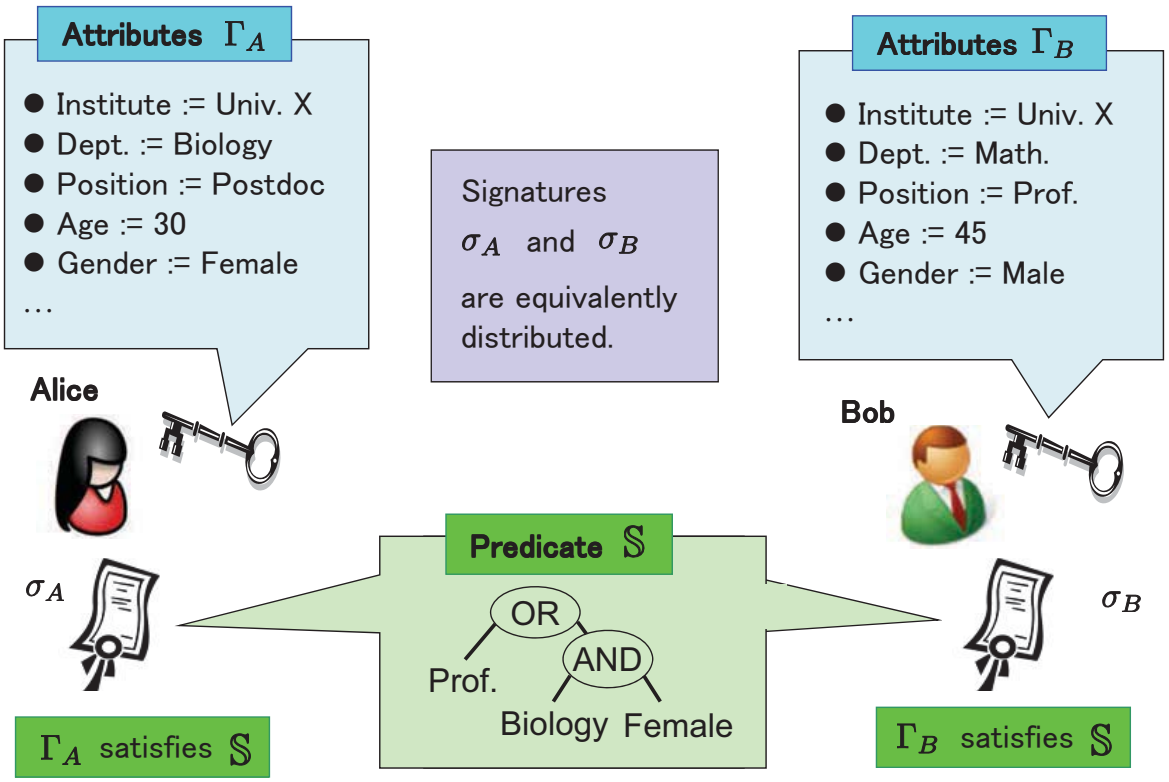


- Group members share the same key
  - ⇒ × No fine grained access control.
  - ⇒ × Re-distribution of all the keys when one key is lost.
- Each member has his own specialized attributes key
  - ⇒ ○ Fine-grained access control with attributes
  - ⇒ ○ Revocation of some persons

Attribute-Based Signatures (ABS)

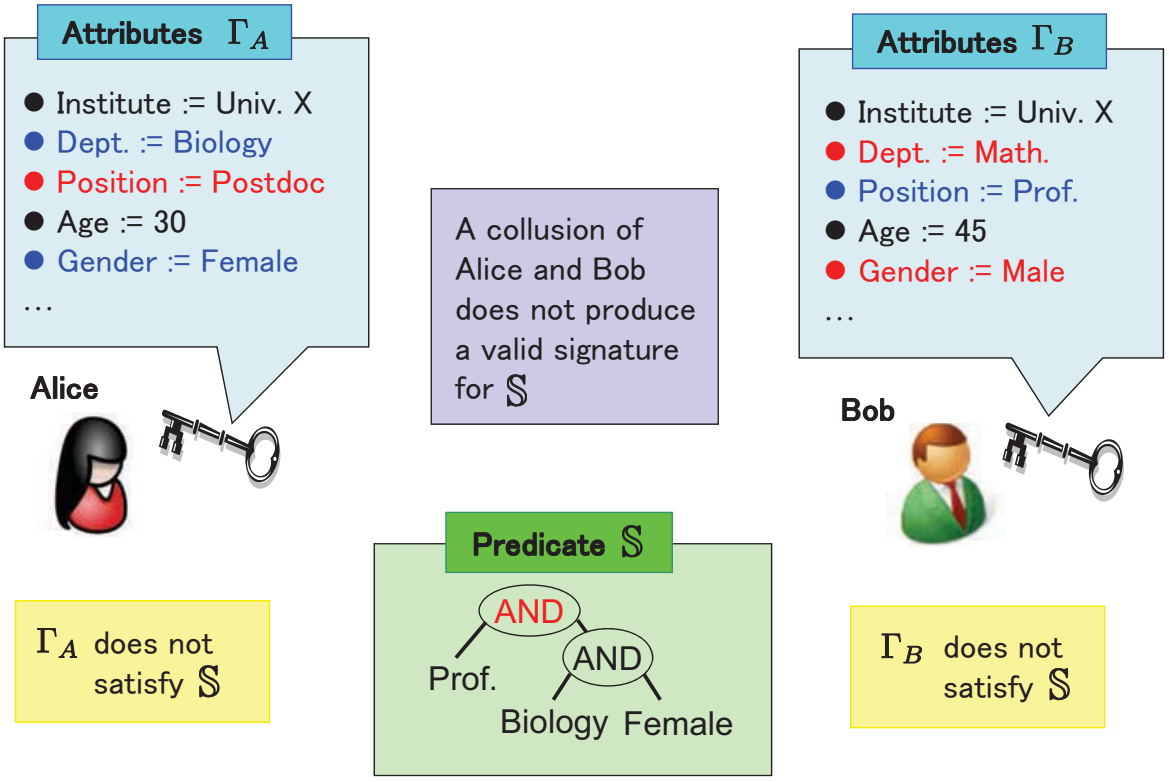


(Perfect) Privacy



9

Unforgeability : Collusion Resistance



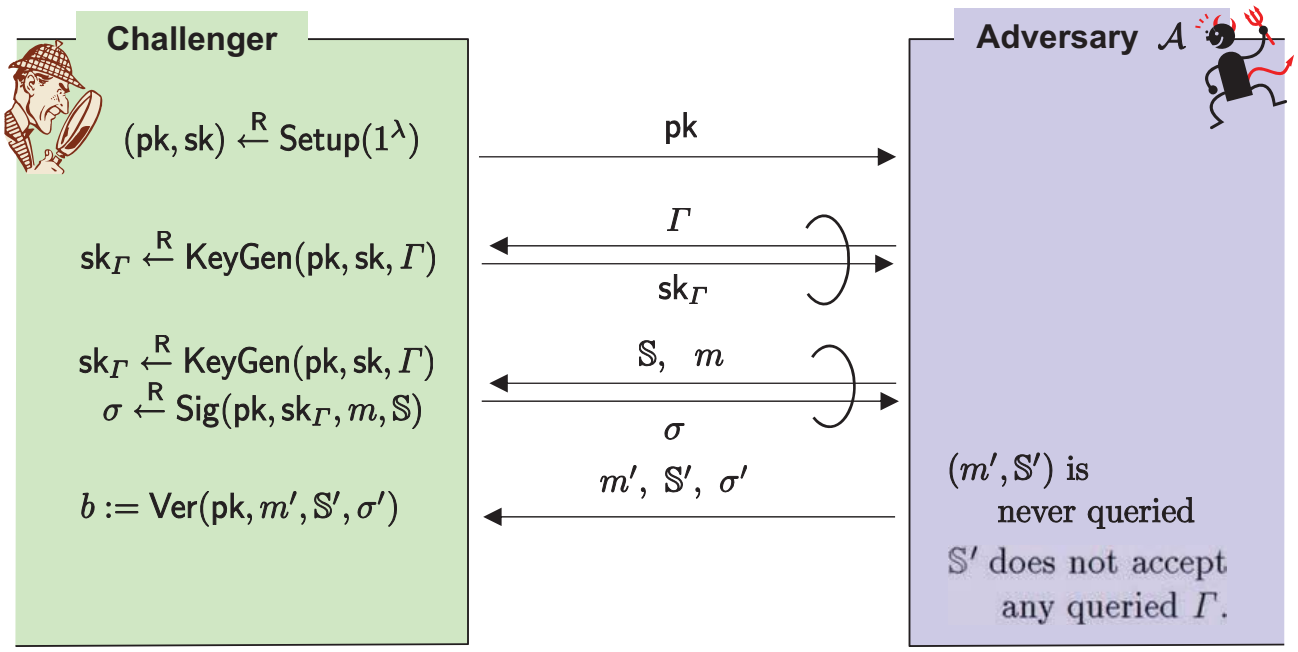
10

Attribute-Based Signatures (ABS)

- Setup: pk: (master) public key, sk: (master) secret key
- KeyGen(pk, sk, Γ): sk_Γ: secret key for Γ
- Sig(pk, sk_Γ, m, $\mathbb{S}$): $\sigma_{\mathbb{S}}$: signature for $\mathbb{S}$
  such that $\mathbb{S}$ accepts Γ
- Ver(pk, m, $\mathbb{S}$, $\sigma_{\mathbb{S}}$): boolean value accept $:= 1$ or reject $:= 0$

11

Adaptive-Predicate Unforgeability of ABS



- Adversary $\mathcal{A}$ wins if $b = 1$.

12

Attribute-Based Signatures (ABS)

- Attribute-Based :
  Flexibility in the relationship between a signing key and a verification key
- Unforgeability :
  Collusion resistance, Adaptive-predicate unforgeability
- Privacy :
  (In this work, we consider prefect privacy.)

We say that an ABS scheme is fully secure
if it satisfies adaptive-predicate unforgeability and privacy.

13

Decentralized Multi-Authority (DMA) CP-ABE and ABS

14

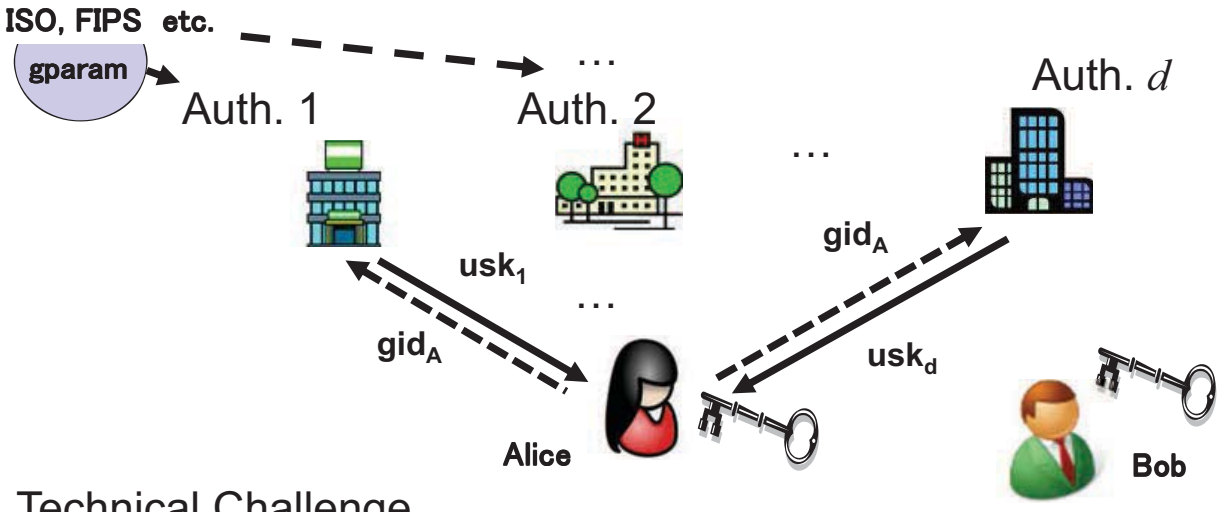
Multi-Authority (MA-) ABE / ABS

- In the basic ABE/ABS, only a single authority exists. It should issue user keys associated with all attributes, and if it is corrupted, the system will be totally broken.
- The concept of multi-authority (MA-) ABE / ABS, was introduced [Cha07, MPR11, OT11].
  The schemes require a special central authority, and if it is corrupted, the system will be totally broken.
- Lewko-Waters [LW11] presented the first DMA system for ABE (but not for ABS). It requires a trusted setup of a parameter, composite number, $N := p_1 p_2 p_3$ since if trapdoor (p_1, p_2, p_3) is compromised, the system is not secure.
- Okamoto-T [OT13] presented the first DMA-ABS with no central authority. It is constructed based on prime order pairing groups without no trapdoor for initial setup.

15

Decentralized Multi-Authority ABE/ABS (DMA-ABE/ABS)

- gparam for a prime order bilinear group & hash function are publicly available from ISO and FIPS.



Technical Challenge

- No central authority exists and no global coordination is required except for the above parameters. No trusted setup is necessary.
- Collusion Resistance

16

Decentralized Multi-Authority (DMA) CP-ABE

- GSetup: gparam: global parameter
- ASetup: apk_t: authority public key, ask_t: authority secret key

- $\mathsf{AttrGen}(\mathsf{gparam}, t, \mathsf{ask}_t, \mathsf{gid}, x_t)$: $\mathsf{usk}_{\mathsf{gid},(t,x_t)}$: secret key for x_t
- $\mathsf{Enc}(\ \mathsf{gparam},\ \mathsf{apk}_t,\ m,\ \mathbb{S} := (M, \rho)\ )$ and
- $\mathsf{Dec}(\ \mathsf{gparam},\ \{\mathsf{apk}_t, \mathsf{usk}_{\mathsf{gid},(t,x_t)}\},\ \mathsf{ct}_{\mathbb{S}}\ )$ are essentially the same as those of single-authority CP-ABE ([OT10]).

17

Our Results [OT13]

- We proposed the first DMA-ABS scheme for a wide class of predicates, non-monotone span programs.
  - ➢ No central authority exists, and no trusted setup is necessary.
  - ➢ Fully secure under the decisional linear (DLIN) assumption in the random oracle model
- The efficiency is comparable to those of the existing ABS schemes [MPR11, OT11].
- As a by-product, we also presented a new DMA-ABE scheme, which is adaptively secure without a trusted setup under the DLIN assumption in the random oracle model.

18

Key Techniques

- To realize DMA-ABE / ABS schemes, the top level strategy is employing dual system encryption developed by Waters for achieving full security of functional encryption.

- To construct DMA-ABE / ABS schemes, we follow several established key ideas: (random oracle) hashing of global identifier gid, etc. Moreover, we employ

  - Use of a new trapdoor, authority secret key X_t , (on dual pairing vector spaces : DPVS)
  - Distribution of a crucial role of the central space $(t = 0)$ to all spaces

    = Distributed form of dual system encryption

Non-monotone Span Programs on Attributes

Attributes Γ

- Institute := Univ. X
- Dept. := Biology
- Position := Postdoc
- Age := 30
- Gender := Female

…

Attributes $\Gamma := \{(t, x_t)\}$

t : Name index (Subuniverse index)	x_t : Attribute Value
1	x_1
⋮	⋮
d	x_d

Predicate (Span Program) $\mathbb{S} := (M, \rho)$

$\rho(i) = (t, v_i)$ or $\neg(t, v_i)$

(t_1, v_1)

$\neg(t_2, v_2)$

$\neg(t_\ell, v_\ell)$

M (r columns, ℓ rows)

M_1

M_2

⋮

M_ℓ

$\mathbb{S}$ accepts Γ ⇔

There exists $I \subset \{1, \ldots, \ell\}$ and $\{\alpha_i\}_{i \in I}$ such that

$\vec{1} = (1, \ldots, 1) = \sum_{i \in I} \alpha_i M_i$ and

$I \subseteq \{\, i \mid [\rho(i) = (t, v_i) \wedge (t, x_t) \in \Gamma \wedge v_i = x_t]$

$\vee \; [\rho(i) = \neg(t, v_i) \wedge (t, x_t) \in \Gamma \wedge v_i \neq x_t] \,\}$

Secret-sharing Scheme for Span Program

$$M \in \begin{array}{c} \overbrace{\qquad\qquad}^{r} \\ \left.\begin{array}{|c|}\hline M_1 \\ \hline M_2 \\ \hline M_3 \\ \hline \\ \hline \vdots \\ \hline M_\ell \\ \hline \end{array}\right. \end{array}$$

$$\begin{pmatrix} s_1 \\ \vdots \\ s_\ell \end{pmatrix} := M \begin{pmatrix} f_1 \\ \vdots \\ f_r \end{pmatrix}$$

$$\vec{f}^{\mathrm{T}} := \begin{pmatrix} f_1 \\ \vdots \\ f_r \end{pmatrix} \xleftarrow{\cup} \mathbb{F}_q^r$$

$$\vec{1} := (1, \ldots, 1)$$

$s_0 := \vec{1} \cdot \vec{f}^{\mathrm{T}}$: secret to be shared

$s_i := M_i \cdot \vec{f}^{\mathrm{T}}$ $(i = 1, \ldots, \ell)$: the share belonging to i

$\mathbb{S}$ accepts Γ $\Leftrightarrow$ $\exists\, I$ and $\{\alpha_i\}_{i \in I}$ s. t. $\vec{1} = \sum_{i \in I} \alpha_i M_i$ and ... $\Rightarrow$ $\vec{1} \cdot \vec{f}^{\mathrm{T}} = \sum_{i \in I} \alpha_i M_i \cdot \vec{f}^{\mathrm{T}}$, $s_0 = \sum_{i \in I} \alpha_i s_i$

CP-ABE in [OT10]

$\mathbb{S} := (M, \rho)$, $\vec{f} \xleftarrow{\cup} \mathbb{F}_q^r$ $\Rightarrow$ $s_0 := \vec{1} \cdot \vec{f}$, $s_i := M_i \cdot \vec{f}$: share of user I for s_0

Ciphertext

c_0 $(-s_0, \zeta)$ c_1 (v_1, s_1) c_2 (v_2, s_2) ... c_ℓ (v_ℓ, s_ℓ) $c_T = g_T^{\zeta} \cdot m$

matching

Key

k_0^* (δ) $k_{t_1}^*$ (x_{t_1}, δ) ... $k_{t_\ell}^*$ (x_{t_ℓ}, δ)

$\vec{1} \cdot \vec{f} = \left(\sum_{i \in I} \alpha_i M_i\right) \cdot \vec{f}$ $\Rightarrow$ $s_0 = \sum_{i \in I} \alpha_i s_i$ $\Rightarrow$ retrieve g_T^{ζ} $\Rightarrow$ retrieve m

Key Ideas for DMA-ABE Construction (I): Use of a Trapdoor Matrix X_t

KeyGen of ABE

Setup: G ← X_t (Basis Matrix)

$\mathbb{B}_t^* := G \cdot X_t$

KeyGen: $\delta, x_t, \ldots$ ← Common Randomness

$\boldsymbol{k}_t^* := (\delta(1, x_t), \ldots)_{\mathbb{B}_t^*}$

Distributed AttrGen of DMA-ABE

$\delta_{\mathsf{gid}} \cdot G := Hash(\mathsf{gid})$

Common Randomness: ← $x_t, \ldots$

$(\delta_{\mathsf{gid}} G, x_t(\delta_{\mathsf{gid}} G), \ldots)$

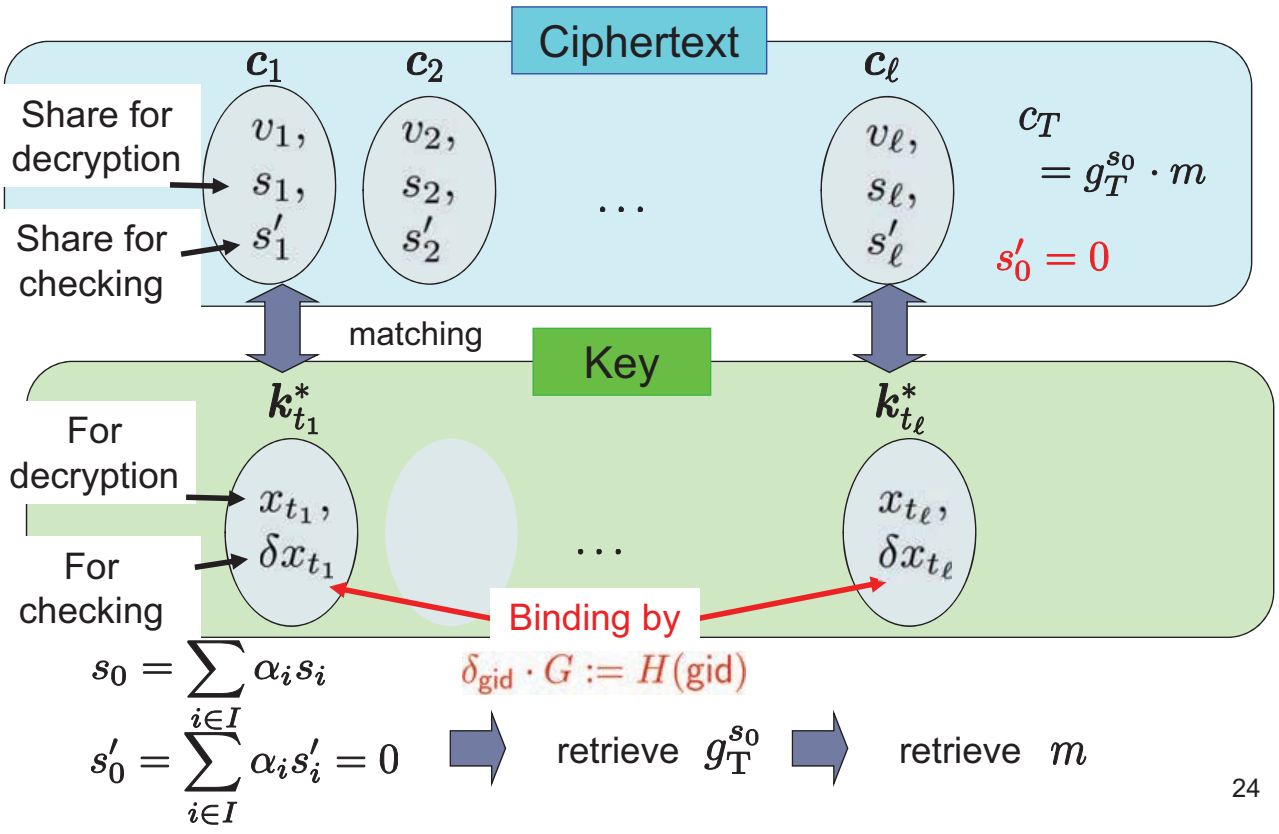
← X_t (Basis Matrix)

$\boldsymbol{k}_t^* := (\delta_{\mathsf{gid}} G, x_t(\delta_{\mathsf{gid}} G), \ldots) X_t$

$= (\delta_{\mathsf{gid}}(1, x_t), \ldots)_{\mathbb{B}_t^*}$

23

Key Ideas for DMA-ABE Construction (II): Distribution of t=0 component



24

DMA CP-ABE Scheme

- GSetup : $\mathsf{gparam} := (\mathsf{param}_{\mathbb{G}},\ H)$
- ASetup : $X_t \xleftarrow{\cup} GL(11, \mathbb{F}_q)$,

  $\widehat{\mathbb{B}}_t := (\boldsymbol{b}_{t,1}, \ldots, \boldsymbol{b}_{t,4}, \boldsymbol{b}_{t,11})$, $\quad \mathsf{apk}_t := (\mathsf{param}_{\mathbb{V}_t}, \widehat{\mathbb{B}}_t)$, $\quad \mathsf{ask}_t := X_t$
- AttrGen(gparam, t, ask_t, gid, x_t) : $\delta_{\mathsf{gid}}\, G_1 := H(\mathsf{gid}) \in \mathbb{G}$,

  By using $\delta_{\mathsf{gid}}\, G_1, X_t$, $\quad \boldsymbol{k}^*_t = (\overbrace{1,\ x_t}^{2}, \overbrace{\delta_{\mathsf{gid}}(1,\ x_t)}^{2}, \overbrace{0^4}^{4}, \overbrace{\varphi_{t,1},\ \varphi_{t,2}}^{2},\ 0\ )_{\mathbb{B}^*_t}$.

- Enc(gparam, $\{\mathsf{apk}_t\}$, m, $\mathbb{S} := (M, \rho)$) :

for $i = 1, \ldots, \ell$,

if $\rho(i) = (t, v_i)$, $\quad \boldsymbol{c}_i := (\overbrace{s_i + \theta_i v_i, -\theta_i}^{2}, \overbrace{s'_i + \theta'_i v_i, -\theta'_i}^{2}, \overbrace{0^4}^{4}, \overbrace{0^2}^{2}, \overbrace{\eta_i}^{1})_{\mathbb{B}_t}$,

if $\rho(i) = \neg(t, v_i)$, $\quad \boldsymbol{c}_i := (\overbrace{s_i(v_i, -1)}^{2}, \overbrace{s'_i(v_i, -1)}^{2}, \overbrace{0^4}^{4}, \overbrace{0^2}^{2}, \overbrace{\eta_i}^{1})_{\mathbb{B}_t}$,

$c_{d+1} := g_T^{s_0} m$, where $g_T := e(\boldsymbol{b}_i, \boldsymbol{b}^*_i)$

Summary and Further Problem

- We presented an adaptively secure DMA-ABE scheme, in which no central authority and no trusted setup required (secure in the random oracle model).

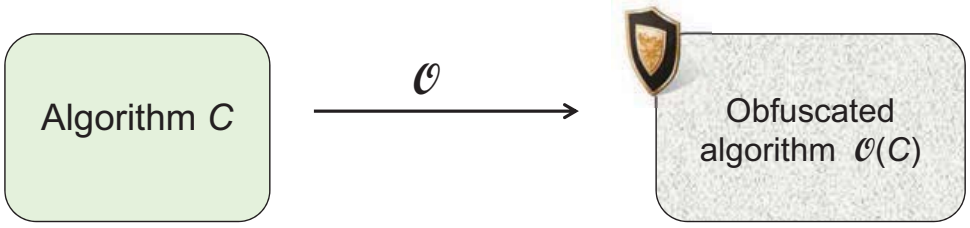
  The first DMA-ABS scheme with no trusted setup was also presented in [OT13].

One of the most important remaining problems is to construct a DMA-ABE (and DMA-ABS) scheme in the standard model (without random oracles).

DMA ABE / ABS without RO from iO

27

(Theoretical) Obfuscation



- Correctness: $\mathcal{O}(C)$ computes C exactly
- Efficiency: $\mathcal{O}(C)$ is at most polynomially larger than C
- Security: $\mathcal{O}(C)$ is "unintelligible"
  - ➢ Virtual black box (VBB) security :

    A cannot do much more with $\mathcal{O}(C)$ than running it on various inputs !

    $$\forall A\ \exists S\ \forall C \qquad A(\mathcal{O}(C)) \approx_{\text{ppt}} S^{C}(1^{|C|})$$

✓ **There exist functions for which VBB obfuscation is impossible** (Barak et al. 01)

28

Indistinguishability Obfuscation (iO)

- Indistinguishability Obfuscation (iO) for poly size circuits is constructed from computational assumptions (existence of Mmap, FHE), weaker security than VBB obfuscation.

• Definition **[iO]**

If C_1 and C_2 compute the same function and $|C_1| = |C_2|$,

$$i\mathcal{O}(C_1) \approx_{\mathrm{ppt}} i\mathcal{O}(C_2)$$ (indistinguishable by ppt TM)

• Inefficient iO is always possible（ canonicalization ）

$i\mathcal{O}(C) :=$ lexicographically 1st circuit computing the same function as C

C_1 and C_2 computes the same function & $|C_1| = |C_2|$ ⇨ $i\mathcal{O}(C_1) = i\mathcal{O}(C_2)$

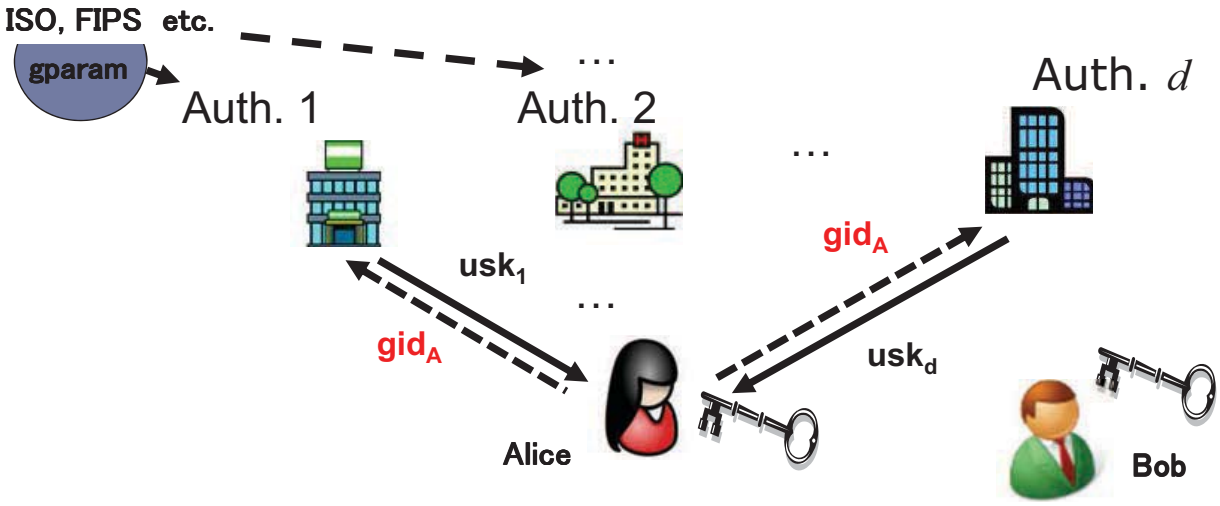
Applications of iO to cryptography

- Functional encryption
- Witness encryption
- Round-efficient multi-party computation
- Deniable encryption
- Broadcast encryption, Traitor tracing
- Multi-input functional encryption
- Functional encryption for randomized functionalities
- Random-oracle free construction (FDH sig.) [HSW13, …]
- Polynomial number of hardcore bits for one-way functions

….

Decentralized Multi-Authority ABE/ABS (DMA-ABE/ABS)

- gparam for a prime order bilinear group & hash function are publicly available from ISO and FIPS.



Previous DMA-ABE/ABS [LW11, OT13] can be proved only in the Random Oracle Model（ROM）. It was an important open problem to remove ROM !

Our Approach and Results

Proposals of DMA-ABE / DMA-ABS from iO without ROM

- Polynomial number of gid hash applications
  Selectively payload-hiding DMA-ABE / DMA-ABS

  Assump： existence of iO, DLIN (SXDH) assump., existence of puncturable PRF

- Constant number of gid hash applications
  Adaptively payload-hiding DMA-ABE / DMA-ABS

  Assump： existence of iO, SXDH assump., n-DDHI type assump.

✓ While an access structure of CP-ABE is given by a span program, the proposed technique can be applied to general circuit CP-ABE.

Hash function constructed from iO

KeyGen of ABE

Setup: G, X_t (Basis Matrix) $\rightarrow \mathbb{B}_t^* := G \cdot X_t$

KeyGen: $\delta, x_t, \ldots$ (Common Randomness) $\rightarrow \boldsymbol{k}_t^* := (\delta(1, x_t), \ldots)_{\mathbb{B}_t^*}$

Distributed AttrGen of DMA-ABE

$\delta_{\mathsf{gid}} \cdot G := H(\mathrm{gid})$

Common Randomness: $x_t, \ldots$ $\rightarrow (\delta_{\mathsf{gid}} G, x_t(\delta_{\mathsf{gid}} G), \ldots)$

X_t (Basis Matrix)

$\boldsymbol{k}_t^* := (\delta_{\mathsf{gid}} G, x_t(\delta_{\mathsf{gid}} G), \ldots) X_t$
$= (\delta_{\mathsf{gid}}(1, x_t), \ldots)_{\mathbb{B}_t^*}$

While H is a random oracle in [OT13],
new $H := i\mathcal{O}(\mathcal{H})$ is constructed from iO

Puncturable Pseudorandom Function [BW13,..]

- **Def [Puncturable PRF]**

$$F := (\mathsf{Key}_F, \mathsf{Punc}_F, \mathsf{Eval}_F) \quad \text{3 algorithms}$$

➢ Functionality preservation under puncturing

For any $S \subseteq \{0,1\}^{n(\lambda)} (\xleftarrow{R} \mathcal{A}(1^\lambda))$, any $x \in \{0,1\}^{n(\lambda)} \setminus S$

If $K \xleftarrow{R} \mathsf{Key}_F(1^\lambda), K_S \xleftarrow{R} \mathsf{Punc}_F(K, S), \quad \mathsf{Eval}_F(K, x) = \mathsf{Eval}_F(K_S, x).$

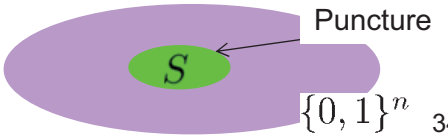
➢ Pseudorandomness at punctured points

For any adversary $\mathcal{A}$, any $K \xleftarrow{R} \mathsf{Key}_F(1^\lambda), \quad K_S \xleftarrow{R} \mathsf{Punc}_F(K, S)$

$$|\Pr[\mathcal{A}(K_S, S, \mathsf{Eval}_F(K, S)) = 1] - \Pr[\mathcal{A}(K_S, S, U_{m(\lambda \cdot |S|)}) = 1]|$$

is negligible.

- $\mathsf{Eval}_F(K, S) := \{\mathsf{Eval}_F(K, x) \mid x \in S\}$
- U_ℓ : ℓ bit uniform sequence



Selectively-Secure DMA CP-ABE from iO

- GSetup : $\mathsf{gparam} := (\mathsf{param}_{\mathbb{G}}, H)$

  with puncturable $H := iO(\mathcal{H}) : \{0,1\}^* \to \mathbb{G}$

- ASetup : $X_t \xleftarrow{\cup} GL(11, \mathbb{F}_q)$,

  $\widehat{\mathbb{B}}_t := (\boldsymbol{b}_{t,1}, \ldots, \boldsymbol{b}_{t,4}, \boldsymbol{b}_{t,11})$, $\quad \mathsf{apk}_t := (\mathsf{param}_{\mathbb{V}_t}, \widehat{\mathbb{B}}_t)$, $\quad \mathsf{ask}_t := X_t$

- AttrGen(gparam, t, ask_t, gid, x_t) : $\delta_{\mathsf{gid}} G_1 := H(\mathsf{gid}) \in \mathbb{G}$,

  By using $\delta_{\mathsf{gid}} G_1, X_t$, $\quad \boldsymbol{k}^*_t = (\overbrace{1,\ x_t}^{2},\ \overbrace{\delta_{\mathsf{gid}}(1,\ x_t)}^{2},\ \overbrace{0^4}^{4},\ \overbrace{\varphi_{t,1},\ \varphi_{t,2},\ 0}^{2})_{\mathbb{B}^*_t}$.

- Enc(gparam, $\{\mathsf{apk}_t\}$, m, $\mathbb{S} := (M, \rho)$) :

for $i = 1, \ldots, \ell$,

if $\rho(i) = (t, v_i)$, $\quad \boldsymbol{c}_i := (\overbrace{s_i + \theta_i v_i, -\theta_i}^{2},\ \overbrace{s'_i + \theta'_i v_i, -\theta'_i}^{2},\ \overbrace{0^4}^{4},\ \overbrace{0^2}^{2},\ \overbrace{\eta_i}^{1})_{\mathbb{B}_t}$,

if $\rho(i) = \neg(t, v_i)$, $\quad \boldsymbol{c}_i := (\overbrace{s_i(v_i, -1)}^{2},\ \overbrace{s'_i(v_i, -1)}^{2},\ \overbrace{0^4}^{4},\ \overbrace{0^2}^{2},\ \overbrace{\eta_i}^{1})_{\mathbb{B}_t}$,

$c_{d+1} := g_T^{s_0} m$, where $g_T := e(\boldsymbol{b}_i, \boldsymbol{b}^*_i)$

Partitioning Security Proof for iO Based DMA-ABE

- $H := iO(\mathcal{H}) : \{0,1\}^* \to \mathbb{G}$
  - Pre-obfuscation hash with punc. PRF $F\ (= \mathsf{eval}_F) : \{0,1\}^* \to \mathbb{F}_q$

    $\mathcal{H}(\mathsf{gid}) := F(K, \mathsf{gid}) \cdot G$ $\quad$ with $\quad K \xleftarrow{R} \mathsf{Key}_F(1^\lambda)$

- Summary of (selective-gid) security proof

1. Change of $\mathcal{H}$ using iO security

   Punctures are given by used gid: $S := \{\mathsf{gid}_i\}_{i=1,\ldots,\nu}$

   $\mathcal{H}^*(\mathsf{gid})$: If $\mathsf{gid} = \mathsf{gid}_i \in S$, output $T_i := F(K, \mathsf{gid}_i) \cdot G$

   Otherwise, output $F(K_S, \mathsf{gid}) \cdot G$

   Functionality preservation under puncturing

2. Change of $\mathcal{H}^*$ using pseudorandomness at punctured points of F

   $\mathcal{H}^{**}(\mathsf{gid})$: If $\mathsf{gid} = \mathsf{gid}_i \in S$, output $T_i := \delta_{\mathsf{gid}_i} G$ with $\delta_{\mathsf{gid}_i} \xleftarrow{\cup} \mathbb{F}_q$

   Otherwise, output $F(K_S, \mathsf{gid}) \cdot G$

3. Since $\delta_{\mathsf{gid}_i} \xleftarrow{\cup} \mathbb{F}_q$, we can apply the DLIN assumption and finish the proof.

- We gave an outline of
  - ✓ (CP-)ABE and ABS
  - ✓ Decentralized MA-ABE and MA-ABS

    [OT13] T. Okamoto and K. Takashima,
    Decentralized attribute-based signatures.
    In PKC 2013, ePrint: http://eprint.iacr.org/2011/701.
  - ✓ DMA-ABE / DMA-ABS without ROM

    [Tak14] K. Takashima,
    Decentralized attribute-based cryptosystems
    from indistinguishability obfuscation.
    In SCIS 2014

Thank you !

37

Dynamic Threshold Public-Key Encryption with Decryption Consistency from Static Assumptions

Goichiro Hanaoka (Joint work with Yusuke Sakai, Keita Emura, Jacob C.N. Schuldt, and Kazuo Ohta)

AIST

hanaoka-goichiro@aist.go.jp

Dynamic threshold public-key encryption (dynamic TPKE) is a natural extension of ordinary TPKE which allows decryption servers to join the system dynamically after the system is set up, and allows the sender to dynamically choose the authorized set and the decryption threshold at the time of encryption. Currently, the only known dynamic TPKE scheme is a scheme proposed by Delerablee and Pointcheval [3]. This scheme is proven to provide message confidentiality under a q-type assumption, but to achieve decryption consistency, a random oracle extension is required.

In this paper we show conceptually simple methods for constructing dynamic TPKE schemes with decryption consistency from only static assumptions (e.g., the decisional linear assumption in bilinear groups) without relying on random oracles. Our first construction is a purely generic construction from public-key encryption with non-interactive opening (PKENO) formalized by Damgard et al. [2]. However, this construction achieves a slightly weaker notion of decryption consistency compared to the random oracle extension of the Delerablee and Pointcheval scheme, which satisfies the notion defined by Boneh, Boyen and Halevi [1]. Our second construction uses a specific PKENO scheme based on the decisional linear assumption in combination with the efficient zero-knowledge proofs by Groth and Sahai. In contrast to our first construction, our second construction achieves the stronger notion of decryption consistency defined by Boneh, Boyen and Halevi.

References

[1] D. Boneh, X. Boyen, and S. Halevi. Chosen Ciphertext Secure Public Key Threshold Encryption Without Random Oracles. CT-RSA 2006, pp.226-243.

[2] I. Damgrd, D. Hofheinz, E. Kiltz, and R. Thorbek. Public-Key Encryption with Non-interactive Opening. CT-RSA 2008, pp.239-255.

[3] C. Delerable and D. Pointcheval. Dynamic Threshold Public-Key Encryption. CRYPTO 2008, pp.317-334.

Dynamic Threshold Public-key Encryption with Decryption Consistency from Static Assumptions

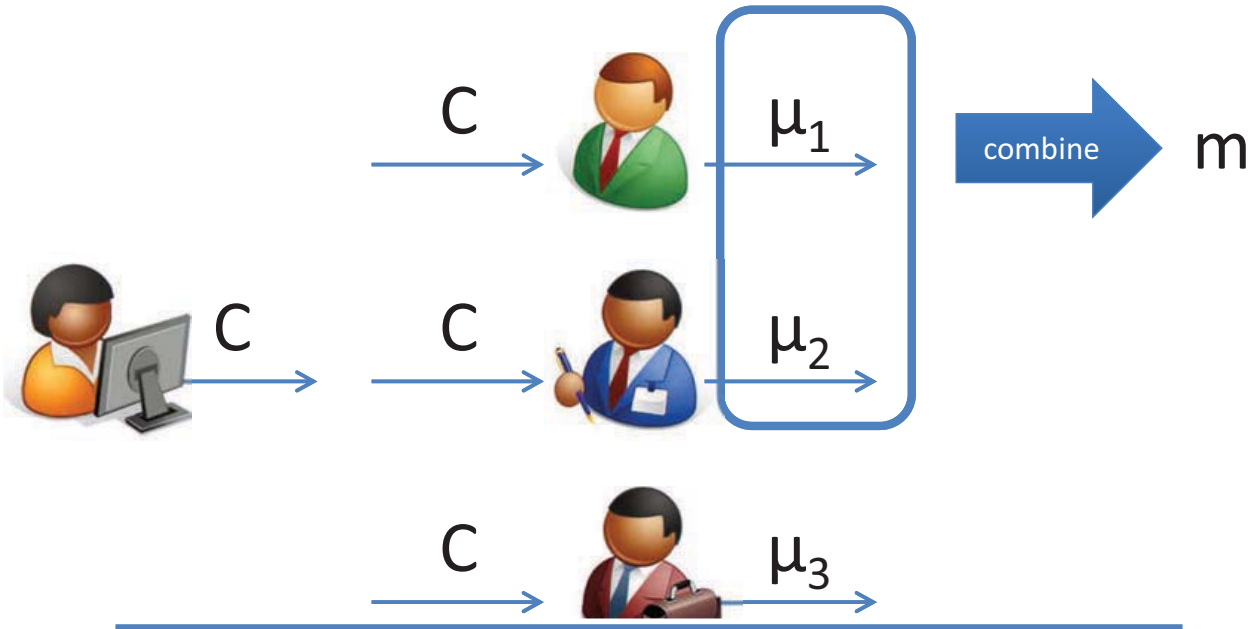
Yusuke Sakai[1], Keita Emura[2], Jacob C.N. Schuldt[1], Goichiro Hanaoka[1] and Kazuo Ohta[3]

1. National Institute of Advanced Industrial Science and Technology (AIST)
2. National Institute of Information and Communications Technology (NICT)
3. The University of Electro-Communications

Contents

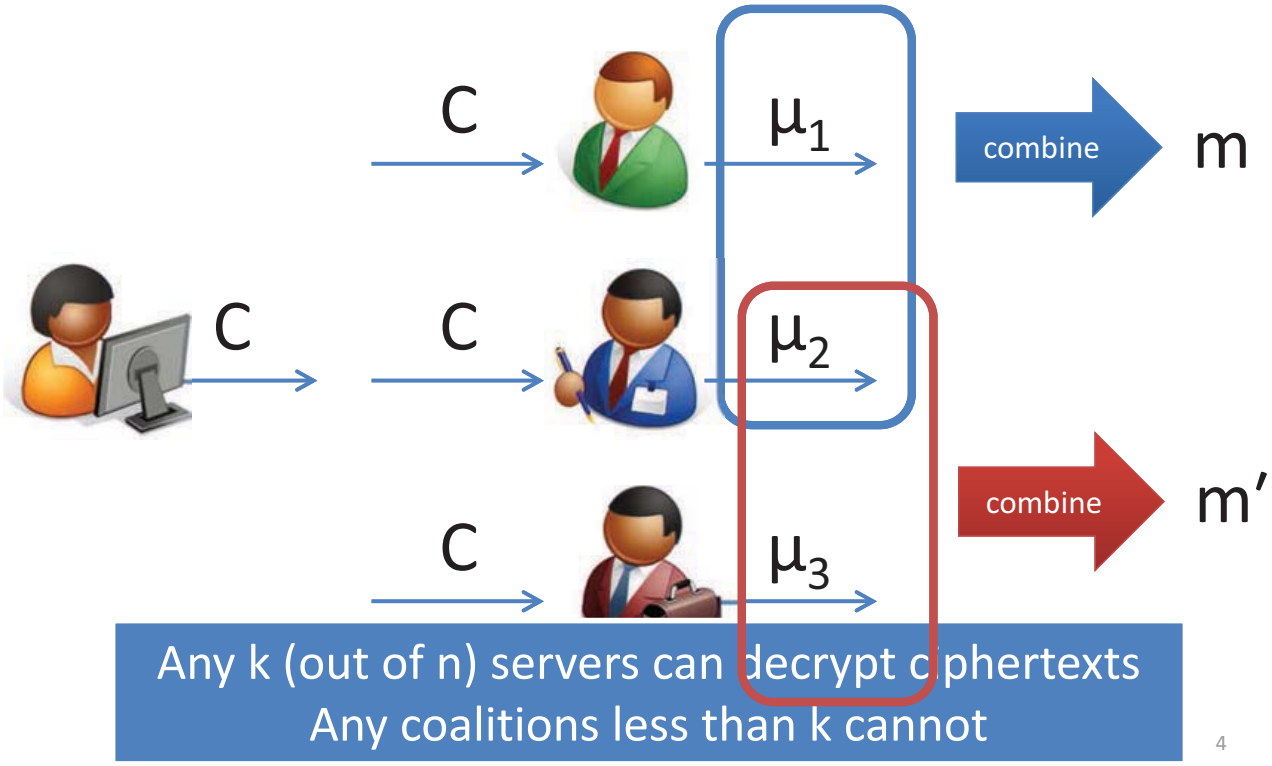
- Threshold Public Key Encryption (TPKE)
- Dynamic TPKE
  - Delerablee and Pointcheval (CRYPTO 2008)
    - Secure under a q-type assumption called the multi-sequence of exponent Diffie-Hellman (MSE-DDH) assumption
    - Random oracle model (for decryption consistency)
- Our Contribution
  - Two dynamic TPKE schemes supporting decryption consistency without relying on random oracles or q-type assumptions
    - A generic construction of dynamic TPKE (with weak decryption consistency) from PKENO in a black-box manner
      - PKENO: Public Key Encryotion with Noninteractive Opening (Damgard et al. CT-RSA 2008) with a tag-based variant. Give tag-based PKENO schemes from the DLIN or DBDH assumptions
      - The Dodis-Katz multiple encryption technique (TCC 2005)
      - Verifiable Secret Sharing
    - A dynamic TPKE scheme with strong decryption consistency
      - Secure under the DLIN assumption without using random oracles
      - The Groth-Sahai proofs (EUROCRYPT 2008)
      - The first dynamic TPKE scheme with strong decryption consistency w/o RO
  - A black-box construction of (non-dynamic) TPKE from PKENO with strong decryption consistency
    - Multiple-assignment secret sharing (Ito-Saito-Nishizeki, Journal of Cryptology 1993)
    - It can be seen as confirmation of the conjecture by Galindo et al. (Africacrypt 2010)

k-out-of-n threshold PKE



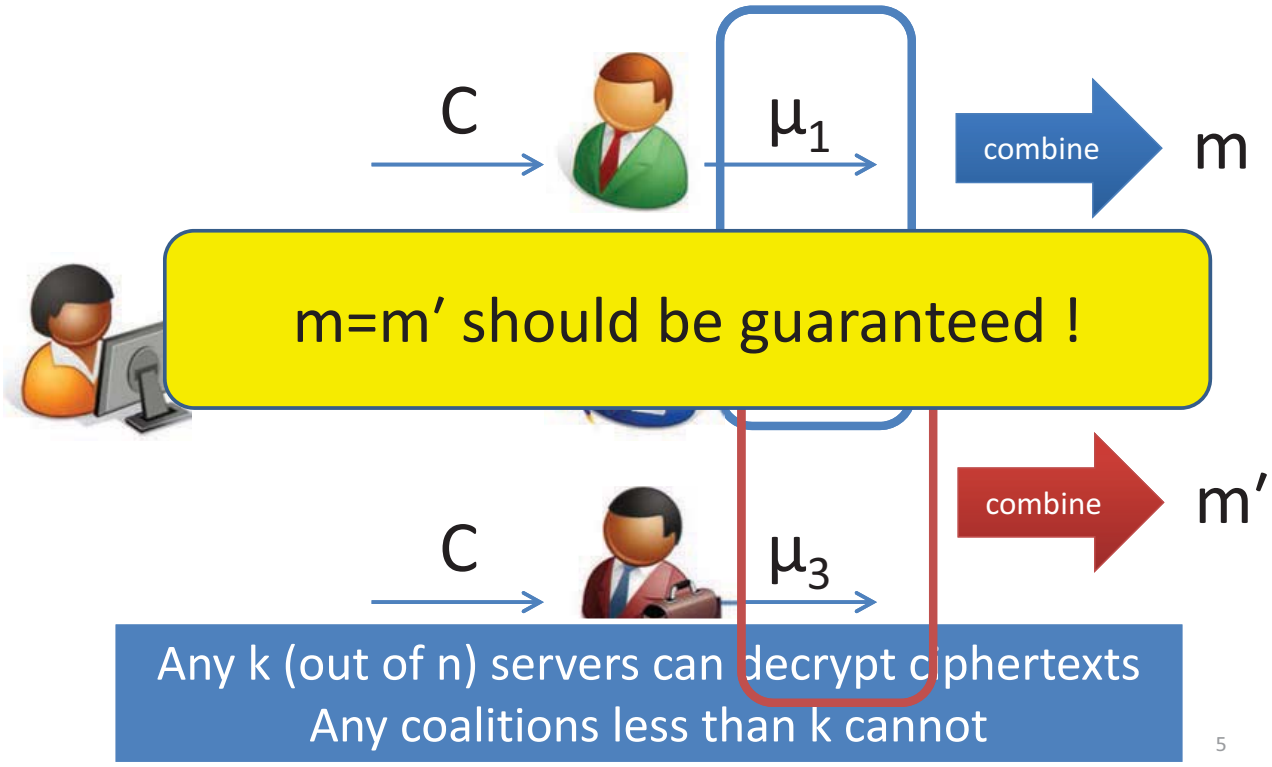
3

k-out-of-n threshold PKE



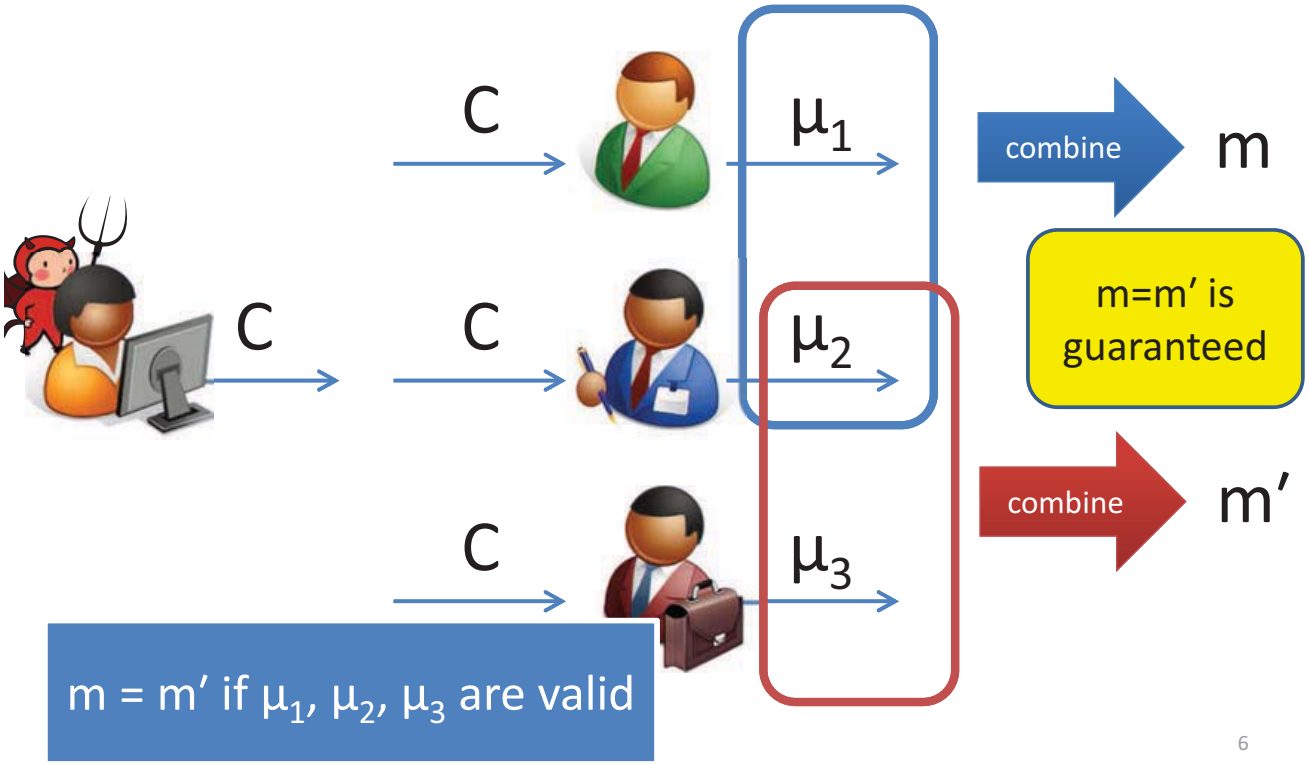
4

k-out-of-n threshold PKE



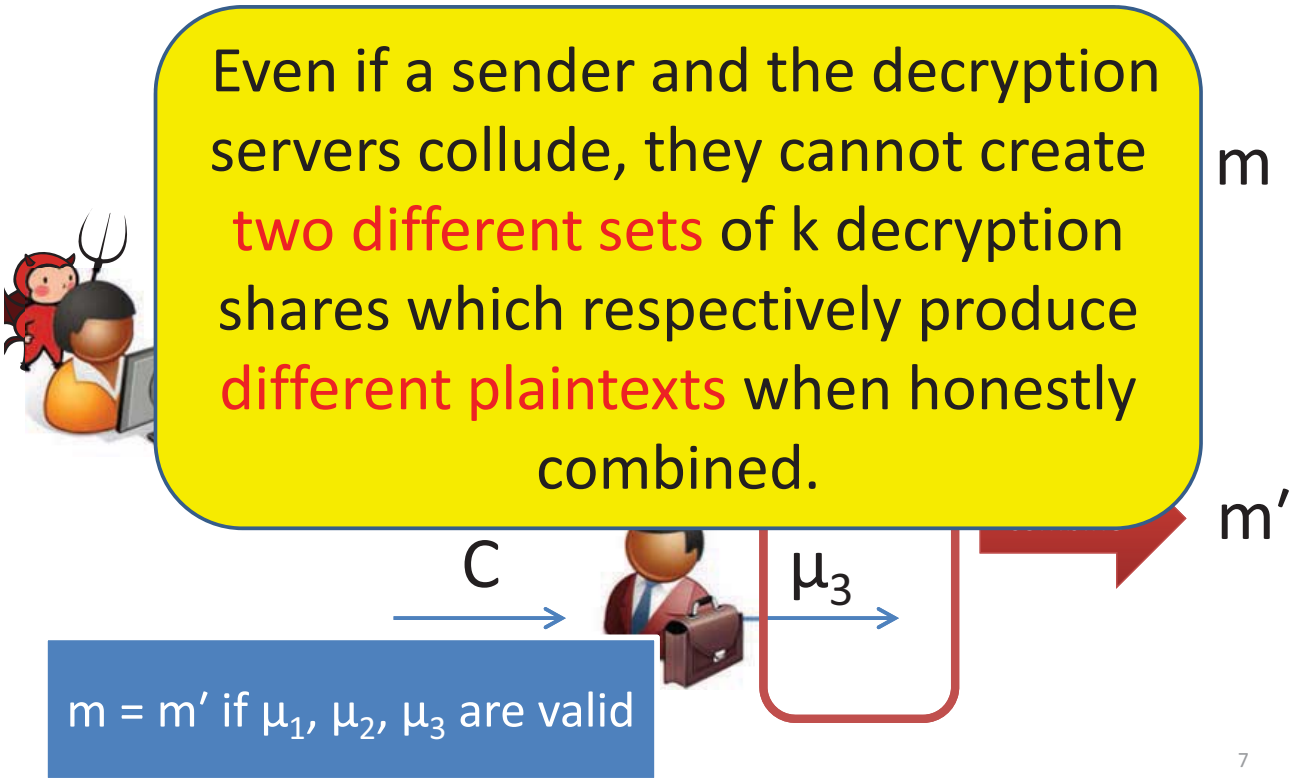
5

Decryption Consistency [BBH06][SG98]



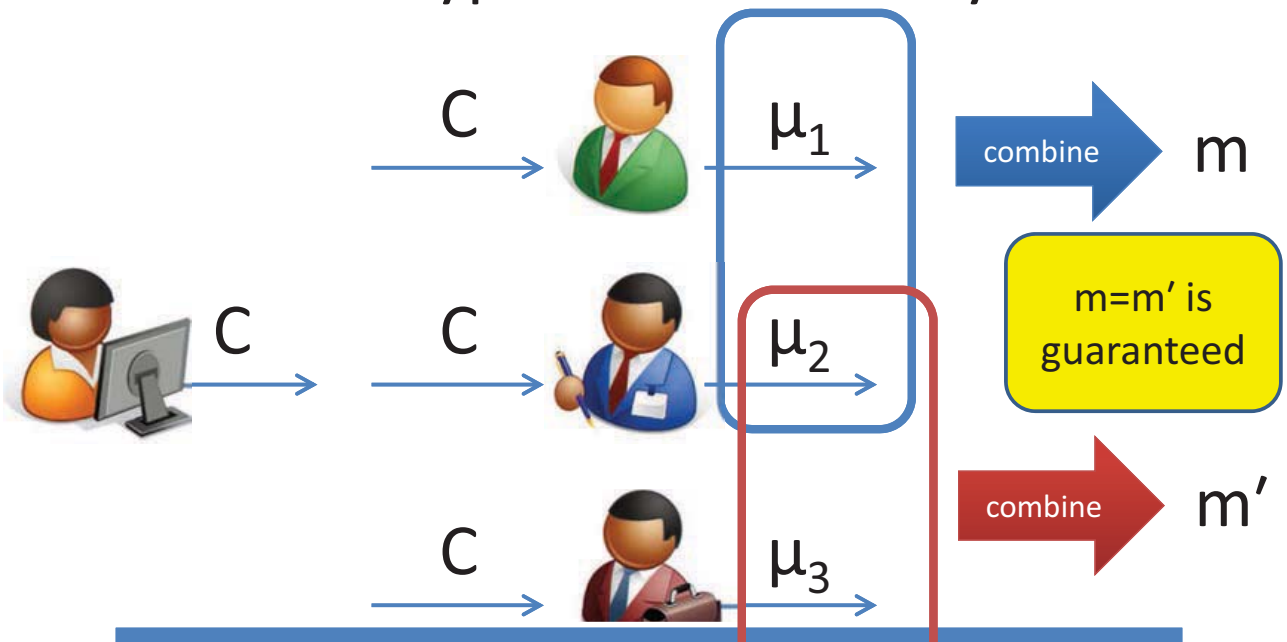
6

Decryption Consistency [BBH06][SG98]



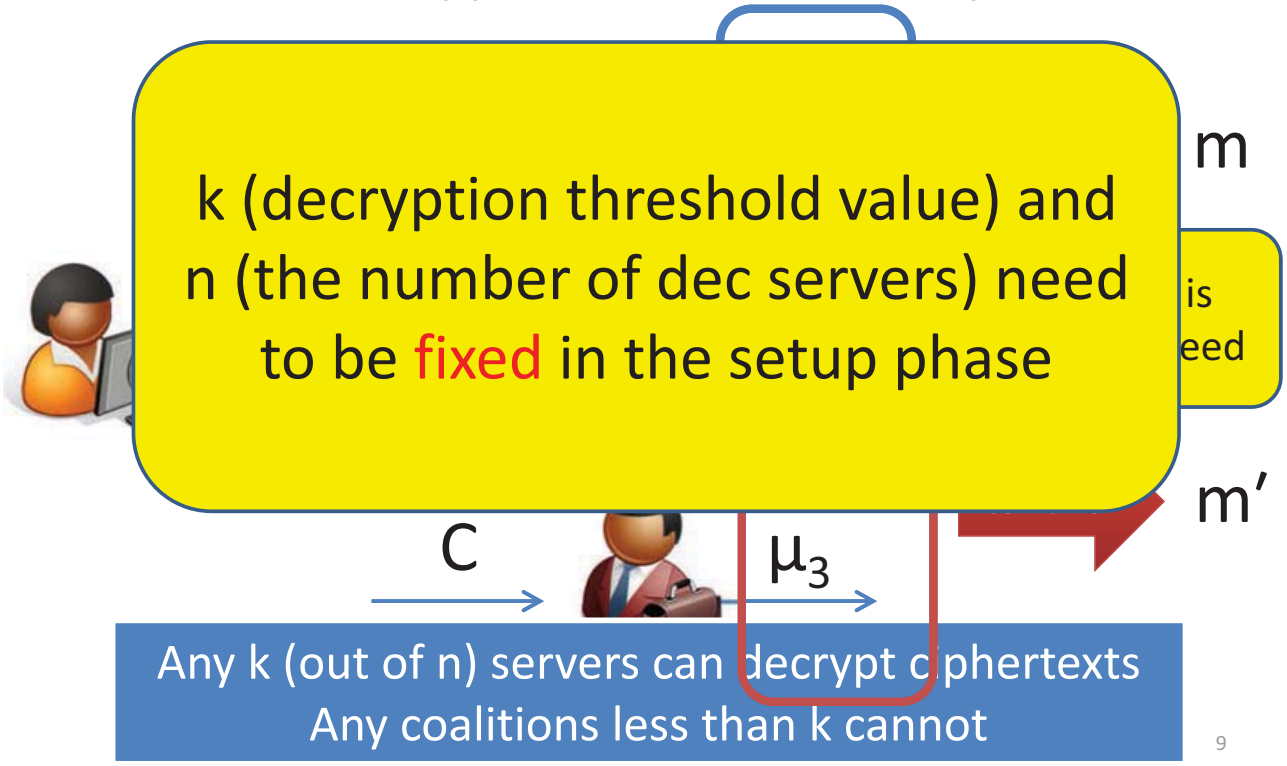
7

k-out-of-n threshold PKE with Decryption Consistency



8

k-out-of-n threshold PKE with Decryption Consistency



Dynamic TPKE [DP08]

- Allow decryption servers to join the system dynamically after the system is set up
- Allow the sender to dynamically choose the authorized set and the decryption threshold k at the time of encryption.

- The DP scheme
  - Secure under a q-type assumption called the multi-sequence of exponent Diffie-Hellman (MSE-DDH) assumption
  - Random oracle model (for decryption consistency)

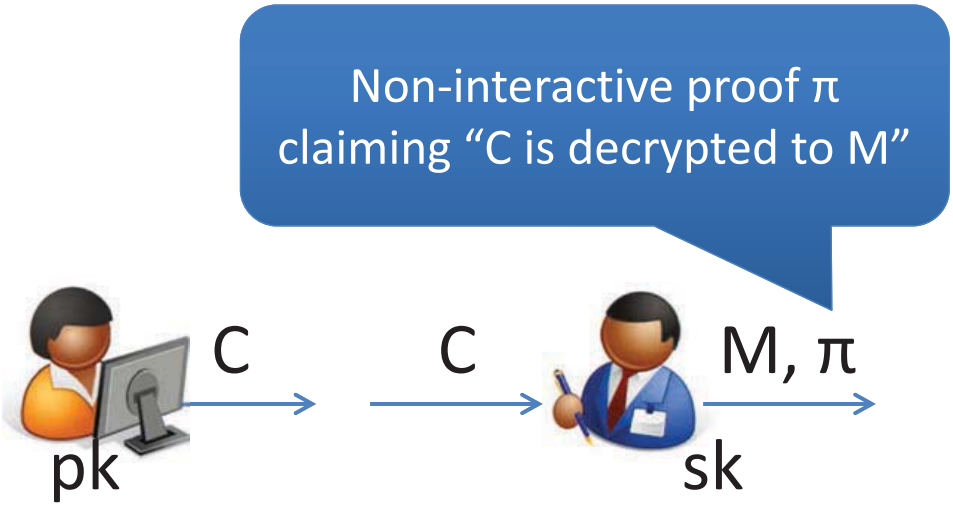
Our Contribution

- Two dynamic TPKE schemes supporting decryption consistency without relying on random oracles or q-type assumptions
  - A generic construction of dynamic TPKE (with weak decryption consistency) from PKENO in a black-box manner
    - PKENO: Public Key Encryotion with Noninteractive Opening (Damgard et al. CT-RSA 2008) with a tag-based variant.
      - Give tag-based PKENO schemes from the DLIN or DBDH assumptions
    - The Dodis-Katz multiple encryption technique (TCC 2005)
    - Verifiable Secret Sharing
  - A dynamic TPKE scheme with strong decryption consistency
    - Secure under the DLIN assumption without using random oracles
    - The Groth-Sahai proofs (EUROCRYPT 2008)
    - The first dynamic TPKE scheme with strong decryption consistency w/o RO

- (A black-box construction of non-dynamic TPKE)

11

PKE with Non-interactive Opening [DHKT08]



Proof for some ciphertext C doesn't help recovering the plaintext of another ciphertext C'

12

PKE with Non-interactive Opening [DHKT08]

In our construction, the underlying PKENO needs to be "tag-based".

Proof for some ciphertext C doesn't help recovering the plaintext of another ciphertext C'

Instantiation of tag-based PKENO (1)

M: plaintext
t: tag

From DBDH:

$pk = (g, z, u, h)$

$Enc(pk, t, M) = (\underbrace{g^r}_{C_1}, \underbrace{(z^t u)^r}_{C_2}, \underbrace{e(z,h)^r M}_{C_3})$

Proof: (M, z^r)

Verify: $e(C_1, z) = e(g, z^r)$
$e(C_2, g) = e(z^t u, C_1)$
$C_3/M = e(z^r, h)$

Instantiation of tag-based PKENO (2)

M: plaintext
t: tag

From DLIN (with pairing):

$pk = (g_1, g_2, z, u_1, u_2)$

$Enc(pk, t, M) = (\underbrace{g_1^r}_{C_1}, \underbrace{g_2^s}_{C_2}, \underbrace{(z^t u_1)^r}_{C_3}, \underbrace{(z^t u_2)^s}_{C_4}, \underbrace{z^{r+s} M}_{C_5})$

Proof: (M, z^r, z^s)

Verify: $e(C_1, z) = e(g_1, z^r)$ $\quad e(C_2, z) = e(g_2, z^s)$
$e(C_3, g_1) = e(z^t u_1, C_1)$ $\quad e(C_4, g_2) = e(z^t u_2, C_2)$
$C_5/M = z^r z^s$

Naive construction w/o decryption consistency

C = (Enc(pk, f(1)), Enc(pk, f(2)), Enc(pk, f(3))

f is a random degree-1 polynomial with f(0) = m

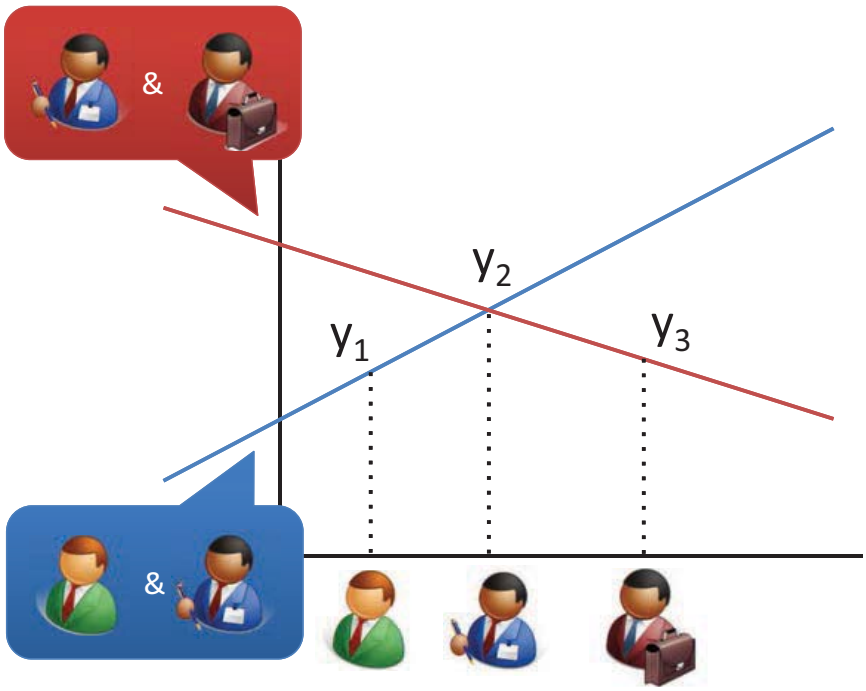
… does not provide decryption consistency!

(Counter)example:

C = (Enc(pk, y_1), Enc(pk, y_2), Enc(pk, y_3))

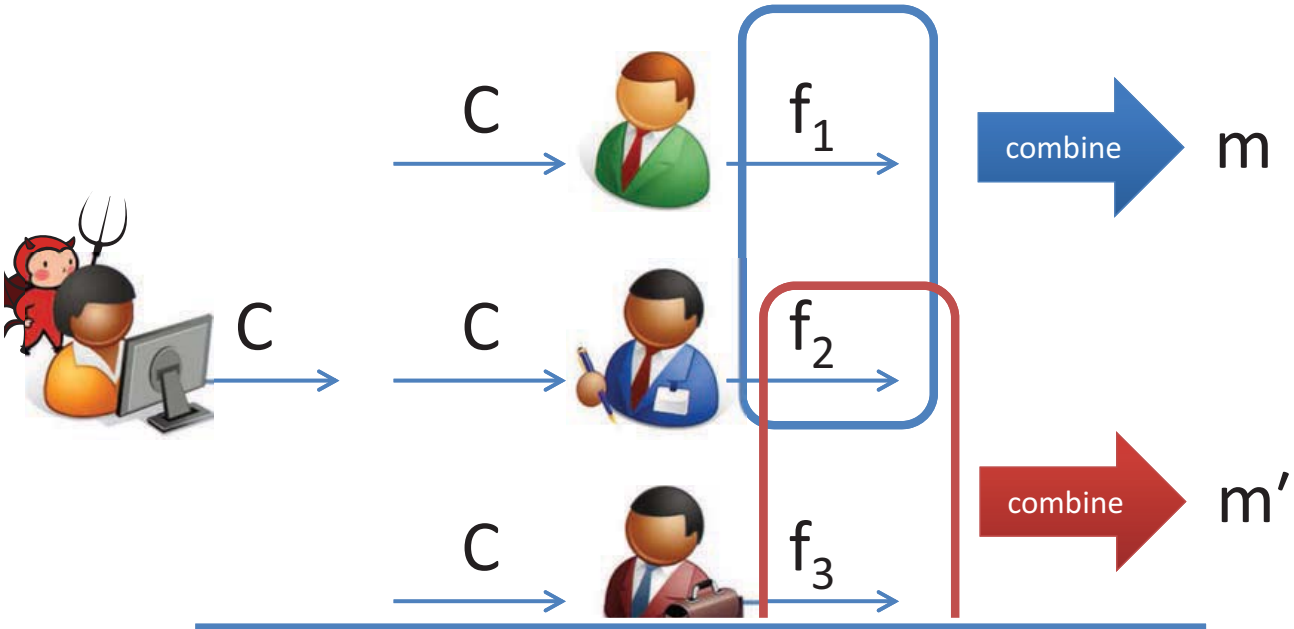
where $(1,y_1)$, $(2,y_2)$, $(3,y_3)$ don't lie in a single line

Counterexample (cont'd)



17

Observation



If f_1, f_2, f_3 are collinear, decryption consistency is ensured

18

Proposed Construction 1

f(x,y) is a random symmetric polynomial (i.e. f(x,y) = f(y,x))
with f(0,0) = m

$$C = \left(\begin{array}{lll} \mathrm{Com}(f(1,1), r_{11}) & \mathrm{Com}(f(1,2), r_{12}) & \mathrm{Com}(f(1,3), r_{13}) \\ & \mathrm{Com}(f(2,2), r_{22}) & \mathrm{Com}(f(2,3), r_{23}) \\ & & \mathrm{Com}(f(3,3), r_{33}) \end{array} \right.$$

$$\left. \begin{array}{l} \mathrm{Enc}(\, pk_{\text{Alice}}, \langle f(1,1), f(1,2), f(1,3), r_{11}, r_{12}, r_{13} \rangle \,) \\ \mathrm{Enc}(\, pk_{\text{Bob}}, \langle f(2,1), f(2,2), f(2,3), r_{12}, r_{22}, r_{23} \rangle \,) \\ \mathrm{Enc}(\, pk_{\text{Carol}}, \langle f(3,1), f(3,2), f(3,3), r_{13}, r_{23}, r_{33} \rangle \,) \end{array} \right)$$

Proposed Construction 1

f(x,y) is a random symmetric polynomial (i.e. f(x,y) = f(y,x))
with f(0,0) = m

Theorem. The construction is CCA secure and has weak decryption consistency
if (1) the tag-based PKENO is selective-tag CCA secure,
(2) the commitment is binding and hiding, and
(3) the one-time sig. is strongly unforgeable.

$$C = \left(\begin{array}{lll} \mathrm{Com}(f(1,1), r_{11}) & \mathrm{Com}(f(1,2), r_{12}) & \mathrm{Com}(f(1,3), r_{13}) \\ & \mathrm{Com}(f(2,2), r_{22}) & \mathrm{Com}(f(2,3), r_{23}) \\ & & \mathrm{Com}(f(3,3), r_{33}) \end{array} \right.$$

$$\left. \begin{array}{l} \mathrm{Enc}(\, pk_{\text{Alice}}, \langle f(1,1), f(1,2), f(1,3), r_{11}, r_{12}, r_{13} \rangle \,) \\ \mathrm{Enc}(\, pk_{\text{Bob}}, \langle f(2,1), f(2,2), f(2,3), r_{12}, r_{22}, r_{23} \rangle \,) \\ \mathrm{Enc}(\, pk_{\text{Carol}}, \langle f(3,1), f(3,2), f(3,3), r_{13}, r_{23}, r_{33} \rangle \,) \end{array} \right)$$

Proposed Construction 1

f(x,y) is a random symmetric polynomial (i.
with f(0,0) = m

Explain later

Theorem. The construction is CCA secure and has weak decryption consistency
if (1) the tag-based PKENO is selective-tag CCA secure,
(2) the commitment is binding and hiding, and
(3) the one-time sig. is strongly unforgeable.

C = (

Enc(pk, ⟨f(2,1), f(2,2), f(2,3), r_{12}, r_{22}, r_{23}⟩)

Enc(pk, ⟨f(3,1), f(3,2), f(3,3), r_{13}, r_{23}, r_{33}⟩)

Decryption (1/3)

C = (

Com(f(1,1), r_{11}) Com(f(1,2), r_{12}) Com(f(1,3), r_{13})
Com(f(2,2), r_{22}) Com(f(2,3), r_{23})
Com(f(3,3), r_{33})

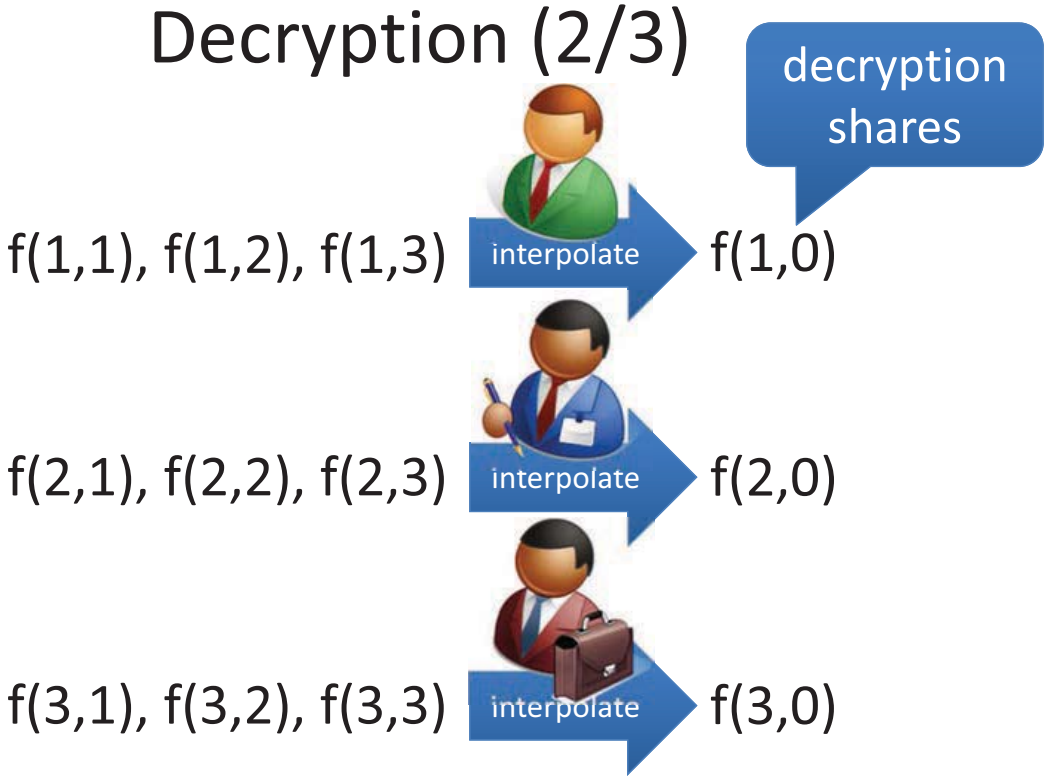
Enc(pk, ⟨f(1,1), f(1,2), f(1,3), r_{11}, r_{12}, r_{13}⟩)

Enc(pk, ⟨f(2,1), f(2,2), f(2,3), r_{12}, r_{22}, r_{23}⟩)

Enc(pk, ⟨f(3,1), f(3,2), f(3,3), r_{13}, r_{23}, r_{33}⟩)

1. Verifies decommitment and confirm f(i,1), …, f(i,n) constitute a degree-(k-1) curve

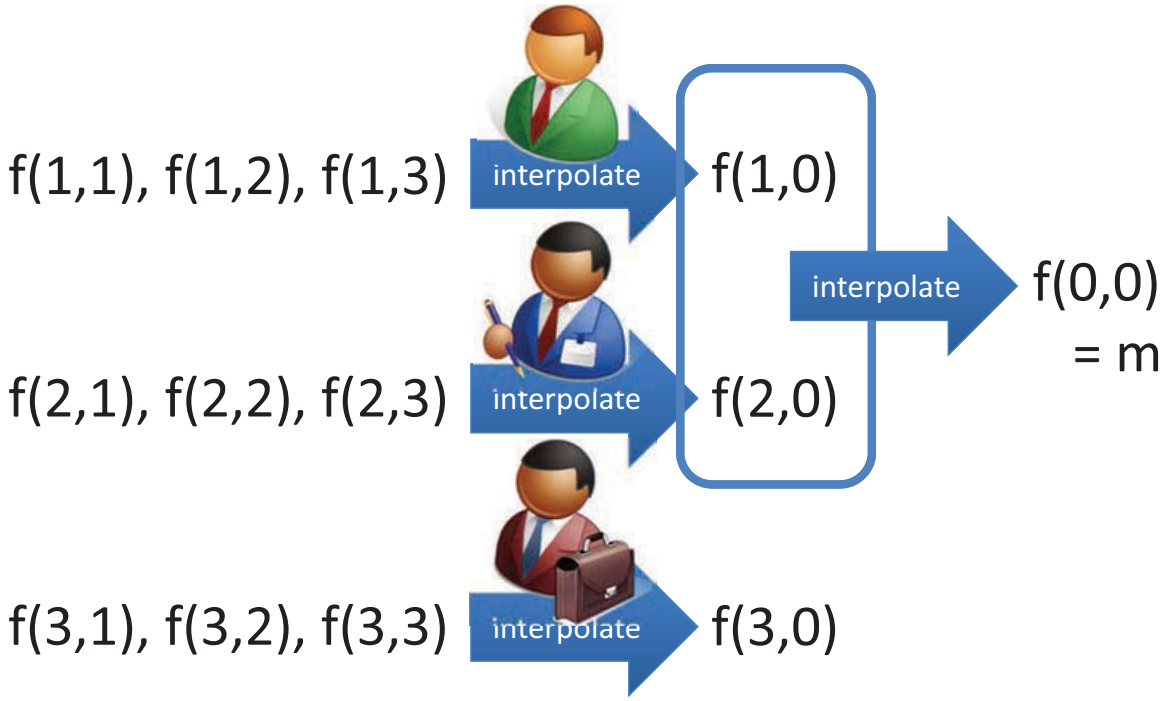
Decryption (2/3)



2. Each server interpolates f(i,0) from f(i,1), …, f(i,n)

23

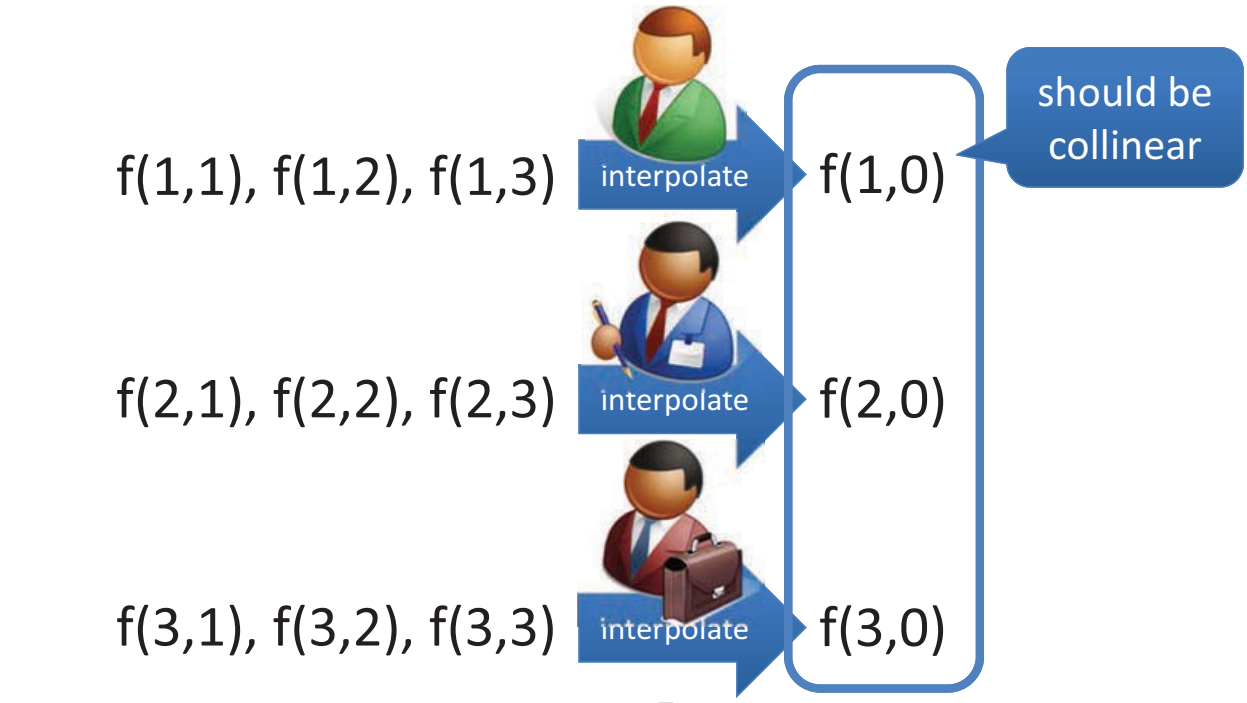
Decryption (3/3)



3. Combine shares by interpolating k shares

24

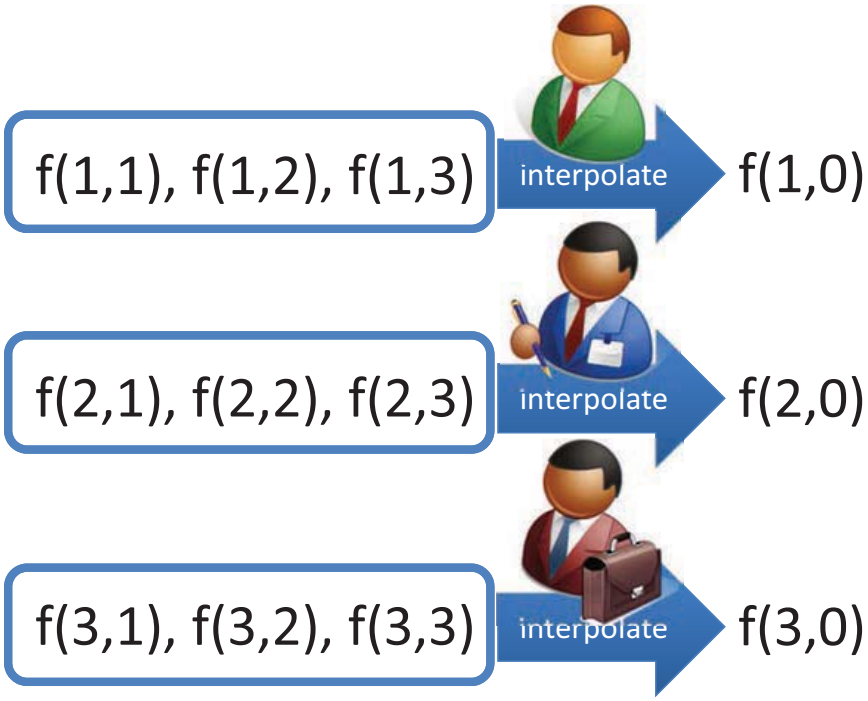
Decryption Consistency of the construction



f(1,0), …, f(n,0) is degree-(k-1) ➔ the scheme achieves decryption consistency

25

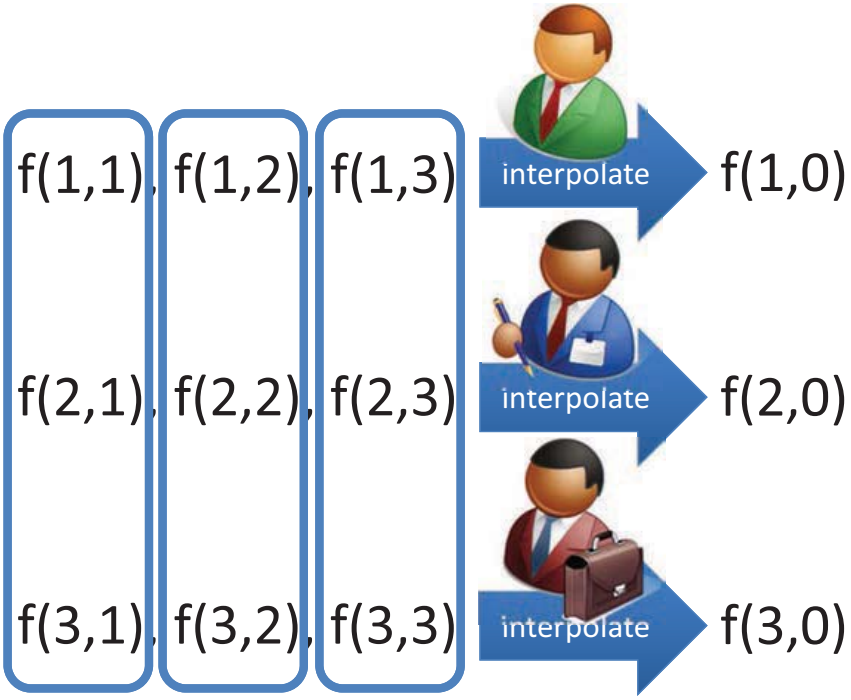
Decryption Consistency of the construction



f(i,1), …, f(i,n) is degree-(k-1) (verified by each server)

26

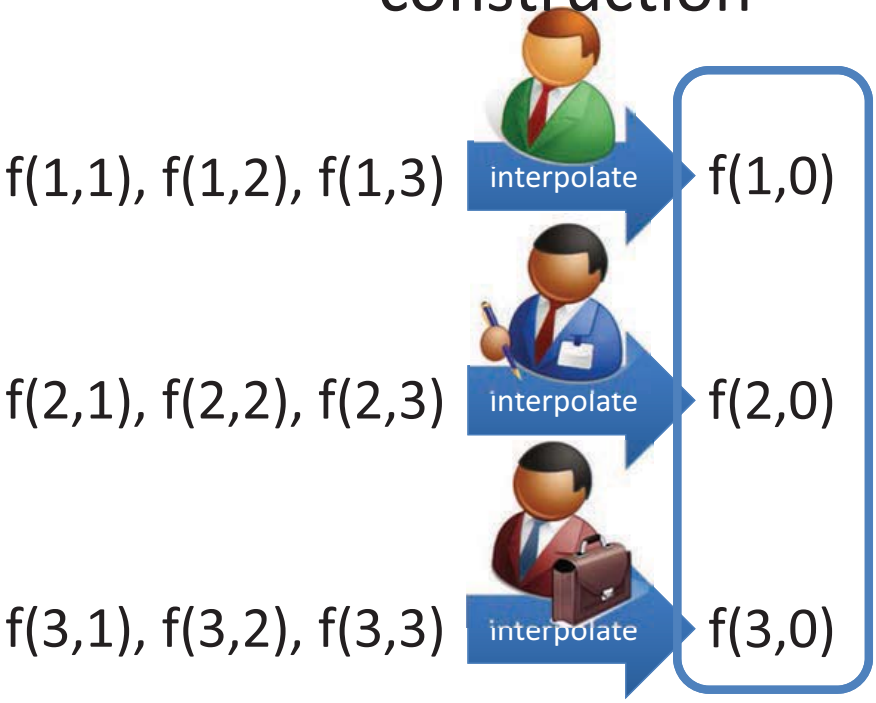
Decryption Consistency of the construction



f(1,i), …, f(n,i) is also degree-(k-1) (by symmetry of f)

27

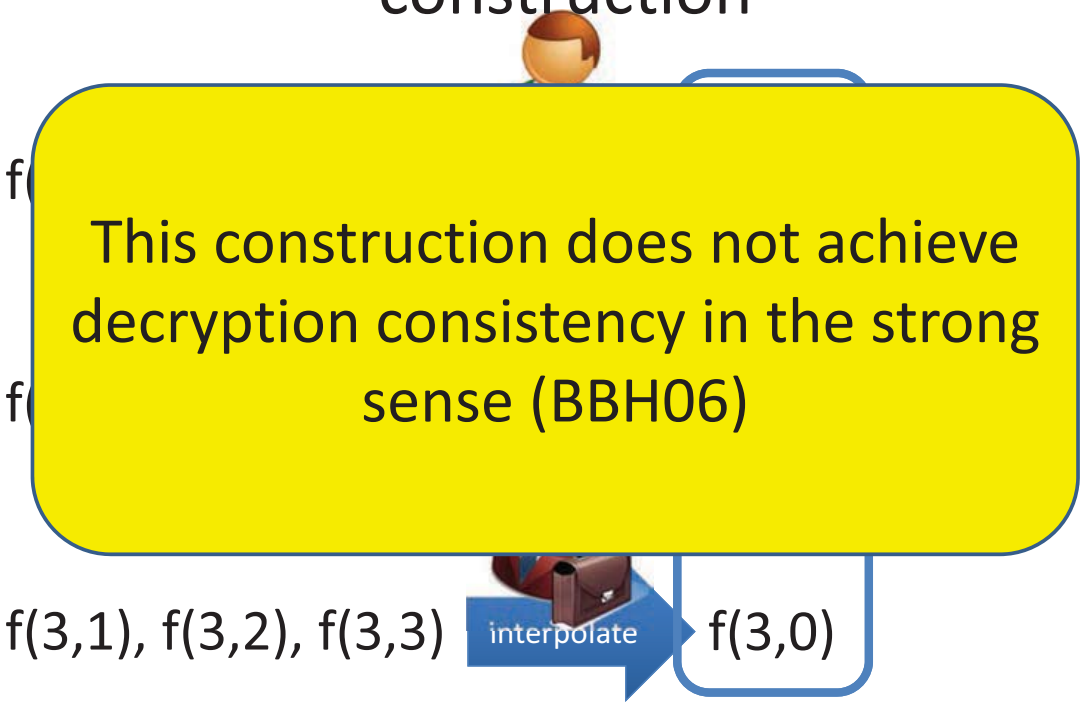
Decryption Consistency of the construction



f(1,0), …, f(n,0) are also degree-(k-1)

28

Decryption Consistency of the construction



f(1,0), …, f(n,0) are also degree-(k-1)

The BBH model

Completeness:

$\forall C \in \{0,1\}^*$

$\textit{ShareVerify}(pk, C, \text{ShareDec}(pk, sk_i, C)) = T$

Decryption Consistency:

$\forall C \in \{0,1\}^*, \{\mu_i\}, \{\mu'_i\}$

$\textit{ShareVerify}(pk, C, \mu_i) = T$

$\textit{ShareVerify}(pk, C, \mu'_i) = T$

$\Rightarrow \text{Combine}(pk, C, \{\mu_i\}) = \text{Combine}(pk, C, \{\mu'_i\})$

[BBH06] D. Boneh, X. Boyen, S. Halevi. Chosen Ciphertext Secure Public Key Threshold Encryption Without Random Oracles. In Proc. of CT-RSA '06.

When  receives an invalid share…

Enc(pk  ,⟨f(1,1), f(1,2), f(1,3)⟩)  interpolate f(1,0)

Enc(pk  ,⟨f(2,1), f(2,2), f(2,3)⟩)  interpolate f(2,0)

Enc(pk  ,⟨0000000000000000⟩)  interpolate ???

To claim *"I receive an invalid share!"* (using the power of PKENO) seems suffice…

When  receives an invalid share…

Completeness:

$\forall C \in \{0,1\}^*$

ShareVerify(pk, C, ShareDec(pk, sk_i, C)) = T

Probably no other choice than claiming that the shares are invalid…

Enc(pk  ,⟨0000000000000000⟩)  interpolate ???

When  receives an invalid share…

Decryption Consistency:

$\forall C \in \{0,1\}^*, \{\mu_i\}, \{\mu'_i\}$

ShareVerify$(pk, C, \mu_i) = T$

ShareVerify$(pk, C, \mu'_i) = T$

$\Rightarrow$ Combine$(pk, C, \{\mu_i\})$ = Combine$(pk, C, \{\mu'_i\})$

Enc(pk  ,⟨f(2,1), f(2,2), f(2,3)⟩) interpolate f(2,0)

…, but he is required to output something successfully combined

Enc(pk  ,⟨000000000000000⟩) interpolate ???

Our relaxed requirements

The verification algorithm outputs three values T_{valid}, $T_{invalid}$, and bot

Completeness:

$\forall C \in \{0,1\}^*$

ShareVerify$(pk, C, \text{ShareDec}(pk, sk_i, C)) = T_{valid}/T_{invalid}$

Weak Decryption Consistency:

$\forall C \in \{0,1\}^*, \{\mu_i\}, \{\mu'_i\}$

ShareVerify$(pk,C,\mu_i) = T_{valid}$ and *ShareVerify*$(pk,C,\mu'_i) = T_{valid}$

$\Rightarrow$ Combine$(pk, C, \{\mu_i\})$ = Combine$(pk, C, \{\mu'_i\})$

$\mu, \mu' \in \{\mu_i\} \cup \{\mu'_i\}$ (which are both attributed to the same dec server)

ShareVerify$(pk,C,\mu) = T_{valid}$ and *ShareVerify*$(pk,C,\mu') = T_{invalid}$

Proposed Construction 2

- This is the first dynamic TPKE scheme with
  - Strong decryption consistency in the sense of BBH06
  - Secure under the DLIN assumption without using random oracles.

- DLIN-based tag-PKENO
- Groth-Sahai NIZK proofs (DLIN) for decryption consistency

$g^{r+s}\, m\, g^{a1\, i+a2\, i\hat{}2+\ldots.+ak-1 i\hat{}k-1} = g^{r+s}\, m\, (g^{i})^{a1}+(g^{i\hat{}2})^{a2}+\ldots(g^{i\hat{}k-1})^{ak-1}$

The Dodis-Katz multiple enc

Prove m, r, s, a1,…,ak-1 by the Groth-Sahai proof for each i in [1,n]

Black-box construction of Non-dynamic TPKE w/ decryption consistency

- A generic construction of dynamic TPKE (with weak decryption consistency) from PKENO in a black-box manner
- A dynamic TPKE scheme with strong decryption consistency which is constructed by a non-black-box manner (Groth-Sahai proof)

- Can we construct TPKE with strong decryption consistency from PKENO in a black-box manner?

Black-box construction of Non-dynamic TPKE

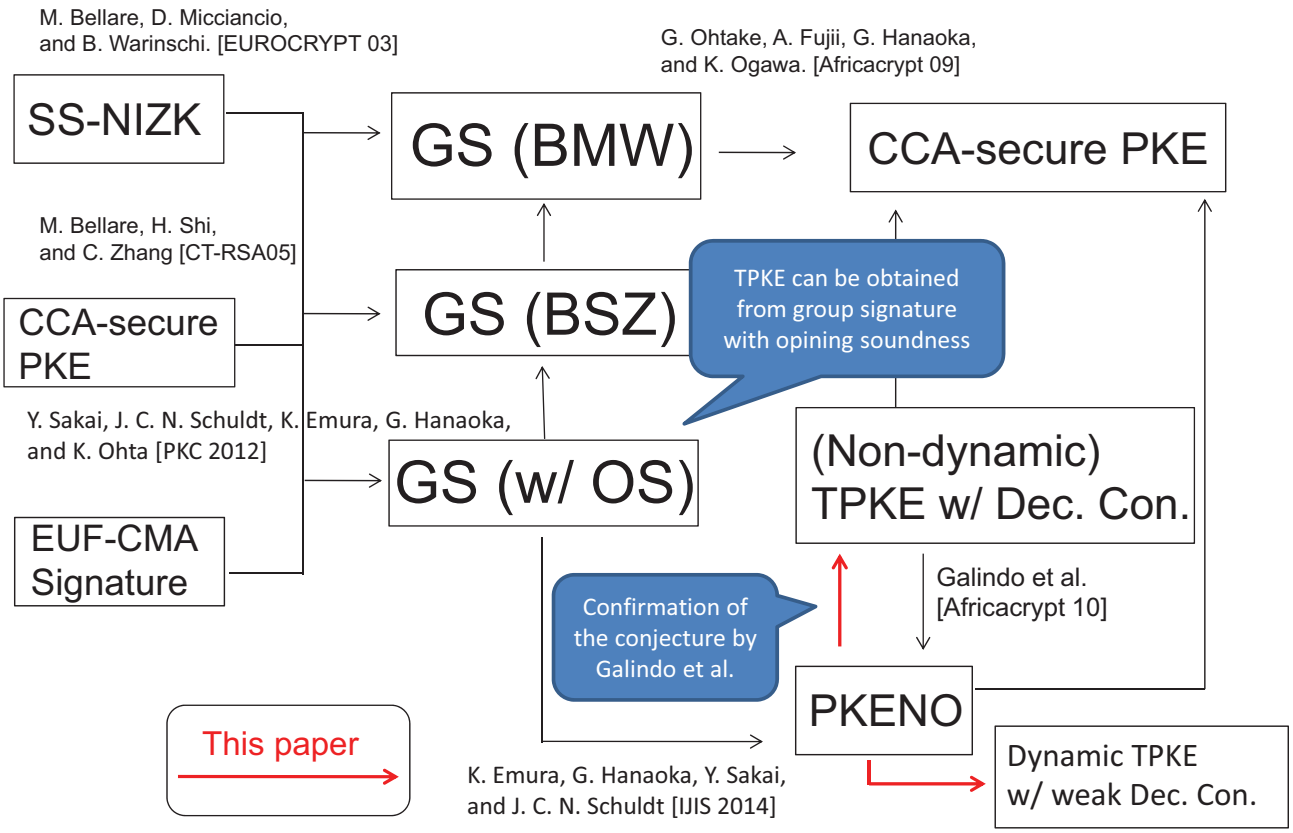
YES!

Employ the Ito-Saito-Nishizeki multiple-assignment secret sharing [JoC93]

See full ver. of this paper

- consistency from PKENO in a black-box manner?

Summary of the black-box construction



Comparison

	Dynamic	Decryption Consistency	Black-box Construction	Ciphertext size	Assumption	Corruption Model
Shoup-Genaro 98	No	Yes	No	O(1)	CDH(RO) DDH(RO)	Static
Boneh-Boyen-Halevi06	No	Yes	No	O(1)	DBDH	Static
Delerablee-Pointcheval 08	Yes	Yes	No	O(1)	q-MSE-DDH (RO)	Static
Libert-Yung 11	No	Yes	No	O(1)	Assumption 1, 2, 3	Adaptive
Libert-Yung 12	No	Yes	Yes (All-But-One Perfectly Sound Threshold Hash Proof Systems)	O(1)	Subgroup Decision DLIN/SXDH	Adaptive
Ours 1	Yes	Yes (Weak)	Yes (PKENO)	$O(n^2)$	-	Static
Ours2	Yes	Yes	No	O(n)	DLIN	Static
Ours 3	No	Yes	Yes (PKENO)	$O(2^n)$	-	Static

39

Conclusion

- Two dynamic TPKE schemes supporting decryption consistency without relying on random oracles or q-type assumptions
  - A generic construction of dynamic TPKE with weak decryption consistency from PKENO in a black-box manner
  - A dynamic TPKE scheme with strong decryption consistency which is the first dynamic TPKE scheme with strong decryption consistency w/o RO
- A black-box construction of (non-dynamic) TPKE with strong decryption consistency from PKENO
  - Confirmation of the conjecture by Galindo et al.

- Future work
  - Dynamic TPKE with strong decryption consistency (from PKENO or other primitive) with constant ciphertext size in a black-box manner
  - Dynamic TPKE with adaptive corruption

40

On the Application of Clique Problem for Proof-of-Work in Cryptocurrencies

Samiran Bag (Joint work with Sushmita Ruj and Kouichi Sakurai)

Kyushu University
samiran.bag@gmail.com

In this work we propose a scheme that could be used as an alternative to the existing proof of work(PoW) scheme for mining in Bitcoin P2P network. Our scheme ensures that the miner must do at least a non-trivial amount of computation for solving the computational problem put forth in the paper and thus computing a PoW. Here, we have proposed to use the problem of finding the largest clique in a big graph as a replacement for the existing Bitcoin PoW scheme. In this paper, we have dealt with a graph having $O(2^{30})$ vertices and $O(2^{48})$ edges which is constructed deterministically using the set of transactions executed within a certain time slot. We have discussed some algorithms that can be used by any Bitcoin miner to solve the PoW puzzle. Then we discuss an algorithm that could perform this task by doing $O(2^{80})$ hash calculations. We have also proposed an improvement to this algorithm by which the PoW puzzle can be solved by calculating $O(2^{70.5})$ hashes and using $O(2^{48})$ space. This scheme is better than the existing proof of work scheme that uses Hashcash, where a lucky miner could manage to find a solution to the proof of work puzzle by doing less amount of computation. In our proposed scheme only the miner needs to invest a sizable amount of computational resources for solving the proof of work puzzle.

References

[1] Noga Alon, Raphael Yuster, and Uri Zwick. Finding and counting given length cycles (extended abstract). In *Proceedings of the Second Annual European Symposium on Algorithms*, ESA '94, pages 354–364, London, UK, UK, 1994. Springer-Verlag.

[2] Adam Back. Hashcash - a denial of service counter-measure. Technical report, August 2002. (implementation released in mar 1997).

[3] Coen Bron and Joep Kerbosch. Algorithm 457: Finding all cliques of an undirected graph. *Commun. ACM*, 16(9):575–577, September 1973.

[4] Norishige Chiba and Takao Nishizeki. Arboricity and subgraph listing algorithms. *SIAM J. Comput.*, 14(1):210–223, February 1985.

[5] Nicolas T Courtois and Lear Bahack. On subversive miner strategies and block withholding attack in bitcoin digital currency. *arXiv preprint arXiv:1402.1718*, 2014.

[6] Friedrich Eisenbrand and Fabrizio Grandoni. On the complexity of fixed parameter clique and dominating set. *Theoretical Computer Science*, 326(13):57 – 67, 2004.

[7] G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 77:313–324, 3 1975.

[8] Tang Jian. An o(20.304n) algorithm for solving maximum independent set problem. *Computers, IEEE Transactions on*, C-35(9):847–851, Sept 1986.

[9] Ton Kloks, Dieter Kratsch, and Haiko Mller. Finding and counting small induced subgraphs efficiently. *Information Processing Letters*, 74(34):115 – 121, 2000.

[10] Loi Luu, Ratul Saha, Inian Parameshwaran, Prateek Saxena, and Aquinas Hobor. On power splitting games in distributed computation: The case of bitcoin pooled mining. Cryptology ePrint Archive, Report 2015/155, 2015. http://eprint.iacr.org/.

[11] D.W. Matula. On the complete subgraph of random graph. In *Comb. Math and its Appl*, pages 356–369, Chappel Hill, N.C., 1970.
[12] J.W. Moon and L. Moser. On cliques in graphs. *Israel Journal of Mathematics*, 3(1):23–28, 1965.
[13] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *Consulted*, 1(2012):28, 2008.
[14] Jaroslav Nešetřil and Svatopluk Poljak. On the complexity of the subgraph problem. *Commentationes Mathematicae Universitatis Carolinae*, 26(2):415–419, 1985.
[15] Bharath Pattabiraman, Md.MostofaAli Patwary, AssefawH. Gebremedhin, Wei-keng Liao, and Alok Choudhary. Fast algorithms for the maximum clique problem on massive sparse graphs. In Anthony Bonato, Michael Mitzenmacher, and Pawe Praat, editors, *Algorithms and Models for the Web Graph*, volume 8305 of *Lecture Notes in Computer Science*, pages 156–169. Springer International Publishing, 2013.
[16] J.M Robson. Algorithms for maximum independent sets. *Journal of Algorithms*, 7(3):425 – 440, 1986.
[17] John M Robson. Finding a maximum independent set in time o (2n/4). Technical report, Technical Report 1251-01, LaBRI, Université de Bordeaux I, 2001.
[18] Meni Rosenfeld. Analysis of bitcoin pooled mining reward systems. *CoRR*, abs/1112.4980, 2011.
[19] Robert E. Tarjan and Anthony E. Trojanowski. Finding a maximum independent set. Technical report, Stanford University, Stanford, CA, USA, 1976.
[20] Etsuji Tomita, Akira Tanaka, and Haruhisa Takahashi. The worst-case time complexity for generating all maximal cliques and computational experiments. *Theor. Comput. Sci.*, 363(1):28–42, October 2006.
[21] John Tromp. Cuckoo cycle: a memory bound graph-theoretic proof-of-work. Cryptology ePrint Archive, Report 2014/059, 2014. `http://eprint.iacr.org/`.
[22] Virginia Vassilevska and Ryan Williams. Finding, minimizing, and counting weighted subgraphs. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, STOC '09, pages 455–464, New York, NY, USA, 2009. ACM.
[23] Raphael Yuster. Finding and counting cliques and independent sets in r-uniform hypergraphs. *Information Processing Letters*, 99(4):130 – 134, 2006.

Application of Clique Problem for Proof-of-Work in Cryptocurrencies

Samiran Bag Sushmita Ruj Kouichi Sakurai

[1]Faculty of Information Science and Electrical Engineering
Kyushu University

[2]Applied Statistics Unit
Indian Statistical Institute

[3]Faculty of Information Science and Electrical Engineering
Kyushu University

Workshop on Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques

channel hopping

WHAT IS BITCOIN?

- Bitcoin has been the first successful Cryptocurrency.
- Bitcoin was proposed by 'Satoshi Nakamoto' in 2008.
- The first software of Bitcoin was released in 2009.
- In April 2015 the price of a Bitcoin reached $235.
- The total value of all Bitcoins in circulation is more than 3 billion US dollars.

BITCOIN:HOW IT WORKS

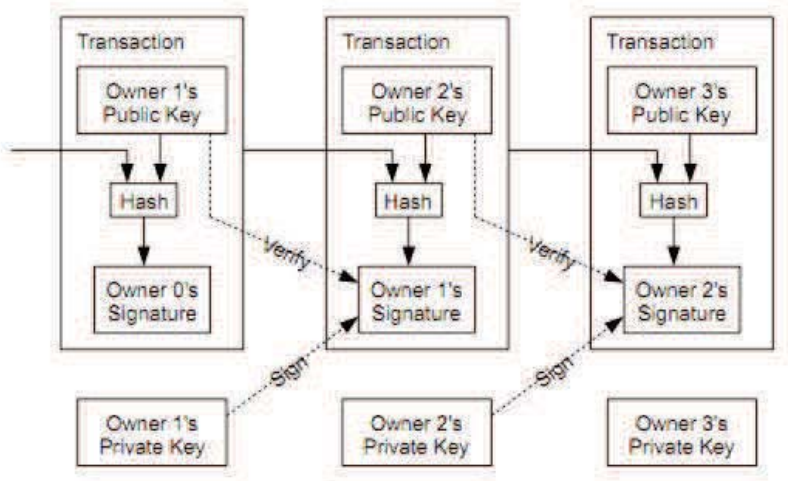
- Bitcoin Network
- Bitcoin Transaction
- Bitcoin Block Chain

Bitcoin Network

Bitcoin is a peer to peer electronic cash system that does not depend on a centralized trusted server such as a bank.

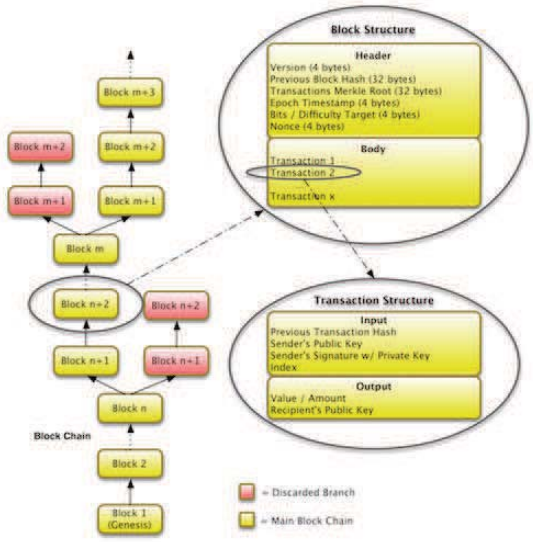
Bitcoin Transaction

A user makes electronic transactions by digitally signing a message with her secret key.



Bitcoin Block Chain

The network maintains a publicly auditable ledger called the Bitcoin block chain. It consists of all valid transaction executed so far.



A Bitcoin Block

version	02000000
previous block hash (reversed)	17975b97c18ed1f7e255adf297599b55 330edab87803c8170100000000000000
Merkle root (reversed)	8a97295a2747b4f1a0b3948df3990344 c0e19fa6b2b92b3a19c8e6badc141787
timestamp	358b0553
bits	535f0119
nonce	48750833
transaction count	63
coinbase transaction	
transaction	
...	

Block hash
0000000000000000 e067a478024addfe cdc93628978aa52d 91fabd4292982a50

CONSTRUCTION OF A BITCOIN BLOCK

HOW BITCOIN BLOCKS ARE CONSTRUCTED

- Bitcoin blocks are constructed by miners.
- Massive computing power is required to construct a Bitcoin block.
- A miner needs to solve a puzzle in order to construct a Bitcoin block.
- Every miner gets some reward for constructing a block.

The Bitcoin mining puzzle is based on Proof of Work scheme. Presently Bitcoin network uses a mining puzzle based on Adam Back's hashcash(2002).

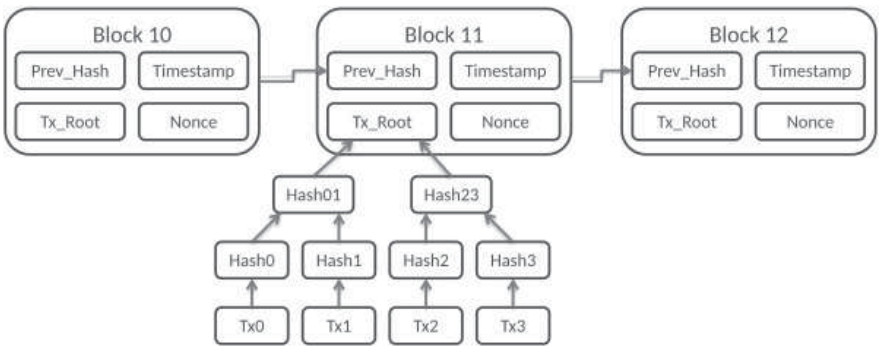
PROPERTIES OF PROOF OF WORK PUZZLE

- The Bitcoin proof of work puzzle should be non-precomputable.
- The proof of work puzzle should be feasible to compute using its input.
- The solution of the proof of work puzzle should be hard to compute.

SOLVING THE BITCOIN MINING PUZZLE

- A Bitcoin mining puzzle could be solved by computing a nonce such that the hash of the entire block header including the nonce happens to be less than a pre-defined target value.
- This nonce is computed by doing a brute force search over the set of values the nonce could take.
- This requires calculation of huge number of hashes.
- Currently the Bitcoin network calculates 400 Million GigaHashes.

A BITCOIN BLOCK



THE SHORTCOMING OF EXISTING POW SCHEME

- The proof of work scheme has high variance of solving time.
- A lucky miner could solve it even at the first attempt.

ALTERNATIVE APPROACHES

Tromp 2014

- Tromp proposed first graph theoretic proof of work.
- Tromp used detection of a subgraph in a random graph.
- He used detection of cycle in large random graph for proof of work.
- He did not mention how to apply it to the Bitcoin system.

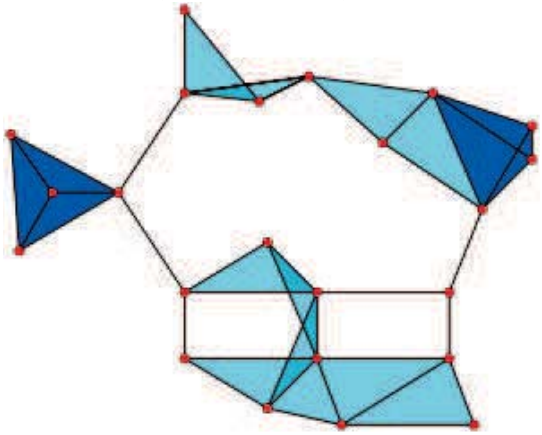
Can we do better?

Our Solution

- We propose to use clique problem as a replacement Bitcoin proof of work scheme.

CLIQUE

A clique is a complete subgraph of a graph



OUR PROPOSED WORK

CONSTRUCTION OF PUZZLE: RANDOM GRAPH

Let, $\tau = \{T_i : 1 \leq i \leq 2^\nu\}$ be the set of transactions that are yet to be included in a Bitcoin block in a certain epoch ε_0. Let $\mathcal{T}$ be a set such that $\mathcal{T} = \{L_{(i-1)*2^{30-\nu}+j} : L_{(i-1)*2^{30-\nu}+j} = T_i || j, 1 \leq i \leq 2^\nu, 0 \leq j \leq 2^{30-\nu} - 1\}$. Hence, $|\mathcal{T}| = 2^\nu * 2^{30-\nu} = 2^{30} = n$(say). We construct a graph $\mathcal{G}_{n,p} = (V, E)$ such that $V = \mathcal{T}, n = |\mathcal{T}|$. We define the set of edges as $E = \{(u, v) : u, v \in V, H(u||v) = 0^{12}||\{0,1\}^*\}$, where $H(\cdot)$ is a hash function. Hence, a single hash needs to be calculated to check the existence of an edge between a pair of vertices. Also, the probability of occurrence of an edge between any two vertices is given by $p = \frac{1}{2^{12}}$.

SOLUTION OF PUZZLE

Our proposed proof of work scheme is to compute the largest clique in the random graph $\mathcal{G}_{n,p}$.

RATIONALE BEHIND SELECTING CLIQUE PROBLEM

- Finding largest clique is proven to be NP-Hard.
- Only algorithms based on heuristic can solve it in polynomial time.

SIZE OF MAXIMUM CLIQUE OF A RANDOM GRAPH

THEOREM

Let $Z(n,p)$ denote the size of largest complete subgraph of a graph having n vertices and if p be the probability of occurrence of an edge between any two vertices, The sequence $\{Z(n,p)\}$ of random variables satisfies

$$\lim_{n\longrightarrow\infty} Pr[Z(n,p) \rightarrow \frac{2\log n}{\log\frac{1}{p}} + O(\log n)] = 1$$

THEOREM

For any $\epsilon > 0$,

$$\left\{\sum_{j=\max(0,2k-n)}^{k} \frac{\binom{n-k}{k-j}\binom{k}{j}}{\binom{n}{k}} p^{-j(j-1)/2}\right\}^{-1} \leq Prob[Z(n,p) \geq k] \leq \binom{n}{k} p^{k(k-1)/2}$$

channel hopping
Proposed Scheme

NEW POW SCHEME FOR BITCOIN MINING

- **Setup** $\tau = \{T_1, T_2, \ldots, T_{2^\nu}\}$ are the 2^ν transactions occurring in a particular epoch. $\mathcal{T} = \{L_{(i-1)*2^{30-\nu}+j} : L_{(i-1)*2^{30-\nu}+j} = T_i||j, 1 \leq i \leq 2^\nu, 0 \leq j \leq 2^{30-\nu}-1\}$. $V = \mathcal{T}$.
- **Proof of Work Computation** In a scratch off attempt a miner computes a $V \times V$ matrix B as follows:

$$B[i][j] = B[j][i] = \begin{cases} * & \text{if } i = j \\ 1 & \text{if } H(v_i||v_j) \in \{0^{12}||\{0,1\}^*\} \\ 0 & \text{elsewhere} \end{cases}$$

  where $v_i, v_j \in V, i \leq j$. The user finds the largest square submatrix B' of B such that each and every entry of B' is a 1 except the diagonal elements. If she finds such a submatrix then she broadcasts it along with the index set I.
- **Verification** The verifier checks whether $\forall i, j \in I, i \neq j$

$$H(v_i||v_j) \in \{0^{12}||\{0,1\}^*\}$$

  If the Verification is successful then the user caches the proof of work and the block generated by the prover. After a predefined time has elapsed the verifier checks the cache and selects a proof with the highest size of B'. Ties are broken by selecting the proof that came first and the corresponding block is accepted and added to the Bitcoin block chain.

FIGURE: An alternative proof of work scheme for Bitcoin mining

NEW POW SCHEME FOR BITCOIN MINING

- **Setup** Same as Algorithm 1.
- **Proof of Work Computation** In a scratch off attempt a miner computes a $V \times V$ matrix B as follows: For $i \leq j$

$$B[i][j] = B[j][i] = \begin{cases} * & \text{if } i = j \\ 1 & \text{if } H(v_i||v_j) \in \{0^{12}||\{0,1\}^*\} \\ 0 & \text{elsewhere} \end{cases}$$

where $v_i, v_j \in V, i \leq j$. The user finds the largest square submatrix B'_0 of B such that each and every entry of B'_0 is a 1 except the diagonal elements. If she finds such a submatrix then she broadcasts it along with the index set I_0.
Step $r = 1, \ldots$ do {
if the epoch has ended then exit
else find another square submatrix from B such that $[B'_r] = [B'_0]$. If such a matrix could be found then broadcast it along with the index set I_r.
}

- **Verification** The verifier holds separate caches for storing the solutions from different miners. The verifier checks every submatrix for being a valid clique as shown in Algorithm 1. If the Verification is successful then the user caches the proof of work along with the index set which corresponds to the set of vertices The verifier rejects all PoW whose index sets have a non-null intersection with a PoW present in its cache. After the epoch has expired the verifier checks its cache and selects a miner who has so far sent the highest number of proofs(largest cliques). Ties are broken arbitrarily and the corresponding block is accepted and added to the Bitcoin block chain.

FIGURE: An alternative proof of work scheme for Bitcoin mining

SIZE OF LARGEST CLIQUE

THEOREM

Let $Z(n,p)$ denote the size of largest complete subgraph of a graph $\mathcal{G}_{n,p}$. The sequence $\{Z(n,p)\}$ of random variables satisfies

$$\lim_{n \longrightarrow \infty} Pr[Z(n,p) \rightarrow \frac{2\log n}{\log \frac{1}{p}} + O(\log n)] = 1$$

In our graph $\mathcal{G}_{n,p}$, $n = 2^{30}$, $p = \frac{1}{2^{12}}$.
So, the clique number $= 5 + 30c, c \ll 1$.

Solving our PoW using Brute Force Search

Lemma

If k be the size of the largest clique in our graph then the expected amount of computation needed to find the first largest clique in brute force search method is of the order of $2^{6k(k-1)}$.

Proof.

Let us assume that the clique number k is known to the miner. Also let N^k be the number of iteration required to find the first clique of size k. The miner's strategy as follows:
Step 1: $\Omega = \emptyset$.
For each Step i, The miner chooses a set Γ_i of k vertices randomly from the set of vertices V such that $\Gamma_i \notin \Omega$. If the vertices in Γ_i form a clique then she outputs it. Else she updates Ω as $\Omega = \Omega \cup \Gamma_i$. Now, The Expected number of iterations needed to find the first clique is

$$E(N^k) = \sum_{i=1}^{\infty} i(1-r)^{i-1} r,$$

where $r = p^{\frac{1}{2}k(k-1)}$. So, $E(N^k) = r\sum_{i=1}^{\infty} i(1-r)^{i-1} = \frac{1}{r}$. Since, in our scheme $p = \frac{1}{2^{12}}$, the lemma follows. ■ □

Average amount of computation needed to find a largest clique

in our model, the value of k roughly equals to 5 or 6, so, the average amount of computation required is of the order of 2^{120} or 2^{180} if brute force search is applied.

A SIMPLE PARALLEL ALGORITHM TO FIND A LARGEST CLIQUE

INPUT : Let, $P_m, 1 \leq m \leq c$ be the id of a miner of some pool having c members. $LH_T : \{0,1\}^{T+x} \rightarrow \{0,1\}^T$ is a function that outputs the most significant T bits of any binary bit string of length at least T. n is the number of vertices in the graph $\mathcal{G}_{n,p}$. Each vertex is represented by $\lceil \log n \rceil$ bits. $\mathcal{H}(\cdot)$ is SHA-256 hash function.

OUTPUT : Find a clique of size k in the graph $\mathcal{G}_{n,p}$.

P_m chooses a random nonce s_m.
while epoch has not expired **do**
  Calculate $\Lambda = LH_{k \log n}(\mathcal{H}(m||s_m))$ as the left most $k\lceil \log n \rceil$ bits of $\mathcal{H}(m||s_m)$.
  Choose k vertices $\langle V_1, V_2, \ldots, V_k \rangle$ using the bitstring Λ.
  if $V_1, V_2, \ldots, V_k$ form a clique **then**
    output it and exit.
  end if
end while

SOME RESULT

LEMMA

The Algorithm 3 takes $O(2^{6k(k-1)})$ time to find the first clique of size k if $k < 6$.

TOWARDS BETTER ALGORITHMS...

LEMMA

The computation needed to find all cliques of size $\kappa+1$, if it exists given the set of all cliques of size κ is $O(2^{30+6k(6-\kappa)}\kappa)$.

PROOF.

The expected number of cliques of size κ is $\omega = \binom{n}{\kappa} p^{\frac{1}{2}\kappa(\kappa-1)}$. In our case, $n = 2^{30}, p = \frac{1}{2^{12}}$. So,
$\omega = \binom{2^{30}}{\kappa} / 2^{6\kappa(\kappa-1)} \approx \frac{2^{30\kappa}}{2^{6\kappa(\kappa-1)}} = 2^{6\kappa(5-\kappa+1)} = 2^{6\kappa(6-\kappa)}$. Now, from this set of candidate cliques of size κ, a clique of size $\kappa+1$ can be found using the above method. So, the total computation needed is $\omega * 2^{30} * \kappa$. Hence, proved. □

INPUT : $\mathcal{C}_3$.
OUTPUT : List all largest clique of the random graph.

```
for i := 3; C_i ≠ ∅; i++ do
  while C_i is not empty do
    Choose a clique ⟨v_1, v_2, ..., v_i⟩ ∈ C_i
    for all v ∈ V \ {v_1, v_2, ..., v_i} do
      if ⋃_{j=1}^{i} {v_i, v} ⊂ E then
        C_{i+1} = C_{i+1} ∪ ⟨v_1, v_2, ..., v_i, v⟩.
      end if
    end for
    C_i = C_i \ ⟨v_1, v_2, ..., v_i⟩.
  end while
end for
```

FIGURE: An Algorithm to list all largest cliques

RESULT

Lemma

The computational complexity of Algorithm 4 is $O(2^{85.58})$

Proof.

The proof can be found in the paper. □

IMPROVEMENT OF COMPLEXITY OF THE ALGORITHM

Further reduction of the computational complexity of finding $\mathcal{C}_6$ can be done by computing all K_4s of the graph as follows:
1) Find all quadrangles of the graph using the Chiba & Nishizeki Algorithm.
2) Find all K_4s using the set of quadrangles.
3) Then using the set of all K_4s ($\mathcal{C}_4$) try to find the set of all K_5s $\mathcal{C}_5$.
4) Then using the set of all K_5s ($\mathcal{C}_6$) try to find the set of all K_6s $\mathcal{C}_6$.

Lemma

Let $A[]$ be a sorted array of all K_4s such that for every $i \in [|\mathcal{C}_4|]$, if $A[i] = \{v_1, v_2, v_3, v_4\}$, then $v_1 < v_2 < v_3 < v_4$ and for every $i, j \in |\mathcal{C}_4|, i < j$ if $A[i] = \{v_1, v_2, v_3, v_4\}$ and $A[j] = \{v'_1, v'_2, v'_3, v'_4\}$, then $\exists l \in \{1, 2, 3, 4\}$ such that $v_m = v'_m, \forall 1 \leq m < l$ and $v_l < v'_l$. Then finding every pairs $(i, j), i \neq j$ such that $A[i] = \{v_1, v_2, v_3, v_4\}$ and $A[j] = \{v_1, v_2, v_3, v'_4\}$ takes $O(|A|)$ time.

Proof.

For any arbitrary $i \in [|\mathcal{C}_4|]$, Let $A[i] = \{v_1, v_2, v_3, v_4\}$. Therefore, all other K_4s like $\{v_1, v_2, v_3, z\}$ if there are any will be adjacent to $A[i]$ in the array. If there are Δ such cliques $A[j]$ whose first three vertices are same as that of $A[i]$, then all such pairs (i, j) could be found in $O(\Delta^2)$ time. Whether any of these pairs could be extended to form a K_5 can be ascertained by computing a single hash per such pair. Now, for any three vertices $\{u, v, w \in V\}$, $\Delta = |\{i : i \in [|\mathcal{C}_4|], \{u, v, w\} \subset A[i]\}|$. So, $E(\Delta) = 2^{30} * (\frac{1}{2^{12}})^3 = \frac{1}{2^6}$. Now, the total time needed to find every pairs $(i, j), i \neq j$ such that $A[i] = \{v_1, v_2, v_3, v_4\}$ and $A[j] = \{v_1, v_2, v_3, v'_4\}$ is $O\left(|A| \left\{\max(1, E^2(\Delta))\right\}\right)$ or $O(|A|)$. □

Theorem

There exists an Algorithm that outputs the set of maximum cliques in $O(2^{70.5})$ time and $O(2^{48})$ space.

Proof.

The proof of this theorem can be found in the paper. □

PREVENTING THEFT OF PROOF OF WORK

- This PoW scheme is vulnerable to theft of solution.
- Any malicious miner may steal the PoW solution and relay it as her own solution.

Addressing the shortcoming

Construction of puzzle: random graph

Let, $\tau = \{T_i : 1 \leq i \leq 2^{\nu}\}$ be the set of transactions that are yet to be included in a Bitcoin block in a certain epoch ε_0. Let $\mathcal{T}$ be a set such that $\mathcal{T} = \{L_{(i-1)*2^{30-\nu}+j} : L_{(i-1)*2^{30-\nu}+j} = T_i||j, 1 \leq i \leq 2^{\nu}, 0 \leq j \leq 2^{30-\nu} - 1\}$. Hence, $|\mathcal{T}| = 2^{\nu} * 2^{30-\nu} = 2^{30} = n(\text{say})$. We construct a graph $\mathcal{G}_{n,p} = (V, E)$ such that $V = \mathcal{T}, n = |\mathcal{T}|$. We define the set of edges as $E = \{(u, v) : u, v \in V, H(u||v) = 0^{12}||\{0,1\}^*\}$, where $H(\cdot)$ is a hash function. Hence, a single hash needs to be calculated to check the existence of an edge between a pair of vertices. Also, the probability of occurrence of an edge between any two vertices is given by $p = \frac{1}{2^{12}}$.

channel hopping
Proposed Scheme

Addressing the shortcoming

Construction of puzzle: random graph

Let, $\tau = \{T_i : 1 \leq i \leq 2^{\nu}\}$ be the set of transactions that are yet to be included in a Bitcoin block in a certain epoch ε_0. Let $\mathcal{T}$ be a set such that $\mathcal{T} = \{L_{(i-1)*2^{30-\nu}+j} : L_{(i-1)*2^{30-\nu}+j} = T_i||pk||j, 1 \leq i \leq 2^{\nu}, 0 \leq j \leq 2^{30-\nu} - 1\}$. Hence, $|\mathcal{T}| = 2^{\nu} * 2^{30-\nu} = 2^{30} = n(\text{say})$. We construct a graph $\mathcal{G}_{n,p} = (V, E)$ such that $V = \mathcal{T}, n = |\mathcal{T}|$. We define the set of edges as $E = \{(u, v) : u, v \in V, H(u||v) = 0^{12}||\{0,1\}^*\}$, where $H(\cdot)$ is a hash function. Hence, a single hash needs to be calculated to check the existence of an edge between a pair of vertices. Also, the probability of occurrence of an edge between any two vertices is given by $p = \frac{1}{2^{12}}$.

NEW POW SCHEME FOR BITCOIN MINING

- **Setup** $\tau = \{T_1, T_2, \ldots, T_{2^\nu}\}$ are the 2^ν transactions occurring in a particular epoch. $\mathcal{T} = \{L_{(i-1)*2^{30-\nu}+j} : L_{(i-1)*2^{30-\nu}+j} = T_i||pk||j, 1 \leq i \leq 2^\nu, 0 \leq j \leq 2^{30-\nu} - 1\}$. $V = \mathcal{T}$.
- **Proof of Work Computation** In a scratch off attempt a miner computes a $V \times V$ matrix B as follows:

$$B[i][j] = B[j][i] = \begin{cases} * & \text{if } i = j \\ 1 & \text{if } H(v_i||v_j) \in \{0^{12}||\{0,1\}^*\} \\ 0 & \text{elsewhere} \end{cases}$$

where $v_i, v_j \in V, i \leq j$. The user finds the largest square submatrix B' of B such that each and every entry of B' is a 1 except the diagonal elements. If she finds such a submatrix then she broadcasts it along with the index set I.

- **Verification** The verifier checks whether $\forall i, j \in I, i \neq j$

$$H(v_i||v_j) \in \{0^{12}||\{0,1\}^*\}$$

If the Verification is successful then the user caches the proof of work and the block generated by the prover. After a predefined time has elapsed the verifier checks the cache and selects a proof with the highest size of B'. Ties are broken by selecting the proof that came first and the corresponding block is accepted and added to the Bitcoin block chain.

FIGURE: An alternative proof of work scheme for Bitcoin mining

channel hopping
Proposed Scheme

LEMMA

Let C be the incidence matrix of a random graph $R(2^{30}, \frac{1}{2^{12}})$ and B be the incidence matrix of the graph of the PoW puzzle. For every probabilistic polynomial time distinguisher Δ,

$$|Pr[\Delta(B) = 1] - Pr[\Delta(C) = 1]| \leq \epsilon$$

provided $H()$ is a secure hash function.

Proof.

If the above lemma is not true, then the hash function will be such that for any two vertices, $v_1, v_2, Pr[e(v_1, v_2)] \neq \frac{1}{2^{12}}$. So, for a pair of any $T_1, T_2 \in \tau, 0 \leq m, n \leq 2^{30-\nu} - 1, Pr\left[H((T_1||pk||m)||(T_2||pk||n)) = 0^{12}||\{0,1\}^*\right] \neq \frac{1}{2^{12}}$. So, there exists x, y such that $Pr[H(x) = a] \neq Pr[H(y) = a]$, for some a belonging to the domain of the output of the hash function. This shows that $H()$ is not a secure hash function. So, our assumption is incorrect.

□

Lemma

The miner cannot compute transactions $T_1, T_2, \ldots, T_5$,(not all necessarily distinct), such that $H((T_i||pk||j)||(T_k||pk||l)) = 0^{12}||\{0,1\}^*, 1 \leq i, k \leq 5, i < k, 0 \leq j, l \leq 2^{30-\nu}, (T_i, j) \neq (T_k, l)$.

Proof.

The miner will need to find transactions $T_1, T_2, \ldots, T_5$(not all distinct) such that there exists $\{j_1, j_2, \ldots, j_5\}$, $0 \leq j_i \leq 2^{30-\nu}$, such that $H((T_i||pk||j_i)||(T_k||pk||j_k)) = 0^{12}||\{0,1\}^*, \forall i, k \in \{1,2,3,4,5\}, i < k$. For any $i, k \in \{1,2,\ldots,5\}$ and $0 \leq j_i, j_k \leq 2^{30-\nu}$, $Pr\left[H((T_i||pk||j_i)||(T_k||pk||j_k)) = 0^{12}||\{0,1\}^*\right] = \frac{1}{2^{12}}$. So, finding two transactions T_i, T_k such that $H((T_i||pk||j_i)||(T_k||pk||j_k)) = 0^{12}||\{0,1\}^*, 1 \leq i, k \leq 5, i < k, 0 \leq j_i, j_k \leq 2^{30-\nu}, (T_i, j_i) \neq (T_k, j_k)$ takes 2^{12} hash computations. So, finding a set of transactions $\{T_1, T_2, \ldots, T_5\}$(not all necessarily distinct) and indices $\{j_1, j_2, \ldots, j_5\}$ (not all necessarily distinct) such that $H((T_i||pk||j_i)||(T_k||pk||j_k)) = 0^{12}||\{0,1\}^*$, $1 \leq i, k \leq 5, i < k, j_i, j_k \in \{j_1, j_2, \ldots, j_5\}, (T_i, j_i) \neq (T_k, j_k)$ takes on an average $(2^{12})^{\binom{5}{2}} = 2^{120}$ hash computations. As we shall see later, this is higher than the computation required to find a maximum clique in the graph.

□

CONCLUSION

IN THIS WORK

- we have proposed a new PoW scheme.
- we have shown how it can be used for mining Bitcoin.
- we have discussed algorithms to solve the PoW.
- we have made this scheme resilient against theft of solution.

Thank You

Panel Discussion
Decentralized Control : Reason and Objective

Masashi Une
Center for Information Technology Studies (CITECS)
Institute for Monetary and Economic Studies (IMES)
Bank of Japan

Before short presentation

- The views expressed throughout this panel discussion are those of the speaker and do not necessarily reflect the official views of the Bank of Japan.

Introduction

- IMES is
  - an internal organization of the Bank of Japan to conduct a wide range of studies on monetary and economic issues with the aim of establishing an appropriate background for the Bank's policies.
- CITECS is
  - in charge of studies on information security issues relating to the financial services.
    - Topics of interest: Internet/mobile banking security, biometrics-based authentication, smart card security, artifact-metrics, …

Minimum security requirements

- Objective:
  - Develop next-generation cryptographic schemes (NGC) for decentralized control in such a way to widely and continuously used in various applications for a long time.
- Users may need to
  - Understand the security properties of NGC in the corresponding application
    - In general, almost all of the users are likely to be non-experts.
  - Confirm how to deal with serious problems such as security incidents
- The following are minimum security requirements from the users' viewpoints.

1. Clarification of security properties
2. Clarification of impact of security incidents
3. Clarification of how to deal with security events

Clarification of security properties

- (1) Define operational environment
  - Define entities and their roles in the application
  - Define threats to be assumed
- (2) Evaluate the security properties
  - Evaluate the security of cryptographic primitives and schemes
    - Digital signature, hash function, random number generator, …
    - Unforgeability, non-malleability, …
  - Evaluate the security of a specific application system in which the cryptographic primitives and schemes are utilized

Clarification of impact of security incidents

- (3) Identify security incidents
  - Decrease of security of cryptographic primitives over time
    - Improvement of performance of computer and cryptanalysis techniques
  - Compromise of secret signing key
    - Inappropriate operation, malware, unauthorized access from the Internet, …
  - Scope of this process depends on criticality, information asset and system life of the crypto application.
- (4) Clarify impact of the security incidents
  - All of previous transactions are still secure?
  - Asset is also still valid?
  - Security risk should be evaluated.

Clarification of how to deal with security events

- (5) Develop a system migration plan
  - Replace current crypto primitives with more secure ones
  - Modify security parameters such as key length
  - Update software and hardware utilized in the application
  - Discuss when the migration should be completed
- (6) Develop a security recovery plan
  - Detect and trace malicious behavior
  - Discuss and apply countermeasure
    - Including immediate replacement of crypto primitives

My interest: discussion points

- Which requirement can be achieved in the existing cryptographic schemes?
  - Requirement 1: already discussed?
  - Requirements 2 and 3: future topics?
- Who plays a role to discuss and achieve these requirements?
  - Does a single entity play the role under “decentralized control”?
    - A group of entities do so?
  - Does a single authority play the role among “multi-authority”?
    - A group of authorities do so?

「マス・フォア・インダストリ研究」シリーズ刊行にあたり

本シリーズは，平成 23 年 4 月に設立された九州大学マス・フォア・ インダストリ研究所(IMI)が，平成 25 年 4 月に共同利用・共同研究拠点「産業数学の先進的・基礎的共同研究拠点」として，文部科学大臣より認定を受けたことにともない刊行するものである．本シリーズでは，主として，マス・フォア・インダストリに関する研究集会の会議録，共同研究の成果報告等を出版する． 各巻はマス･フォア･インダストリの最新の研究成果に加え，その新たな視点からのサーベイ及びレビューなども収録し，マス・フォア・インダストリの展開に資するものとする．

平成 26 年 10 月
マス・フォア・インダストリ研究所
所長 福本康秀

Next-generation Cryptography for Privacy Protection and Decentralized Control and Mathematical Structures to Support Techniques

マス・フォア・インダストリ研究 No.4, IMI, 九州大学
ISSN 2188-286X
発行日　2016 年 1 月 29 日
編　集　穴田啓晃, 安田貴徳, 櫻井幸一, 寺西勇
発　行　九州大学マス・フォア・インダストリ研究所
〒819-0395 福岡市西区元岡 744
九州大学数理・IMI 事務室
TEL 092-802-4402　FAX 092-802-4405
URL http://www.imi.kyushu-u.ac.jp/
印　刷　社会福祉法人　福岡コロニー
〒811-0119　福岡県糟屋郡新宮町緑ケ浜 1 丁目 11 番 1 号
TEL 092-962-0764　FAX 092-962-0768

シリーズ既刊

Issue	Author ／ Editor	Title	Published
マス・フォア・インダストリ研究　No.1	穴田 啓晃 安田 貴徳 Xavier Dahan 櫻井 幸一	Functional Encryption as a Social Infrastructure and Its Realization by Elliptic Curves and Lattices	26 February 2015
マス・フォア・インダストリ研究　No.2	滝口 孝志 藤原 宏志	Collaboration Between Theory and Practice in Inverse Problems	12 March 2015
マス・フォア・インダストリ研究　No.3	筧 三郎	非線形数理モデルの諸相：連続，離散，超離散，その先 (Various aspects of nonlinear mathematical models : continuous, discrete, ultra-discrete, and beyond)	24 March 2015

Institute of Mathematics for Industry
Kyushu University

九州大学マス・フォア・インダストリ研究所

〒819-0395 福岡市西区元岡744
URL http://www.imi.kyushu-u.ac.jp/