2010 Japan-Indo Joint Workshop on Cryptology and Related Areas

Date: July 8, 2010

Venue: At AIST Akihabara Site, Tokyo Japan

Sponsors: JST and DST

PREFACE

This document is a record of the second joint workshop of the following three Japan-India projects:

- **"Analysis of Cryptographic Algorithms and Evaluation on Enhancing Network Security Based on Mathematical Science"** leaded by Bimal Roy, Professor, Applied Statistics Unit, Indian Statistical Institute and Kouichi SAKURAI, Professor, Computer Science and Comm. Engineer, Kyushu University;
- **"Security Proofs and Multidisciplinary Evaluation for Dynamic Hierarchical Key Assignment Schemes"** leaded by Kanta MATSUURA, Associate Professor, Institute of Industrial Science, University of Tokyo and Anish Mathuria, Professor, Dhirubhai Ambani Institute of ICT;
- **"Security Evaluation and Design of Components and Cryptographic Primitives for RFID and Sensor Networks"** leaded by Hajime WATANABE, Professor, Research Center for Information Security, National Institute of Advanced Industrial Science and Technology and Sugata Gangopadhyay, Assistant Professor, Department of Mathematics, Indian Institute of Technology Roorkee.

The projects are parts of the Strategic Japanese-Indian Cooperative Programme on "Multidisciplinary Research Field, which combines Information and Communications Technology with Other Fields (Multidisciplinary ICT)" sponsored by Japan Science and Technology Agency ("JST") and Department of Science and Technology ("DST") of the Government of India for the three years period 2009-2011. The aim of the programme is to strengthen the collaboration between Japan and India within the field of "Multidisciplinary ICT" to achieve world-class scientific results, leading towards new innovative technologies

The three projects are dedicated to research activities within a hot topic of establishing trustful and secure information society based on information-communication technologies which is one of the international priorities as well as of Japan and India. Additional information on the projects is available at:

http://www.rcis.aist.go.jp/project/JST-DST/index-en.html and
http://itslab.csce.kyushu-u.ac.jp/JIP/en/index.html ;

The first joint workshop was in New Delhi, India, on December 12, 2009. The second workshop held in Tokyo on July 08, 2010, was a forum for further exchange of the research ideas and results, fruitful research discussions, as well as stimulation for further joint research activities between the institutions in India and Japan.

Program

2010 Japan-Indo Joint Workshop on Cryptology and Related Areas

July 8, 2010, At AIST Akihabara Site, Tokyo Japan, Sponsored by JST and DST

10:00 - 10:30 Opening

10:30 - 12:00 Session 1:

Analysis of Cryptographic Algorithms and Evaluation on Enhancing Network Security Based on Mathematical Science

1. Information Theoretic Discussions on Perfectly Secure Multi-use Multi-secret Sharing Scheme, Avishek Adhikari (Univ. of Calcutta)
2. Improved Subset Difference Method based on Ternary Tree, Kazuhide Fukushima (KDDI Labs)
3. On Deployment of Sensors, Bimal Roy (ISI, Kolkata)
4. Non-committing Encryption Scheme Based on DDH Assumption, Takashi Nishide (Kyushu Univ.)

12:00 - 13:00 Lunch

13:00 - 14:30 Session 2:

Security Proofs and Multidisciplinary Evaluation for Dynamic Hierarchical Key Assignment Schemes

1. A Framework for Choosing Security Modules, Kanta Matsuura (Univ. of Tokyo)
2. Nominative Signatures, Jacob Schuldt (Univ. of Tokyo)
3. On the Security Proof of Wu-Wei Hierarchical Key Assignment Scheme, Murali Medisetty (DA-IICT)

14:30 - 15:00 Coffee Break

15:00 - 16:30 Session 3:

Security Evaluation and Design of Components and Cryptographic Primitives for RFID and Sensor Networks

1. A Generic Weakness of the k-normal Boolean Functions Exposed to Dedicated Algebraic Attack, Miodrag Mihaljevic (AIST), Goutam Paul (Jadavpur Univ.) and Sugata Gangopadhyay (IIT Roorkee)
2. RFID Authentication Using Quasi-Dyadic Fix Domain Shrinking, Yang. Cui (AIST)
3. Computationally Secure Communication in the Wire-tap Scenario, Kirill Morozov (AIST)
4. A Low Complexity Encryption Technique Based on Joint Employment of Pseudorandomness, Randomness and Coding, Miodrag Mihaljevic (AIST)

16:30 - 17:00 Coffee Break

17:00 - 18:00 Round Table Discussion and Closing

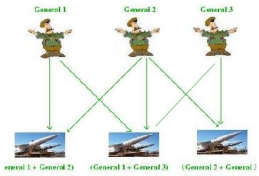
Information theoretic discussions on perfectly secure multi-use multi-secret sharing scheme

Avishek Adhikari

Research Team Members :
Partha Sarathi Roy, Angsuman Das

Department of Pure Mathematics
University of Calcutta, Kolkata, India.

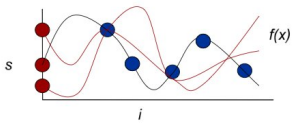
What is secret sharing?



Formally, a *secret sharing scheme for general access structure* is a method of sharing a secret K among a finite set of participants $\mathcal{P} = \{P_1, P_2, \ldots, P_n\}$ in such a way that

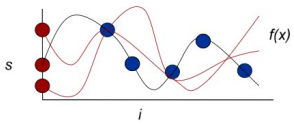
1. if the participants in $\mathcal{A} \subseteq \mathcal{P}$ are qualified to know the secret, then by pooling together their partial information, they can reconstruct the secret K,
2. any set $\mathcal{B} \subset \mathcal{P}$ which is not qualified to know K, cannot reconstruct the secret K.

Shamir's (k, n)-Secret Sharing Scheme



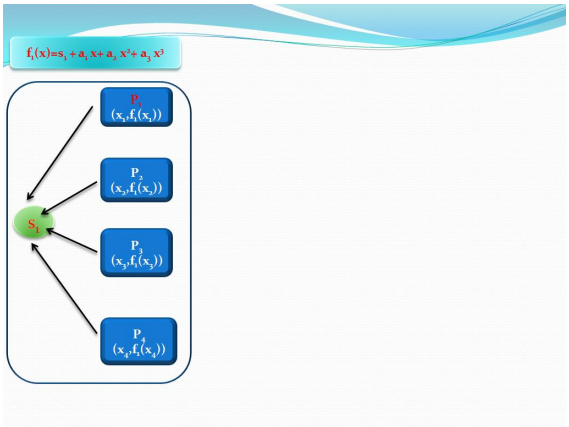
- It takes two points to define a straight line, three points to fully define a quadratic, four points to define a cubic, and so on.
- One can fit a unique polynomial of degree $(t-1)$ to any set of t points that lie on the polynomial.

Shamir's (k, n)-Secret Sharing Scheme

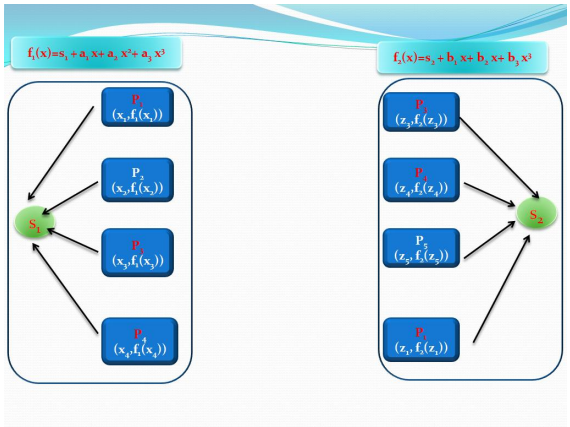


- It takes two points to define a straight line, three points to fully define a quadratic, four points to define a cubic, and so on.
- One can fit a unique polynomial of degree $(t-1)$ to any set of t points that lie on the polynomial.

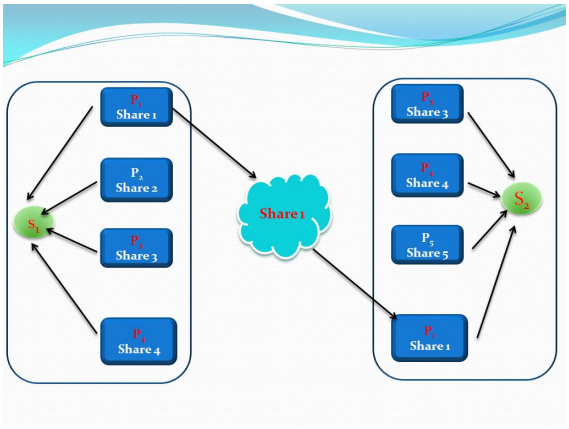
Shamir's Sharing Scheme



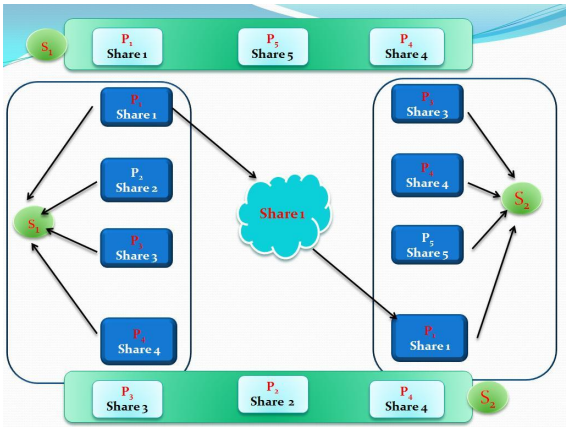
Multi-Secret Sharing Scheme



Multi-Secret Sharing Scheme with one share



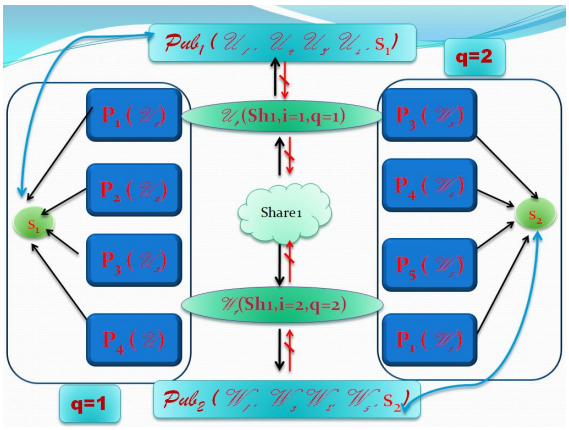
Multi-Secret Sharing Scheme



Issues to be discussed

- Multiple secrets.
- Re-usability of the same shares.
- Renewable of the secrets, participants.
- Verifiable.
- Size of the Shares.

Multi-use Multi-Secret Sharing Scheme



Information Theoretical Discussions: The Model

- In this talk, an information theoretical framework for perfectly secure multi-use multi-secret sharing schemes, in which each participant has to carry only one share, is build.
- In this model the shares carried by each participant is independent of the secrets.
- Qualified set of participants can reconstruct the corresponding secret with the knowledge of pseudo shares which are generated from shares with the help of some public entities that depend on the secrets and qualified set of participants.
- a k-tuple of secrets $(s_1, s_2, \ldots, s_k) \in S_1 \times S_2 \times \ldots \times S_k$ are shared in k-tuple of access structures $\Gamma_{s_1} \times \Gamma_{s_2} \times \ldots \times \Gamma_{s_k}$ on $\mathcal{P}$ in such a way that, for each $i = 1, 2, \ldots, k$, the access structure $\Gamma_{s_i} = \{A_1^{s_i}, A_2^{s_i}, \ldots, A_{l_i}^{s_i}\}$, where $A_q^{s_i} = \{P_1^{i_q}, P_2^{i_q}, \ldots, P_{m_{i_q}}^{i_q}\} \subseteq \{P_1, P_2, \ldots, P_n\}$ and $q = 1, 2, \ldots, l_i$, is the collection of the set of all subsets of $\mathcal{P}$ that can recover the secret $s_i \in S_i$, where $i = 1, 2, \ldots, k$.

Information Theoretical Discussions: The Notations

- A boldface capital letter, say $\mathbf{X}$, denotes a random variable that takes values on a set, denoted by the corresponding capital letter X according to some probability distribution $\{Pr_{\mathbf{X}}(x)\}_{x \in X}$.
- The values that a random variable can take are denoted by the corresponding lower-case letter.
- Given a random variable $\mathbf{X}$, let $H(\mathbf{X})$ denote the Shannon entropy of $\{Pr_{\mathbf{X}}(x)\}_{x \in X}$.
- Let d be an arbitrary positive integer and let $\mathbf{X_1}, \ldots, \mathbf{X_d}$ be d random variables taking values on the sets $X_1, \ldots, X_d$, respectively. For any subset $V = \{i_1, \ldots, i_V\} \subseteq \{1, \ldots, d\}$, we denote with X_V , the set $X_{i_1} \times \ldots \times X_{i_V}$.

Information Theoretical Discussions: The Notations

- $U_{P_b^{i_q}}$ denotes the set of all possible pseudo shares of the participant $P_b^{i_q} \in \mathcal{A}_q^{s_i}$ corresponding to the q-th qualified set of the i-th secret
- $U_{\mathcal{A}_q^{s_i}}$ denotes the Cartesian product $U_{P_1^{i_q}} \times \ldots \times U_{P_{m_{i_q}}^{i_q}}$.
- Note that as pseudo share depends on share as well as the access structure, probability distribution on share space SH naturally induces a probability distribution on $U_{\mathcal{A}}$ and it is denoted by $\{Pr_{\mathbf{U}_{\mathcal{A}}}(u)\}_{u \in U_{\mathcal{A}}}$, where $\mathcal{A} \subseteq \mathcal{P}$.
- Finally, $H(\mathbf{U}_{\mathcal{A}})$ denotes the entropy of $\{Pr_{\mathbf{U}_{\mathcal{A}}}(u)\}_{u \in U_{\mathcal{A}}}$.

Definition of perfectly secure multi-use multi-SSS

Definition

A perfectly secure multi-use multi-SSS for $(\Gamma_{s_1}, \ldots, \Gamma_{s_k})$ is a sharing of the secrets $(s_1, \ldots, s_k) \in S_1 \times \ldots \times S_k$ in such a way that

1. Any subset of participants qualified to recover a secret can compute the secret. Formally, if $\mathcal{A}_q^{s_i} \in \Gamma_{s_i}$, then for all $u \in U_{\mathcal{A}_q^{s_i}}$ with $Pr_{\mathbf{U}_{\mathcal{A}_q^{s_i}}}(u) > 0$, it holds $Pr(s_i|u) = 1$, i.e., the values of pseudo shares held by $\mathcal{A}_q^{s_i} \in \Gamma_{s_i}$ completely determine the secret s_i.
2. Any forbidden set of participants, even knowing other secrets, has no more information about the secret other than the information given by the known secrets. Formally, if $\mathcal{A} \notin \Gamma_{s_i}$, then for all $u \in U_{\mathcal{A}}$ and $t \subseteq \{s_1, \ldots, s_k\} \setminus \{s_i\}$, it holds that $Pr(s_i|ut) = Pr(s_i|t)$, i.e., the probability that a secret is equal to s_i given any subset t of secrets excluding s_i and the set u of pseudo shares held by $\mathcal{A} \notin \Gamma_{s_i}$ is same as the probability of the secret s_i given t.

Entropy approach: Definition

Definition

A perfectly secure multi-use multi-secret sharing scheme for $(\Gamma_{s_1}, \ldots, \Gamma_{s_k})$ is a sharing of the secrets $(s_1, \ldots, s_k) \in S_1 \times \ldots \times S_k$ in such a way that

1. Any subset of participants qualified to recover a secret can compute the secret. Formally, for all $\mathcal{A}_q^{s_i} \in \Gamma_{s_i}$, it holds $H(\mathbf{S_i}|\mathbf{U}_{\mathcal{A}_q^{s_i}}) = 0$, where $i = 1, 2, \ldots, k$, $q = 1, 2, \ldots, l_i$, $b = 1, 2, \ldots, m_{i_q}$ i.e., set of values of pseudo shares in $U_{\mathcal{A}_q^{s_i}}$ corresponds to a unique value of the secret.
2. Any subset of participants not qualified to recover a secret, even knowing other secrets, has no more information about the secret other than the information given by the known secrets. Formally, for all $\mathcal{A} \notin \Gamma_{s_i}$ and $T \subseteq \{S_1, \ldots, S_k\} \setminus \{S_i\}$, it holds that $H(\mathbf{S_i}|\mathbf{U}_{\mathcal{A}}\mathbf{T}) = H(\mathbf{S_i}|\mathbf{T})$, where $i = 1, 2, \ldots, k$ i.e., $\mathbf{S_i}$ and $\mathbf{U}_{\mathcal{A}}$ are statistically independent given the secrets in T, for $q = 1, 2, \ldots, l_i$.

Bounds on the Size of Shares and Pseudo Shares

Lemma

For all $\mathcal{A}_q^{s_i} \in \Gamma_{s_i}$, it holds $H(\mathbf{S_i}|\mathbf{X}_{\mathcal{A}_q^{s_i}}) = 0$, where $i = 1, 2, \ldots, k$, $q = 1, 2, \ldots, l_i$.

Lemma

For all $\mathcal{A} \notin \Gamma_{s_i}$ and $T \subseteq \{S_1, \ldots, S_k\} \setminus \{S_i\}$, it holds that $H(\mathbf{S_i}|\mathbf{X}_{\mathcal{A}}, \mathbf{T}) = H(\mathbf{S_i}|\mathbf{T})$, where $i = 1, 2, \ldots, k$.

Theorem: Bounds on the Size of Shares

Theorem

Let $(\Gamma_{s_1}, \ldots, \Gamma_{s_k})$ be a k-tuple of access structures on the set of participants $\mathcal{P}$. Assume that for all $S_i \in \{S_1, \ldots, S_k\}$ and $T \subseteq \{S_1, \ldots, S_k\} \setminus \{S_i\}$, it holds $H(S_i|T) > 0$. If there exist a participant $P \in \mathcal{P}$ and subsets of participants $\mathcal{A}_{i_1}, \ldots, \mathcal{A}_{i_j} \subseteq \mathcal{P}$, such that $\{P\} \cup \mathcal{A}_{i_g} \in \Gamma_{s_{i_g}}$ and $\mathcal{A}_{i_g} \notin \Gamma_{s_{i_g}}$, for $1 \leq g \leq j$ and $j \leq k$, then in a multi-use multi-secret sharing scheme for $(\Gamma_{s_1}, \ldots, \Gamma_{s_k})$ the entropy of a share given to P satisfies

$$H(\mathbf{X_P}) \geq H(\mathbf{S_{i_1}}, \ldots, \mathbf{S_{i_j}}).$$

Corollary: Bounds on the Size of Shares

Corollary

Suppose $s_1, s_2, \ldots, s_k$ to be shared among n participants $P_1, \ldots, P_n$ in such a way that any set of participants with cardinality t_i is qualified to reconstruct the secret s_i. In this threshold structure it is clear from the Theorem 5 that

$$H(\mathbf{X_P}) \geq H(\mathbf{S_1}, \ldots, \mathbf{S_k}).$$

Theorem: Bounds on the Size of Pseudo Shares

Theorem

Let $(\Gamma_{s_1}, \ldots, \Gamma_{s_k})$ be a k-tuple of access structures on the set of participants $\mathcal{P}$. Assume that for all $S_i \in \{S_1, \ldots, S_k\}$ and $T \subseteq \{S_1, \ldots, S_k\} \setminus \{S_i\}$, it holds $H(\mathbf{S_i}|\mathbf{T}) > 0$. Let $\{P\}, Y \subset \mathcal{P}$ be such that $\{P\}, Y \notin \Gamma_{s_i}$ but $\{P\} \cup Y = \mathcal{A}_q^{s_i} \in \Gamma_{s_i}$, for some q. Then in any multi-use multi-secret sharing scheme for $(\Gamma_{s_1}, \ldots, \Gamma_{s_k})$, it holds

$$H(\mathbf{U_P}) \geq H(\mathbf{S_i}).$$

Corollary

Suppose $s_1, s_2, \ldots, s_k$ to be shared among n participants $P_1, \ldots, P_n$ in such a way that any set of participants with cardinality t_i is qualified to reconstruct the secret s_i. Then

$$H(\mathbf{U_{P^{t_i}}}) \geq H(\mathbf{S_i}).$$

where, $H(\mathbf{U_{P^{t_i}}})$ means the uncertainty of the pseudo share of the participant P corresponding to the i-th secret.

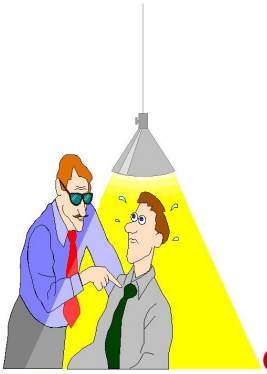
Ideal Multi-use Multi- SSS and future work

Definition

A perfectly secure multi-use multi-secret sharing scheme is said to be ideal if the size of shares and pseudo shares of each participant attain the bounds, as mentioned in Theorem 5 and Theorem 7, with equality.

- Construction of an ideal multi-use multi-secret sharing scheme for general access structure.

Questions



Questions???

Improved Subset Difference Method based on Ternary Tree

Kazuhide Fukushima[1], Shinsaku Kiyomoto[1], Toshiaki Tanaka[1], and Kouichi Sakurai[2, 3]

1. KDDI R&D Laboratories, Inc.
2. Faculty of Information Science and Electrical Engineering, Kyushu University
3. Institute of Systems, Information Technologies and Nanotechnologies

Our Results

	Subset Difference Method (by Naor et al.)	Proposed Scheme
Communication Cost	$2 \ln 2 \cdot r$ $\sim 1.38r$	$3\ln 2 \cdot r$ $\sim 2.08\ r$
Storage Size	$(\log_2 n)^2/2$ $= 0.5 \log_2 n$	$(\log_3 n)^2 /2$ $\sim 0.2 \log_2 n$
Computational Cost	O(log n)	O(log n)
Resistance to Coalition Attacks	✓	✓

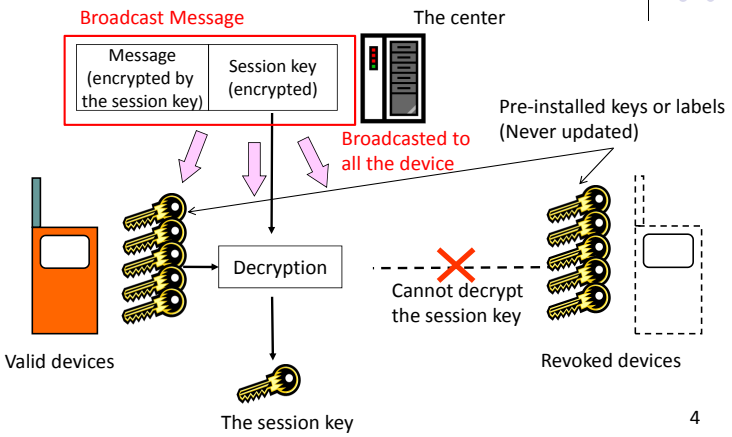
n: The number of client devices
r: The number of revoked devices

Background

- High-speed Internet has been deployed.
  - Pay content broadcasting services have been major.
- Copyright protection is an issue for these services.
  - Copying of digital content is easy and requires little effort.
- Content should be encrypted so that only valid client devices can use.

➔ **A key-management scheme is required to share the same key between the center and valid client devices.**

Overview of Broadcast Encryption Schemes



Related Work (1/2)

- The first broadcast encryption scheme was proposed by Berkovits (1991).
- Naor et al. (2001) proposed the Complete Subtree Method and Subset Difference Method:
  - An improved Complete Subtree method is used in Marlin Broadband.
    - Marlin is a DRM technology for consumer electronics devices and multimedia service.
  - The Subset Difference Method is used in AACS.
    - AACS is a DRM technology for the next generation of optical discs and DVDs.

Related Work (2/2)

- Many improvements to these Subset Difference Methods have been proposed:
  - Pseudo Random Number Generator-based schemes:
    - Halevy et al.(2002), Goodrich et al.(2004), Jho et al.(2005), Hwang et al.(2005), and Attrapadung et al.(2007)
  - RSA-based schemes:
    - Asano(2002), Attrapadung et al.(2003), and Gentry et al.(2004)
  - Pairing-based scheme:
    - Boneh et al.(2005)

Challenging Issue

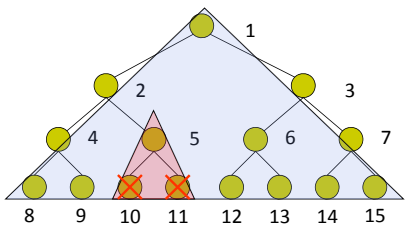
- The schemes based on an RSA cryptosystem and pseudo-random number generator involve a tradeoff between communication cost and storage size.
- The scheme based on Pairing imposes high computational cost on client devices.

➔ We construct a new broadcast encryption scheme:

- It has resistance against coalition attacks.
- It imposes a **feasible computational cost.**
- It **simultaneously reduces the communication cost and storage size.**

Subset Difference Method

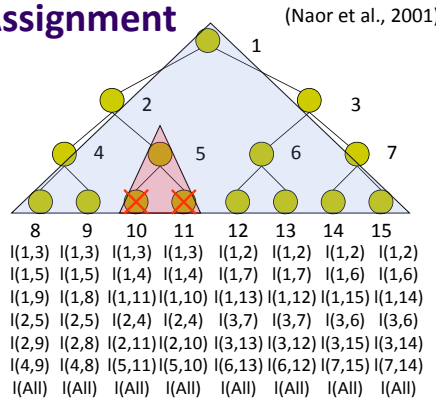
(D. Naor, M. Naor, and J. Lotspiech, Crytpto2001)



- Broadcast encryption scheme based on the binary structure:
  - All the valid devices are covered with a complete subtree with a crack.
  - A key is assigned to each subtree.

Label Assignment

(Naor et al., 2001)



Pseudo Random Number Generator $f:\{0,1\}^n \rightarrow \{0,1\}^{3n}$

$f(l(u,v)) = f_L(u,v) \,||\, f_K(u,v) \,||\, f_R(u,v) = l(u,v) \,||\, k(u,v) \,||\, l(u,v)$

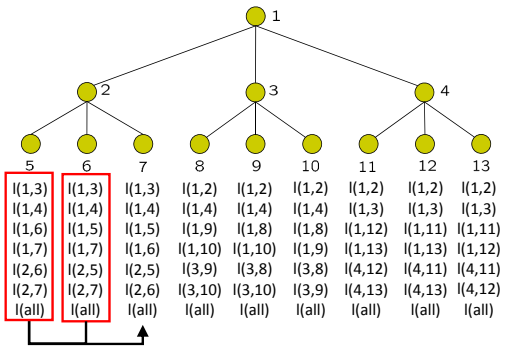
Label Assignment Algorithm

(Naor et al., 2001)

1. Assign client devices to leaf nodes of the binary tree.
2. Generate initial labels for all the nodes.
   - Label l(u, u) is assigned to node u.
   - Other labels are derived according to the rule:
     - f(l(u,v)) = l(u,2v) || k(u,v) || l(u,2v+1).
3. Assigns labels l(u,v) to each client device.
   - Client device at node u has labels l(u,v):
     - u is on the path from the root node to the node to which this device is assigned.
     - v is a node that hangs on the path.

Straightforward Expansion of the Subset Difference Method

- Label assignment algorithm can be applied to the ternary tree structure; however, ...



Issue

- This straightforward expansion cannot prevent the coalition attacks
  - The client device at node 5 has following labels: **l(1,3)**, **l(1,4)**, **l(1,6)**, l(1,7), **l(2,6)**, l(2,7), and **l(all)**
  - The client device at node 6 has following labels: **l(1,3)**, **l(1,4)**, **l(1,5)**, l(1,7), **l(2,5)**, l(2,7), and **l(all)**
  - ➔ If they collude, they can collect all the label the client device at node 7 has:
    l(1,3), l(1,4), l(1,5), l(1,6), l(2,5), l(2,6), l(all)

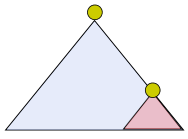
We need new label assign algorithm to design coalition resistant Subset Difference Method

Our Idea

- We introduce new key assignment algorithm and encryption algorithm to support two-way revocation.
  - Labels and hashed label are assigned to client devices in order to prevent coalition attacks.

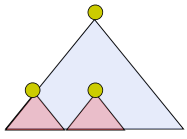
Two-Way Revocation

- Tree with one crack:



- The corresponding key is derived from a hashed label:
- ➔ Client devices have the label or the hashed label can derive the key.
- ➔ One subtree is revoked.

- Tree with two cracks:



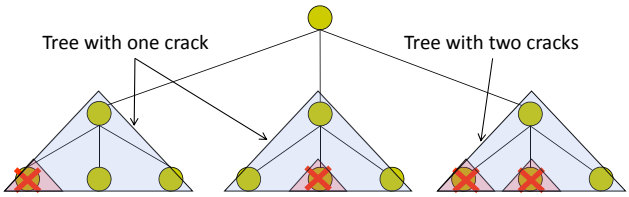
- The corresponding key is derived from a label:
- ➔ Client devices have the label can derive the key.
- ➔ Two subtrees are revoked.

Our Proposed Scheme

- Our scheme uses a complete ternary tree structure for label assignment.
- Our scheme uses the following primitives:
  - **An encryption function F_K** to encrypt message M.
  - **An encryption function E_L** to encrypt session key K.
  - **Independent hash functions h, h_l, h_c, h_r, and h_{kg}**:
    - h_l, h_c, h_r, and h are used to derive label l(u,w) from label l(u, v) or hashed label h(l(u, v)) where node w is a child of node v.
    - h_{kg} is used to derive a key from a label or hashed label.

Subsets in Ternary Subset Difference Method

- Valid client devices are covered with subsets that are:
  - Set of leaf nodes of a tree with one crack.
  - Set of leaf nodes of a tree with two cracks.
- A key to encrypt the session key is assigned to each subset:



Cover Finding Algorithm

- **<u>New algorithm is required to realize the ternary Subset Difference Method</u>**
  - The algorithm in the conventional Subset Difference Method cannot be applied
- Input: The set of revoked client device R
- Output: The collection of disjoint subset S_t such that: $\cup_{t=1}^{w} S_t = N \backslash R$

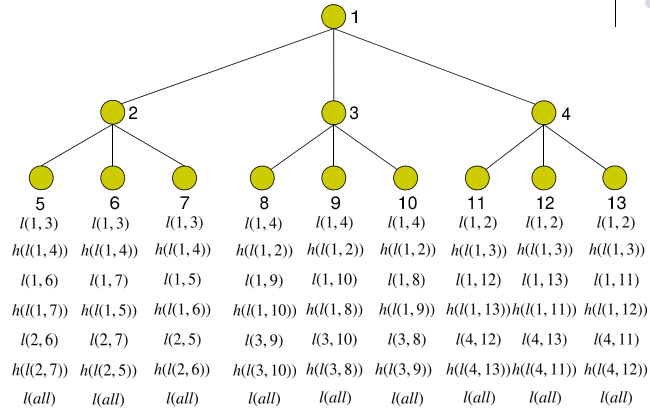
Label Assignment Algorithm

1. Assign client devices to leaf nodes of the ternary tree.
2. Generate initial labels for all the nodes:
   - Label l(u, u) is assigned to node u.
   - Other labels are derived according to the rule:
     - $l(u,w) = h_l(h(l(u, v)))$, $h_c(h(l(u, v)))$, or $h_r(h(l(u, v))$ where w is the left, center, or right child of v.
3. Assign labels and hashed labels to client devices.
   - Client device at the node u has the following labels $l(u,v_R)$ and the hashed label $l(u,v_L)$:
     - u is on the path from the root node to the node to which this device is assigned ($u \in Path_u$).
     - v_R is a node that hangs on right of the path ($v_R \in RightHang_u$).
     - v_L is a node that hangs on left of the path ($v_L \in LeftHang_u$).

Notations

- $Path_u$: the set of nodes that are on the path from the root node to u
- $LeftHang_u$: the set of nodes that hang on the left of the $Path_u$
  - If $Path_u$ contains the leftmost node, the rightmost sibling is in $LeftHang_u$
- $RightHang_u$: the set of nodes that hang on the right of the $Path_u$
  - If $Path_u$ contains that rightmost node, the leftmost sibling is in $RightHang_u$
- right(u): immediate right sibling of u
  - If u is the rightmost node, right(u) means the leftmost sibling

Example of Label Assignment



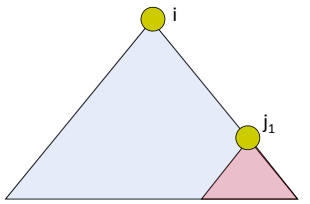
Encryption Algorithm

1. Encrypt message M with (randomly selected) session key K.
2. Cover all the valid client with subsets $S_1,\ldots,S_w$ using the cover finding algorithm.
3. Assign keys $L_1,\ldots,L_w$ to each subset (see the next slide).
4. Encrypt session K with each of these keys.
5. Distribute the following broadcast message.

$$\langle [S_1,\ldots,S_w, E_{L_1}(K),\ldots,E_{L_w}(K)], F_K(M)\rangle$$

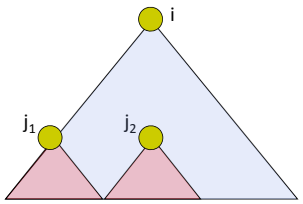
Key Assignment for Subsets

- Subset with one crack
  - One crack rooted at j_1



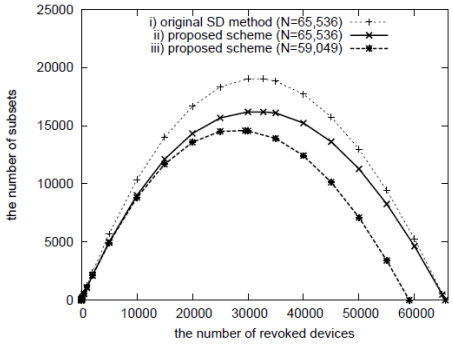
$$L_{i,j_1} = h_{key}(h(l(i,j_1)))$$

- Subset with two cracks
  - Cracks rooted at j_1 and j_2
  - j_2=right(j_1)



$$L_{i,j_1 \oplus j_2} = h_{key}(l(i,j_1))$$

Simulation Results



Theoretical Evaluation of Header Size in Subset Deference Method

- The precise evaluation of the average header size in the SD method:
  $\frac{r(n-r)}{n-1} + \frac{2n}{\binom{n}{r}} \sum_{h=2}^{\log_2 n} \left[\binom{n-2^{h-1}}{r} - \binom{n-2^{h-1}-2^{h-2}}{r} + \binom{n-2^h}{r} \right] / 2^h$
  is given by Okuaki et al. (in Japanese domestic conference.)
- Can we obtain the closed-form expression of the evaluation?
- Can we expand the result to the evaluation of our scheme with a ternary tree?

Coverage Problem of Sensor Networks

Bimal Roy
Indian Statistical Institute, Kolkata, India

Mrinal Nandi
West Bengal State university, Barasat, India

PROBLEM

To place sensors at the nodes of a grid.

First deploy sensors from air by helicopter.

Robot will rearrange the sensors to place at least one sensor at each node.

Problem is to develop algorithms for Robot and compare these algorithms in context with some pre-assigned parameters.

THE GRID

(m+2) x (n+2) rectangular grid whose nodes are labeled as (i,j) for i=0,1,…,m+1 and j=0,1,…, n+1

m, n are parameters of our problem.

Distance between two adjacent nodes are same.

At least one sensor has to be placed at each node (i,j) for i=1,2,…,m and j=1,2,…, n.

SENSORS

Sensors are deployed from air by helicopter.

Sensors may not be placed in the proper node.

May be placed in one of the four adjacent nodes.

Each sensor have an ID number.

SENSOR DEPLOYMENT

Helicopter will deploy at each node:
two sensors with probability μ, or
one sensor with probability (1- μ)

Here μ is a parameter of our problem.

Let the ID number(s) of the sensor(s) deployed on the node (i,j) be ID(i,j).

SENSOR PLACEMENT

The sensor(s) with ID number ID(i,j) are deployed at
node (i,j) with probability p, or
any adjacent node with probability q each

Here q = (1-p)/4 is a parameter of our problem.

THE ROBOT

After deployment of sensors robot will go to the node (1,1).

Robot travels according to some pre-assigned algorithms.

Robot can carry at most one sensor while traveling.

Robot travels along horizontal or vertical paths of the grid.

If robot is standing at node (i,j) then it can recognize the number of sensors and their ID numbers which are placed at that node and also at the adjacent four nodes.

ALGORITHMS

One based on the ID number and another which does not depend the ID number.

For first one assume that sensor with ID number ID(i,j) should be placed by the Robot either at the node (i,j) or one of the four adjacent nodes.

For second and third algorithms we assume that the sensors have no ID numbers.

TRAVERSED LENGTH

Traversed length is ‘L’, which is traveled by the Robot. This is the most important parameter.

L is the length travel by the robot starting from the node (1,1) to the node (m,n) or (m,1) according as m is odd or even.

The distance traveled by the Robot from one node to its adjacent node is one unit of traversed length.

PARAMETERS

Grid size	m, n
Probability	q
Repetition probability	μ
Error Traversed length	L
Traversed length with sensor	Ls
Empty nodes after deployment	N

ASSUMPTIONS

m and n are large such that the product of m and n is near about 10000.

Probability of deploying ID(I,j) at a neighboring node: $q < 0.15$

Given m, n and q can find the relation between μ and L and can minimize μ when some upper bound on L is given.

OUR WORK

Developed three different algorithms for Robot.

Compare L and Ls obtained from three different algorithms for several different values of parameters by simulation.

Find the Expected value of L in terms of other parameters for the first algorithm.

Find the Expected value, an approximate distribution and some theoretical results of N for the first algorithm.

THREE ALGORITHMS

Robot start from node (1,1) and end at (m,n) or (m,1) according as n is odd or even.

Robot move from a node (i, j) toward the next node (i', j')

where (i’, j’) = (i, j+1) if i is odd and $j < n$
(i, j-1) if i is even and $j > 1$
(i+1, j) otherwise

(i‘‘, j'') is the previous node of the node (i, j).

SOME NOTATIONS

X(i,j) is the number of sensor(s) with ID number (i,j) which are placed at the node (i, j).

X(i,j)↑ is the number of sensor(s) with ID number (i,j) which placed at the node (i-1, j).

Similarly, define X(i,j)→, X(i,j) ↓, X(i,j) ←

X'(i,j) is X(i,j)→ or X(i,j) ↓or X(i,j) ← according as (i',j')= (i,j+1) or (i+1,j) or (i,j-1).

SOME NOTATIONS

X"(i', j') is X(i', j')← or X(i', j') ↑ or X(i', j') →
according as (i',j')= (i,j+1) or (i+1,j) or (i,j-1).

H(i,j) is the number of sensors with the Robot when it first comes at the node (i,j).

T(i,j) is the number of sensor(s) which were placed at the node (i,j).

FIRST ALGORITHM

FOR i = 1 to m and j = 1 to n DO:

When the Robot is standing on the node (i,j), the Robot will do the following:

IF (H(i,j)=1)
place the sensor there, and
do the same job as in the case H(i,j)=0

IF (H(i,j)=0)
do the following … (contd.)

FIRST ALGORITHM

IF (X(i,j) = 2, X(i',j') = 0)

move to next node (i',j') with one sensor

IF (X(i,j) = 2, X(i',j') > 0 or X(i,j) = 1, X(i',j') > 0)

move to the next node with no sensor

IF (X(i,j) = 1, X(i',j') = 0, X"(i',j') > 0)

move to the next node with one sensor whose ID number is ID(i',j')

FIRST ALGORITHM

IF (X(i,j) = 1, X(i',j') = 0, X"(i',j') = 0)

move to the next node with no sensor

IF (X(i,j) = 0, X'(i'',j'') > 0 or X(i-1,j) ↓ > 0)

move to the next node with no sensor

ELSE

go to the node where the sensor with ID(i,j) is placed, take the sensor, place it at node(i,j), and do the same job as in case of X(i,j)=1.

SECOND ALGORITHM

FOR i = 1 to m and j = 1 to n DO:

When the Robot is standing on the node(i,j), the Robot will do the following:

IF (T(i,j) > 1 and H(i,j) = 1)
Robot will move to the next node (i',j') with that sensor

IF (T(i,j) > 1 and H(i,j) = 0 and T(i',j') < 2)
Robot will take one sensor and move to the next node

SECOND ALGORITHM

IF (T(i,j)>1 and H(i,j)=0 and T(i',j')>1)
move to the next node without any sensors

IF (T(i,j)=1 and H(i,j)=1)
move to the next node (i',j') with that sensor

IF (T(i,j)=1 and H(i,j)=0 and T(i',j')=0)

check other three adjacent nodes whether the numbers of sensors placed there is greater than 1 (or 0 for external nodes). If so, go there, take one sensor and move to the next node with that.

SECOND ALGORITHM

IF (T(i,j)=1 and H(i,j)=0 and T(i',j')>0)
move to the next node with no sensor

IF (T(i,j)=0 and H(i,j)=1)
place the sensor on the node (i,j), and
do the same job as in case of T(i,j)=1 and H(i,j)=0

IF (T(i,j)=0 and H(i,j)=0)

check other three nodes for more than one sensors.
If so, go there, take one, place it on present node (i,j),
and do the same job as in case of T(i,j)=1 and H(i,j)=0

THIRD ALGORITHM

These is nearly same as the second algorithm.

In addition,

Robot goes back to the two distanced nodes, and
placed sensor if it vacant, or
take one sensor if there are more than one sensors

SIMULATION RESULT

Here, m=100, n=100, and empty nodes means the number of empty nodes after placement for the second algorithm.

p	μ	L for first algorithm	L for second algorithm	Empty nodes
0.5	0.2	14845	11509	103
0.5	0.3	14105	11175	53
0.5	0.4	13557	10847	18
0.5	0.5	12863	10635	10
0.5	0.6	12403	10445	9
0.6	0.2	13959	11277	83
0.6	0.3	13411	10919	39
0.6	0.4	12721	10629	14
0.6	0.5	12119	10471	3

FURTHER SIMULATION RESULT

		initial empty nodes	expected values		traversed length			final empty nodes		traversed with sensor		
p	μ	N	E(N)	E(L1)	L1	L2	L3	N2	N3	LS 1	LS 2	LS 3
0.5	0.2	2373	2387	15284	14949	11437	11955	113	2	3584	7170	7240
0.5	0.3	2171	2138	14688	14083	11187	11425	52	0	3280	7376	7376
0.5	0.4	1996	1909	14122	13731	10945	11001	15	1	3051	7543	7511
0.5	0.5	1708	1698	13588	12843	10603	10635	8	0	2700	7708	7639
0.5	0.6	1528	1501	13084	12447	10449	10465	4	0	2454	7798	7741
0.6	0.2	2148	2130	14206	14057	11257	11633	81	2	2993	7270	7273
0.6	0.3	1969	1905	13660	13355	11071	11163	37	2	2713	7549	7514
0.6	0.4	1679	1694	13150	12813	10647	10713	17	0	2426	7854	7808
0.6	0.5	1471	1496	12676	12173	10443	10509	4	0	2176	7922	7865
0.6	0.6	1308	1311	12236	11755	10363	10375	3	0	1938	8077	8017

THE EXPECTED VALUE OF L

$$E(L) = (mn-1) + 2\,(t1 + 3t2 + 2\,(m + n - 4)\,t3 + (m - 2)(n - 2)\,t4)$$

with
$t1 = d - 2ab + cb2$
$t2 = d – a – ab – de + cb + ae + abe – cbe$
$t3 = d – a - 2ab – de + 2cb + ae + cb2 + 2ab2e - 2cbe - cb2e2$
$t4 = d – a - 3ab – de + 2cb + 3ab + ae + 3cb2 - 3abe - 3cbe - 3cb2e$

and
$a = 7\mu qq + (1- \mu)\,q,$
$b = \mu(2-q)q + (1- \mu)q,$
$c = 2\mu qq,$
$d = 4q(1- \mu(1-4q)),$
$e= \mu(1-4q)\,(1-4q)$

SOME RESULTS ON T(i,j)

Probability that any node (i,j) is empty is P(T(i,j)=0).

This depends on
$P(X(i,j)=0) = \{\mu(4q)2+(1- \mu)(4q)\} = (1-p)(1- \mu p)$, and
$P(X(i,j)\uparrow=0) = \{\mu(1-q)2+(1- \mu)(1-q)\} =(1-q)(1- \mu q)$

$$P\,(\,T(i,j) = 0\,) < 0.25 \quad \text{for } p > 0.6$$

and

$$P\,(\,T(i,j) = 0\,) < 0.2 \quad \text{for } p > 0.8$$

SOME RESULTS ON N

Let N(i,j) = 1 if T(i,j) = 0
= 0 otherwise

Then, N(i,j)'s are Bernoulli random variables with parameter P(T(i,j)=0) but they are dependent

P(T(i,j)=0)'s are equal for $i \neq 1,m$ and $j \neq 1,n$,
P(T(i,j)=0)'s are equal for $i = 1,m$ and $j = 1,n$, and
P(T(i,j)=0)'s are equal for other i and j.

Important: $N = \sum N(i,j)$

SOME RESULTS ON N

The Expectation of N is:

$$E(N) = E\left(\sum N(i,j)\right) = \sum P\left(T(i,j)=0\right)$$

This depends on the parameters of the problem.

Expectation of N for large m,n is:

$$E(N) \approx mnp(1-p)(1-\mu+4\mu^2pq)$$
$$< mn(1-\mu+\mu^2/4)/4$$

SOME RESULTS ON N

Important to note that:

$P(T(i,j)=0) < 0.23$, for $p > 1/2$, $\mu > 0.2$

From this, one obtains:

N is approximately Normally distributed for large m, n and for $p > 1/2$, $\mu > 0.2$

FUTURE WORKS

Develop some randomized algorithm and compare them in a similar way.

Impose more conditions on the Robot or relax the same.
Ex: We can assume that Robot can carry k many sensors when it travels.

Develop some optimal algorithm in some classes of algorithms to optimized L.

Find more results on individual parameters of the problem.

Thank You

ありがとう.

Non-committing Encryption Scheme Based on DDH Assumption

Takashi Nishide
joint work with
Huafei Zhu, Tadashi Araragi, Kouichi Sakurai
2010 July 8th

Motivation

- Security against more powerful adversary is more preferable.
- However, constructing protocols that withstand a wider class of adversaries is harder to achieve…
- We consider to construct a secure channel protocol against an adaptive (more powerful) adversary in the UC framework.

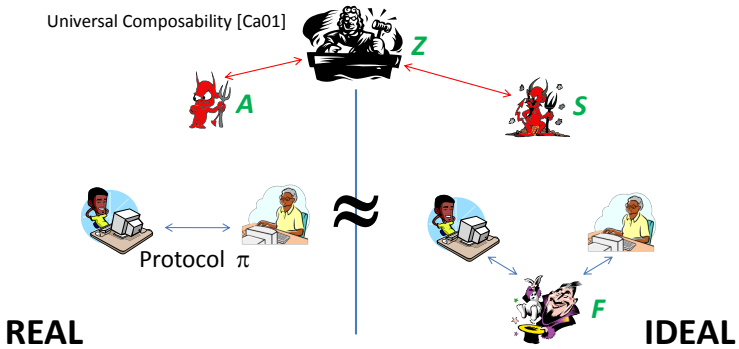
Adversarial Models in Cryptographic Protocol

- Static vs Adaptive
  - Static adversary
    - needs to decide the set of players to corrupt prior to the execution of the protocol
  - Adaptive adversary
    - can corrupt players during the execution of the protocol arbitrarily
    - More flexible and realistic
- Erasure vs Non-Erasure
  - In the erasure model, players are supposed to erase the past data when corrupted by an adversary
    - So the adversary cannot get the computation history even after the corruption occurs
  - The erasure model is not realistic and may be impossible…
- Adversarial models have a large influence on security proof
- In particular, an adaptive adversary in the non-erasure model makes it hard to construct a secure channel

Adaptive Security for Secure Channel

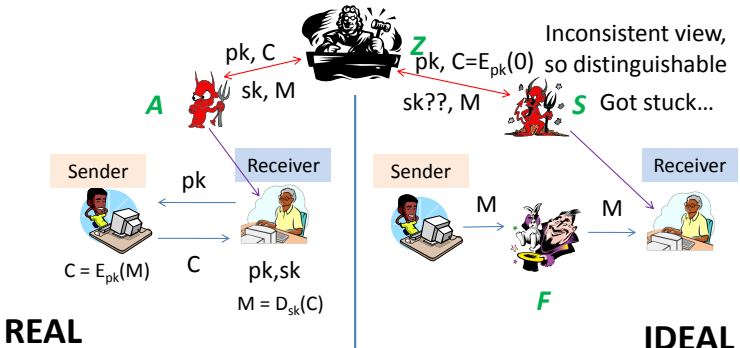
- Secure channel is a basic cryptographic primitive.
- However, to construct a secure channel against an adaptive adversary, **traditional public key encryption** is not sufficient…
- [Nie02] proved that no non-interactive communication protocol can achieve adaptive security without the random oracle model(RO).
- So we need an interactive protocol to realize a secure channel against an adaptive adversary w/o RO model.

Security Definition in UC Framework



Protocol π is a secure realization of an ideal functionality F if
for every real adversary A
there exists a simulator S s.t.
ideal & real worlds are indistinguishable to any environment Z

Secure Channel with Adaptive Adversary?



Non-committing Encryption needed instead of Public Key Encryption

Non-committing Encryption

- With non-committing encryption(NCE), we can construct a secure channel protocol against an adaptive adversary.
- Simulator can run an NCE protocol and create a fake ciphertext that can be opened to any chosen plaintext (0 or 1).
- Encryption is done for each bit of message M
  - inefficient, but same efficiency as other schemes in the non-erasure model
  - Price for adaptive security…

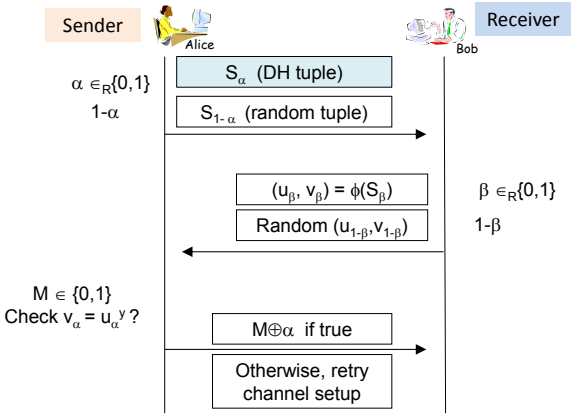
Building Block

- Naor-Pinkas randomizer (NPR) ϕ [NP01]
- Setup: $p = 2q+1$
  $G \subseteq Z_p^*$ is a subgroup of order q
- $\phi((g_1, g_2, h_1, h_2) \times (s,t))$ defined as
  $(u, v) = (g_1^s g_2^t \bmod p,\ h_1^s h_2^t \bmod p)$
  where $s,t \in_R Z_q$,and $g_i, h_i \in G$
  - If $(g_1, g_2, h_1= g_1^y, h_2=g_2^y)$ is a random Diffie-Hellman tuple, we have $v = u^y \bmod p$
  - If (g_1, g_2, h_1, h_2) is a non-DH random tuple, (u,v) is a random tuple in G^2.

Building Block cont'd

- Canetti-Fischlin oblivious sampling & faking algorithms [CF01]
- By using the faking algorithm, the simulator can construct a fake transcript (computation history) to the environment ***Z***
  - in such a way that a Diffie-Hellman tuple looks completely random

Sketch of Construction



More Formal Construction

- Sender generates with secret $\alpha \in_R \{0,1\}$, y
  - $S_0 = (g_{1,0}, g_{2,0}, h_{1,0}, h_{2,0})$
  - $S_1 = (g_{1,1}, g_{2,1}, h_{1,1}, h_{2,1})$
  - where S_α is a DH tuple, $S_{1-\alpha}$ is a random tuple, and $h_{1,\alpha} = g_{1,\alpha}^y$, $h_{2,\alpha} = g_{2,\alpha}^y$
- Receiver generates with secret $\beta \in_R \{0,1\}$
  - $w_\beta = (u_\beta, v_\beta)$ from S_β with Naor-Pinkas randomizer
  - $w_{1-\beta} = (u_{1-\beta}, v_{1-\beta})$ at random
  - Sends w_β , $w_{1-\beta}$ to the sender
- Sender checks $v_\alpha = u_\alpha^y \bmod p$?
  - If true, ciphertext $c = m \oplus \alpha$ where $\alpha = \beta$
  - Otherwise, $\alpha \neq \beta$ and retries the channel setup

Proof in UC Framework

- We define a simple ideal functionality for non-committing encryption.
- Case analysis based on when the corruption occurs
- Simulator uses the Canetti-Fischling oblivious faking algorithm to show the randomness used in the corrupted player to the environment ***Z***.
- Indistinguishability based on DDH assumption

Summary

- Non-committing encryption protocol secure against an adaptive adversary with the DDH assumption
- Proof given in the UC framework and non-erasure model
- Can be used as a building block realizing secure channel in other protocols that need to be secure against an adaptive adversary

A Framework for Choosing Security Modules

Kanta Matsuura
(The University of Tokyo)

Economics of information security: Analysis, Analysis, & Analysis.

- Why information security is hard?
- Why free-riding problems happen?
- Why software vendors prefer the patch-after-patch approach?

- Economics of Information Security (EIS) can give possible reasons.
- Early works of raising problems are in 1990's.
- Many early works of the current trend of EIS are between 2000-2004.
- WEIS (Workshop on the EIS) started in 2002.

Next Trend would be: Synthesis

- Many people are noticing this (e.g. the panel at WEIS2009): Oh, as long as we take analysis-only approaches, we are having a slump.
- Industry would ask what they can do with the help of such analyses.
- A possible approach:
  - Use implications from analytical models to refine best-practice documents.

Parameters & functions in the base model (Gordon-Loeb 2002)

ACM TISSEC

- The loss when breached: λ
- The probability of a threat occurring: t
- The potential loss: $L = t\lambda$
- The conditional probability of the threat being successful (conditional on the occurrence), called "vulnerability" in the model: v
- The information-security investment: z
- The conditional probability after the investment (security-breach probability function): $S(z,v)$
  - Class I: $S(z,v) = v/(\alpha z+1)^{\beta}$
  - Class II: $S(z,v) = v^{\alpha z+1}$

This α is called the productivity of information security.

K. Matsuura: Productivity Space of Information Security in an Extension of the Gordon-Loeb's Investment Model. In: *Managing Information Risk and the Economics of Security*, Springer, pp.99-119, 2009.

An extension (Matsuura 2008)

- Let us assume the investment z reduces not only v but also t (down to $T(z,t)$).
- Optimal investment z^* is determined by

$$ENBIS(z) = \underbrace{vt\lambda - S(z,v)T(z,t)\lambda}_{=B} - \underbrace{z}_{=C} \to \max.$$

- If the marginal benefit at z=0 is less than or equal to the marginal cost of such investment, z^* equals zero. $\left.\frac{\partial B}{\partial z}\right|_{z=0} \le \left.\frac{\partial C}{\partial z}\right|_{z=0}$

Closed-form solution for $S(z,v)=v^{\alpha z+1}$, $T(z,t)=t^{\beta z+1}$

- The condition for having a zero investment as the optimum:

$$F(v) \equiv v\ln v + \frac{\beta \ln t}{\alpha}\cdot v + \frac{1}{\alpha L} \ge 0$$

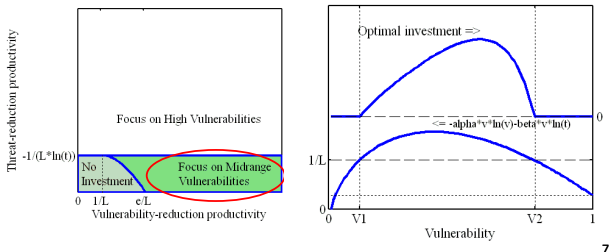
That is, $-\alpha v \ln v - \beta v \ln t \le \frac{1}{L}$.

- When $F(v)<0$, we have

$$z^* = \frac{\ln\{-1/(vt\lambda\ln(v^{\alpha}t^{\beta}))\}}{\ln(v^{\alpha}t^{\beta})} = \frac{\ln\frac{1}{-vL\{\alpha(\ln v)+\beta(\ln t)\}}}{\alpha(\ln v)+\beta(\ln t)}$$

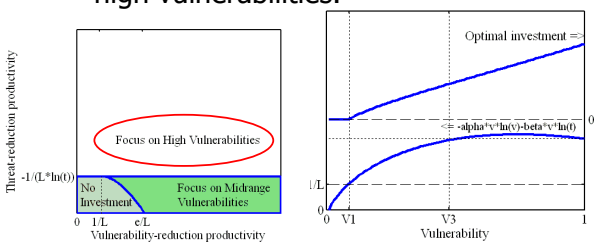
Type (A): Mid-vulnerability intensive area

- A firm may be better off concentrating its efforts on midrange vulnerabilities.



Type (B): High-vulnerability intensive area

- If the threat-reduction productivity is sufficiently high, a firm should focus on high vulnerabilities.

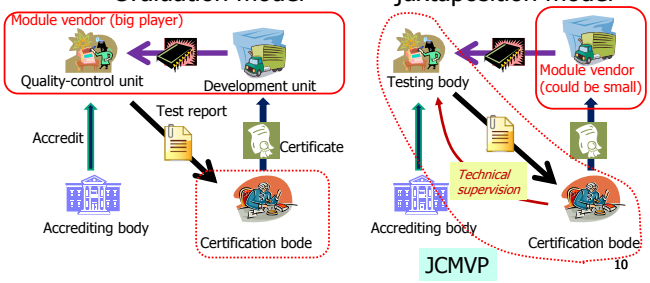


Apply the model for JCMVP

- Japan Cryptographic Module Validation Program
- Full operation started in April 2007.
- Follows certificate-authority juxtaposition model (as opposed to independence evaluation model).
- Level of a module: level1, 2, 3, or 4.

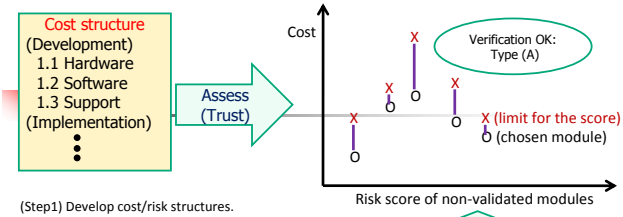
Product-validation system

- Independence evaluation model
- Certificate-authority juxtaposition model

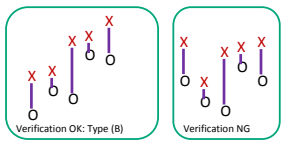


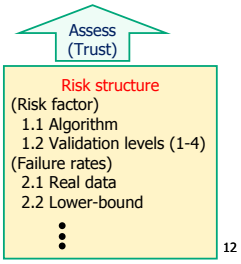
Users (=system vendors) need help

- System vendors must choose appropriate modules.
- Obviously, the higher level, the higher security. However, there are problems of cost, usability, and so on.
- In principle, JCMVP pays attention to small vendors; they need help for efficient and reliable design.



(Step1) Develop cost/risk structures.
(Step2) Assess parameters for alternatives.
(Step3) Choose a module (*heuristics*).
<Do Steps1-3 for all the building blocks in the system.>
(Step4) *Trust* the results so far, and plot them.
(Step5) If the plot implies a curve of type (A) or (B), accept the results (*Verified*). If not, go back to Step2.





The guideline was released on 20 May, 2010 (Sorry, only in Japanese).

- Available at http://kmlab.iis.u-tokyo.ac.jp/resources/guideline_1_0.pdf
- Not exclusively for JCMVP; we use a generalized description (so that we can consider other validation systems in the future).
- Current version (not an official guideline but a reference document) is a *minimal* set; it does not use many (potentially applicable) theories.
- Further development and case studies will come.

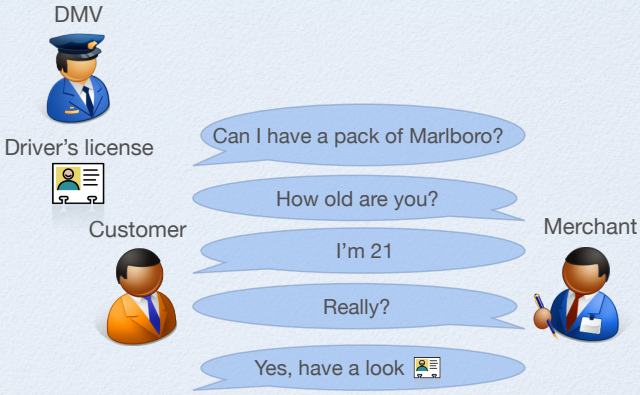
Concluding remarks: Implications of the general description

- *General* in terms of module classes
  - Biometrics modules.
  - Computer/network security building blocks (E.g., The core topic of our project: Provable security of Hierarchical Key Assignment Schemes).
- *General* in terms of validation methods
  - Product validation.
  - Theoretical proofs.

Nominative Signatures

Jacob Schuldt
Ph.D. student, Matsuura Lab
University of Tokyo

Certification of Information



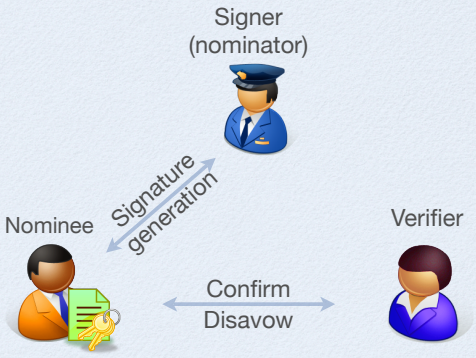
Certification of Information

- The merchant is convinced that the customer is 21 years old since
  - he believes the driver's license belongs to the customer
  - he believes the driver's license is issued by the DMV
  - he trusts the DMV
- Note that the information being certified (the customer's age) belongs to the customer and not the DMV
- The customer controls who is able to verify his age
  - he will only show his driver's license to intended verifiers
  - verification is non-transferable

Digital Equivalent?

- The properties of a digital system are different from the properties relied upon in the previous example
  - An exact copy of any evidence presented to the merchant can easily be made
  - => controlled verifiability is a concern
- Nominative signatures have been proposed to address this type of scenario

Nominative Signatures



Nominative Signatures

- Many related schemes
  - Undeniable signatures
  - Designated confirmer signatures
  - (Universal) designated verifier signatures
- Unlike the above, a nominative signature scheme guarantee that
  - both signer and nominee have been involved in signature generation
  - no information leaked from the signer will reveal information about a signature

Brief History

- Nominative signatures were introduced by Kim, Park and Won [KPW96]
  - No security model defined
  - Scheme does not provide full invisibility
- Formal security model defined by Liu et al. [LWHWHMS07]
  - However, invisibility against malicious signers is not captured
  - Scheme vulnerable to such attacks
- Liu et al. [LHW07] and Zhao et al. [ZLY09] update security model
  - Allows deterministic schemes -- only limited invisibility is provided
  - Scheme from [ZLY09] is insecure

This Talk

- We define stronger security models without artificial restrictions
- We propose a provable secure standard model scheme
  - which is as efficient as the RO scheme from [LHW07]
  - but requires key registration

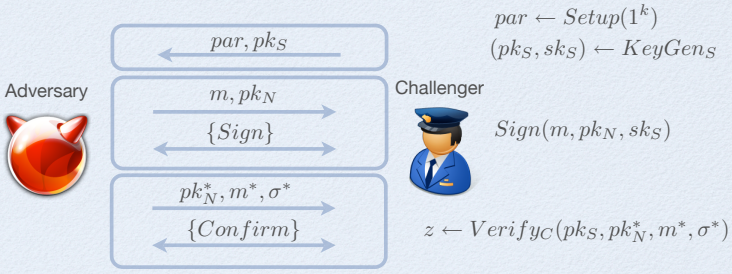
Definition of Nominative Signatures

- $Setup(1^k) \rightarrow par$
- $KeyGen_S(par) \rightarrow (pk_S, sk_S)$
- $KeyGen_N(par) \rightarrow (pk_N, sk_N)$
- $Sign(par, pk_N, m, sk_S) \leftrightarrow Receive(par, pk_S, m, sk_N)$
  - Output of $Receive$ is σ
- $Valid(par, pk_S, sk_N, m, \sigma) \rightarrow 0/1$
- $Confirm(par, pk_S, m, \sigma, sk_N) \leftrightarrow Verify_C(par, pk_S, pk_N, m, \sigma)$
- $Disavow(par, pk_S, m, \sigma, sk_N) \leftrightarrow Verify_D(par, pk_S, pk_N, m, \sigma)$

Security Requirements

- Unforgeability
  - Against malicious nominees
  - Against malicious signers
- Invisibility
  - Even against malicious signers
- Zero-knowledge Confirm and Disavow protocols

Unforgeability against Malicious Nominees



$par \leftarrow Setup(1^k)$
$(pk_S, sk_S) \leftarrow KeyGen_S$

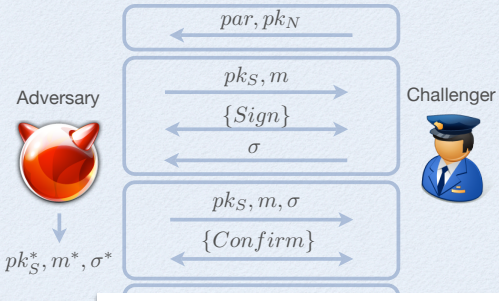
$Sign(m, pk_N, sk_S)$

$z \leftarrow Verify_C(pk_S, pk_N^*, m^*, \sigma^*)$

Adversary wins if $z = \texttt{accept}$, and (pk_N^*, m^*) was not queried to the sign oracle

11

Unforgeability against Malicious Signers



$par \leftarrow Setup(1^k)$
$(pk_N, sk_N) \leftarrow KeyGen_N$

$\sigma \leftarrow Receive(pk_S, m, sk_N)$

$Confirm(pk_S, m, \sigma, sk_N)$

Adversary wins if $Valid(par, pk_S^*, sk_N, m^*, \sigma^*) = 1$, and (pk_S^*, m^*, σ^*) was not obtained in a sign query.

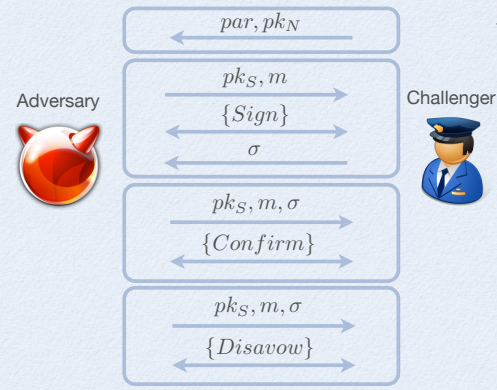
12

Unforgeability

- Note that there is a difference in strength between the above two definitions
  - Unforgeability against malicious nominees is similar to normal unforgeability
  - Unforgeability against malicious signers is similar to strong unforgeability
- To prove invisibility of our concrete scheme, strong unforgeability against malicious signers is needed

13

Invisibility



$par \leftarrow Setup(1^k)$
$(pk_N, sk_N) \leftarrow KeyGen_N$

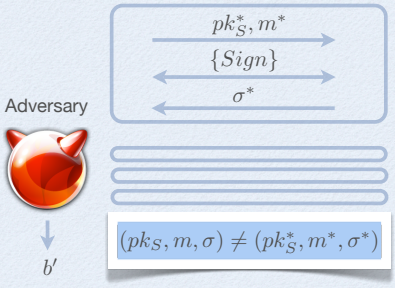
$\sigma \leftarrow Receive(pk_S, m, sk_N)$

$Confirm(pk_S, m, \sigma, sk_N)$

$Disavow(pk_S, m, \sigma, sk_N)$

14

Invisibility



$\sigma' \leftarrow Receive(pk_S^*, m^*, sk_N)$

$b \leftarrow \{0,1\}$ $\quad b = 0: \ \sigma^* \leftarrow \sigma'$ $\quad b = 1: \ \sigma^* \leftarrow \mathcal{S}$

Adversary wins if $b' = b$

15

Concrete Scheme

- Our concrete scheme is based upon the signature scheme by Waters [W05]
- The scheme is defined in a group equipped with a bilinear map => the parameters of our scheme are defined as follows:

$par = (\mathbb{G}, \mathbb{G}_T, p, e, g, CR) \leftarrow Setup(1^k)$

$|\mathbb{G}| = |\mathbb{G}_T| = p \qquad e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$

$\langle g \rangle = \mathbb{G}$

$CR : \{0,1\}^* \rightarrow \mathbb{Z}_p$

16

Concrete Scheme - Key Generation

Signer

$pk_S = (g^{\alpha_s}, h_s, H_s) \qquad sk_S = (\alpha_s, \{v_i\})$

$h_s \in \mathbb{G} \qquad H_s(m) = V_0 \prod_i V_i^{m_i}$

$V_i \leftarrow g^{v_i}$

Nominee

$pk_N = (g^{\alpha_n}, h_n, k, u, v, H_n) \qquad sk_N = (\alpha_n, x, y, \{u_i\})$

$h_n, k \in \mathbb{G} \qquad H_n(m) = U_0 \prod_i U_i^{m_i}$

$u \leftarrow g^{x^{-1}} \qquad v \leftarrow g^{y^{-1}} \qquad U_i \leftarrow g^{u_i}$

Concrete Scheme - Signing Protocol

Signer

$pk_S = (g^{\alpha_s}, h_s, H_s) \qquad sk_S = (\alpha_s, \{v_i\})$

$\hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2) \leftarrow (g^r, h_s^{\alpha_s} H_s(pk_N||m)^r)$

$\hat{\sigma}$

Nominee

$pk_N = (g^{\alpha_n}, h_n, k, u, v, H_n) \qquad sk_N = (\alpha_n, x, y, \{u_i\})$

$(\sigma_1', \sigma_2') \leftarrow (\hat{\sigma}_1 g^{r'}, \hat{\sigma}_2 H_s(pk_N||m)^{r'})$

$\sigma_1 \leftarrow (\sigma_1')^{x^{-1}} \qquad \sigma_2 \leftarrow (\sigma_1')^{y^{-1}}$

$t \leftarrow CR(\sigma_1||\sigma_2||pk_S||m) \qquad M \leftarrow g^t k^{r''}$

$\sigma_3 \leftarrow \sigma_2' h_n^{\alpha_n} (\sigma_1')^{u_0 + \sum_i u_i M_i}$

$\sigma = (\sigma_1, \sigma_2, \sigma_3, r'')$

Concrete Scheme - Confirm/Disavow

$Confirm:$

$ZKPK\{(x,y) : u^x = g \wedge v^y = g \wedge e(g,\sigma_3) = e(g^{\alpha_s}, h_s)e(g^{\alpha_n}, h_n)$
$e(\sigma_1, H_s(pk_N||m)H_n(M))^x e(\sigma_2, H_s(pk_N||m)H_n(M))^y\}$

$Disavow:$

$ZKPK\{(x,y) : u^x = g \wedge v^y = g \wedge e(g,\sigma_3) \neq e(g^{\alpha_s}, h_s)e(g^{\alpha_n}, h_n)$
$e(\sigma_1, H_s(pk_N||m)H_n(M))^x e(\sigma_2, H_s(pk_N||m)H_n(M))^y\}$

Security

- The following results are in the registered key model

Security Notion	Assumptions
Unforgeability (malicious nominees)	Unforgeability of Waters signatures, Confirm protocol is POK
Strong unforgeability (malicious signers)	Unforgeability of Waters signatures, DL in $\mathbb{G}$, collision resistance of CR
Invisibility	Strong unforgeability (malicious signers), DLIN in $\mathbb{G}$

Extensions

- Conversion of signatures
  - Allows the nominee to convert a signature to a publicly verifiable signature
  - Warrants additional security requirements
  - Easy to achieve for our concrete scheme
- Security without key registration
  - Seems possible at the expense of larger signatures

Summary

- Nominative signatures provide a solution to the problem of information certification
- We have proposed a scheme which
  - provides a high level of security (no artificial restrictions in used security models)
  - is provably secure in the standard model
  - is fairly efficient (matches efficiency of similar RO schemes)

References

- [KPW96] Kim, Park, Wong: Zero-knowledge Nominative Signatures, PragoCrypt, 1996
- [LWHWHMS07] Liu, Wong, Huang, Wang, Huang, Mu, Susilo: Formal Definition and Construction of Nominative Signature, ICICS 2007
- [LHW07] Liu, Huang, Wong: An Efficient One-move Nominative Signature Scheme, eprint report 2007/260
- [W05] Waters: Efficient Identity-based Encryption without Random Oracles, Eurocrypt, 2005
- [ZLY09] Zhao, Lin, Ye: Provably Secure Convertible Nominative Signature Scheme, Inscrypt 2009

Questions?

On the security proof of Wu-Wei hierarchical key assignment scheme

Indo-Japan joint workshop, 8th June 2010

Murali K Medisetty
DA-IICT

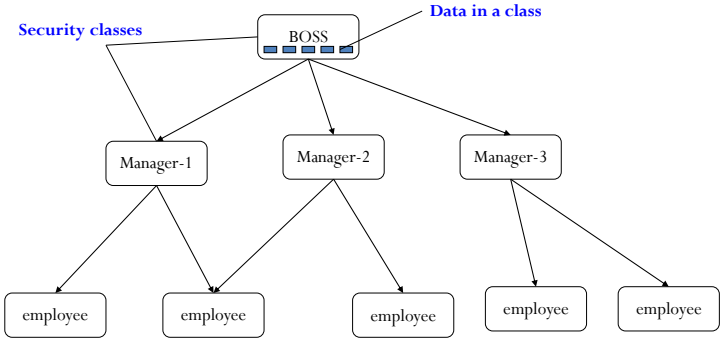
DST-JST Project

- Goal: Security evaluation of key assignment schemes
  - Partner institutions: DA-IICT and Univ. of Tokyo
  - Co-PIs: Anish Mathuria and Kanta Matsuura
- DA-IICT team
  - Anish Mathuria (Faculty)
  - Manik Lal Das (Faculty)
  - Naveen Kumar Chaudhary (RA)
  - Pooja Hegde (JRF)
  - Murali Medisetty (UG RA)

Outline

- Introduction
  - What's a hierarchical key assignment scheme (HKAS)?
- Wu-Wei's scheme[1]
- Proofs for security of HKAS
  - Outline of Wu-Wei's proof
- Our work
  - Problems identified in the proof
  - Fixes to the proof
- Conclusions & future work

Hierarchy in real life!



- Boss can access everyone's data
- A manager can access data of the employees under him

Hierarchical Key Assignment Scheme (HKAS)

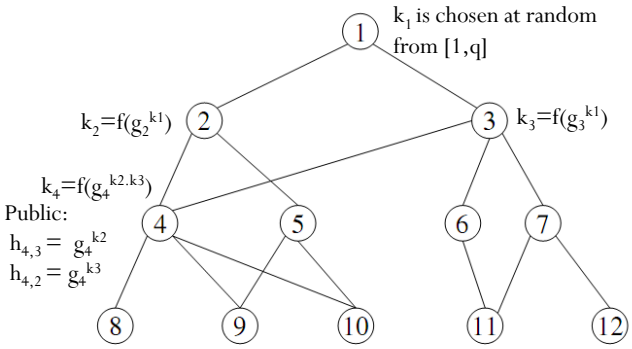
- A HKAS assigns secret keys to all the security classes
- Each class key encrypts the data of that class
- Knowing the secret key of a particular security class, one can derive the keys of all its successor classes using public information
- First such scheme was proposed by Akl and Taylor in 1983 [6]

Key Generation[1]

- CA (Central Authority) chooses a group $\mathbf{Z_p}^*$ where p= 2.q +1, p and q are large primes
- **G** be subgroup of Z_p^* of order q
- CA assigns a unique generator $\mathbf{g_x}$ for all classes x∈ P
- Define an Auxiliary function $f : G \rightarrow [1,q]$

$$f(x) = \begin{cases} x; & x \leq q \\ p - x; & x > q \end{cases}$$

Example: Key generation



k_x denotes the key of class x

Key Derivation

- To compute the key of any successor, the node finds a path from itself to the successor
- Starting from the immediate successor in this path, the key k_i for every node p_i along the path is found using the algorithm below :
  - If p_i ***has only one predecessor*** p_j then
    - $k_i = f(g_i^{k_j})$
  - else {comment: p_j is the predecessor of p_i that is on the path}
    - $k_i = f(h_{i,j}^{k_j})$

Secure HKAS

- There are two notions of security in key hierarchies[5]
  - **Security against key recovery (KR-secure)**
  - Security against key distinguishability
- Wu-Wei's proof attempts to establish the former one
- A KR-secure HKAS (Informal Definition)
  - A key hierarchy is said to be secure, if any polynomial time adversary can not compute the key of a node even after colluding with all the nodes other than the attacking node and its predecessors

HKASes with proofs

- Despite the existence of multitude of HKASes, only the following schemes are proved secure
  - Wu-Wei's scheme [1]
    - **first provably secure scheme (KR-secure)**
    - ***But, problems in the proof***
  - Atallah et al's scheme [5,6]
    - Two schemes secure in two notions of security exist
  - Akl-Taylor's scheme (proof in [4])

Motivation

- Wu-Wei were the first to propose a provably secure HKAS
- Their proof model could be used by other dependent key HKAS* for providing security proofs
- But, their model and proof suffers from some problems
- So the fix to their model would produce a security model for dependent keyed HKAS

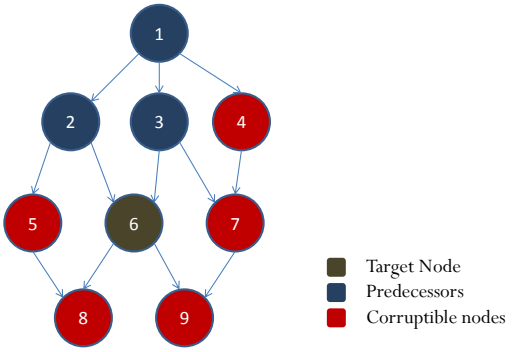
* keys used are dependent on their immediate predecessors (eg. Wu-Wei's scheme)

NOTATIONS

- g is a generator of group **G**
- p_t is the target node
- k_x is the key of node x
- g_x is the generator assigned to node x
- A is the set of immediate predecessors of p_t
- χ is the set of keys of immediate predecessors of p_t
- $\Pi(S)$ is the product of all elements in set S

Outline of Wu-Wei's proof

Corruptible Nodes



Set partitioning

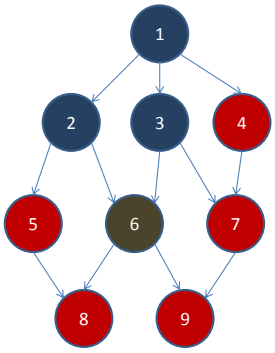
- Wu-Wei divided the corruptible nodes into 3 sets in order to prove the result incrementally in cases
  - **B :** contains all those nodes which have predecessors of target node as direct predecessors in it
  - **D :** contains immediate successors of target node
  - **R :** rest of the corruptible nodes
- *B - Definition* [4]: set of nodes in P-A, which have no predecessors in P-A and which are not p_t

Set partitioning (2)

- The proof then considers the following cases
  - Users in **B alone** collude
  - Users in **B U D** collude
  - All users collude
- But, their definition of B fails to achieve the actual aim, leading to the problems with some of the arguments in the proof
- Hence a new definition is proposed as a fix

Problem with the Definition

- Acc. to Wu-Wei's definition
  - B = {4,5}
- But, node 7 which has node 3 as common immediate predecessor with target node is missed
- Their definition fails in some cases



Users in B collude

- Claim [4] :When all the users in B collude the following is the whole information possibly held by them, which is related to k_{pt}*
  - Through public values related to p_t :
    $$\{g_t, g_t^{\prod(S)} | S \subset \mathcal{X} \text{ and } |S| = |\mathcal{X}| - 1\}$$
  - Through secret keys of nodes in B which share a **immediate predecessor** with p_t :
    $$\{g_{b_i}, g_{b_i}^{\prod(S)} | S \subseteq \chi \text{ and } b_i \in B\}$$

* we will see that this is not the case (only common immediate predecessors considered)

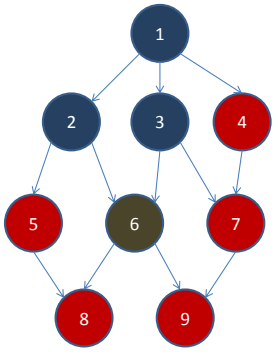
Incomplete information

- The proof considers only the info available to adversary when nodes in B **share a immediate predecessor** with p_t
- The information available to the nodes in B which share non-immediate predecessor(s)* is neglected

* Common predecessor, which is non-immediate to at least one amongst target node and node in B

Example

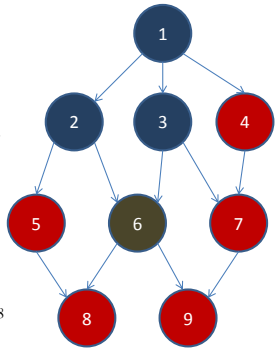
- In the example shown, 4's parent 1 is a **non-immediate predecessor** of the target node (i.e., 6)
- Also, $k_6 = f(g_6^{f(g_2^{k_1}) \cdot f(g_3^{k_1})})$
- 4 has the critical information $k_4 = f(g_4^{k1})$ in form of its own key
- But the **info. held by 4 is neglected**



Information available to the colluders is remodeled to cover all the possible data and the proof is fixed. We will not consider those details in this talk

Incorrect claim by Wu-Wei

- Claim [4] : When users in B collude, they can not distinguish between k_{pt} and a random nonce
- Consider the case where B = {4,5} collude
  - Adversary derives k_8 using k_5 and $h_{8,5}$
  - Now, when asked to determine whether 'challenge x' is k_{pt} or random nonce, adversary checks if $k_8 = f(h_{8,6}^{x})$ or not
  - Hence, we see that the above claim is incorrect



Conclusions

- Wu-Wei's scheme is the first provably secure scheme in the literature
- The security model and the proof of Wu-Wei suffer from few problems
- Our work so far has focused on producing a correct proof

Future work and goals

- The discussion so far is related to the provable security for static hierarchies
- No scheme in the literature is proved to be secure in the dynamic case
- Notions of security in dynamic key hierarchies is to be established
  - Proofs for dynamic HKAS should be provided using these established notions

Thank You

References

1. Wu, J., Wei, R.: An access control scheme for partial ordered set hierarchy with provable security. Selected Areas in Cryptography 2005, LNCS 3897 (2006) 221–232
2. J. Crampton, K. Martin, P. Wild, On key assignment for hierarchical access control, in: Proceedings of 19th Computer Security Foundations Workshop, 2006, pp. 98–111
3. P. Vadnala, A. Mathuria, An Efficient Key Assignment Scheme for Access Control in a Hierarchy, ICISS 2006: 205-219
4. P.D'Arco, A. De Santis, A.L. Ferrara, B. Masucci, Variations on a theme by Akl and Taylor: Security and tradeoffs, TCS 411(2010) 213-227
5. M.J. Atallah, M. Blanton, N. Fazio, K.B. Frikken, Dynamic and efficient key management for access hierarchies, ACM trans. Inf. Sysst. Secur. 12(3)(2009) Article 18
6. Selim G. Akl, Peter D. Taylor, Cryptographic Solution to a Problem of Access Control in a Hierarchy ACM Trans. Comput. Syst. 1(3): 239-248 (1983)

A Generic Weakness of the k-normal Boolean Functions Exposed to Dedicated Algebraic Attack

Miodrag Mihaljevic, Goutam Paul and Sugata Gangopadhyay

The Second Workshop on Cryptology and Information Security of Three Japan-India Projects

Tokyo, July 08, 2010

Roadmap

- Introduction and Motivation for the Work
- Underlying Ideas and the Framework for mounting Algebraic Cryptanalysis
- A Generic Weakness of the k-Normal Boolean functions Exposed to **Dedicated Algebraic Attacks**
- Concluding Remarks

I. Introduction

k-Normal Boolean Functions and motivation for the work

k-normal Boolean functions

Definition. Let $k \leq n$. A Boolean function f on $\mathcal{F}_2^n$ is called k-normal if there exists a k-dimensional flat on which f is constant.

A Toy Example.

$f(x_1, x_2, x_3, x_4, x_5, x_6) = x_1 \oplus x_2 \oplus x_3 \oplus x_1x_4 \oplus x_2x_5 \oplus x_3x_6$

$f(x_1 = 0, x_2 = 0, x_3 = 0, x_4, x_5, x_6) = 0$ independetly of x_4, x_5, x_6.

Statements of Claude Carlet regarding k-normal Boolean Functions

- "The complexity criterion we are interested in is non-k-normality with small k (smaller is k, harder is the criterion)."
- "**This complexity criterion is not yet related to explicit attacks on ciphers**."
- "The situation of the degree and of the nonlinearity, when they were first considered, was similar."
- "For instance, the linear attack has been discovered by Matsui sixteen years after Rothaus introduced the idea."

Motivation and Goals

- **Consideration of vulnerabilities** of cryptographic primitives which employ k-normal Boolean Functions.
- **Cryptanalysis of particular stream ciphers** which employ k-normal Boolean Functions**.**
- Developing of **dedicated algebraic** which employ a weakness of k-normal Boolean Functions**.**

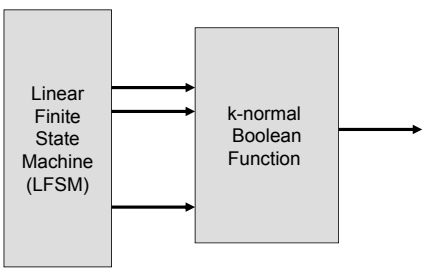
II. Cryptanalysis of Certain Keystream Generators Employing a Weakness of k-Normal Boolean Functions

Nonlinear Filter Generator and Combination Generator with k-Normal Boolean Functions

Boolean Functions and NF

- **Nonlinear Filter** (**NF**) is a textbook keystream generator but also can be considered as approximations of certain more complex generators.
- Design criteria and cryptographic complexity consideration of Boolean functions is usually related to their employment in NF.

Nonlinear Filter (NF)



III. Underlying Ideas and Theoretical Framework for the Cryptanalysis

mounting an attack for secret key recovery

Preliminary Considerations (1)

Suppose in the keystream we observe a run of zeroes of length m. We denote this event by $\mathrm{y} = 0$. Clearly $\Pr(\mathrm{y} = 0) = 2^{-m}$. In order to obtain a keystream of length m, we need a sequence of length $\Delta + (m \quad 1)$ from the $LFSR_0$.

Let I denote the set of all $LFSR_0$ sequences x of length $\Delta + (m-1)$ such that the following system of equations is satisfied.

$$\begin{aligned} X^{(t)}_{i_{k+1}} &= \bigoplus_{j=1}^{k} a_{k+1j} X^{(t)}_{i_j} \oplus b_{k+1} \\ X^{(t)}_{i_{k+2}} &= \bigoplus_{j=1}^{k} a_{k+2j} X^{(t)}_{i_j} \oplus b_{k+2} \\ \cdot \quad & \cdot \quad \cdot \\ \cdot \quad & \cdot \quad \cdot \\ \cdot \quad & \cdot \quad \cdot \\ X^{(t)}_{i_n} &= \bigoplus_{j=1}^{k} a_{nj} X^{(t)}_{i_j} \oplus b_n \end{aligned}$$

for all t such that $0 \le t \le \Delta + (m-1)$.

Preliminary Considerations (2)

Lemma. Let L be the set of all $LFSR_0$ sequences $\mathbf{x}$ of length $\Delta + (m-1) \le N$, and I as defined above. Then there exists δ such that

$$\Pr(\mathbf{x} \in I) = 2^{-(\Delta + m - 1 - \delta)}$$

$$\Pr(\mathbf{x} \in I | \mathrm{y} = 0) = 2^{-(\Delta - \delta - 1)}.$$

Underlying Ideas for Mounting Algebraic Attack

Development of the algebraic technique for cryptanalysis originates from the following:
- The k-normality provides that appearance of a run of m zeros in the keystream sequence implies that certain hypothesis on the corresponding input pattern to the Boolean function are correct with a probability significantly higher in comparison with the random guessing.

The main underlying idea for the cryptanalysis is to employ this heavily biased probability for mounting an algebraic attack for recovering the initial state of the generator which corresponds to the first bit in the considered run.

Two Phases Framework for Cryptanalysis

Phase I:

- Pre-Processing: **Independent of any Secret Key or Sample**
- **Should be done only once.**
- A Preparation for the secret key recovery

Phase II:

- **Generator Internal state and Secret Key Recovery** for a given sample.

IV. Dedicated Algebraic Attack

IV.1 Algorithm of Pre-Processing

A Preparation Phase:
Should be Performed Only Once

- Pre-compute all possible solutions for the FSM state which yields the first bit of a run of m zeros in the output of the keystream generator;

- For each candidate state generate the corresponding N-bit output from the generator subsequent to the run of zeroes of length m and store the state and the generator output in a row of two-columns table $\mathbf{T}$.

- *Input*: The parameters m and N.

- *Construction of the Pairs State-Keystream* Perform the steps (1)-(3) from the next slide.

- *Output*: Table $\mathbf{T}$ with collection of the following pairs: (i) state of FSM; (ii) corresponding output sequence from the keystream generator.

1. Employing the state transition function $\mathcal{F}(\cdot)$, rewrite the system of equations so that the all unknown variables belong to the state which correspond to the beginning of the run.

2. Determine and list all possible solutions of the system of above equations. The number of solutions of the system of equations is 2^{N-r}, where r is a parameter which depends on the state transition function and the taps (from the state to the k-normal Boolean function).

3. For each FSM candidate state generate the corresponding N-bit output sequence of the generator, and memorize the pair (state, output sequence) as a row of the two-columns table $\mathbf{T}$.

IV.2 Algorithm for the Internal State and Secret Key Recovery

For a Given Sample Recovers the Secret Key

- Search for runs of zeroes of length m in the keystream.

- For each run compare the N-bit segment of the keystream after the considered run with the memorized N-bit segments in the second column of the table $\mathbf{T}$.

- If the sample segment match with a segment in the table, read the corresponding state and based on the given state recover the secret key.

- *Input*:
  - The table $\mathbf{T}$, and the parameters m and N;
  - the sample $\{y_t\}_{t=1}^{N}$.

- *Processing Steps*
  Perform the steps (1)-(3) from the next slide

- *Output*:
  (a) recovered secret key;
  (b) flag that the algorithm has failed to recover the secret key.

1. Search the sequence $\{y_t\}_t$ for the next not considered run of m zeros; Let this run has been found for $t \in \{\tau, \tau+1, ..., \tau+m-1\}$. If there are no more runs of m zeros, go to Output (b).

2. Compare N bits of the keystream following the run of zeroes of length m with the second column of the table $\mathbf{T}$, and do the following:
- if the matching does not exist, go to the processing Step 1;
- if the matching is found, read the state in the pair with the keystream segment and go to the Step 3.

3. Via an inverse transformation (solving the corresponding system of equations) recover the secret key which corresponds to the considered FSM state. Go to Output (a).

V. Concluding Notes

Main Messages of This Talk (1)

- This talk points out some **possible vulnerabilities of cryptographic primitives which employ k-normal Boolean functions**.
- Particularly, this talk confirms that the **Non-Normality is an important design criteria for Boolean functions**
- The framework for exploiting weaknesses of k-normal Boolean functions employing **dedicated algebraic attacking approache** is pointed out.

Main Messages of This Talk (2)

- An approach for cryptanalysis of the considered keystream generators is proposed which is based on the possibility for pre-computing a table of the state-keystream pairs via solving certain system of algebraic equations as a consequence of the employed k-normal Boolean function.
- This pre-computed table is the main origin for mounting the cryptanalysis and it is independent of a the sample for cryptanalysis and the secret key employed for generating the sample.
- **An algorithm is proposed for construction of the table employing a system of equations corresponding to an m-run of zeros in the keystream sequence and implied by the k-normality**.
- Higher k-normality implies a smaller dimension of the table and a lower complexity of the pre-processing as well as complexity of entire cryptanalysis.

VI. References

Illustrative References

- C. Carlet, “The complexity of Boolean functions from cryptographic point of view”, in Complexity of Boolean Functions, *Dagestuhl Seminar Proceedings 06111*, 2006.
- C. Carlet, “On the degree, nonlinearity, algebraic thickness and nonnormality of Boolean functions, with developments on symmetric functions”, *IEEE Transactions on Information Theory*, vol. 50, pp. 2178-2185, 2004.
- C. Carlet, H. Dobbertin and G. Leander, “Normal Extensions of Bent Functions”, *IEEE Trans. on Information Theory*, vol. 50} no. 11, pp. 2880 – 2885, 2004.
- P. Charpin, “Normal Boolean functions”, *Journal of Complexity*, vol. 20, pp. 245 – 265, 2004.
- E. Pasalic, "On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers", IEEE Trans. Inform. Theory}, vol. 55, pp. 3398-3406, July 2009.

Thank You Very Much for the Attention,

and

QUESTIONS Please!

National Institute of Advanced Industrial Science and Technology AIST

RSIS Research Center for Information Security

RFID Authentication Using Flexible Quasi-Dyadic Fix Domain Shrinking

Yang Cui
Joint Work with Tomohiro Sekino, Kazukuni Kobara and Hideki Imai

2010.07.08

Contents

- Introduction
- Preliminaries
- Our Proposal
- Conclusion

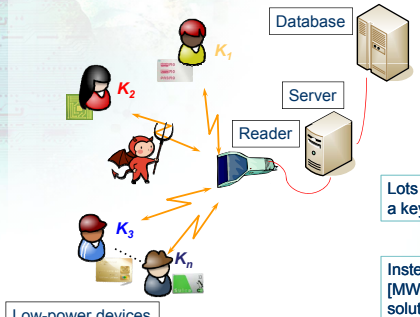
Light-Weight Crypto Applications

Contents
•Small commodities
•Significant biographical information

Threats
•Identify Friend or Foe
•Sniffing and Tracking
•Denial of Service

Authenticity!

Problem of Symmetric Authentication

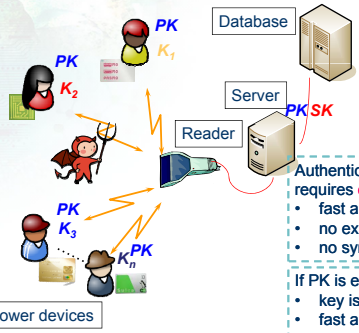


	Key	Item	Time
ID_1	K_1		
ID_2	K_2		
⋮	⋮	⋮	⋮
ID_n	K_n		

Lots of keys involved, lead to a key management problem

Instead of exhaustive testing, [MW04]…[CHT09] provide solutions w/t computational cost around $O(n^{1/2}\log_2 n)$

Advantage of Asymmetric Authentication



	Key	Item	Time
ID_1	K_1		
ID_2	K_2		
⋮	⋮	⋮	⋮
ID_n	K_n		

Authentication in pervasive devices, requires efficient key management
- fast and low-cost computation
- no exhaustive search for all keys
- no synchronization

If PK is employed, then
- key is easier to store, since public
- fast authentication (1 decryption)
- but high-cost computation!!!

Motivation

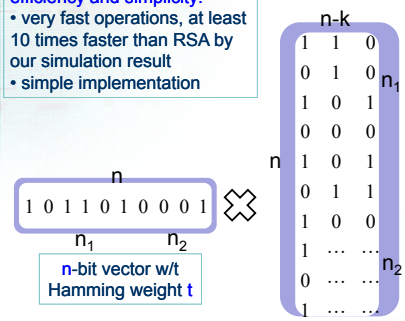
- Can we build a light-weight authentication protocol, which satisfies the following,
  - efficient key management (in server side)
  - fast computation (in low-power device side)
  - acceptable for low-power devices
    - Public-key encryption, such as RSA is difficult to realize on low-power device, due to expensive modular exponentiation calculation
    - Symmetric key encryption, such as AES is considered to achieve a reasonable computation cost
    - A typical RFID chip only has a few kilo gates, very restricted.

Our Idea: Niederreiter Encryption

efficiency and simplicity:
- very fast operations, at least 10 times faster than RSA by our simulation result
- simple implementation

$mK^T = c$

security:
- relies on a random linear decoding problem
- a successful brute-force attack will solve NP-C problem
- stay alive for more than 20 years



PK: n × (n-k) parity check matrix K^T
SK: decoding algorithm of K

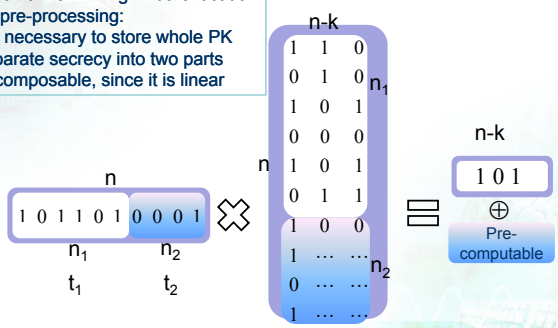
Is that enough?

- Unfortunately, PK n × (n-k) is too large!
  - A typical parameter,n=1024,k=524,|PK|≈500kb
    - RSA: 1024 bit, ECC: 320 bit
  - We solve one problem (computation), but introduce one (storage) at the same time
- However, notice that:
  - computation bottleneck is crucial
  - Memory becomes much cheaper

Our Technique 1

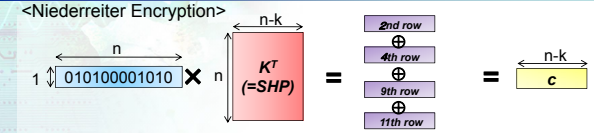
Fix domain shrinking - Authentication with pre-processing:
- not necessary to store whole PK
- separate secrecy into two parts
- decomposable, since it is linear

$mK^T = c$

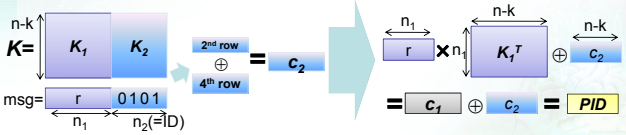


Improved Fast Encryption

<Niederreiter Encryption>



<Authentication by Fix Domain Shrinking>

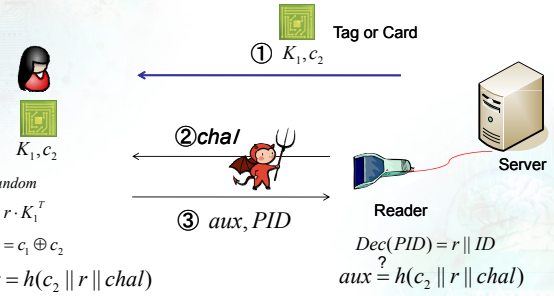


<Decryption> output r 0101

Sufficient to store K_1 and c_2 for authentication

A Challenge-Response Authentication Protocol

- Low-power device, reader, server



r : random
$c_1 = r \cdot K_1^T$
$PID = c_1 \oplus c_2$
$aux = h(c_2 \| r \| chal)$

$Dec(PID) = r \| ID$
$aux \stackrel{?}{=} h(c_2 \| r \| chal)$

Storage Cost by Using Fix Domain Shrinking (not optimized)

Security (log)	n_1	n_2	k	t	t_1	#Id(log)	|PK|	Total Storage (bit)
80	871	153	824	20	18	13.5	134,200	134,400
82	861	163	774	25	22	19.4	152,750	153,000
83	854	170	724	30	26	30.3	166,200	166,500
81	849	175	674	35	30	30.3	174,650	175,000

Compared with 500kb PK of direct use of Niederreiter, we obtain a shorter PK.
*Data evaluated in 2007, some are not good due to new attacks

Our Technique 2

PK is still big, how small can we get while keeping the security?

Idea: to use error-correcting codes with more structures.

Dyadic Matrix

- symmetric property:
  For any Dyadic M, there is $M_{ij} = h_{i \oplus j}$

eg. t × t Dyadic Matrix (t=2)

$M = \begin{bmatrix} h_0 & h_1 \\ h_1 & h_0 \end{bmatrix}$ =

Red needs to store

Quasi-Dyadic Matrix

- composed of Dyadic Matrix
- but not necessary to be all Dyadic

Quasi-Dyadic Matrix =

Red needs to store

Quasi-Dyadic Goppa Code

- e.g. (n=14,k=4,t=2) Quasi-Dyadic Goppa Code; m=5

  reduce H from (n-k) × k ⇒ m × k

n-k H=

k													
0	1	0	1	1	0	0	0	0	0	0	0	0	0
1	0	1	0	0	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	1	1	0	0	0	0	1	0	0	0	0	0
0	0	1	1	0	0	0	0	0	1	0	0	0	0
0	1	1	0	0	0	0	0	0	0	1	0	0	0
1	0	0	1	0	0	0	0	0	0	0	1	0	0
1	1	0	0	0	0	0	0	0	0	0	0	1	0
1	1	0	0	0	0	0	0	0	0	0	0	0	1

PK:10 × 4
(n-k) × k=mt × k

QD ($n=2^m-t$)
× 1/t

H =

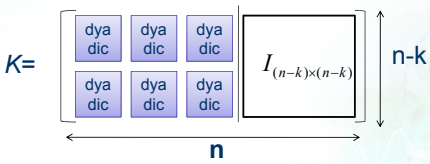
k			
0	1	0	1
1	0	1	0
0	1	0	0
1	0	0	0
0	0	1	1
0	0	1	1
0	1	1	0
1	0	0	1
1	1	0	0
1	1	0	0

(n-k)/t

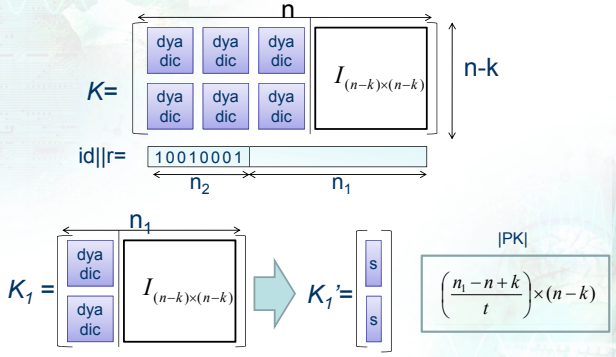
PK':10 × 2
(n-k)/t × k=m × k

To Generate PK

1. To generate parity check matrix n × (n-k) with Quasi-Dyadic properties
2. Randomized without breaking QD structure
3. Use Gaussian elimination to simplify matrix
4. Different from other work, we make use of a specific "Flexible Quasi-Dyadic" generation



Flexible Quasi-Dyadic with Fix Domain Shrinking



Storage Cost by using Fix Domain Shrinking & Flexible Quasi-Dyadic

Security (log)	n	n_1	k	t	t_1	#Id(log)	c_2	|K_1'|	Total Storage (bit)
82	2016	864	1664	32	30	19	352	5632	5984
83	1984	1088	1280	64	60	34	704	4224	4928
91	1920	1440	864	96	80	97	1056	4224	5280
81	1920	1568	688	35	112	72	1232	3696	4928

Compared with 500kb PK of direct use of Niederreiter, and 134kb of Fix Domain Shrinking, we are able to achieve more efficient 4928-bit storage cost.
*Evaluation has included the latest attacks.

Conclusion

- To provide a light-weight authentication protocol with
  - Efficient key management
  - Fast computation
  - Small storage cost
- Our proposal makes use of
  - Coding theory based encryption
  - Fix domain shrinking
  - Flexible Quasi-Dyadic code

Computationally Secure Communication in the Wire-tap Scenario

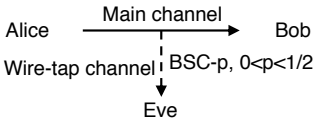
(ongoing research)

Kirill Morozov
(RCIS, AIST)

2010 Japan-Indo Joint Workshop on Cryptology and Related Areas

July 8, 2010

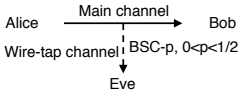
Wire-Tap Scenario



- *Wiretap system*: Adversary Eve is eavesdropping communication between legal players Alice and Bob, over an *independent channel*

- Our focus on the following (for now simple) setting: noiseless main channel (say, binary), a non-trivial BSC (error rate p: 0<p<1/2) as wiretap channel
- Goal: Secure communication from Alice to Bob
- Standard approach: A key generation paradigm introduced by Wyner [1] and generalized by Maurer [2] can be used

Standard Approach



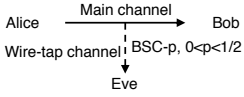
- An information-theoretically secure key generation paradigm introduced by Wyner [1] and generalized by Maurer [2] can be used
- Maurer's paradigm:
  - Random data exchange
  - Information reconciliation (error correction)
  - Privacy amplification (randomness extraction)
- Secure communication are achieved as follows: the shared is used as one-time pad for messages
- Main disadvantage: Precise knowledge of channel parameters (error rate, in our case) is required
  - Otherwise, no clear security guarantee is devised
- Other disadvantages: Privacy amplification relies on randomness extraction either by universal hashing (which needs local randomness of size of the whole data) or extractors (which have complicated implementation)

Our Motivation

- Noise in communication media is present "for free"
  - We only need to know how to utilize it
- In reality, noise can be *anything*
  - Say, in terms of error rate: from 0 to ½
- A "ramp scheme" is desirable, where security vanishes *gradually and smoothly* with the error rate
  - E.g. *some guarantee* is better than no guarantee at all!
- To achieve the above objective, we will (for now) give up information-theoretic security and focus on…
- … Secure communication with computational security
  - I.e. based on some hardness assumption

Preliminaries

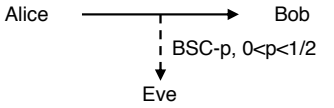
- Remark 1: In order to exhibit the power of noise, and also to keep this presentation easy-to-understand, we focus on the simplest possible setting



- Let Eve (as well as Alice and Bob) be a PPT algorithm
- Eve has no control over the main channel (which is authenticated)
- (Hidden) Assumption: Messages must be long enough for the noise to kick in (whatever it means)

- Remark 2: Alice and Bob have no pre-shared key

Trivial Scheme

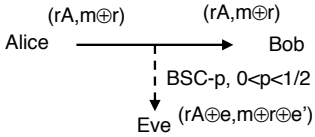


- Let $m \in \{0,1\}^n$ be a message
- Let m be *uniformly distributed* $\{0,1\}^n$
- n is "long enough"

- Alice sends m which is received by Bob
- Security: Eve learns m⊕e, where e is the noise vector distributed as Bernoulli(p), i.e. having pn errors with high probability (w.h.p.) in n
- W.h.p. the scheme is one-way (meaning, it is hard for Eve to recover the whole m)
  - Not so high security and also uniform distribution of messages is required
- We would prefer at least IND-CPA (meaning, any message of some pair can be chosen by the adversary)
  - Then, we will not care about the distribution of plaintexts

Basic Scheme

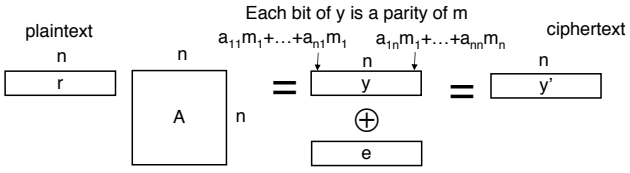
- Let $A \in \{0,1\}^{n \times n}$ be invertible, randomly chosen, *public matrix*



- Alice generates $r \in_R \{0,1\}^n$ and sends (rA,m⊕r)
- Bob computes $rAA^{-1}=r$ and m⊕r⊕r=m
- Eve receives (rA⊕e,m⊕r⊕e'), where e,e'∈Bernoulli(p)
  - I.e. both e and e' have Hamming weight d=pn w.h.p.
  - d is the security parameter
- For IND-CPA, i.e. computational indistinguishability of the ciphertext, it is enough to prove that the distribution of Ar+e is computationally indistinguishable from random

Proof Roadmap

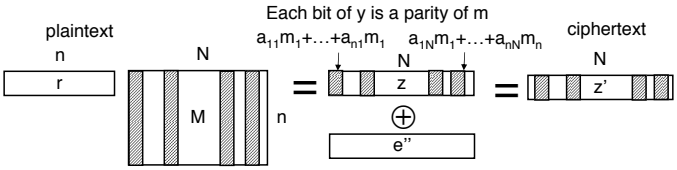
- At the moment, there is no formal proof, but only an intuition why it may work
- Suppose that there is "enough noise for security" (whatever it means)
- We got the following situation:



- This reminds very much of the Learning Parity with Noise (LPN) problem

Learning Parity with Noise

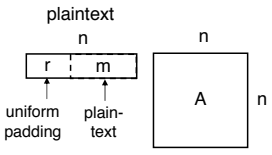
- LPN oracle: $r \in_R \{0,1\}^n$ is secret, $M \in_R \{0,1\}^{n \times N}$ is public



- If r is uniform, then recovering of r is believed to be hard
  - This is the Bounded Distance Decoding problem known to be NP-hard
- Suppose that I got an inverter for MyProblem (the one above)
- If n is a linear fraction of N (typical case), then M is rank-n w.h.p.
- Collect n columns of M which make up a rank-n (shown as dashed), and the corresponding bits of y'
- Submit the resulting matrix and vector to the inverter as A and y' and return the output of the inverter

Details

- The output of the LPN oracle is proven pseudorandom by Katz and Shin [3]
- In order to complete our proof, we must mimic their proof, while have the setting of MyProblem at hands
- If the proof goes through, then we may try to save on local randomness using random padding of the message a la "semantically secure McEliece encryption in the standard model" by Nojima et al [4]



= . . .

- This is just my speculation, there is proof intuition

Conclusion

- We have proposed to investigate the effects of physical (low-weight) noise to (computational) security of information transmission
- Our analysis shows that even an extremely simple primitives may provide (at least computational) security in this case

Possible Applications

- Such systems may be particularly suitable for light-weight systems, however, any communication with noise is eligible
- Also, we note some similarities with side-channel attacks
  - Typically, side-channel adversary experience noisy measurement, hence our techniques might be applicable there as well

Extensions

- Noise-based key exchange with hybrid ramp security
  - It is interesting to combine our scheme with the known information-theoretically secure key exchange systems to get a scheme whose security gradually reduces while noise is decreasing
  - The security guarantee must switch from information-theoretic to computational at some point
  - This would greatly improve practical attractiveness of noise-based schemes
  - This will require dealing with noisy main channel and also with continuous noise
- Cryptographic primitives enhanced by low-weight noise
  - Effects of noise on popular cryptographic schemes will be considered
  - Example: How much adding a low-weight noise increases security of RSA encryption?
  - What is the optimal way to utilize noise in this setting?
  - What about other primitives?

References

1. A. Wyner: The wiretap channel, Bell Syst Tech. J. 54: 1355-1387, October (1975)
2. Ueli M. Maurer: Secret key agreement by public discussion from common information. IEEE Transactions on Information Theory 39(3): 733-742 (1993)
3. J. Katz, J.S. Shin: Parallel and Concurrent Security of the HB and HB+ Protocols, EUROCRYPT 2006: 73-87 (2006)
4. R. Nojima, H. Imai, K. Kobara, K. Morozov: Semantic security for the McEliece cryptosystem without random oracles. Des. Codes Cryptography 49(1-3): 289-305 (2008)

A Low Complexity Encryption Technique Based on Joint Employment of Pseudorandomness, Randomness and Coding

Miodrag Mihaljevic and Hideki Imai

The Second Workshop on Cryptology and Information Security of Three Japan-India Projects

Tokyo, July 08, 2010

I. Roadmap

- A Practical Stream Cipher Based on Randomness and Dedicated Coding
- Algebraic Representation of Encryption and Security Evaluation Approaches
- Computational Complexity Evaluation
- Comparison with Some Recently Reported Results
- Concluding Remarks

II. A Stream Cipher Based on Randomness and Dedicated Coding

a generic scheme

and

a particular instantiation

Recent Achievements in RCIS on Stream Ciphers Based on Randomness&Coding

- M. Mihaljevic and H. Imai, **"An Approach for Stream Ciphers Design Based on Joint Computing over Random and Secret Data"**, *COMPUTING*, vol. 85, no. 1-2, pp. 153-168, June 2009.
- M. Mihaljevic, **"A Framework for Stream Ciphers Based on Pseudorandomness, Randomness and Error-Correcting Coding"**, in *Enhancing Cryptographic Primitives with Techniques from Error Correcting Codes,* Eds. *B. Preneel et al.,* Vol. 23 in the Series Information and Communication Security, pp. 117-139, IOS Press, Amsterdam, June 2009.
- M. Mihaljevic and H. Imai, "**A Stream Cipher Design Based on Embedding of Random Bits**", *IEEE 2008 Int. Symp. on Inform. Theory and its Appl. - ISITA2008*, Auckland, New Zealand, Dec. 7-10, 2008, Proceedings, pp. 1497-1502.

Motivation and Goals

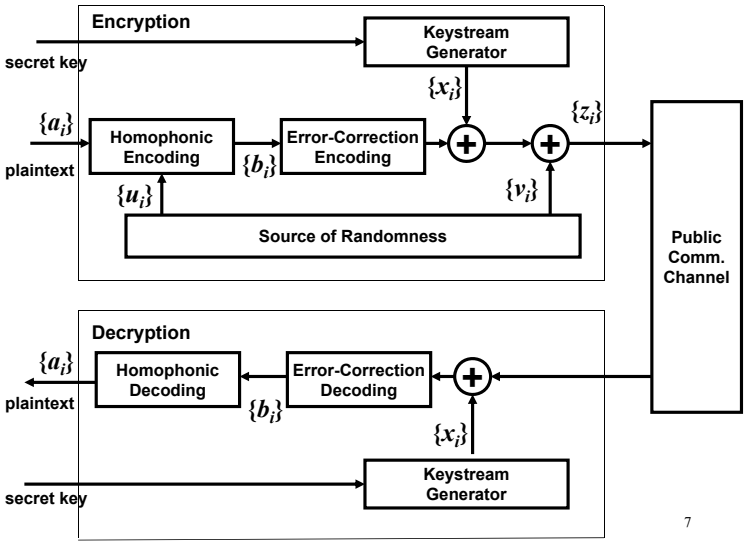
- Developing a **novel framework for stream ciphers design based on randomness and dedicated coding** which could provide security close to the maximal possible one determined by the employed secret key.
- Based on the already achieved results**, design of an implementation practical and secure** stream ciphers.
- **Security evaluation of the design**.
- Evaluation of implementation complexity and communications overhead.

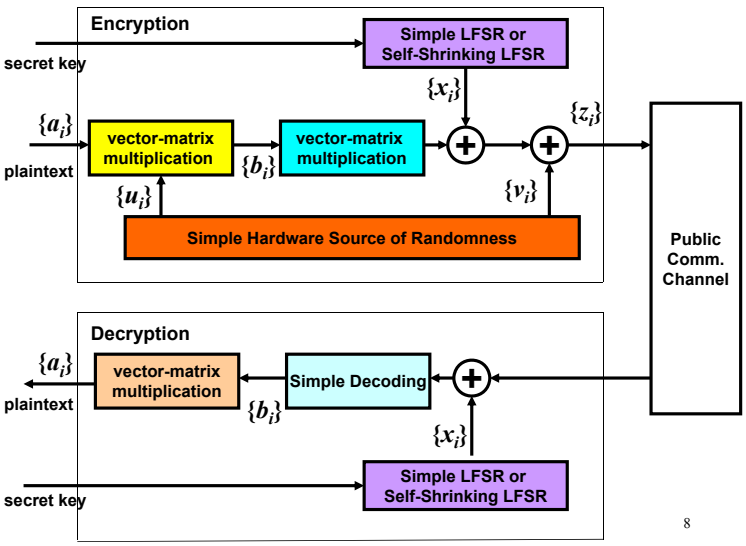
II. A Stream Cipher Based on Randomness and Dedicated Coding

a generic scheme

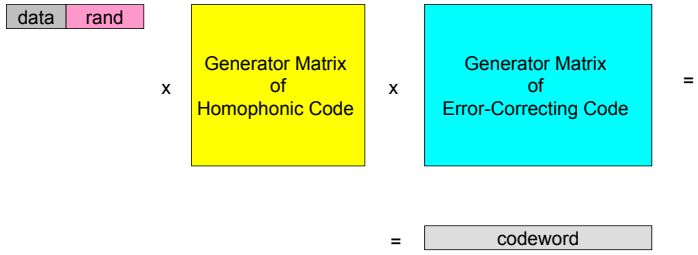
and

a particular instantiation





Homophonic (Wire-Tap Channel) and Error-Correction Encoding



Origins of for the Enhanced Security

- **Effects of involvement randomness**.
- **Hardness of decoding without secret key.**

III. Algebraic Representation of the Encryption

Word-level and Bit-level Representations

When the employed homophonic (wire-tap channel) and error-correcting codes are linear the encoding operations in the both cases are vector-matrix multiplications. Accordingly, the encoded version of $\mathbf{a}||\mathbf{u}$ is given by the following:

$$C_E(C_H(\mathbf{a}||\mathbf{u})) = [\mathbf{a}||\mathbf{u}]\mathbf{G}_H\mathbf{G}_E = [\mathbf{a}||\mathbf{u}]\mathbf{G}$$

where $\mathbf{G}_H$ is $m \times m$ binary matrix corresponding to $C_H(\cdot)$, $\mathbf{G}_E$ is $m \times n$ binary matrix corresponding to $C_E(\cdot)$, and $\mathbf{G} =$ is $m \times n$ binary matrix. Let $\mathbf{G} = [g_{i,j}]_{i=1}^{m}{}_{j=1}^{n}$.

$$\mathbf{z} = [\mathbf{a}||\mathbf{u}]\mathbf{G} \oplus \mathbf{x} \oplus \mathbf{v}$$

$$= [\mathbf{a}||\mathbf{u}] \begin{bmatrix} \mathbf{g_1} \\ \mathbf{g}_2 \\ . \\ . \\ . \\ \mathbf{g}_m \end{bmatrix} \oplus \mathbf{x} \oplus \mathbf{v}$$

$$= (a_1\mathbf{g}_1 \oplus ... \oplus a_\ell \mathbf{g}_\ell) \oplus (u_1\mathbf{g}_{\ell+1} \oplus ... \oplus u_{m-\ell}\mathbf{g}_m) \oplus \mathbf{x} \oplus \mathbf{v}$$

When $a_i = 0$, $i = 1, 2, ..., \ell$, we have

$$\mathbf{x} = \mathbf{z} \oplus (u_1\mathbf{g}_{\ell+1} \oplus ... \oplus u_{m-\ell}\mathbf{g}_m) \oplus \mathbf{v}$$

When $\mathbf{x} = \mathbf{x}_0\mathbf{S}$, we have

$$\mathbf{x}_0\mathbf{S} = \mathbf{z} \oplus (u_1\mathbf{g}_{\ell+1} \oplus ... \oplus u_{m-\ell}\mathbf{g}_m) \oplus \mathbf{v}$$

Let $\mathbf{z} = [z_i]_{i=1}^{n}$. Then,

$$z_i = (\bigoplus_{k=1}^{\ell} g_{k,i}a_k) \oplus (\bigoplus_{k=1}^{m-\ell} g_{\ell+k,i}u_k) \oplus x_i \oplus v_i \ , \ i = 1, 2, ..., n,$$

implying that under the known plaintext attack we have

$$x_i \oplus (\bigoplus_{k=1}^{m-\ell} g_{\ell+k,i}u_k) \oplus v_i = z_i \oplus (\bigoplus_{k=1}^{\ell} g_{k,i}a_k) \ , \ i = 1, 2, ..., n,$$

where the right-hand side of the equation has known value.

IV. Security Evaluation

Computational Complexity Approach

Some Background References for Computational Complexity Security Evaluation

- M. Mihaljevic and H. Imai, "**An approach for stream ciphers design based on joint computing over random and secret data**", *Computing*, vol. 85, no. 1-2, pp. 153-168, June 2009.
- M. Fossorier, M. Mihaljevic and H. Imai, "**Modeling Block Encoding Approaches for Fast Correlation Attack**", *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4728-4737, Dec. 2007.
- M. Mihaljevic, M. Fossorier and H. Imai, "**Security Evaluation of Certain Broadcast Encryption Schemes Employing a Generalized Time-Memory-Data Trade-Off**", *IEEE Communications Letters*, vol. 11, no. 12, pp. 988-990, Dec. 2007.

Security Implied by Hardness of Recovering Secret Key Based on the Algebraic Representation of Encryption

- The Computational Complexity -

"Parity-Checks" of the Random Bits (1)

Note that $n > m-\ell$ and there are just $m-\ell$ independent realizations $\{u_i\}_{i=1}^{m-\ell}$ of the random variables U_i, $\Pr(U_i = 1) = \Pr(U_i = 1) = 1/2$, $i = 1, 2, ..., m-\ell$. Accordingly, always exists a vector $[\alpha_1, \alpha_2, ..., \alpha_n] \in \{0,1\}^n$ such that

$$\bigoplus_{i=1}^{n} \alpha_i \Big(\bigoplus_{k=1}^{m-\ell} g_{\ell+k,i} u_k \Big) = 0$$

"Parity-Checks" of the Random Bits (2)

The previous implies that for each $i = 1, 2, ..., n$, we have the following equation:

$$x_i \textcircled{\scriptsize |} \bigoplus_{k=1,\, k\neq i}^{n} \alpha_k x_k = (z_i \textcircled{\scriptsize |} \bigoplus_{k=1,\, k\neq i}^{n} \alpha_k z_k) \textcircled{\scriptsize |} (v_i \textcircled{\scriptsize |} \bigoplus_{k=1,\, k\neq i}^{n} \alpha_k v_k) \, .$$

Basic System of Equations Related to a **Single Word** when the Plaintext Consists of all Zeros

$$\begin{array}{ccccccc}
x_1^{(t)} & = & z_1^{(t)} & \oplus & \mathcal{L}_1(\{u_i^{(t)}\}_i) & \oplus & v_1^{(t)} \\
x_2^{(t)} & = & z_2^{(t)} & \oplus & \mathcal{L}_2(\{u_i^{(t)}\}_i) & \oplus & v_2^{(t)} \\
\vdots & & & & & & \\
x_{m-\ell}^{(t)} & = & z_{m-\ell}^{(t)} & \oplus & \mathcal{L}_{m-\ell}(\{u_i^{(t)}\}_i) & \oplus & v_{m-\ell}^{(t)} \\
\hline
x_{m-\ell+1}^{(t)} & = & z_{m-\ell+1}^{(t)} & \oplus & \mathcal{L}_{m-\ell+1}(\{u_i^{(t)}\}_i) & \oplus & v_{m-\ell+1}^{(t)} \\
x_{m-\ell+2}^{(t)} & = & z_{m-\ell+2}^{(t)} & \oplus & \mathcal{L}_{m-\ell+2}(\{u_i^{(t)}\}_i) & \oplus & v_{m-\ell+2}^{(t)} \\
\vdots & & & & & & \\
x_n^{(t)} & = & z_n^{(t)} & \oplus & \mathcal{L}_n(\{u_i^{(t)}\}_i) & \oplus & v_n^{(t)}
\end{array}$$

Processed System of Equations Related to a Single Word when the Plaintext Consists of all Zeros

$$\begin{array}{ccccccc}
x_1^{(t)} & = & z_1^{(t)} & \oplus & \mathcal{L}_1(\{u_i^{(t)}\}_i) & \oplus & v_1^{(t)} \\
x_2^{(t)} & = & z_2^{(t)} & \textcircled{\scriptsize |} & \mathcal{L}_2(\{u_i^{(t)}\}_i) & \textcircled{\scriptsize |} & v_2^{(t)} \\
\vdots & & & & & & \\
x_{m-\ell}^{(t)} & = & z_{m-\ell}^{(t)} & \oplus & \mathcal{L}_{m-\ell}(\{u_i^{(t)}\}_i) & \oplus & v_{m-\ell}^{(t)} \\
\hline
\mathcal{L}^*_{m-\ell+1}(\{x_i^{(t)}\}_i) & = & \mathcal{L}^*_{m-\ell+1}(\{z_i^{(t)}\}_i) & \oplus & 0 & \oplus & \mathcal{L}^*_{m-\ell+1}(\{v_i^{(t)}\}_i) \\
\mathcal{L}^*_{m-\ell+2}(\{x_i^{(t)}\}_i) & = & \mathcal{L}^*_{m-\ell+2}(\{z_i^{(t)}\}_i) & \oplus & 0 & \oplus & \mathcal{L}^*_{m-\ell+2}(\{v_i^{(t)}\}_i) \\
\vdots & & & & & & \\
\mathcal{L}^*_n(\{x_i^{(t)}\}_i) & = & \mathcal{L}^*_n(\{z_i^{(t)}\}_i) & \oplus & 0 & \oplus & \mathcal{L}^*_n(\{v_i^{(t)}\}_i)
\end{array}$$

Aggregated System of Equations when the Plaintext Consists of all Zeros

$$
\begin{array}{lcl}
\vdots & & \\
\mathcal{L}^*_{m-\ell+1}(\{x_i^{(t)}\}_i) & = & \mathcal{L}^*_{m-\ell+1}(\{z_i^{(t)}\}_i) \oplus \mathcal{L}^*_{m-\ell+1}(\{v_i^{(t)}\}_i) \\
\mathcal{L}^*_{m-\ell+2}(\{x_i^{(t)}\}_i) & = & \mathcal{L}^*_{m-\ell+2}(\{z_i^{(t)}\}_i) \oplus \mathcal{L}^*_{m-\ell+2}(\{v_i^{(t)}\}_i) \\
\vdots & & \\
\mathcal{L}^*_n(\{x_i^{(t)}\}_i) & = & \mathcal{L}^*_n(\{z_i^{(t)}\}_i) \oplus \mathcal{L}^*_n(\{v_i^{(t)}\}_i) \\
\mathcal{L}^*_{m-\ell+1}(\{x_i^{(t+1)}\}_i) & = & \mathcal{L}^*_{m-\ell+1}(\{z_i^{(t+1)}\}_i) \oplus \mathcal{L}^*_{m-\ell+1}(\{v_i^{(t+1)}\}_i) \\
\mathcal{L}^*_{m-\ell+2}(\{x_i^{(t+1)}\}_i) & = & \mathcal{L}^*_{m-\ell+2}(\{z_i^{(t+1)}\}_i) \oplus \mathcal{L}^*_{m-\ell+2}(\{v_i^{(t+1)}\}_i) \\
\vdots & & \\
\mathcal{L}^*_n(\{x_i^{(t+1)}\}_i) & = & \mathcal{L}^*_n(\{z_i^{(t+1)}\}_i) \oplus \mathcal{L}^*_n(\{v_i^{(t+1)}\}_i) \\
\vdots & &
\end{array}
$$

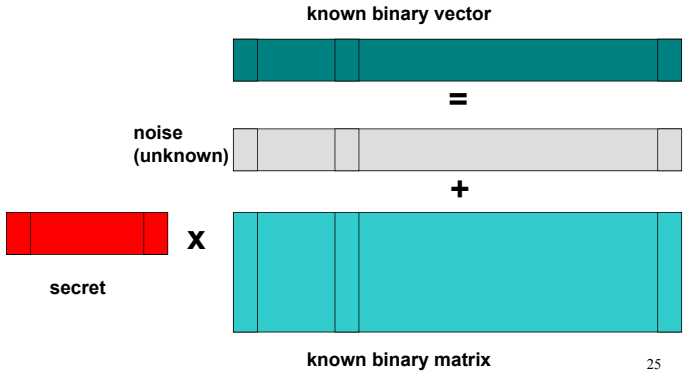
A Particular Model of Keystream Generation

Let $\mathbf{x}_0$ be an unknown n-dimensional binary vector and $\mathbf{S} = [\mathbf{S}_i]_{i=1}^n$ be a known $n \times n$-dimensional binary matrix where each $\mathbf{S}_i$ is an n-dimensional column vector. Let $\mathbf{S}^t = [\mathbf{S}_i^{(t)}]_{i=1}^n$ denotes the t-th power of the matrix $\mathbf{S}$. When $x_i^{(t)} = \mathbf{x}_0\mathbf{S}_i^{(t)}$, $i = 1, 2, ..., n$, $t = 1, 2, ...$, the above system of equations becomes the following one:

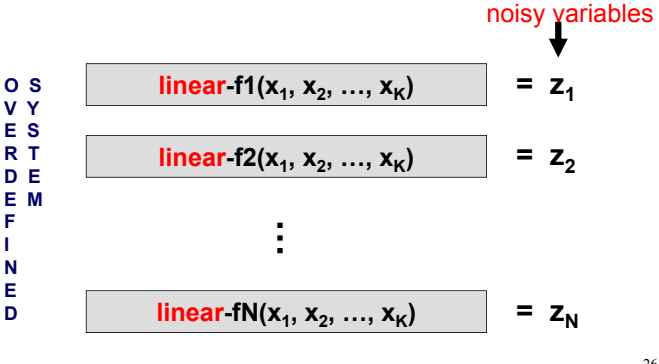
The **Aggregated System** under a **Particular Keystream Model**

$$
\begin{array}{lcl}
\vdots & & \\
\mathcal{L}^*_{m-\ell+1}(\{\mathbf{x}_0\mathbf{S}_i^{(t)}\}_i) & = & \mathcal{L}^*_{m-\ell+1}(\{z_i^{(t)}\}_i) \oplus \mathcal{L}^*_{m-\ell+1}(\{v_i^{(t)}\}_i) \\
\mathcal{L}^*_{m-\ell+2}(\{\mathbf{x}_0\mathbf{S}_i^{(t)}\}_i) & = & \mathcal{L}^*_{m-\ell+2}(\{z_i^{(t)}\}_i) \oplus \mathcal{L}^*_{m-\ell+2}(\{v_i^{(t)}\}_i) \\
\vdots & & \\
\mathcal{L}^*_n(\{\mathbf{x}_0\mathbf{S}_i^{(t)}\}_i) & = & \mathcal{L}^*_n(\{z_i^{(t)}\}_i) \oplus \mathcal{L}^*_n(\{v_i^{(t)}\}_i) \\
\mathcal{L}^*_{m-\ell+1}(\{\mathbf{x}_0\mathbf{S}_i^{(t+1)}\}_i) & = & \mathcal{L}^*_{m-\ell+1}(\{z_i^{(t+1)}\}_i) \oplus \mathcal{L}^*_{m-\ell+1}(\{v_i^{(t+1)}\}_i) \\
\mathcal{L}^*_{m-\ell+2}(\{\mathbf{x}_0\mathbf{S}_i^{(t+1)}\}_i & = & \mathcal{L}^*_{m-\ell+2}(\{z_i^{(t+1)}\}_i) \oplus \mathcal{L}^*_{m-\ell+2}(\{v_i^{(t+1)}\}_i) \\
\vdots & & \\
\mathcal{L}^*_n(\{\mathbf{x}_0\mathbf{S}_i^{(t+1)}\}_i) & = & \mathcal{L}^*_n(\{z_i^{(t+1)}\}_i) \oplus \mathcal{L}^*_n(\{v_i^{(t+1)}\}_i) \\
\vdots & &
\end{array}
$$

LPN Problem (an equivalent formulation)



Underlying Problem of the LPN



The Corrupting Noise

The system of equations implies the following. Assuming that each $v_i^{(t)}$ is a realization of a random binary variable $V_i^{(t)}$, such that $\Pr(V_i^{(t)} = 1) = 1 - \Pr(V_i^{(t)} = 0) = p$ $i = 1, 2, ..., n$, $t = 1, 2, ...$, we have the following:

$$\Pr(\mathcal{L}_j^{(t)}(\{V_i\}_i) = 1) = \frac{1 - (1-2p)^{m/2}}{2},$$

when the parameter m is an even number.

Security and LPN Problem

Accordingly, the considered system of equations implies the following:
- Asymptotically, the security of the considered stream cipher corresponds to hardness of the LPN problem;
- In the non-asymptotic scenarios the security corresponds to solving the LPN problem (i.e. decoding) or in particular cases performing fast correlation attack when the involved noise has the probability of ones equal to $\frac{1-(1-2p)^{m/2}}{2}$ (where m is a parameter).

V. Comparison with Recently Reported Related Schemes at

ICALP2008 and CRYPTO2009

"Targets" of Comparison

- H. Gilbert, M.J.B. Robshaw, and Y. Seurin, "**How to Encrypt with the LPN Problem**", *ICALP 2008, Part II, Lecture Notes in Computer Science*, vol. 5126, pp. 679-690, 2008.
- B. Applebaum, D. Cash, C. Peikert and A. Sahai, "**Fast Cryptographic Primitives and Circular-Secure Encryption Based on Hard Learning Problems**", *CRYPTO 2009, Lecture Notes in Computer Science*, vol. 5677, pp. 595-618, Aug. 2009.

Substantial Differences in Comparison with the Proposals at **ICALP2008 and CRYPTO2009**

- **The encryptions reported at ICALP2008 and CRYPTO2009 do not involve the keystream generators and t**he considered scheme provides a generic framework for strengthening of any keystrem generator and the related stream cipher. On the other hand, **.**
- Instead of employing a huge secret key, the matrix **S,** a simple **keystream generator** is employed in conjunction with a **dedicated homophonic coding scheme**.
- **The homophonic encryption scheme provides involvement of the pure randomness into each bit of the ciphertext which can be easily removed when the secret key is known but removing these pure random bits from the ciphertext without knowledge of a secret key is as hard as solving certain LPN problem**.

Comparison of the Main Security Features

	# of secret key bits	# of secret key bits involved in a ciphertext bit	# of unknown pure random bits involved in ciphertext bit	# of unknown biased random bits involved in ciphertext bit
encryption ICALP2008	$m \cdot n$	m	0	n
symmetric encryption CRYPTO2009	$\ell \cdot N >> m \cdot n$	ℓ	0	m
considered encryption	$k << m \cdot n$	$\approx k$	$m - \ell$	n

VI. Concluding Notes

Main Messages of This Talk

- The talk points out to **a design of a stream ciphers family** which involves randomness and dedicated coding.
- The design provides **provable security** and it is **suitable for implementation**.
- Security evaluation has been performed employing **computational complexity** approaches.
- The security claims have been compared with the related recently reported ones showing the **advantages** of the considered stream ciphers.

Thank You Very Much for the Attention,

and

QUESTIONS Please!