3 Countermeasures Against Information Leakage

3-1 Subgroup Membership Problem and Its Applications to Information Security

YAMAMURA Akihiro and SAITO Taiichi

The widely used algorithmic problems, the quadratic residue problem and the decision Diffie-Hellman problem, are characterized as the subgroup membership problem. Several cryptographic schemes are realized assuming the hardness of the subgroup membership problem. We apply the subgroup membership problem to several information security schemes: a probabilistic encryption, a bit commitment and a private information retrieval.

Keywords
Subgroup membership problem, Decision Diffie-Hellman problem, Quadratic residue problem, Probabilistic encryption, Private information retrieval

1 Introduction

Open networks have been developed and internet is inevitable in the business activities. The digital document law was started in April of 2005 and it accelerates the digital society. Digitalization of information brings convenience of our life however it also causes risk of information leakage. It is easy to copy digital data and no traces are left. Supplying sufficient security technologies is a pressing need in order to avoid abuse of information systems. Cryptographic technologies are essential in information security, it is impossible to construct a secure system without them. Adversaries attack weak part of the information systems, and so, cryptographic technologies have been seldom attacked because cryptography is theoretically proven strong. However, the ability of adversaries is always improving in view of recent attacks on several hash functions. It is important to continue the research on cryptographic technologies. In this paper, we introduce secure protocols using an algorithmic problem as one of trials for information security.

2 Subgroup membership problem

It is well known that the subgroup membership problem for a finitely presented group is not decidable in general, because Novikov-Boone theorem claims the existence of finitely presented group whose word problem is unsolvable. This implies that for a certain finitely presented group, there exists no procedure to check whether or not an element given as a word is equal to the identity element of the group. The subgroup membership problem is often called the generalized word problem in the literature of combinatorial group theory. On the other hand, the word problem is always solvable for the class of finite groups or finitely generated abelian groups.

However, if we consider more practical computation, that is, the bounded probabilistic polynomial time algorithms (or equivalently the computation class BPP), the membership problem is not trivial even for the class of finite abelian groups. When we consider a mathematical object, the object is described by finite data. The effectiveness is measured by the asymptotical behavior of algorithm to carry out certain tasks like deciding a mathematical proposition (equivalently calculating a Boolean predicate) and computing functions. In the case of decision problems for finitely presented groups, we consider the class of recursive functions. In specific cases like automatic groups and word hyperbolic groups, the word problem can be solved in polynomial time with respect to the word length. In the case of finite groups, the description of groups has simple structure, and any decision problem is solvable. We are interested in the effectiveness of such algorithmic problems. The behavior of algorithm is related to the size of data structure of a group family.

Several algorithmic problems used in cryptography are characterized as the subgroup membership problem. We note that there exists no known probabilistic polynomial time algorithm for the integer factorization or the discrete logarithm problem for some class of finite cyclic groups. So these problems are not in the class of BPP. The quadratic residue (QR for short) problem and the decision Diffie-Hellman (DDH for short) problem have numerous applications in cryptography, and hence, they have been studied in detail. In [17], the similarity of QR and DDH is discussed. We now give more formal approach to generalize and formalize cryptographic hard problems as the subgroup membership problem, and show many other algorithmic problems, which are used in public key cryptography, are characterized as the subgroup membership problem as well. Such a unification of algorithmic problems used in cryptography has not been appeared up to date as far as the authors know.

Widely used assumptions in cryptography are divided into two groups: the algorithmic assumptions related to the integer factoring (and the QR) and the algorithmic assumptions related to the discrete logarithm problem (and the DDH). The first is originated from the RSA cryptosystem [15] and the second from the Diffie-Hellman key exchange protocol [6]. These two look different and are usually discussed separately. The unified approach to the integer factoring problem and the discrete logarithm problem shed light on the fundamental properties of algorithms required to provide the security. Therefore, we can get better understanding of the algorithmic problems by unified treatment of subgroup membership problems. To apply the membership problem to cryptographic schemes such as asymmetric cryptosystems, we require the efficiency of computation for legal participants and the existence of a trapdoor. Once we prove that the subgroup membership problem is applicable to a certain scheme in general, then any primitive based on the subgroup membership problem concerning a specific group is applicable to the scheme in principle. As an example, in this paper, we show that any subgroup membership problem can be employed to construct a computational PIR system by constructing a PIR system using the subgroup membership problem in a general manner.

Determining the membership of a given element of a certain group in its subgroup is not always easy. As a matter of fact, the membership problem of a subgroup in a finitely presented group is not recursive in general. To apply the membership problem to cryptographic schemes such as asymmetric cryptosystems, we require the efficiency of computation for legal participants and the existence of a trapdoor. In this section we consider the subgroup membership problem with a trapdoor, and show that several problems widely used in cryptography are characterized as the subgroup membership problem.

Let G be a group, and let H be its subgroup. The membership problem is to decide whether or not a given element g in G belongs to H. Furthermore, we consider a fam-

ily of finite groups indexed by a parameter and the asymptotic behavior according to computation. In such a case, the subgroup membership is described as a computation problem to decide the membership when given an element, a subgroup and a group indexed by a parameter. A computation problem is hard if no efficient algorithms. The efficiency is characterized by the asymptotic behavior of an algorithm

2.1 Subgroup membership assumption

We suppose that every element in G has a binary representation of size k, where k is the security parameter. The membership can be decided within polynomial time in k if a certain information, called a trapdoor, is provided. The membership of an element g in G in H can be decided provided the trapdoor, however, the membership cannot be decided with a probability substantially larger than one half without the trapdoor. We now formalize the subgroup membership problem.

Let k be the security parameter. For the input 1^k, a probabilistic polynomial time algorithm IG outputs the description of a group G, the description of a subgroup H of G and the trapdoor that provides a polynomial time algorithm for the subgroup membership problem of H in G. The algorithm IG is called the instance generator. Every element of G is represented as a binary sequence of length k. Computation of the multiplication in G is performed in polynomial time in k.

The predicate for the membership of a subgroup is denoted by *Mem*, that is, *Mem* is defined as follows.

$$Mem(G,H,x) = \begin{cases} 1 & \text{if x lies in H} \\ 0 & \text{if x lies in S} \end{cases}$$

where IG outputs the pair (G, H) for 1^k, x is in G, and $S = G\backslash H$. The subgroup membership problem is to compute *Mem* in polynomial time in k when we inputs 1^k and obtain a pair of groups (G, H) and an element g in G, which is uniformly and randomly chosen from H or G according to the coin toss $b \xleftarrow{R} \{0,1\}$. If there does not exist a probabilistic polynomial time algorithm that computes *Mem* with a probability substantially larger than $\frac{1}{2}$, then we say that the membership problem is intractable. We also assume that one can choose uniformly and randomly an element from both H and G. This is significant to apply to cryptographic schemes.

The following is trivial, however, it is useful for the construction of a PIR system based on the subgroup membership problem.

Proposition 1

Let G be a group, and let H be a subgroup of G. For any g in G and h in H, gh lies in H if and only if g lies in H.

Subgroup membership assumption I

For every constant c, and every family $\{C_k | k \in N\}$ of circuits of polynomial size in k, there is an integer K such that for all $k > K$ we have

$$\text{Prob}(C_k(G,H,g) = Mem(G,H,g)) < \frac{1}{2} + \frac{1}{k^c}. \quad (2.1)$$

The assumption claims that there exists no polynomial size circuit family to compute the predicate *Mem*. The following is equivalent to the assumption above.

Subgroup membership assumption II

For every constant c, and every family $\{C_k | k \in N\}$ of circuits of polynomial size in k, there is an integer K such that for all $k > K$ we have

$$\left| P_H - P_S \right| < \frac{1}{k^c} |, \quad (2.2)$$

where the probabilities P_H and P_S are defined as follows;

$$P_H = \text{Prob}_{(G,H) \leftarrow IG(1^k); g \xleftarrow{R} H}(C_k(G,H,g) = 1)$$

and

$$P_S = \text{Prob}_{(G,H) \leftarrow IG(1^k); g \xleftarrow{R} S}(C_k(G,H,g) = 1)$$

2.2 Examples

We exhibit several subgroup membership problems: the DDH problem, the QR problem, the r th residue (RR for short) problem. Recall that the assumption that the QR problem is

intractable (QR assumption) is employed to prove the semantic security of the Goldwasser-Micali cryptosystem[8], and the assumption that the DDH problem is intractable (DDH assumption) is employed to prove the semantic security of the ElGamal cryptosystem. These two have many other applications. The assumption that one of problems above is intractable is employed to prove the semantic security of the corresponding cryptosystem[10][13][14] respectively.

Quadratic residue problem

Let p, q be prime integers. Set $N = pq$. The primes p and q are trapdoor information for the quadratic residue problem, on the other hand, the number N is public information. Let G be the subgroup of $(Z/(N))^*$ consisting of the elements whose Jacobi symbol is 1, and let H be the subgroup of G consisting of quadratic residues of G, that is,

$$H = \left\{ x \in G \middle| x = y^2 \bmod \mathrm{N} \text{ for } y \in (\mathrm{Z}/(\mathrm{N}))^* \right\}.$$

The quadratic residue problem of H in G is to decide whether or not, a given element g belongs to H. We can effectively determine the membership of g in H provided that the information p and q are available. No polynomial time algorithm is known for the membership of a randomly chosen element of G in H without the information p and q. Hence, if we define an instance generator for the QR problem as a probabilistic algorithm, then the QR problem is considered as the subgroup membership problem.

Decision Diffie-Hellman problem

Let C be a cyclic group of prime order p. The group C may be the multiplication group of a finite field or the group of rational points of an elliptic curve. Let g be a generator of C. The decision Diffie-Hellman problem is to decide whether or not $h_2 = g_2^a$ for the given quadruple (g_1, h_1, g_2, h_2) of elements in C with $h_2 = g_2^a$ for some $1 \le a \le p - 1$. If so, we say that (g_1, h_1, g_2, h_2) is a Diffie-Hellman quadruple. The integer a is the trapdoor of the decision Diffie-Hellman problem. Knowing the trapdoor a, we can efficiently decide whether or not $h_2 = g_2^a$.

The DDH problem can be characterized as the subgroup membership problem for a certain group as follows. We set G to be the direct product $C \times C$. Then the input to the DDH problem is (x, y), $x, y \in G$, that is, $x = (g_1, h_1)$ and $y = (g_2, h_2)$. It is obvious that (g_1, h_1, g_2, h_2) is a Diffie-Hellman quadruple if and only if y belongs to the subgroup $< x >$ of G generated by x. It follows that the DDH problem for the cyclic group C is equivalent to the subgroup membership problem of the group $H = < x >$, $x = (g_1, g_1^a)$, in the group $G = C \times C = < g_1 > \times < g_1 >$. Note that, when a generator x of H is given, it is possible to choose uniformly and randomly elements from H without the trapdoor information.

We summarize the examples in Table 1. We note that the table is not exhaustive at all. We mentioned about algorithmic problems equivalent to the subgroup membership problem in[17][18] in detail. See[17][18] for further information.

2.3 Probabilistic encryption

Goldwasser and Micali[8] introduce a semantic secure probabilistic encryption scheme, whose security is based on the QR assumption. An encryption is called semantic secure if the information leaked to a passive enemy is computationally negligible. This concept is a computational version of Shannon's perfect secrecy. The concept is significant in modern cryptography.

The subgroup membership problem is applied to a probabilistic encryption. See[16] for a probabilistic encryption based on the decision Diffie-Hellman problem.

Key generation:

Bob inputs 1^k to a probabilistic polynomial time algorithm IG, called instance generator, and gets a pair (G, H) of groups and the trapdoor for the subgroup membership problem of H in G, where k is the security parameter. Every element of G is represented by a binary sequence of length k. We assume the subgroup membership assumption of H in G. Therefore, Alice can generates elements in both and uniformly and randomly. Bob publicizes G and

Table 1 Subgroup Membership Problems

	Related Problem	Group Subgroup	Applications
DDH	DLP DH	$C \times C$: Direct Product of Cyclic Groups <(g, h)>: Subgroup Generated by (g, h)	ElGamal
QR	$FACT(pq)$	$\left\{x \in \mathrm{Z}_N^* \middle\| \left(\frac{x}{N}\right) = 1\right\}$ $\left\{x^2 \bmod N \middle\| x \in \mathrm{Z}_N^*\right\}$	Goldwasser-Micali[8]
RR	$FACT(pq)$	Z_N^* $\left\{x^\tau \bmod N \middle\| x \in \mathrm{Z}_N^*\right\}$	Kurosawa-Tsujii[10]
PSUB	$FACT(p^2q)$	$\{x \mid x = g^m y^N \bmod N$ *for* $m \in \mathrm{Z}/(p), y \in (\mathrm{Z}/((N))^*\}$ $\left\{y^N \bmod N \middle\| y \in \mathrm{Z}_N^*\right\}$	Okamoto-Uchiyama[13] Naccache-Stern[11]
DCR	$FACT(pq)$	$\{x \mid x = g^m y^N \bmod N^2$ $m \in \mathrm{Z}/(N), y \in (\mathrm{Z}/((N^2))^*\}$ $\left\{y^N \bmod N \middle\| y \in \mathrm{Z}_N^*\right\}$	Paillier[14]

H, but keeps the trapdoor information for the subgroup membership problem of H secret.

Encryption:

Suppose Alice encrypts a message $M=b_1b_2b_3...n_l$, where b_i belongs to $\{0, 1\}$ for every $i=1, 2, 3, ..., l$. For every $b_i(1 \le i \le l)$, Alice generates random element r_i, where r_i belongs to H if $b_i = 1$, and b_i belongs to $G\backslash H$ otherwise. Then the sequence of group elements $(r_1, r_2, r_3, \cdots, r_l)$ is an encrypted message for M. We note that the encrypted message is a random element in the direct product $S_1 \times S_2 \times S_3 \times ...S_l$, where $S_i = H$ if $b_i = 1$, and $S_i = G\backslash H$ otherwise. So the encryption is probabilistic.

Decryption:

Bob knows the trapdoor for the subgroup membership problem of H in G. Hence, he can decides whether or not each element belongs to H in polynomial time in the security parameter k.

Security:

An encryption scheme is semantic secure, if any adversary cannot computationally distinguish two ciphertexts of two messages of the same length. This means that no probabilistic polynomial time algorithm can distinguish two ciphertexts C_1 and C_2. It follows that the encryption above is semantic secure if and only if no probabilistic polynomial time algorithm can distinguish two direct products $S_1 \times S_2 \times S_3 \times ...S_l$ and $S_1 \times S_2 \times S_3 \times ...S_l$. Thus, the encryption is semantic secure under the subgroup membership assumption for H in G.

2.4 Bit commitment

Another possible application of the subgroup membership problem is the bit commitment scheme. We briefly describe a bit commitment scheme based on the subgroup membership problem. See[16] for a bit commitment scheme based on the decision Diffie-Hellman problem.

Key generation:

Alice inputs 1^k to an instance generator IG, and gets a pair (G, H) of groups and the trapdoor for the subgroup membership problem of H in G, where k is the security parameter. We assume the subgroup membership assumption of H in G. Alice publicizes G and H, but keeps the trapdoor information for the subgroup membership problem of H secret.

Committing:

Alice commits her bit b in (0, 1). She also generates uniformly and randomly an element r according to her bit b so that r belongs to H if $b = 1$ and r belongs to $G\backslash H$ otherwise.

Verifying:

Alice confesses her bit b to Bob, and gives the trapdoor for the subgroup membership problem. Bob can verify the membership of the element r. Thus, we can use any subgroup membership problem to construct a bit commitment protocol. We note that the bit commitment protocol can be used to construct a coin flipping protocol as well.

3 Private information retrieval

Chor, Goldreich, Kushilevitz and Sudan[3] introduced the private information retrieval scheme for remote database access, in which the user can retrieve the data of user's choice without revealing it. Their scheme attains information theoretic security, however, the database must be replicated in several locations where the managers are not allowed to communicate each other. The computational private information retrieval scheme was introduced by Chor and Gilboa[4]. Their scheme attains more efficient communication than Chor, Goldreich, Kushilevitz and Sudan's model by sacrificing the information theoretic security, nevertheless, their scheme enjoys computational security by assuming the existence of pseudorandom generators. However, their scheme still needs replication of the database. Kushilevitz and Ostrovsky[9] introduced a computational private information retrieval scheme in which only one database is needed. Their scheme depends on the intractability of the quadratic residue problem. More efficiency, polylogarithmic communication complexity, is attained by Cachin, Micali and Stadler [2]. They assume a number theoretic hypothesis, which they call the Φ assumption, and sacrifice one-round communication and then obtain polylogarithmic communication complexity. However, a rigorous proof of the intractability of the Φ assumption or its equivalence to a widely used assumption like the quadratic residue assumption or the integer factorization is not given in[2]. We summarize the known results on private information retrievals in Table 2.

We briefly review the general scheme of a private information retrieval (PIR for short) scheme. A computational PIR scheme with a single database is a protocol for two players, a user U and a database manager DB. Both are able to perform only probabilistic polynomial time computation. The database manager DB maintains a database, which is a binary sequence $X = x_0x_1x_2...x_{n-1}$. The goal of the protocol is to allow U to obtain the i th bit x_{i+1} of X without leaking any information on x_i to DB. The protocol runs as follows.

Step 1

U computes a query Query(i) using his random tape (coin toss), which U keeps secret. Then he sends Query(i) to DB.

Step 2

DB receives Query(i). He performs a polynomial-time computation for the input X, Query(i) and his random tape. The computation yields the answer Answer(Query(i)). He sends Answer(Query(i)) back to U.

Step 3

U receives Answer(Query(i)). He performs a polynomial-time computation using the answer Answer(Query(i)) and his private information (his random tape). The computation yields the i th bit x_{i+1} of the database.

Correctness

For any database sequence X and for any query for i th bit of X, U obtains x_{i+1} at the end.

Privacy

DB cannot distinguish a query for the i th bit and a query for the j th bit by a polynomial-time (probabilistic) computation with non-negligible probability. Formally, for all constants c, for all database of length n, for any two $1 \le i, j \le n$, and all polynomial-size family of circuits C_k, there exists an integer K such that for all $k > K$ we have

$$\left|\mathrm{Prob}(C_k(\mathrm{Query(i)}) = 1) - \mathrm{Prob}(C_k(\mathrm{Query(j)}) = 1)\right| < \sigma, \quad (3.1)$$

where k is the security parameter of the protocol and $\sigma = \frac{1}{(Max(k,n))^c}$.

Computation

Computations of both DB and U are bounded above by a polynomial in the size n of the database and the security parameter k.

Table 2 *Several Private Information Retrieval Schemes*

Scheme	Round Number	Security Assumption	Communication Complexity	Number of DBs
Chor, Coldreich, Kushilevitz, Sudan[3]	1	Information Theoretical	$o(n^{1/3})$	≥ 2
Ambainis[1]	1	Information Theoretical	$o(n^{2k-1})$ *for* $k(>1)DB$	≥ 2
Chor and Gilboa[4]	1	Existence of Pseudo Number Generators	$O(n^c)$ $c>0$	≥ 2
Kushilevitz and Ostrovsky[9]	1	Quadratic Residue Problem Assumption	$O(n^c)$ $c>0$	1
Ostrovsky and Shoup[12]	Multiple	Reduction to Read only scheme		
Cachin, Micali and Stadler[2]	2	Φ Assumtion	Polylogarithmic	1
Proposed Scheme	1	Subgroup Membership Assumption (e.g. DDH assumption)	$O(n^c)$ $c>0$	1

3.1 PIR scheme based on the subgroup membership problem

We show that the subgroup membership problem can be applied to a PIR scheme by modifying Kushilevitz and Ostrovsky's scheme[9]. The proposed scheme has the same communication complexity as Kushilevitz and Ostrovsky's scheme whose security depends on the QR assumption. On the other hand, the security of the private information retrieval scheme proposed in this paper is based on the subgroup membership assumption. Therefore, we can construct a private information retrieval scheme based on any algorithmic problems in Section **2.2**, in particular, we can use groups of rational points on elliptic curves or multiplicative groups of finite fields under the corresponding DDH assumption. We should remark that all the private information retrieval schemes proposed so far depend on either the existence of pseudorandom number generators or intractability assumption related to the integer factorization. No private information retrieval scheme based on the DDH has been proposed, yet as far as the authors know. Modifying[9], we construct a PIR scheme based on the subgroup membership problem.

3.2 Basic idea

First of all, we explain the basic idea of the scheme by a simple model. Suppose *DB* has the database $X = x_0x_1x_2...x_{n-1}$ and that *U* wishes to know the i th bit x_{i-1} . *U* chooses group elements $g_0, g_1, g_2, ..., g_{i-1}, ..., g_{n-1}$, so that g_j in H for $j \neq i-1$ and g_{j-1} in $S = G \backslash H$. Then *U* sends them all to *DB*. *DB* computes the group element $g = g_0^{x_0} g_1^{x_1} g_2^{x_2}, ...g_{i-1}^{x_{i-1}}, ...g_{n-1}^{x_{n-1}}$ sends it back to *U*. *DB* cannot get to know which of $g_0, g_1, g_2, ..., g_{i-1}, ..., g_{n-1}$ comes from *S* if the subgroup membership problem of *H* in *G* is intractable. Since *U* possesses the trapdoor, he can determine whether or not g lies in *H*. By Proposition 1, g lies in *H* if and only if $x_{i-1} = 0$. Therefore, *U* can obtain the i th bit x_{i-1}. This simple model illustrates the idea of using the subgroup membership problem, but the communication complexity is still large. We need the trick by[9] to reduce the communication complexity.

3.3 Scheme

We now describe the private information retrieval scheme using the subgroup membership problem.

Step 0

The user U inputs 1^k to the instance generator IG and then gets a pair (G, H) of groups and the trapdoor for the subgroup membership problem of H in G, where k is the security parameter and every element of G is represented by a binary sequence of length k. We assume the subgroup membership assumption of H in G. The group G is shared by both DB and U. On the other hand, U keeps the trapdoor information for the subgroup membership problem of H secret. Computations of both DB and U are performed in the group G. Let X be the database managed by DB. We suppose that $X = x_0x_1x_2...x_{n-1}$, where x_i lies in (0, 1), and that $n = t^l$, where t, l are positive integers.

Step 1

U computes a query Query(i) for his desired bit x_{i-1}, where $1 \le i \le n$, in the following manner. First, U computes the t-adic expansion of i. Let $i = \alpha_0$. Then the t-adic expansion of i is $\beta_l\ \beta_{l-1} ... \beta_2\ \beta_1$, where

$$\begin{aligned}
&\alpha_0 = \alpha_1 t + \beta_1 \quad 0 \le \alpha_0 \le t^{l-1} \text{ and } 0 \le \beta_1 \le t-1\\
&\alpha_1 = \alpha_2 t + \beta_2 \quad 0 \le \alpha_1 \le t^{l-2} \text{ and } 0 \le \beta_2 \le t-1\\
&\alpha_2 = \alpha_3 t + \beta_3 \quad 0 \le \alpha_2 \le t^{l-3} \text{ and } 0 \le \beta_3 \le t-1\\
&\qquad\qquad \cdots\cdots\\
&\alpha_{l-2} = \alpha_{l-1} t + \beta_{l-1} \quad 0 \le \alpha_{l-2} \le t-1 \text{ and } 0 \le \beta_{l-1} \le t-1\\
&\qquad\qquad 0 \le \alpha_{l-1} = \beta_l \le t-1 \quad \alpha_l = 0
\end{aligned} \tag{3.2}$$

For each $u\,(1 \le u \le l)$, U chooses uniformly and randomly t-1 elements $g_{(u,0)}, g_{(u,1)}, ..., g_{(u,\beta_u-1)}, g_{(u,\beta_u+1)}, ..., g_{(u,t-1)}$ from H. He also chooses uniformly and randomly $g_{(u,\beta_u)}$ from $S = G\backslash H$. U defines Q(u) by

$$(g_{(u,0)}, g_{(u,1)}, ..., g_{(u,\beta_u-1)}, g_{(u,\beta_u)}, g_{(u,\beta_u+1)}, ..., g_{(u,t-1)}) \tag{3.3}$$

that is, Q(u) is a sequence of group elements of G such that the u β_u th component is uniformly and randomly chosen from $S = G\backslash H$ and the others are uniformly and randomly chosen from H. Then, Q(1), Q(2), ..., Q(l) comprise a query (denoted by Query}(i)) for the i th bit x_{i+1} of X, and U sends Query(i) to DB. Since each Q(u) consists of t group elements from G, Q(u) is represented by $k \times t$ bits. Thus, Query}(i) consists of $k \times t \times l$ bits.

Step 2

Receiving Query(i), DB constructs child databases recursively from the original database X. We regard X as the $t^{l-1} \times t$ binary matrix

$$D(0,\lambda) = \begin{pmatrix} x_0 & x_1 & x_2 \ ... & x_{t-1} \\ x_t & x_{t+1} & x_{t+2} & x_{2t-1} \\ & & ... & \\ x_{t^l-t} & x_{t^l-t+1} & ... \ ... & x_{t^l-1} \end{pmatrix}$$

where λ denotes the empty sequence in {0, 1, 2, ..., k-1}*. We note that the target bit x_{i-1} is the (α_1, β_1) entry of $D(0, \lambda)$ (α_1 and β_1 are obtained in (3.2)). Denote it by Target $D(0, \lambda)$.

We recursively define child databases $D(u, s)$, where $1 \le u \le l$ and s belongs to {0, 1, 2, ..., k-1}^{u}. Suppose that we have defined the databases $D(u, s)$ and their target bits Target $D(u, s)$ and s in {0, 1, 2, ..., k-1}^{u} for $0 \le u < l$-1. Then we define the databases $D(u+1, s0)$, $D(u+1, s1)$, ... $D(u+1, s(k-1))$.

The database $D(u, s)$ is a binary sequence of length t^{l-u}. We regard $D(u, s)$ as a $t^{l-u-1} \times t$ binary matrix. Suppose that

$$D(0,\lambda) = \begin{pmatrix} y_0 & y_1 & y_2 \ ... & y_{t-1} \\ y_t & y_{t+1} & y_{t+2} & y_{2t-1} \\ & & ... & \\ y_{t^{l-u}-t} & y_{t^{l-u}-t+1} & ... \ ... & y_{t^{l-u}-1} \end{pmatrix}$$

We now construct k child databases, $D(u+1, s0), D(u+1, s1), ... D(u+1, s(k-1))$.

Recall that Q(u) consists of t group elements $g_{(u,0)}, g_{(u,1)}, ..., g_{(u,\beta_u-1)}, g_{(u,\beta_u+1)}, ..., g_{(u,t-1)}$ in G (defined in (3.3)). We define a group element g_v for each row $v = 0, 1, 2, ..., t^{l-u-1}$ as follows.

We set

$$f_{(v,w)} = \begin{cases} g_{(u,w)} & \text{if } D_{(u,s)(v,w)} = 1 \\ 1 & \text{if } D_{(u,s)(v,w)} = 0 \end{cases}, \tag{3.4}$$

where $D(u, s)(v, w)$ denotes the (v, w) entry of $D(u, s)$. Then we set

$$f_{D(u,s),v} = \prod_{w=0,1,2,...,t-1} f_{(v,w)} \tag{3.5}$$

for each row $v = 0, 1, 2, ... t^{l-u-1}$-1. Note

that the group element $f_{D(u,\,s),\,v}(0 \le v \le t^{l-u-1}-1)$ is of size k, and that $f_{D(u,\,s),\,v}$ belongs to H if and only if $D(u, s)(v, \beta_u) = 0$ by Proposition1 The r th child database $D(u+1, sr)(0 \le r \le k-1)$ is defined to be the sequence consisting of $g_0(r)$, $g_1(r)$, $\cdots$, $g_{t^{l-u-1}-1}$, where $g_v(r)$ denotes the r th bit of the representation of $f_{D(u,\,s),\,v}$. Hence, we have the following matrix equation:

$$\begin{pmatrix} f_{D(u,s),0} \\ f_{D(u,s),1} \\ \cdots \\ f_{D(u,s),t^{l-u-1}-1} \end{pmatrix} = (D(u+1,s0)\ \ D(u+1,s1)\ \ \ldots\ \ D(u+1,s(k-1))) \quad (3.6)$$

where each $f_{D(u,\,s),\,v}$ is a row vector and each $D(u+1, sr)$ is a column vector. Thus, $D(u+1, sr)$ is a binary sequence of length t^{l-u-1}. We regard it as a $t^{l-u-2} \times t$ binary matrix. Then the target bit for it (denoted by Target$(D(u+1, sr))$ is defined to be the $(\alpha_{u+1}, \beta_{u+1})$ entry of $D(u+1, sr)$ for every r in $\{0, 1, 2, \ldots, k-1\}$ (α_{u+1} and β_{u+1} are obtained in (3.2)).

Step 3

In the last stage of constructing child databases, DB obtains k^{t-1} databases

$D(l-1, s)(s$ lies in $\{0, 1, 2, \ldots, k-1\}^{t-1})$. Note that each $D(l-1, s)$ contains t bits. We regard $D(l-1, s)$ as a $1 \times t$ matrix. For each $D(l-1, s)$, we define a group element $A(s)$ as follows. First, we define

$$f_{(0,w)} = \begin{cases} g_{(u,w)} & if\ \ D(l-1,s)(0,w) = 1 \\ 1 & if\ \ D(l-1,s)(0,w) = 0. \end{cases}$$

Then, we set

$$f_{D(l-1,s),0} = \prod_{w=0,1,2,\ldots,t-1} f_{(0,w)} = A(s).$$

The group element $A(s)$ is of size k for every s in $\{0, 1, 2, \ldots, k-1\}^{t-1}$. Then the group elements $A(s)(s$ lies in$\{0, 1, 2, \ldots, k-1\}^{t-1})$ form the answer Answer(Query(i)) to the query Query(i), and DB sends Answer(Query(i)) to U.

Step 4

U receives Answer(Query(i)) consisting of $A(s)$, where s belongs to $\{0, 1, 2, \ldots, k-1\}^{t-1}$. U can retrieve the target bit x_i = Target $(D_{(0,\,\lambda)})$ in polynomial time in k, n. In fact, the following holds in general.

Theorem 2

For every database $D_{(u,\,s)}$, where $0 \le u \le l-2$ and s in $\{0, 1, 2, \ldots, k\}^u$, U can compute Target $(D_{(u,\,s)})$ in polynomial time in n, k if Target$(D_{(u+1,\,s0)})$, Target$(D_{(u+1,\,s1)})$, ... Target$(D_{(u+1,\,s(k-1))})$ are given.

See[17][18] for proof.

3.4 Privacy

In the proposed scheme, the query Query(i) consists of Q(1), Q(2), ..., Q(l), and each Q(u) consists of

$$(g_{(u,0)}, g_{(u,1)}, \ldots, g_{(u,\beta_u-1)}, g_{(u,\beta_u)}, g_{(u,\beta_u+1)}, \ldots g_{(u,t-1)}),$$

where one of the components is chosen uniformly and randomly from $S = G \backslash H$ and the others are chosen uniformly and randomly from H. The privacy is assured by the inequality

$$\left|\mathrm{Prob}(C_k(\mathrm{Query(i)})=1) - \mathrm{Prob}(C_k(\mathrm{Query(j)})=1)\right| < \sigma$$

where $\sigma = \frac{1}{(Max(k,n))^c}$. Hence, the privacy of the proposed scheme is guaranteed by the subgroup membership assumption by (3.1).

3.5 Communication complexity

In the first step, U sends

$$\mathrm{Query(i)} = (\mathrm{Q(1)}, \mathrm{Q(2)}, \cdots, \mathrm{Q(l)}).$$

Each Q(u) consists of t group elements in G. Since every element in G is represented by a binary sequence of length k, the total bits sent in this stage is $l \times t \times k$. In the second step, DB sends Answer(Query(i)) consisting of k^{l-1} group elements in G. Therefore, the total bits sent in this stage is $k^{l-1} \times k = k^l$. Consequently, the communication complexity is $ltk + k^l = \ln^{\frac{1}{7}} k + k^l$. Suppose that $k = n^c$ and $l = \mathrm{O}\left(\frac{\log n}{\log k}\right)$, then the communication complexity is $O(n^c)$. See [17][18].

3.6 Conclusion

We formalize QR problem and DDH problem as a membership problem and show that several cryptographic protocols can be implemented using the subgroup membership problem.In particular, we show that it can be

applied to the private information retrieval schemes. This gives a first private information scheme based on the DDH problem.A small example of such a scheme is given in[18]. We would like to apply the subgroup membership problem to other cryptographic protocols.

References

1 A.Ambainis, "Upper Bound on the Communication Complexity of Private Information Retrieval, Automata", Languages and Programming, LNCS, Vol.1256, Springer-Verlag, pp.401-407, 1997.

2 C.Cachin, S.Micali, and M.Stadler, "Computationally Private Information Retrieval with Polylogarithmic Communication", Advances in Cryptology, LNCS, Vol.1592, Springer-Verlag, pp.402-414, 1999.

3 B.Chor, O.Goldreich, E.Kushilevitz, and MM.Sudan, "Private Information Retrieval", IEEE Symposium on Foundations of Computer Science, pp.41-50, 1995.

4 B.Chor and MN.Gilboa, "Computationally Private Information Retrieval", ACM Symposium on Theory of Computing, pp.304-313, 1997.

5 R.Cramer and MV.Shoup, "A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack", Advances in Cryptology, LNCS, Vol.1462, Springer-Verlag, pp.13-25, 1998.

6 W.Diffie and MM.E. Hellman, "New directions in cryptography", IEEE Transactions on Information Theory, Vol.22, pp.644-654, 1976.

7 T.ElGamal, "A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms", IEEE Transactions on Information Theory, Vol.31, pp.469-472, 1985.

8 S.Goldwasser and MS.Micali, "Probabilistic Encryption", J. Computer and System Science, Vol.28, pp.270-299, 1984.

9 E.Kushilevitz and R.Ostrovsky, "Replication Is not Needed : Single Database, Computationallyprivate Information Retrieval", IEEE Symposium on Foundations of Computer Science, pp.364-373, 1997.

10 K.Kurosawa and S.Tsujii, "A General Method to Construct Public Key Residue Cryptosystems", Transactions of the IEICE E-73, pp.1068-1072, 1990.

11 D.Naccache and J.Stern, "A New Public-key Cryptosystem", Advances in Cryptology, LNCS, Vol.1233, Springer-Verlag, pp.27-36, 1997.

12 R.Ostrofsky and V.Shoup, "Private Information Storage", ACM Symposium on Theory of Computing, pp.294-303, 1997.

13 T.Okamoto and S.Uchiyama, "A New Public-key Cryptosystem as Secure as Factoring", Advances in Cryptology, LNCS, Vol.1403, Springer-Verlag, pp.308-318, 1998.

14 P.Paillier, "Public-key Cryptosystems Based on Composite Degree Residuosity Classes", Advances in Cryptology, LNCS, Vol.1592, Springer-Verlag, pp.223-238, 1999.

15 R.L.Rivest, A.Shamir, and L.Adlemman, "A method for obtaining digital signatures and public key cryptosystems", Communications of the ACM, Vol.21, pp.120-126, 1978.

16 T.Saito, T.Koshiba, and A.Yamamura, "The Decision Diffie-Hellman assumption and the Quadratic Residuosity Assumption", IEICE Transactions on Fundamentals of Electronics (1) E84-A, pp.165-171, 2001.

17 A.Yamamura and T.Saito, "Private Information Retrieval Based on the Subgroup Membership Problem", Information Security and Privacy, LNCS Vol2119. Springer-Verlag, pp.206-220, 2001.

18 A.Yamamura and T.Saito, "Subgroup membership problems and applications to information security", Scientiae Mathematicae Japonicae, Vol.57, pp.25-41, 2003.

***YAMAMURA Akihiro**, Ph.D.*

Group Leader, Security Fundamentals Group, Information and Networks Systems Department

Information security, Cryptography, Algebraic systems and their algorithms

***SAITO Taiichi**, Dr. Eng.*

Associate Professor, Department of Information and Communication Engineering, Tokyo Denki University

Cryptography