Asymmetric secret key ciphers and encrypted data retrieval schemes *

Akihiro Yamamura
National Institute of Information and Communications Technology
4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan

1 Introduction

Traditionally cryptosystems are mainly divided into two categories; symmetric encryptions and asymmetric encryptions. Symmetric ciphers are often called a conventional cipher, a secret key cipher or a common key cipher. The cryptosystems in this category share the property that the legitimate users share a common key in advance and the key are used to both encrypt plaintexts and decrypt ciphertexts. On the other hand, the asymmetric ciphers are often called a public key cryptosystem. The cryptosystem in this category share the property that the message receiver publicize his encryption key and keeps his decryption key. Any message sender can encrypt a message using the public encryption key. It is clear that the decryption key is different from the public key otherwise the anybody can decrypt any message. Thus, the first category of ciphers is called a *symmetric key cipher* and the second is called *asymmetric key cipher*. There is another possible category of cryptosystems that have not been studied so far. We examine cryptosystems whose encryption keys are different from the decryption keys and the both are kept secret. On the other hand, we shall show that the class of asymmetric secret key cryptosystem makes sense in a certain occasion. Briefly speaking, we need extra properties for secret key ciphers, which the traditional secret key cipher usually does not possess. To construct a valiant of oblivious transfer scheme, we need *commutative property* for the family of secret key ciphers. Then we shall show that both commutativity property and security cannot be achieved together by a family of symmetric secret key ciphers. Therefore, it is essential to construct a family of asymmetric secret key ciphers that satisfies the commutative property.

We also consider the family of encryption functions. A family of encryption functions $\{f_i \mid i \in I\}$, where each f_i is an encryption function of the set of messages M onto M, is called *commutative* if $f_i(f_j(m)) = f_j(f_i(m))$ for every m in M and $i, j \in I$. We shall propose a commutative family of asymmetric secret key ciphers. The commutative property is implicitly used to construct a blind signature. Our method is similar to the construction of blind signature, however, ours is based on different assumptions on algorithmic problems and attacking models.

Our motivation comes from the data management in the ubiquitous network, in particular, retrieving data from ciphertext embedded in an RFID. As we will see such scheme is closely related to the oblivious transfer schemes and private information retrievals. Oblivious transfer schemes and

*This paper is an extended abstract and the detailed version will be published elsewhere.

all-or-nothing disclosure schemes attain the ability of retrieving secretly data from the database server, however, both schemes suffer from the communication complexity. If the database is massive, the communication complexity is also large and this is fatal for some applications. In this paper, we propose a concept of *encrypted data retrieval scheme* that solves the problem and give a concrete method to construct such schemes employing a commutative family of asymmetric secret key ciphers.

2 Private encrypted data retrieval

2.1 Oblivious transfer and private information retrieval

Oblivious Transfer (OT) refers to several types of two party protocols, where one party, the sender, transmits part of its input to another party, the chooser, in a way that protects both parties: the sender is assured that the chooser does not get more information than it is entitled, and chooser is assured that the sender does not learn which part of the inputs it received. The notion of *1-out-of-2 oblivious transfer* (OT_1^2 for short) was introduced in [8], as generalization of Rabin's concept of OT [11]. Brassard, Crépeau, and Robert in [3] generalized the notion further to *1-out-of-N oblivious transfer* (OT_1^N) under the name *all-or-nothing disclosure* (ANDOS). ANDOS allows the sender, who holds several secrets, to disclose one of them to the receiver, with the guarantee that no information about other secrets will be revealed. Furthermore, the receiver has the guarantee the sender will not be able to find out which secret was picked.

Private Information Retrieval (PIR) *schemes* [5] allow a user to access a database consisting of $N$ data $m_1, m_2, \dots, m_N$ (usually data are just a bit) and read any elements without a database manager learning which element was accessed. The emphasis in PIR is on communication complexity which must be $o(N)$. PIR schemes do not protect the owner of the database, because they do not prevent the user from learning more than a single element. Currently, the question of protecting the database was addressed as well. A PIR scheme where a user does not learn more than a single data is called a *Symmetric* PIR (SPIR) [9].

An all-or-nothing disclosure is a two party protocol in which the vendor, who holds several secrets, to disclose one of them to the buyer, with the guarantee that no information about other secrets will be revealed. Furthermore, the buyer has the guarantee the vendor will not be able to find out which secret was picked. In the literature only 1-out-of-$t$ all-or-nothing disclosure schemes have been studied as far as the authors know. J.Stern [12] proposed a 1-out-of-$t$ all-or-nothing disclosure scheme based on homomorphic encryptions. Suppose that a vendor possesses $t$ data $d_1, d_2, \dots, d_t$. A buyer wishes to obtain $s$ data out of $t$ data without informing which data the buyer tries to retrieve. Suppose the indices of the buyer's choice are $i_1, i_2, \dots, i_s$ $(1 \le i_1 < i_2 < \cdots < i_s \le t)$, that is, the buyer wishes to obtain the data $d_{i_1}, d_{i_2}, \dots, d_{i_s}$.

We propose a similar data retrieval scheme using our asymmetric secret key ciphers. In the scheme, the database is encrypted by the server's encrypted key and each user has to ask the server to decrypt it in the way that the server does not notice which data he is decrypting. In our approach, the database is encrypted and publicized and this reduce the communication complexity of both directions. This is extremely ideal to ubiquitous setting where wireless transmission is limited and required to reduce the amount of data transmission.

Suppose that $n$ is the number of the data in the database and $k$ is the size of group element in the data retrieval schemes. The we summarize these approaches as follows. Clearly our scheme is suitable to application in which data transmission is limited such as RFID embedded systems because communication complexity of our method is much smaller than the others. In particular, if the number of data is large, then the communication complexity of the other two will gets large whereas the communication complexity of ours does not depend on the number of data.

	Comm. complexity from user to server	Comm. complexity from server to user	Status of data
Oblivious transfer	k	nk	Secret
ANDOS	nk	k	Secret
Our approach	k	k	Encrypted database is public

Table 1: Data retrieval

3 Asymmetric secret key ciphers

3.1 Asymmetric secret key cipher based on RSA modulus

Examples in [1, 2] are explained. Let $p$ and $q$ be distinct primes of the same size. Set $n = pq$. We define an encryption function enc of the set of messages $\mathbb{Z}/n$ (denoted by $M$) onto itself. Suppose $e \in \mathbb{Z}/(p-1)(q-1)$ such that $e$ and $(p-1)(q-1)$ are coprime. Then the set $K$ of keys is $\mathbb{Z}/(p-1)(q-1)$. Then there exists $d \in \mathbb{Z}/(p-1)(q-1)$ such that $m^{ed} = m$ for every $m \in M$.

Secret key The pair $(e, d)$ is a secret key.

Public information The primes $p$ and $q$ are publicized.

Encryption Take message $m \in M$. The function $f$ of $M \times K$ into $M$ is defined by $f(m, e) = m^e$, where $(m, e) \in M \times K$. The ciphertext $enc_e(m)$ of $m$ is given by $f(m, e) = m^e$, that is, $enc_e(m) = f(m, e)$

Decryption For any $C \in M$, the decryption related to the key $d$ is given by $f(c, d)$, that is, $dec_d(c) = f(c, d) = c^d$.

Commutative property The family of encryption functions $\{enc_e \mid e \in K\}$ is commutative because $enc_{e_1}(enc_{e_2}(m)) = m^{e_1 e_2} = m^{e_2 e_1} = enc_{e_2}(enc_{e_1}(m))$ for all $m \in M$ every $e_1, e_2 \in K$.

Security We suppose that the discrete logarithm problem for $(\mathbb{Z}/p)^*$ and $(\mathbb{Z}/p)^*$ are intractable.

Remark The encryption is not a public key cryptosystem although we employ completely same ingredient as the RSA public key cryptosystem. We note that the factorization of the modulus $n$ is secret in RSA whereas the factorization is public in our cipher. On the other hand, the encryption parameter $e$ is public in RSA, whereas both $e$ and $d$ are secret in our cipher. Publicizing the primes

p and q, anybody can create the encryption and decryption keys in our cipher. Thus, this cipher algorithm related to the modulus n can be employed by multiple users.

The textbook RSA is not randomized cryptosystem and so the scheme is not semantic secure. We shall discuss randomized asymmetric secret key ciphers which are more secure than the ones in the full version of the paper.

4 Private encrypted data retrieval protocols

We shall briefly discuss how to construct secret data retrieval schemes using asymmetric secret key ciphers.

4.1 Basic scheme

Suppose $\{enc_k \mid k \in K\}$ is a family of commutative symmetric ciphers.

The server $\mathcal{S}$ encrypts the data $m_1, m_2, \ldots, m_N$ by his secret encryption key e_s and publicizes the ciphertexts $enc_{e_s}(m_1), enc_{e_s}(m_2), enc_{e_s}(m_3), \ldots, enc_{e_s}(m_N)$.

The receiver $\mathcal{U}$ wishes to obtain one of the data (say m_α) by decrypting $enc_{e_s}(m_\alpha)$ in the way that $\mathcal{S}$ cannot obtain any information on α while $\mathcal{U}$ gets only m_α. Note that $\mathcal{U}$ has access to the ciphertexts $enc_{e_s}(m_1), enc_{e_s}(m_2), enc_{e_s}(m_3), \ldots, enc_{e_s}(m_N)$, which makes a difference between the usual oblivious transfer protocol. The second condition implies that $\mathcal{U}$ cannot obtain the secret key e_s. A general one round private encrypted data retrieval protocol runs as follows:

Step 1 $\mathcal{U}$ generates the system parameters. $\mathcal{U}$ computes $Q = enc_{e_u}(enc_{e_s}(m_\alpha))$. Then $\mathcal{R}$ sends Q to$\mathcal{S}$.

Step 2 $\mathcal{S}$ receives Q and computes $dec_{e_s}(Q) = dec_{e_s}(enc_{e_u}(enc_{e_s}(m_\alpha)))$
$= dec_{e_s}(enc_{e_s}(enc_{e_u}(m_\alpha))) = enc_{e_u}(m_\alpha)$. Then $\mathcal{S}$ sends $enc_{e_u}(m_\alpha)$ to $\mathcal{U}$.

Step 3 $\mathcal{U}$ receives $enc_{e_u}(m_\alpha)$ and computes $dec_{e_u}(enc_{e_u}(m_\alpha)) = m_\alpha$.

Correctness
If both party play honestly, $\mathcal{U}$ obtains m_α.

Privacy for $\mathcal{U}$
$\mathcal{S}$ cannot distinguish a query for the αth and the βth data for all α and β.

Privacy for $\mathcal{S}$
$\mathcal{U}$ cannot obtain any information on the other data. This implicitly implies that the protocol guarantees that $\mathcal{U}$ cannot obtain any information on the secret key e_S.

Computation
Computations of both $\mathcal{U}$ and $\mathcal{S}$ are bounded above by a polynomial in the size N of the database and the security parameter k.

4.2 Proposed scheme

In the basic scheme, there is a security issue. The family of commutative symmetric ciphers may satisfy the homomorphic property as well. In such a case, $\mathcal{U}$ may be able to obtain information of both m_1 and m_2 by only one query. For example, $\mathcal{U}$ computes $\beta = enc_{e_s}(m_1)enc_{e_s}(m_2) = enc_{e_s}(m_1m_2)$ and sends $Q = enc_{e_u}(\beta)$ to $\mathcal{S}$. $\mathcal{S}$ computes $dec_{e_s}(Q) = enc_{e_u}(m_1m_2)$ and sends it to $\mathcal{U}$. Then $\mathcal{U}$ can obtain m_1m_2, which contains information on both m_1 and m_2. Thus the scheme does not satisfy the privacy for $\mathcal{S}$.

To repair this flaw, we operate another encryption for the data. Suppose that R_k is a random permutation with the key space $\{k \mid k \in K\}$ and *totally anti-homomorphic*. This means that for every pair of messages m_1 and m_2, $R_k(m_1)R_k(m_2)$ and $R_k(m_1m_2)$ are not correlated, that is, there is no relation between the distribution $R_k(m_1)R_k(m_2)$ and $R_k(m_1m_2)$.

Then the proposed scheme is described as follows.

The server $\mathcal{S}$ encrypts the data $m_1, m_2, \ldots, m_N$ by his secret encryption key e_s and k and publicizes the ciphertexts $enc_{e_s}(R_k(m_1)), enc_{e_s}(R_(m_2)), enc_{e_s}(R_k(m_3)), \ldots, enc_{e_s}(R_k(m_N))$.

Step 1 $\mathcal{U}$ generates the system parameters. $\mathcal{U}$ computes $Q = enc_{e_u}(enc_{e_s}(R_k(m_\alpha)))$. Then $\mathcal{R}$ sends Q to$\mathcal{S}$.

Step 2 $\mathcal{S}$ receives Q and computes $dec_{e_s}(Q) = dec_{e_s}(enc_{e_u}(enc_{e_s}(R_k(m_\alpha))))$ $= dec_{e_s}(enc_{e_s}(enc_{e_u}(R_k(m_\alpha)))) = enc_{e_u}(R_k(m_\alpha))$. Then $\mathcal{S}$ sends $enc_{e_u}(R_k(m_\alpha))$ to $\mathcal{U}$.

Step 3 $\mathcal{U}$ receives $enc_{e_u}(R_k(m_\alpha))$ and computes $dec_{e_u}(enc_{e_u}(R_k(m_\alpha))) = R_k(m_\alpha)$. Finally computes $D_k(R_k(m_\alpha) = m_\alpha$.

We should note that using R_k prevents $\mathcal{U}$ from obtaining more information from one query. For example, by the strategy above, $\mathcal{U}$ can obtain $R_k(m_1)R_k(m_2)$ that is not related to $R_k(m_1m_2)$. Therefore, $\mathcal{U}$ can obtain no information on m_1m_2.

References

[1] F.Bau, R.Deng, and P.Feng, An efficient and practical scheme for privacy protection in the E-commerce of digital goods, ICISC 2000, LNCS 2015 (2000), 162–170

[2] F.Bau, R.Deng, P.Feng, Y.Guo, and H.Wu, Secure and private distribution of online video and some related cryptographic issues, ACISP2001, LNCS 2119 (2001), 190–205

[3] G. Brassard, C. Crépeau and J.-M. Robert, All-or-Nothing Disclosure of Secrets, CRYPTO'86 LNCS, Vol. 263. Springer-Verlag, (1987) 234–238.

[4] C.Cachin,, S.Micali, M.Stadler, Computationally Private Information Retrieval with Polylogarithmic Communication, Advances in Cryptology. Lecture Notes in Computer Science, Vol. 1592. Springer-Verlag, (1999) 402–414

[5] B.Chor, O.Goldreich, Kushilevitz,E., Sudan,M.: Private Information Retrieval, IEEE Symposium on Foundations of Computer Science. (1995) 41–50

[6] C. K. Chu and W. G. Tzeng. "Efficient k-out-of-n Oblivious Transfer Schemes with Adaptive and Non-Adaptive Queries". In *Proc. of PKC '05*, LNCS 3386, pages 172–183. Springer Verlag, 2005.

[7] B.Chor, N.Gilboa, Computationally Private Information Retrieval ACM Symposium on Theory of Computing. (1997) 304–313

[8] S. Even, O. Goldreich and A. Lempel, A randomized protocol for signing contracts, Communications of the ACM 28 (1985) 637–647.

[9] Y. Gertner, Y. Ishai, E. Kushilevitz, T. Malkin, Protecting data privacy in private data information retrieval schemes, STOC'98 (1998) 151–160.

[10] E.Kushilevitz, R.Ostrovsky, Replication Is not Needed: Single Database, Computationally-private Information Retrieval, IEEE Symposium on Foundations of Computer Science. (1997) 364–373

[11] M. Rabin, How to exchange secerts by oblivious transfer, Technical Report TR-81, Harvard University (1981)

[12] J.Stern, A new and efficien all-or-nothing disclosure of secrets protocol, Advances in Cryptology. LNCS, Vol. 1514. Springer-Verlag, (1998) 357–371

[13] W. G. Tzeng. "Efficient 1-Out-n Oblivious Transfer Schemes". In *Proc. of PKC '02*, LNCS 2274, pages 159–171. Springer-Verlag, 2002.

[14] D.Q.Viet, A.Yamamura, and H.Tanaka, Anoymous Password-based Authenticated Key Exchange, *Progress in Cryptology* (Indocrypt 2005) LNCS **3797**, Springer-Verlag, (2005) 244–257.

[15] A.Yamamura, T,Jajcayova, and T.Kurokawa, Oblivious transfer and private information retrieval based on the p-subgroup assumption, *SOFSEM 2005 Communications*, Slovak Society for Computer Science, (2005) 101–110.

[16] A.Yamamura, T.Kurokawa, and J.Nakazato, Threshold anonymous group identification and zero-knowledge proof, *Information Security and Privacy* (ACISP2007) LNCS **4586**, Springer-Verlag, (2007) 370–384.

[17] A.Yamamura and T.Saito, Private information retrieval based on the subgroup membership problem, (with T.Saito) *Information Security and Privacy* (ACISP2001), LNCS **2119**, Springer-Verlag, (2001) 206–220.

[18] A.Yamamura and T.Saito, Subgroup membership problems and applications to information security, *Scientiae Mathematicae Japonicae*, (1) **57** (2003) 25–41.