

## On the Implementation of Indistinguishable Boxes Needed In Knapsack Zero-Knowledge Interactive Proof Schemes\*

WEN-CHUNG KUO, CHI SUNG LAIH AND M. J. GAU

*Department of Electrical Engineering*

*National Cheng Kung University,*

*Tainan, Taiwan 701, R.O.C.*

*E-mail: laihcs@eembox.ncku.edu.tw*

The concept of the Zero-Knowledge Interactive Proof (ZKIP) scheme was first proposed by Goldwasser, Micali and Rackoff in 1985. Since then, many practical ZKIP schemes have been proposed. One common feature among all these schemes is that the security of the schemes is based on factoring or the discrete logarithm. In 1991, Simmons proposed an alternative practical ZKIP scheme whose security is based on the subset sum problem. However, there is a very strong assumption in the scheme; i.e., Simmons's scheme would be secure under the assumption that an indistinguishable box exists. Unfortunately, nobody, including Simmons, has explained how to implement the indistinguishable box. In this paper, we propose two methods for implementing the indistinguishable box. It is shown that the proposed indistinguishable box is very simple, flexible and secure in the applications of ZKIP schemes.

**Keywords:** ZKIP protocols, identification, cryptography, digital signature, subset sum problem

### 1. INTRODUCTION

In 1985, the concept of the Zero-Knowledge Interactive Proof (ZKIP) scheme was first proposed by Goldwasser, Micali and Rackoff [1]. A ZKIP scheme is a protocol such that prover P can convince verifier V of the validity of the secret (or witness) P knows while V learns nothing (zero knowledge) about P's secret after the protocol is completed. Goldwasser et al. pointed out that a ZKIP scheme must satisfy the following conditions.

1. **Completeness:** If P and V are honest and follow the protocol, then there is a very large probability that V believes P has the secret.
2. **Soundness:** If P is dishonest and V is honest in the protocol, then there is a very large probability that V does not believe P has the secret.

3. **Witness hiding:** P cannot reveal any information about his/her secret to V; i.e., V is not able to learn anything from P even if P is honest.

Many ZKIP schemes have been developed by using different mathematical assumptions, e.g., discrete logarithm or factoring. Among them, Fiat and Shamir [2] proposed the first provable and practical secure identification and signature scheme (the well-known FS scheme), whose security is based on the difficulty of computing square roots modulo of a composite  $n$  when the factorization of  $n$  is unknown. A faster ZKIP scheme whose security is also based on the difficulty of the factorization problem was developed by Guillow and Quisquater [3]. Their scheme is called the GQ scheme. The GQ scheme is faster than the FS scheme by a factor of three in the computation required. Therefore, the GQ scheme is more suitable for smart card applications.

Independently, another practical ZKIP scheme based on the discrete logarithm problem was also developed by Chaum et al. [4]. Chaum et al. presented an improved identification scheme [5] based on some generalizations of the discrete logarithm problem, e.g., the multiple discrete logarithm, relaxed discrete logarithm and simultaneous discrete logarithm. Later, an identification scheme based on the concept of ZKIP and ElGamal scheme [6] was developed by Beth [7]. It was also claimed that the scheme is suitable for smart card applications.

Recently, Simmons [8] proposed a new ZKIP scheme whose security is based on another cryptographic assumption, i.e., the subset sum problem (or knapsack problem). However, Simmons did not explain how to implement his ZKIP scheme since it needs an indistinguishable box satisfying homomorphism under addition operations in the scheme. In this paper, we will propose two methods for implementing the indistinguishable box satisfying homomorphism under addition operations and use the proposed boxes to implement Simmons's ZKIP scheme.

The rest of this paper is organized as follows. In section 2, we will first review Simmons's scheme briefly. In section 3, a concrete implementation of Simmons's ZKIP scheme will be proposed. Security analysis of the proposed ZKIP will also be included. Another effective ZKIP scheme whose security is also based on the knapsack problem is also proposed in section 4. Finally, we will draw some conclusions in section 5.

## 2. REVIEW OF SIMMONS'S SCHEME

Before presenting our implementations, we will give a brief description of Simmons's scheme [8] in this section.

### 2.1 The Subset Sum Problem

Firstly, we will introduce the subset sum problem (or knapsack problem) before we discuss Simmons's scheme. In general, the subset sum problem can be defined as follows: [10] Given a set of values,  $U_1, U_2, \dots, U_n, S$ , compute  $x_1, x_2, \dots, x_n, x_i \in \{0,1\}$  for all  $1 \leq i \leq n$ , such that

$$S = x_1 U_1 + x_2 U_2 + \dots + x_n U_n$$

It is well-known that the knapsack problem is an *NP-complete* problem; i.e., given  $\mathbf{U} = \{U_1, U_2, \dots, U_n\}$  and  $S$ , it is computationally infeasible for anyone to compute  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \in \{0,1\}$  for all  $1 \leq i \leq n$ , such that  $S = \mathbf{X} \cdot \mathbf{U}$ .

## 2.2 Simmons's Scheme [8]

Given a public sequence  $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$ , assume P (the prover) has secret  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$  with weight  $W(\mathbf{X}) = \lfloor \frac{n}{2} \rfloor$  such that P's public key  $S$  satisfying  $S = \sum_i x_i a_i = \mathbf{X} \cdot \mathbf{A}$ . Let  $g$  be an indistinguishable box satisfying  $g(\mathbf{X} \cdot \mathbf{A}, k) = \mathbf{X} \cdot g(\mathbf{A}, k) = \sum_i x_i g(a_i, k)$ , where  $k$  is a random number. Let  $\Omega$  be the set of permutations of  $n$  elements. Obviously, the cardinality of the set  $\Omega$  is  $n!$ . Simmons's ZKIP scheme can be described as follows.

### Simmons's ZKIP scheme

Repeat the following steps  $t$  times.

1. P randomly chooses a permutation  $\pi \in \Omega$  and a random integer  $k$ ; then P finds  $\mathbf{A}' = \{g(a_i, k) \mid a_i \in \mathbf{A}\}$ .
2. P sends  $\pi(\mathbf{A}')$  and  $S' = g(S, k)$  to V (the verifier).
3. V asks P to do the following two things according to  $b = 0$  or  $b = 1$ .
  - When  $b = 0$ : P sends  $k$  to V, and V checks whether  $g(a_i, k) \in \pi(\mathbf{A}')$  for all  $a_i \in \mathbf{A}$  and  $g(S, k) = S'$  are satisfied or not.
  - When  $b = 1$ : P sends  $\pi(\mathbf{X})$  to V, and V checks whether  $\pi(\mathbf{X}) \cdot \pi(\mathbf{A}') = S'$  is satisfied or not.

If the conditions are upheld for all  $t$  times, then V accepts that P is honest; i.e., P knows the secret  $\mathbf{X}$ . Otherwise, P is an impostor. The probability that V accepts P, an impostor, is at most  $2^{-t}$ . The proof that Simmons's scheme is indeed a ZKIP scheme can be found in [8].

## 3. THE IMPLEMENTATION OF SIMMONS'S ZKIP SCHEME

Although Simmons has shown that his scheme is a ZKIP scheme in [8], the security of Simmons's scheme is based on the assumption that the indistinguishable box exists. However, nobody, including Simmons, has explained how to implement the indistinguishable box, i.e., the function  $g(\cdot, \cdot)$  such that  $g(\mathbf{X} \cdot \mathbf{A}, k) = \sum_i x_i g(a_i, k)$ , and we cannot distinguish  $a_i$  from  $g(a_i, k)$  when  $\mathbf{A} = \{a_1, a_2, \dots, a_n\}$  is given and  $k$  is unknown. Here, we will propose two concrete implementations of the indistinguishable box such that above conditions are satisfied. The first scheme proposed in this paper uses the multiplicative property to realize the indistinguishable box needed in Simmons's scheme. Using the indistinguishable box, we will implement Simmons's ZKIP scheme as follows.

### 3.1 Notations and Setting Up Phase

Let  $p_1, p_2, \dots, p_n$  and  $p$  be large primes known by all users in the system such that  $p > p_i$  for all  $i$  and  $|p| \geq |p_i|$ , where  $|a|$  represents the length of  $a$ . For security purposes,  $n \geq 200$  is sufficient to avoid the attack from solving the subset sum problem directly. Furthermore,  $p_i > \sqrt{p}$ ,  $1 \leq i \leq n$ , must be imposed to avoid the attack by factoring from  $\prod_{i=1}^n p_i^{x_i} \bmod p$ . P computes  $M = \prod_{i=1}^n p_i^{x_i} \bmod p$  as its public key, where  $\underline{\mathbf{X}} = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \in \{0, 1\}$  for all  $i$ , is his secret. It is assumed that P's public key  $M$  is authenticated by V and the weight of  $\underline{\mathbf{X}}$  is  $\lfloor \frac{n}{2} \rfloor$ , which is known by all users in advance.

#### Protocol 1:

Repeat the following steps  $t$  times.

1. P randomly selects a permutation  $\pi \in \Omega$  and an integer  $r \in Z_p \setminus \{0\}$  and computes  $M' = M^r \bmod p$  and  $q_i = p_i^r \bmod p$  for  $i = 1, 2, \dots, n$ . Then, P sends  $\underline{\mathbf{Q}} = \{q'_1, q'_2, \dots, q'_n\} = \pi(q_1, q_2, \dots, q_n)$  and  $M'$  to V.
2. V gives P a challenge with  $b = 0$  or  $b = 1$ .
3. P sends some required elements to V, which depend on the status of  $b$ .
  - If  $b = 0$ , then P answers  $r$  to V.
  - If  $b = 1$ , then P replies  $\underline{\mathbf{X}}' = \{x'_1, x'_2, \dots, x'_n\} = \pi(x_1, x_2, \dots, x_n)$  to V.
4. If  $b = 0$ , V checks whether  $M^r \bmod p = M'$  and  $p_i^r \bmod p \in \underline{\mathbf{Q}}$  for all  $1 \leq i \leq n$  are satisfied or not. If they are satisfied, then V accepts P's proof.

If  $b = 1$ , V checks whether  $M' = \prod_{i=1}^n (q'_i)^{\pi(x_i)} \bmod p$  is identical or not. If it is satisfied, then V accepts P's proof.

If the conditions are upheld for all  $t$  times, then V accepts that P is honest; i.e., P knows the secret  $\underline{\mathbf{X}}$ . Otherwise, P is an impostor. The probability that V accepts P, an impostor, is at most  $2^{-t}$ .

**Remark:** One of anonymous referees proposed a special kind of attack to protocol 1. Now, we will introduce this special case briefly as follows.

- (i) P' chooses a number  $r$  such that  $r \mid (p-1)$  and  $k = (p-1) \mid r < \frac{W}{2}$ .
- (ii) P' computes  $M' = M^r \bmod p$  and  $q_i = p_i^r \bmod p$  for all  $i = 1, 2, \dots, n$ . However, there are only  $k$  possible values; i.e.,  $q_1, q_2, \dots, q_k$  are there  $k$  possible values. Thus,  $M' \in \{q_1, q_2, \dots, q_k\}$ . Hence, it sends  $\underline{\mathbf{Q}} = \{q_1, q_2, \dots, q_k, 1, 1, \dots, 1\}$  to V.
- (iii) V gives a challenge with  $b = 0$  or  $b = 1$ .
- (iv) P' sends some required elements to V, which depend on the status of  $b$ .
  - If  $b = 0$ , then P' answers  $r$  to V.
  - If  $b = 1$ , then P' replies  $\underline{\mathbf{X}}' = \{x_1, x_2, \dots, x_n\}$  to V. Without losing the generality, we let  $q_1 = M^r \bmod p$ ,  $x_1 = 1$ ,  $x_2 = x_3 = \dots = x_k = 0$ . Furthermore, we choose  $\lfloor \frac{n}{2} \rfloor - 1$  elements from  $\{x_{k+1}, x_{k+2}, x_{k+3}, \dots, x_n\}$  and let these chosen elements be 1; others are 0.
- (v) V checks whether these conditions are satisfied for  $b = 0$  ( $b = 1$ ). If they are satisfied, then V accepts that P' is not an impostor.

Intuitively, the most obvious case is  $r = \frac{(p-1)}{2}$ . Consequently, the security of protocol 1 seems to be insecure for the special case above. However, if  $p$  is chosen to be  $2u + 1$ , where  $u$  is a prime and step 2 is replaced by step 2', then this attack can be avoided easily. Step 2' : V checks whether  $q'_i = q'_j$  for all  $i \neq j$ . If it is, repeat step 1; otherwise, V gives P a challenge with  $b = 0$  or  $b = 1$ .

### 3.2 Security of Protocol 1

Now, we will use Lemma 1 to show the completeness of Protocol 1.

**Lemma 1:** If P and V are honest and follow the protocol, then V always accepts that the proof is valid with overwhelming probability.

**Proof:** We divide Protocol 1 into two classes in order to discuss it with respect to  $b = 0$  or  $b = 1$ .

(i) P sends  $r$  to V when the challenge is “ $b = 0$ ”. It is obvious that

$$M^r \bmod p = M', \quad (1)$$

$$\text{and that } p_i^r \bmod p \in \pi(q_1, q_2, \dots, q_n) \text{ for all } i. \quad (2)$$

If Eq. (1) and Eq. (2) are upheld, then V can corroborate that P knows the secret exactly; otherwise, V assumes that P is an impostor.

(ii) If  $b = 1$ , then V will receive the information  $\underline{\mathbf{X}}' = \pi(x_1, x_2, \dots, x_n)$  from P, and it is obvious that

$$\begin{aligned} \underline{Q}^{\underline{\mathbf{X}}'} &= \prod_{i=1}^n (q_i^{x_i}) \bmod p \\ &= q_1^{x_1} \cdot q_2^{x_2} \dots q_n^{x_n} \bmod p \\ &= \left( \prod_{i=1}^n p_i^{x_i} \right)^r \bmod p \\ &= M^r \bmod p \\ &= M'. \end{aligned} \quad (3)$$

If Eq. (3) is upheld, then V is certain that P knows the secret; otherwise, P is an impostor. Thus, V accepts that the proof is valid with overwhelming probability if P and V are honest and follows Protocol 1.  $\square$

**Lemma 2:** Let  $p_1, p_2, \dots, p_n$  and  $p$  be primes as defined above, and integer  $M \in GF(p)$ .

Finding  $\underline{\mathbf{X}} = (x_1, x_2, \dots, x_n)$  and the weight of  $\underline{\mathbf{X}}$  is  $\left\lfloor \frac{n}{2} \right\rfloor$ , where  $x_i \in \{0, 1\}$ , such that

$$M = \prod_{i=1}^n p_i^{x_i} \bmod p, \quad (4)$$

is equivalent to solving both the discrete logarithm in  $GF(p)$  and the subset sum problem.

**Proof:** Let  $\alpha$  be a primitive root modulo  $p$ . If the discrete logarithm problem over  $GF(p)$  is feasible, then there exist integers  $a_1, a_2, \dots, a_n$  and  $s$  such that

$$p_i = \alpha^{a_i} \bmod p, \text{ for all } 1 \leq i \leq n,$$

and  $M = \alpha^s \bmod p$ .

Then Eq. (4) can be rewritten as

$$\begin{aligned} M &= \alpha^s = \prod_{i=1}^n p_i^{x_i} \bmod p \\ &= \alpha^{\sum_{i=1}^n a_i x_i} \bmod p. \end{aligned} \tag{5}$$

Thus, the exponent in Eq. (5) can be described as that given integers  $a_1, a_2, \dots, a_n$  and its subset sum  $s$ , finding  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$  such that  $s = \sum_{i=1}^n a_i x_i$  is upheld. Obviously, it is a subset sum problem if the discrete logarithm problem over  $GF(p)$  is feasible.

Note that, if only the subset sum problem is feasible, then it can find  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$  when  $a_1, a_2, \dots, a_n$  are given and its subset sum  $s$  such that  $s = \sum_{i=1}^n a_i x_i$  is upheld. However, it still needs an efficient oracle  $O_G(M, \alpha, p) = s$ ; i.e., for input  $M, \alpha$  and  $p$  to be the oracle, the oracle will produce  $s$  as the output such that  $\alpha^s \equiv M \bmod p$  and  $\alpha$  is a primitive root modulo  $p$ . Clearly, it is the discrete logarithm problem over  $GF(p)$  if the subset sum problem is feasible. On the other hand, if only the discrete logarithm problem is feasible, then it still needs to solve the subset sum problem for given  $a_1, a_2, \dots, a_n$  and  $s$ . Thus, solving  $x_i, 1 \leq i \leq n$ , from Eq.(4) is equivalent to solving both the discrete logarithm in  $GF(p)$  and the subset sum problem.  $\square$

In the following, we will use Lemma 3 to show the soundness of Protocol 1.

**Lemma 3:** If P is dishonest and V is honest in the protocol, then V does not believe P knows the secret with nonnegligible probability.

**Proof:** Let us assume that P knows the challenge  $b$  in advance. If  $b = 0$ , then P randomly chooses  $r$  and  $\pi$ , computes  $\mathbf{Q}$  and  $M'$ , and sends  $\mathbf{Q}$  and  $M'$  to V. Since  $b = 0$ , P can pass the protocol this time by sending  $r$  to V in step 3. However, if  $b = 1$ , then P must deal with the problem of finding  $\mathbf{X}'$  such that  $\mathbf{Q}^{\mathbf{X}'} = M' \bmod p$ , where  $\mathbf{Q}$  and  $M'$  are given in advance. From Lemma 2, we can see that this is equivalent to solving both the discrete logarithm and the subset sum problem. On the other hand, if  $b = 1$ , then P randomly chooses  $\mathbf{Q}, \mathbf{X}'$ , and  $M'$  such that  $\mathbf{Q}^{\mathbf{X}'} = M' \bmod p$  is satisfied, and sends  $\mathbf{Q}$  and  $M'$  to V. Since  $b = 1$ , P is able to pass the protocol this time by sending  $\mathbf{X}'$  to V in step 3. However, if  $b = 0$ , then P has to find an integer  $r$  such that  $P_i^r \bmod p \in \mathbf{Q}$  for all  $1 \leq i \leq n$  and  $M' = M^r \bmod p$  is satisfied. We conjecture that this is at least as difficult as solving the discrete logarithm problem over  $GF(p)$ .

Based on the above discussion, if P does not know of the challenge  $b$  in advance, i.e., V is honest, then the probability that P can produce the correct secret  $\mathbf{X}$  for  $b$  is at most  $\frac{1}{2}$ .

secret  $\underline{X}$  is at most  $2^t$  if P is dishonest and V is honest. In other words, if P does not know the secret, then any cheating by P in Protocol 1 will be detected by V with a probability of at least  $1 - 2^t$ .  $\square$

**Theorem 1:** In Protocol 1, if P is honest then V is not able to grasp any information about  $\underline{X}$  from P, where the weight of  $\underline{X}$  is  $\left\lfloor \frac{n}{2} \right\rfloor$ .

**Proof:** V is unable to learn any information about the secret  $\underline{X}$  from P even if P is honest.

- In the case of  $b = 0$ :

V will receive parameter  $r$  from P; then V can get the information about the permutation  $\pi$  from Eq. (2). However,  $\pi$  is a random permutation which is independent of  $\underline{X}$ . Hence, V knows nothing about  $\underline{X}$  from parameters  $r, p_1, p_2, \dots, p_n$  and  $p$  since  $r$  is a random integer and  $p_1, p_2, \dots, p_n, p$  are known by V in advance.

- In the case of  $b = 1$ :

V knows  $\underline{X}' = \pi(x_1, x_2, \dots, x_n)$  from P. However, V does not know the permutation  $\pi$  unless V can solve the discrete logarithm problem from  $p_i$  and  $q_i$  or  $M$  and  $M'$ . Thus, V knows nothing about the secret  $\underline{X}$  from P if the discrete logarithm problem over  $GF(p)$  is infeasible.

Hence, V cannot learn anything from P even if P is honest.  $\square$

From Lemmas 1 and 3 and Theorem 1, it is obvious that Protocol 1 is a ZKIP scheme if solving the subset sum problem and the discrete logarithm problem is infeasible.

### 3.3 Example

We will use Example 1 to illustrate Protocol 1 further.

**Example 1:** Let  $p_1 = 5, p_2 = 7, p_3 = 11, p_4 = 13, p_5 = 17, p_6 = 19$  and  $p = 101$ . If  $\underline{X} = \{1, 1, 0, 1, 0, 0\}$ , then  $M = 5 \times 7 \times 13 = 51 \pmod{101}$ . Now, Protocol 1 runs as follows.

1. P chooses  $r = 3$  and  $\pi = (135246)$ . Then, P computes  $\{q_1, q_2, q_3, q_4, q_5, q_6\} = \{24, 40, 18, 76, 65, 92\}$ ,  $\underline{Q} = \pi(q_1, q_2, q_3, q_4, q_5, q_6) = \{q_6, q_5, q_1, q_2, q_3, q_4\} = \{92, 65, 24, 40, 18, 76\}$ ,  $M' = 51^3 \pmod{101} = 38$  and then sends  $\underline{Q}$  and  $M'$  to V.
2. P sends  $r = 3$  to V when  $b = 0$ . Otherwise, P answers  $\underline{X}' = \{0, 0, 1, 1, 0, 1\}$  to V.
3. If  $b = 0$ , then V checks whether  $(5^3 \pmod{101}) = 24, (7^3 \pmod{101}) = 40, (11^3 \pmod{101}) = 18, (13^3 \pmod{101}) = 76, (17^3 \pmod{101}) = 65, (19^3 \pmod{101}) = 92$  are  $\in \underline{Q}$  and  $(51^3 \pmod{101}) = 38$  or not.

If  $b = 1$ , then V computes whether  $\underline{Q}^{\underline{X}'} = 24 \times 40 \times 76 \pmod{101}$  is equivalent to 38 or not. If all of the above conditions are satisfied, then V accepts P's proof. Otherwise, P is an impostor.

## 4. AN ALTERNATIVE ZKIP SCHEME

In the above section, we proposed a new ZKIP scheme based on the subset sum problem by using the multiplicative property to implement the indistinguishable box in Simmons's scheme [8]. Another ZKIP scheme will be developed below which is also based on the same cryptographic assumption with different moduli to replace the function of using the multiplicative property. The second scheme proposed in this paper uses different moduli to realize the indistinguishable box needed in Simmons's scheme. Therefore, this new scheme not only achieves the same level of security as Simmons's scheme does, but also is a practical and flexible scheme. The second proposed function is that  $g(a_i, k) = (a_i k \bmod p) \bmod q$ , where  $p$  and  $q$  are primes with  $2n$  bits and  $n$  bits, respectively,  $a_i$  ( $1 \leq i \leq n$ ) are public integers with  $n$  bits and  $k$  is a random integer with  $2n$  bits. Using the indistinguishable box, we will implement Simmons's ZKIP scheme as follows.

### 4.1 Notations and Setting up Phase

Let  $|y|$  denote the number of bit of an integer  $y$ , and let  $g(\cdot, \cdot)$  be the function as defined above, i.e., the indistinguishable box needed in Simmons's ZKIP scheme. In this scheme, a random sequence  $\underline{\mathbf{A}} = \{a_1, a_2, \dots, a_n\}$ ,  $a_i \in (1, 2^n)$ ,  $1 \leq i \leq n$ , is known by all users. In addition, a prime  $q$  whose length is  $n$  is public to all users. Note that it is computationally infeasible to find the subset sum problem in  $\underline{\mathbf{A}}$  since the sequence  $\underline{\mathbf{A}}$  does not need any assumption, e.g., the superincreasing sequence needed in Merkle-Hellman scheme [9]. Let  $p$  be a prime such that  $|p| = 2|q| = 2|\max\{a_i\}| = 2n$ , and let  $k$  be a random integer such that  $k \in \{1, 2, \dots, p-1\}$ . Now, we will describe our concrete ZKIP scheme as follows.

Suppose P has its secret  $\underline{\mathbf{X}} = \{x_1, x_2, \dots, x_n\}$  with  $W(\underline{\mathbf{X}}) = \lfloor \frac{n}{2} \rfloor$ , where  $x_i \in \{0, 1\}$  for all  $i$ , and P has its public information  $S = \sum_{i=1}^n x_i a_i$ . P wants to prove that V knows the secret  $\underline{\mathbf{X}}$ .

#### Protocol 2:

Repeat the following steps  $t$  times.

1. P computes  $r_i = (a_i k \bmod p) \bmod q$  for all  $i$  and  $\underline{\mathbf{R}} = \pi(r_1, r_2, \dots, r_n) = \{r'_1, r'_2, \dots, r'_n\}$ , where  $k$  and  $p$  are random integers as defined above and  $\pi \in \Omega$ . P also calculates  $Z = (Sk \bmod p) \bmod q$ .
2. P sends  $\underline{\mathbf{R}}$  and  $Z$  to V.
3. V gives P a challenge with  $b = 0$  or  $b = 1$ .
4. P sends parameters  $k$  and  $p$  to V, if  $b = 0$ . Otherwise, P sends  $\underline{\mathbf{X}}' = \pi(x_1, x_2, \dots, x_n) = \{x'_1, x'_2, \dots, x'_n\}$  to V.
5. If  $b = 0$ , then V checks whether  $(a_i k \bmod p) \bmod q \in \underline{\mathbf{R}}$  for all  $i$  and whether  $Z = (Sk \bmod p) \bmod q$  are upheld or not. If both of them are upheld, then V can confirm that P knows the exact information; otherwise, P is an impostor.  
If  $b = 1$ , then V checks whether  $\underline{\mathbf{X}}' \cdot \underline{\mathbf{R}} = Z + cq$ , where  $0 \leq c \leq \frac{n}{2} - 1$ , is satisfied or not. If this condition is satisfied, then V can be convinced that P knows the secret  $\underline{\mathbf{X}} = (x_1, x_2, \dots, x_n)$  exactly. Otherwise, P is an impostor.

## 4.2 Security of Protocol 2

The security of our scheme is based on the following two problems. One is that the knapsack problem must be *NP-complete* problem, and the other is that the function  $g(\cdot, \cdot)$  must be an indistinguishable box. The former is true since we let  $\underline{\mathbf{A}} = \{a_1, a_2, \dots, a_n\}$  be a random sequence; i.e.,  $\underline{\mathbf{A}}$  has no assumption like the superincreasing structure [9]. The latter ought to be true since given  $\underline{\mathbf{R}} = \pi(r_1, r_2, \dots, r_n)$ , where  $r_i = (a_i k \bmod p) \bmod q$ , finding the corresponding  $a_i$  seems to be infeasible without knowing  $k$  and  $p$ . We have implemented the low-density attack proposed by Lagarias and Odlyzko [10] to break any low-density knapsack problem, e.g., the Merkle-Hellman knapsack public key cryptosystem, [9] in our laboratory. However, it is shown that it cannot be used to attack Protocol 2 successfully. Thus, we have the following conjecture.

**Conjecture:** Let  $\underline{\mathbf{R}} = \pi(r_1, r_2, \dots, r_n)$ , where  $r_i = (a_i k \bmod p) \bmod q$ ,  $a_i \in \underline{\mathbf{A}}$ , where  $\underline{\mathbf{A}}$ ,  $\pi$ ,  $k$ ,  $p$ , and  $q$  are defined as in our scheme. Then given  $\underline{\mathbf{R}}$  and  $q$ , it is infeasible to find the corresponding  $a_i$  when  $k$  and  $p$  are unknown.

Although the conjecture which can resist low-density attack does not imply it is secure, as far as we know, there is no obvious weakness in this conjecture. Based on the above conjecture, we give the following theorem and lemma to prove that Protocol 2 is a ZKIP scheme for technical reasons.

In Lemma 4, we will explain how V can be convinced that P knows the secret  $\underline{\mathbf{X}}$ .

**Lemma 4:** If P and V follow Protocol 2, then V accepts that the proof is valid with overwhelming probability.

**Proof:** Now, we will divide Protocol 2 into two parts to discuss the correctness of Protocol 2 with respect to  $b = 0$  or  $b = 1$ .

- In the case of  $b = 0$ :

When P replies the correct  $k$  and  $p$  to V, then it is obvious that

$$(a_i k \bmod p) \bmod q \in \underline{\mathbf{R}} \quad \text{for all } i, \quad (6)$$

$$\text{and } Z = (Sk \bmod p) \bmod q. \quad (7)$$

If both Eq.(6) and Eq.(7) are upheld, then V can confirm that P knows the exact information. Otherwise, P is an impostor.

- In the case of  $b = 1$ :

After V receives  $\underline{\mathbf{X}}'$  from P, then V can compute

$$\begin{aligned} \underline{\mathbf{X}}' \cdot \underline{\mathbf{R}} &= \pi(x_1, x_2, \dots, x_n) \cdot \pi(r_1, r_2, \dots, r_n) \\ &= \sum_{i=1}^n x_i r_i \\ &= \sum_{i=1}^n x_i [(a_i k \bmod p) \bmod q] \\ &= (\sum_{i=1}^n a_i x_i k \bmod p) \bmod q + cq \end{aligned}$$

Since  $r_i < q$  and  $x_i \in \{0, 1\}$  for all  $1 \leq i \leq n$ , we have  $0 \leq c < wt(\underline{\mathbf{X}}) = \lfloor \frac{n}{2} \rfloor$ , where  $wt(\underline{\mathbf{X}})$  is the weight of  $\underline{\mathbf{X}}$ . If Eq.(8) is upheld, then V can be convinced that P knows the secret  $\underline{\mathbf{X}} = \{x_1, x_2, \dots, x_n\}$  exactly. Otherwise, P is an impostor.  $\square$

**Theorem 2:**

- (a) If P does not know the secret  $\underline{\mathbf{X}}$ , then the probability that V will accept P's identity in this protocol is at most  $2^{-t}$ .
- (b) V cannot learn anything from P after Protocol 2 is finished even if P is honest.

**Proof:**

- (a) Similar to the proof given in Lemma 3, if P does not know the secret  $\underline{\mathbf{X}}$ , then at each time in step 4 of Protocol 2, P cannot answer V with exact information if P cannot guess the status of  $b$  correctly. Hence, the probability that P passes each round in Protocol 2 is at most  $\frac{1}{2}$  if V is honest. Since Protocol 2 contains  $t$  independent rounds, the probability that P passes Protocol 2 is at most  $2^{-t}$ .
- (b) Even though P is honest, P cannot reveal any information about the secret  $\underline{\mathbf{X}}$  to V. We will discuss it with regard to the status of  $b$ .

- In the case of  $b = 0$ :

V can get the information about permutation  $\pi$  from Eq.(6) after V receives parameters  $k$  and  $p$  from P. However, since  $\pi$  is a random permutation and  $k$  and  $p$  are random integers which are independent of  $\underline{\mathbf{X}}$ , thus, V knows nothing about  $\underline{\mathbf{X}}$  from parameters  $\pi$ ,  $k$  and  $p$ .

- In the case of  $b = 1$ :

V knows  $\underline{\mathbf{X}}' = \pi(x_1, x_2, \dots, x_n)$ . However, since V does not know the permutation  $\pi$  unless V can solve the subset sum problem from  $\underline{\mathbf{A}}$  and  $S$  or  $\underline{\mathbf{R}}$ ,  $q$  and  $Z$ , V knows nothing about  $\underline{\mathbf{X}}$  except  $|\underline{\mathbf{X}}'| = |\underline{\mathbf{X}}| = \lfloor \frac{n}{2} \rfloor$  if solving the subset sum problem in  $\underline{\mathbf{A}}$  and  $\underline{\mathbf{R}}$  is infeasible. However,  $|\underline{\mathbf{X}}| = \lfloor \frac{n}{2} \rfloor$  is known by V in advance. Hence, the witness hiding is upheld.

Therefore, V is not able to learn anything from P even if P is honest.  $\square$

From Lemma 4 and Theorem 2, we know that Protocol 2 is indeed a ZKIP scheme if solving the subset sum problem is infeasible and the above conjecture is true.

**4.3 Example**

Now, we will use Example 2 to explain Protocol 2 more clearly.

**Example 2:** Let  $\underline{\mathbf{X}} = \{1, 1, 0, 1, 0, 0\}$ ,  $\underline{\mathbf{A}} = \{39, 17, 32, 41, 28, 50\}$ ,  $p = 61$  and  $q = 7$ . Now, we will use this information to finish Protocol 2.

1. P chooses  $k = 5$  and  $\pi = (135246)$ . Then P computes  $r_1 = 5, r_2 = 3, r_3 = 3, r_4 = 1, r_5 = 4, r_6 = 6$ . Therefore,  $\underline{\mathbf{R}} = \{6, 4, 5, 3, 3, 1\}$ ,  $S = 97$  and  $Z = (97 \times 5 \text{ mod } 61) \text{ mod } 7 = 2$ .
2. P sends  $\underline{\mathbf{R}} = \{6, 4, 5, 3, 3, 1\}$  and  $Z = 2$  to V.

3. P sends  $k=5, p=61$  and  $q=7$  to V when  $b=0$ . Otherwise, P answers  $\underline{\mathbf{x}}' = \{0, 0, 1, 1, 0, 1\}$  and  $q=7$  to V.
4. If  $b=0$ , then V checks  $(39 \times 5 \bmod 61) \bmod 7 = 5$ ,  $(17 \times 5 \bmod 61) \bmod 7 = 3$ ,  $(32 \times 5 \bmod 61) \bmod 7 = 3$ ,  $(41 \times 5 \bmod 61) \bmod 7 = 1$ ,  $(28 \times 5 \bmod 61) \bmod 7 = 4$ ,  $(50 \times 5 \bmod 61) \bmod 7 = 6$ . It can be checked that all  $r_i \in \underline{\mathbf{R}}$  and  $(97 \times 5 \bmod 61) \bmod 7 = 2 = Z$ .  
If  $b=1$ , then V computes

$$\begin{aligned}\underline{\mathbf{x}}' \cdot \underline{\mathbf{R}} &= (0, 0, 1, 1, 0, 1) \cdot (6, 4, 5, 3, 3, 1) \\ &= 5 + 3 + 1 = 9 = 2 + 1 \times 7,\end{aligned}$$

where  $0 \leq c = 1 \leq \frac{6}{2} - 1 = 2$ . If all of these conditions are satisfied, then V accepts P's proof. Otherwise, P is an impostor.

## 5. CONCLUSIONS

In this paper, we have proposed two methods for implementing the indistinguishable box needed in Simmons's ZKIP scheme. Using the proposed indistinguishable boxes, we have given two concrete implementations of Simmons's ZKIP scheme whose security is based on the subset sum problem. To our best knowledge, no such indistinguishable box has been proposed until now. As in the method proposed by Fiat and Shamir [2], it is easy to modify the identification schemes proposed in this paper to obtain digital signature schemes. However, we do not intend to discuss this issue here.

## REFERENCES

1. S. Goldwasser, S. Micali and C. Rackoff, "The knowledge complexity of interactive proof-systems," *17th ACM Symposium on Theory of Computation*, 1985, pp. 291-304.
2. A. Fiat and A. Shamir, "How to prove yourself: practical solutions to identification and signature problems," *Advances in Cryptology: Proceedings of Crypto '85* (Lecture Notes in Computer Science), Vol. 218, Springer-Verlag, Berlin Heidelberg, 1986, pp. 186-194.
3. L. C. Guillow and J. J. Quisquater, "A practical zero-knowledge protocol fitted to security microprocessor minimizing both transmission and memory," *Advances in Cryptology: Proceedings of Eurocrypt '88* (Lecture Notes in Computer Science), Vol.330, Springer-Verlag, Berlin Heidelberg, 1988, pp. 123-128.
4. D. Chaum, J. H. Evertse, J. van de Graff and R. Peralta, "Demonstrating possession of discrete logarithm without revealing it," *Advances in Cryptology: Proceedings of Crypto '86* (Lecture Notes in Computer Science), Vol. 218, Springer-Verlag, Berlin Heidelberg, 1987, pp. 200-212.
5. D. Chaum, J. H. Evertse and J. van de Graff, "An improved protocol for demonstrating possession of a discrete logarithms and some generalizations," *Advances in Cryptology: Proceedings of Eurocrypt '87*(Lecture Notes in Computer Science), Vol. 263, Springer-Verlag, Berlin Heidelberg, 1988, pp. 127-141.
6. T. ElGamal, "A public-key cryptosystem and a signature scheme based on discrete

7. T. Beth, "Efficient zero-knowledge identification scheme for smart cards," *Advances in Cryptology: Proceedings of Eurocrypt '88* (Lecture Notes in Computer Science), Vol. 330, Springer-Verlag, Berlin Heidelberg, 1988, pp. 77-84.
8. G. J. Simmons, "Identification of data, devices, documents and individuals," in *Proceedings of 1991 IEEE International Carnahan Conference on Security Technology*, pp. 197-218.
9. C. Merkle and M.E. Hellman, "Hiding information and signatures in trapdoor knapsack," *IEEE Transactions on Information Theory*, Vol. 24, No. 5, 1978, pp. 525-530.
10. J. C. Lagarias and A. M. Odlyzko, "Solving low-density subset sum problems," in *Proceedings of the 24<sup>th</sup> IEEE Symposium on the Foundations of Computer Science*, 1983, pp. 1-10.
11. C. S. Lai, W. C. Kuo and M. J. Gau, "Zero-knowledge interactive proof schemes based on subset sum problem," in *Proceedings of the 1996 International Conference on Cryptology and Information Security*, 1996, pp. 91-98.

**Wen-Chung Kuo (郭文中)** was born in Chiayi, Taiwan, on March 3, 1963. He received the B.S. degree in Electrical Engineering from National Cheng Kung University and the M.S. degree in Electrical Engineering from National Sun Yat-Sen University and the Ph. D. degree in Electrical Engineering from National Cheng Kung University in 1990, 1992 and 1997, respectively. Now, he is a Section Director at Radio Wave Regulatory Department at Directorate General of Telecommunications, Ministry of Transportation and Communications. His research interests include cryptography, network security and signal processing.

**Chi Sung Lai (賴溪松)** was born on June 4, 1956 in Chiayi, Taiwan, Republic of China. He received his B.S., M.S. and Ph.D. degrees, all in Electrical Engineering from National Cheng Kung University in 1984, 1986 and 1990, respectively.

Since September 1986, he has been on the faculty of the Department of Electrical Engineering at National Cheng Kung University, Tainan, Taiwan, and currently is a professor. From August 1993 to January 1997, he was an adjunct research fellow at the Engineering and Technology Promotion Center of the National Science Council of the Republic of China. From February 1997, he was the director of Project Management, office of Research and Development at National Cheng Kung University. In June 1997, he was elected Chairman of the Chinese Cryptology and Information Security Association (CCISA). His research interests include Cryptology, Information Security, Error Control Codes and Communication Systems.

Dr. Lai is a member of IEEE, ACM and IACR. He was the winner of the 1991 Acer Long Term Award for Outstanding M.S. Thesis Supervision, the winner of the Graduate Team of TI-Taiwan 1994 DSP Design Championship and the winner of the 1997 Outstanding Paper Award of CCISA. He also received the 1997-1998 Outstanding Research Award from the National Science Council of the Republic of China.

**Min-Jea Gau (高銘智)** was born in Taipei, Taiwan, on February 15, 1966. He received the B.S. degree in Mathematics from Mationa Central University and the M.S. degree in Electrical Engineering from National Cheng Kung University in 1992 and 1996, respectively. Now, he is an engineer at the Computer & Communication Research Laboratory at the Industrial Technology Research Institute. His work is focused on Secure Electronic Transaction. His research interests include cryptography and network security.

