

## ANALYSIS OF PRIVACY-PRESERVING ELEMENT REDUCTION OF A MULTISSET

JAE HONG SEO, HYOJIN YOON, SEONGAN LIM, JUNG HEE CHEON,  
AND DOWON HONG

**ABSTRACT.** The element reduction of a multiset  $S$  is to reduce the number of repetitions of an element in  $S$  by a predetermined number. Privacy-preserving element reduction of a multiset is an important tool in private computation over multisets. It can be used by itself or by combination with other private set operations. Recently, an efficient privacy-preserving element reduction method was proposed by Kissner and Song [7]. In this paper, we point out a mathematical flaw in their polynomial representation that is used for the element reduction protocol and provide its correction. Also we modify their over-threshold set-operation protocol, using an element reduction with the corrected representation, which is used to output the elements that appear over the predetermined threshold number of times in the multiset resulting from other privacy-preserving set operations.

### 1. Introduction

Private set operations such as *set intersection*, *set union*, and *element reduction* of multisets are important tools for privacy in many applications. The *element reduction* (by  $d$ ) method for a multiset  $S$ , a set that allows the repetition of elements, is a method to obtain a multiset,  $(Rd_d(S))$  after reducing the repetition number of each element by  $d$ . Whenever one sees an element  $a$  in  $Rd_d(S)$ , then he/she knows that  $a$  appears more than  $d$  times in  $S$ . A private *element reduction* method of a multiset enables the controlled disclosure of private information and it can be combined with other private set operations to support the controlled privacy level of the output of a private set operation. The *element reduction* of a multiset also can be used to develop privacy-preserving

---

Received March 22, 2007; Revised July 14, 2008.

2000 *Mathematics Subject Classification.* 94A60.

*Key words and phrases.* privacy-preserving operations, set operations, element reduction, multi-party computations.

This work was partly supported by the IT R&D program of MEK/IITA [2005-Y001-04, Development of next generation security technology]. The first and the fourth authors were partially supported by the SRC Program of Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (R11-2007-035-01002-0).

techniques for monitoring distributed networks. In a distributed network monitoring service, each node monitors anomalous local traffic, and the distributed nodes collectively identify behaviors that are identified by at least a threshold number of monitors.

In Crypto 2005, Kissner and Song studied privacy-preserving set operations such as set intersection, set union, and element reduction of multisets [7]. By using the polynomial representation of set operations and a public-key encryption scheme with a homomorphic property, they proposed protocols for privacy-preserving set intersection and set union. They also proposed an over-threshold set-union protocol by using their polynomial representation of an element in a multiset.

In this paper, we point out that there is a mathematical flaw in their polynomial representation of element reduction used for a multiset. Due to this mathematical flaw, the protocol may identify elements that appear in the multiset for less than the threshold ( $t$ ) number of times. Hence, when we apply their over-threshold set-union protocol to a network monitoring system, it may identify user with normal behavior as a user with anomalous behavior and this leads to privacy threat of a normal user. Hence, this could be a serious problem in privacy-preserving techniques, and it is necessary to correct the protocol. We provide a correction to Kissner and Song's polynomial representation of an element in a multiset for element reductions. We also modify their over-threshold set-union protocol and propose an over-threshold set-operation protocol on the basis of the corrected polynomial representation. Our over-threshold set-operation protocol can be combined with any privacy-preserving set operation so that the output contains only the elements that appear in the resulting multiset of the set operation over the predetermined threshold number. The security proof of Kissner and Song's protocol can be preserved in our protocol since our protocol differs from their protocol only in the manner of representing the element reduction of a multiset as polynomials. Hence, our modified over-threshold set-operation protocol is provably secure in both standard adversary models: honest-but-curious (HBC) and malicious adversary model.

We remark that a preliminary version of this work was announced through [10]. A protocol similar to that presented in this work was independently suggested in [6], at about the same time.

Our paper is organized as follows. In Section 2, we introduce a multiset and its polynomial representation used in this paper. In Section 3, we describe Kissner and Song's polynomial representation of element reduction of a multiset and point out a mathematical flaw in the element reduction case and errors in their over-threshold set-union protocol. In Section 4, we present a modified over-threshold set-operation protocol on the basis of the correction. We conclude our paper in Section 5.

## 2. Multiset and its polynomial representation

First, we consider the concept of a *multiset*. In contrast to an ordinary “set”, a multiset permits the duplication of its elements. For example, in a multiset, an element is represented more than once, like  $\{a, a, b\}$ , and the multiset is different from  $\{a, b\}$ .

Now, we define the set intersection, set union of multisets, and the element reduction of a multiset as follows:

**Definition 1.** The intersection of multisets  $A$  and  $B$ ,  $A \cap B$ , is the multiset composed of the elements that are in both  $A$  and  $B$ . If an element “ $a$ ” appears  $l_A$  times in  $A$  and  $l_B$  times in  $B$ , then “ $a$ ” appears  $\min\{l_A, l_B\}$  times in  $A \cap B$ .

**Definition 2.** The union of multisets  $A$  and  $B$ ,  $A \cup B$ , is the multiset composed of the elements that are in  $A$  or  $B$ . If an element “ $a$ ” appears  $l_A$  times in  $A$  and  $l_B$  times in  $B$ , then  $a$  appears  $l_A + l_B$  times in  $A \cup B$ .

**Definition 3.** The element reduction by  $d$ ,  $\text{Rd}_d(A)$ , of a multiset  $A$  is the multiset composed of the elements of  $A$  such that for every element “ $a$ ” that appears  $d'$  times in  $A$ ,  $a$  is included  $\max\{0, d' - d\}$  times in  $\text{Rd}_d(A)$ .

Next, we introduce a polynomial representation of a multiset. Let a ring  $R$  be the domain of the homomorphic encryption function and  $P$  be a subset of the ring  $R$ , where the elements in  $P$  are uniformly distributed in  $R$  and the probability that a randomly chosen element of  $R$  is an element of  $P$  is negligible.

- From a multiset  $S$  to a polynomial  $f_S \in R[x]$  :
  - Given a multiset  $S = \{S_j\}_{1 \leq j \leq k}$ ,  $S_j \in P$ , the polynomial  $f_S \in R[x]$  that represents the multiset  $S$  can be constructed as

$$f_S(x) = \prod_{1 \leq j \leq k} (x - S_j).$$

- From a polynomial  $f \in R[x]$  to a multiset  $S$  :
  - Given a polynomial  $f \in R[x]$ , the multiset  $S$  represented by  $f$  can be defined as follows:

$$a \in S \text{ and } a \text{ appears } t \text{ times in } S \iff (x - a)^t | f, (x - a)^{t+1} \nmid f \text{ and } a \text{ is an element in } P$$

If  $f$  and  $g$  are the polynomial representations of multisets  $S$  and  $T$ , respectively, then  $f * g$  and  $\text{gcd}(f, g)$  are the polynomial representations of  $S \cup T$  and  $S \cap T$ , respectively. Furthermore,  $S \cap T$  is represented by  $f * r + g * s$  for random values of  $r$  and  $s$  of a higher to or the same degree as that of  $\text{deg}(f)$  with an overwhelming probability [7, 8].

## 3. Analysis of Kissner and Song’s element reduction methods

In this section, we review the results of Kissner and Song [7, 8] and point out a flaw in their polynomial representation of an element in a multiset for

element reductions. Furthermore, we show that there are critical errors in over-threshold set-union protocol due to incorrect polynomial representations used for element reduction in a multiset.

### 3.1. Flaws in Kissner and Song's polynomial representation for element reduction

Kissner and Song used polynomials to represent multisets and proposed probabilistic polynomial representations corresponding to the element reduction of a multiset. Kissner and Song's incorrect polynomial representation of the element reduction (by  $d$ ) of a multiset is given as follows:

**(Incorrect) Element reduction (by  $d$ ):** Let  $f$  be the polynomial representation of a multiset  $S$ . For random polynomials  $r$  and  $s$  with a greater than or equal to degree  $\deg(f)$  and a random polynomial  $F$  with degree  $d$  whose solutions are not in  $P$ ,  $f^{(d)} * F * r + f * s$  is equal to  $\gcd(f, f^{(d)}) * u$ , where  $f^{(d)}$  is the  $d$ -th derivative of  $f$ ,  $u$  is uniformly distributed in  $R^\alpha[x]$ ,  $R^\alpha[x]$  is the set of all polynomials whose coefficients are in  $R$  and whose degrees are lesser than or equal to  $\alpha = 2 \deg(f) - |Rd_d(S)|$ . The polynomial  $f^{(d)} * F * r + f * s$  is a polynomial representation of the multiset  $Rd_d(S)$  with an overwhelming probability.

In the above explanation, since  $u$  is uniformly distributed in  $R^\alpha[x]$ , the probability of  $u$  having a root in  $P$  is negligible.  $f^{(d)} * F * r + f * s = \gcd(f, f^{(d)}) * u$  is the polynomial representation of the multiset  $Rd_d(S)$  with an overwhelming probability.

The polynomial representation of the element reduction proposed in [7, 8] uses the following lemma.

**Lemma 1** (Lemma 2 in [7]). *Let  $R$  be a ring and  $f(x) \in R[x]$ . Let  $d \geq 1$ .*

**A.** *If  $(x - a)^{d+1} | f(x)$ , then  $(x - a) | f^{(d)}(x)$ .*

**B.** *If  $(x - a) | f(x)$  and  $(x - a)^{d+1} \nmid f(x)$ , then  $(x - a) \nmid f^{(d)}(x)$ .*

By using Lemma 1, Kissner and Song showed that  $\gcd(f, f^{(d)})$  is a polynomial representation of  $Rd_d(S)$ , where  $f$  is the polynomial representation of the multiset  $S$ . However, Lemma 1 is incorrect when  $d > 1$ . We can provide a counter-example for Lemma 1 as follows.

**Example 1.** Let  $a, b$  and  $c$  be distinct elements of ring  $R$ . Let  $f(x) = (x - a)(x - b)(x - c)$ . If Lemma 1 is correct, then the relation

$$(x - a) \nmid f^{(2)}(x)$$

holds since  $(x - a) | f$  and  $(x - a)^3 \nmid f(x)$ . However,  $f^{(2)}(x) = 6x - 2(a + b + c)$  and  $(x - \frac{a+b+c}{3}) | f^{(2)}(x)$ , i.e.,

$$(x - a) | f^{(2)}(x), \text{ when } c = 2a - b.$$

This contradicts Lemma 1.

The mathematical flaw in Lemma 1 results in errors in their polynomial representation of element reduction. In Example 1, we consider an element reduction by 2 for the set  $S = \{a, b, c\}$  of distinct elements, with  $c = 2a - b$ . Clearly, it can be noted that  $Rd_2(S) = \phi$ . However, as shown in the above example,  $\gcd(f, f^{(2)}) = (x - a)$ , which cannot be a polynomial representation of  $Rd_2(S) = \phi$ .

### 3.2. Analysis of Kissner and Song's protocol

In this section, we analyze the *over-threshold set-union protocol* proposed in [7, 8]. This protocol is a multiparty protocol with  $n$  users under the assumption that at most,  $c$  ( $< n$ ) players can dishonestly collude. A user  $i$  (where  $1 \leq i \leq n$ ) generates a multiset  $S_i$  whose elements represent private information and they are in  $P$ . Assume that each individual multiset should have the same cardinality. That is, for all  $i$  such  $1 \leq i \leq n$ ,  $|S_i| = k$  for some value of  $k$ . The  $j$ -th element of the multiset  $S_i$  is represented by  $(S_i)_j$ . At the end of the protocol, all users want to obtain a multiset that consists of the elements greater than the threshold in the set union  $S = S_1 \cup \dots \cup S_n$  of each user's multiset. The goal of the protocol is to solve the *over-threshold set-union problem* which is defined in [7, 8] as follows:

**Definition 4.**<sup>1</sup> All the players know the elements in the union of the each players' private multisets that appear more than a threshold number of times, and the frequency of these elements in the union without any other information. We call the elements of the resulting set as over-threshold elements in the union of the private sets of all the players.

In their *over-threshold set-union protocol*, Kissner and Song used the element reduction method to obtain the over-threshold elements in the set union. Consider a fixed threshold number  $t$  and a polynomial  $p$  that corresponds to the multiset  $S$  of the union of the private sets of all players. Kissner and Song computed  $\gcd(p, p^{(t-1)})$  as the polynomial representation of  $Rd_{t-1}(S)$ .

If we apply their *over-threshold set-union protocol* to the set union  $S = \{a, b, c\}$  with a threshold of 3 as in Example 1, then the protocol outputs the corresponding set  $\{a\}$  as the set with an over-threshold of 3 in set union. However,  $a$  appears only once in the set  $S$ ; hence, Kissner and Song's protocol is not the correct threshold 3 protocol.

Suppose we consider the above example in a distributed network monitoring system with a privacy policy that states "*the monitoring system identifies only users with an anomalous behavior over threshold 3*". Then, the user  $a$  will be identified in the monitoring system; however, since it appears only once, and it should not be identified in the monitoring system. This conflicts the privacy policy adopted by them.

---

<sup>1</sup>This definition can be extended to an over-threshold set-operation problem by replacing the union by a general set operation.

Hence, a correction is required in Kissner and Song's polynomial representation of element reduction.

#### 4. Corrected polynomial representation and modified over-threshold set-operation protocol

In this section, we suggest a corrected polynomial representation of element reduction and propose a modified over-threshold set-operation protocol by using the corrected polynomial representation.

##### 4.1. Corrected polynomial representation of element reduction

We propose a new method for element reduction by correcting Lemma 1. In particular, we prove that  $\gcd(f, f', \dots, f^{(d)})$  is a polynomial representation of  $Rd_d(S)$ .

By correcting Lemma 1, we obtain the following lemma.

**Lemma 2.** *Let  $f(x) \in R[x]$ . Then the following are equivalent.*

- A.  $(x - a)^{d+1} \mid f(x)$ .
- B.  $f(a) = f'(a) = \dots = f^{(d)}(a) = 0$ , i.e.,  $(x - a) \mid f$ ,  $(x - a) \mid f'$ ,  $\dots$ ,  $(x - a) \mid f^{(d)}$ .

*Proof.* (A  $\Rightarrow$  B) Assume that  $(x - a)^{d+1}$  divides  $f(x)$  and  $f(x) = (x - a)^{d+1}g(x)$  for some  $g(x)$ . Then, the general expression for  $f^{(n)}$  is

$$f^{(n)}(x) = (d + 1)d \cdots (d - n + 2)(x - a)^{d+1-n}g(x) + (x - a)^{d+2-n}h_n(x)$$

for some  $h_n(x)$  and  $1 \leq n \leq d$ . Hence,  $f(a) = \dots = f^{(d)}(a) = 0$ .

(B  $\Rightarrow$  A) First, we will show that if  $f^{(n)}(a) = 0$  and  $(x - a)^n \mid f(x)$ , then  $(x - a)^{n+1} \mid f(x)$ . Since  $(x - a)^n \mid f(x)$ , we have  $f(x) = (x - a)^n g(x)$  for some  $g(x)$ .  $f^{(n)}(x) = n!g(x) + (x - a)h_n(x)$  for some  $h_n(x)$ . Since  $f^{(n)}(a) = 0$ , we have  $g(a) = 0$ , which implies that  $g(x) = (x - a)g_1(x)$  for some  $g_1(x)$ . Therefore,  $f(x) = (x - a)^{n+1}g_1(x)$ .

Because  $f^{(1)}(a) = 0$  and  $(x - a) \mid f(x)$  by hypothesis, we have  $(x - a)^2 \mid f(x)$ . Further, by combining  $(x - a)^2 \mid f(x)$  with  $f^{(2)}(a) = 0$ , we have  $(x - a)^3 \mid f(x)$ . By repeating the same procedure with  $f^{(3)}(a) = 0, \dots, f^{(d)}(a) = 0$ , we eventually obtain  $(x - a)^{d+1} \mid f(x)$ .  $\square$

We obtain the following corollary from Lemma 2.

**Corollary 3.**  $(x - a)^{d+1} \mid f(x) \Rightarrow (x - a)^d \mid f'(x)$ .

*Proof.* By Lemma 2,

$$\begin{aligned} (x - a)^{d+1} \mid f(x) &\Leftrightarrow f(a) = f'(a) = \dots = f^{(d)}(a) = 0 \\ &\Rightarrow f'(a) = f''(a) = \dots = f^{(d-1)}(a) = 0 \\ &\Rightarrow (x - a)^d \mid f'(x) \text{ by applying Lemma 2 to } f'. \end{aligned}$$

Thus, Corollary 3 is proved.  $\square$

Next, we will prove that  $\gcd(f, f', \dots, f^{(d)})$  is a correct polynomial representation of  $Rd_d(S)$ .

**Theorem 4.** *Let  $f$  be a polynomial representation of a multiset  $S$ . For  $a \in S$  and a positive integer  $\ell_a$ ,*

$$(x-a)^{\ell_a} \mid \gcd(f, f', \dots, f^{(d)}), \quad (x-a)^{\ell_a+1} \nmid \gcd(f, f', \dots, f^{(d)}) \\ \Leftrightarrow a \text{ appears } \ell_a \text{ times in } Rd_d(S).$$

That is  $\gcd(f, f', \dots, f^{(d)})$  is a polynomial representation of  $Rd_d(S)$ .

*Proof.* For sufficiency, assume that  $\ell_a$  is a positive integer that satisfies

$$(x-a)^{\ell_a} \mid \gcd(f, \dots, f^{(d)}), \text{ and } (x-a)^{\ell_a+1} \nmid \gcd(f, \dots, f^{(d)}).$$

Then, since

$$(x-a)^{\ell_a} \mid \gcd(f, \dots, f^{(d)}),$$

we have

$$(x-a)^{\ell_a} \mid f, (x-a)^{\ell_a} \mid f', \dots, (x-a)^{\ell_a} \mid f^{(d)}.$$

By Corollary 3, we have

$$(x-a)^{\ell_a-1} \mid f^{(d+1)}, \dots, (x-a)^1 \mid f^{(d+\ell_a-1)}.$$

Thus, we have  $(x-a)^{d+\ell_a} \mid f$  by Lemma 2.

If  $(x-a)^{\ell_a+d+1} \mid f$ , then  $f(a) = \dots = f^{(\ell_a+d)}(a) = 0$ . The part  $f(a) = \dots = f^{(\ell_a)}(a) = 0$  implies that  $(x-a)^{\ell_a+1} \mid f$  by Lemma 2. Similarly,  $f^{(i)}(a) = \dots = f^{(\ell_a+i)}(a) = 0$  implies that  $(x-a)^{\ell_a+1} \mid f^{(i)}$  for all  $i$ , with  $1 \leq i \leq d$ , and it means that  $(x-a)^{\ell_a+1} \mid \gcd(f, \dots, f^{(d)})$ . This contradicts to the hypothesis. Therefore, we have  $(x-a)^{\ell_a+d+1} \nmid f$ .

Hence,  $\ell_a$  satisfies

$$(x-a)^{\ell_a+d} \mid f \text{ and } (x-a)^{\ell_a+d+1} \nmid f.$$

Further, we know that  $a$  appears  $\ell_a + d$  times in multiset  $S$  since  $f$  is a polynomial representation of  $S$ . Thus,  $a$  appears  $\ell_a$  times in  $Rd_d(S)$ .

For necessity, assume that  $a$  appears  $\ell_a$  times in  $Rd_d(S)$ . Then, we have  $(x-a)^{\ell_a+d} \mid f$  and  $(x-a)^{\ell_a+d+1} \nmid f$ . By Corollary 3, we have

$$(x-a)^{\ell_a+d-1} \mid f', \dots, (x-a)^{\ell_a} \mid f^{(d)}, \dots, (x-a) \mid f^{(\ell_a+d-1)}, (x-a) \nmid f^{(\ell_a+d)}.$$

This implies that

$$(x-a)^{\ell_a} \mid f, (x-a)^{\ell_a} \mid f', \dots, (x-a)^{\ell_a} \mid f^{(d)}, \text{ but } (x-a)^{\ell_a+1} \nmid f^{(d)}.$$

Thus,  $\ell_a$  is a positive integer that satisfies

$$(x-a)^{\ell_a} \mid \gcd(f, \dots, f^{(d)}) \text{ and } (x-a)^{\ell_a+1} \nmid \gcd(f, \dots, f^{(d)}). \quad \square$$

By Theorem 4, we can correct Kissner and Song's polynomial representation of element reduction as follows:

**Corrected Element Reduction (by  $d$ )** Let  $f$  be the polynomial representation of a multiset  $S$ . For a random polynomial  $r_i$  of degree greater than or equal to  $\deg(f)$  and a random polynomial  $F_i$  with degree  $i$  whose solutions are not in  $P$ ,  $\sum_{i=0}^d f^{(i)} * F_i * r_i$  is equal to  $\gcd(f, f', \dots, f^{(d)}) * u(x)$ . Thus, the polynomial  $\sum_{i=0}^d f^{(i)} * F_i * r_i$  is a polynomial representation of the multiset  $Rd_d(S)$  with an overwhelming probability.

## 4.2. Over-threshold set-operation protocol

In this section, we propose an over-threshold set-operation protocol using our modified polynomial representation of an element in a multiset. As described above, the goal of this protocol is that all players should obtain the multiset of the elements that appear in the result of the set operation of each private multiset more than a predetermined threshold number of times without acquiring any other information.

There are  $n$  ( $\geq 2$ ) Honest-But-Curious players with a private input set  $S_i$  such that  $|S_i| = k$ . We assume that at most  $c$  ( $< n$ ) players can dishonestly collude. The Honest-But-Curious (HBC) players act according to predetermined actions in the protocol. If no player or coalition of  $c$  players obtains information on the private input of the other players other than what they can deduce from the result of the protocol, then the protocol is secure. We can find the formal definitions of this model in [5]. The players share the secret key  $sk$  corresponding to the public key  $pk$  for a additively homomorphic cryptosystem  $E_{pk}$  that satisfies several characteristics: it is additively homomorphic and supports ciphertext re-randomization and threshold decryption. Paillier's cryptosystem [9] satisfies these requirements. Since we use homomorphic encryption to encrypt polynomial, we define the encryption of a polynomial as follows.

$$E_{pk}(f(x)) := (E_{pk}(f[0]), \dots, E_{pk}(f[\deg(f)]))$$

A detailed explanation of the additively homomorphic public-key cryptosystem and the feasible homomorphic operations of encrypted polynomials can be found in [7, 8]. Let the threshold number be  $d$  and  $F_j$  be an arbitrary polynomial of degree  $j$  that has no roots representing the elements of the set  $P$ .

We use the modified element reduction method of the multiset in step 2 and fix the flaw in the original method proposed by Kissner and Song. Steps 1 and 4 of the above mentioned protocol can be varied according to the type of set operation. In [7, 8], protocols for privacy-preserving set operations, set union, and set intersection were proposed. In step 1, if we apply Kissner and Song's privacy-preserving set union and set intersection then we obtain the over-threshold set-union protocol and the over-threshold set-intersection protocol, respectively.



Protocol: over-threshold set-operation - HBC

1. **Set Operation:** Each player  $i = 1, \dots, n$  computes  $f_i(x) = (x - (S_i)_1) \cdots (x - (S_i)_k)$ . Players perform the predetermined set operation protocol, and player 1 obtains the encryption of polynomial  $p$ , corresponding to the result of the set operation ( $E_{pk}(p)$ ). Player 1 distributes  $E_{pk}(p)$  to players  $2, \dots, c + 1$ .
2. **Element Reduction:** Each player  $i = 1, \dots, c + 1$ 
  - (a): computes  $E_{pk}(p'), \dots, E_{pk}(p^{(d)})$  from  $E_{pk}(p)$ .
  - (b): randomly choose  $(d + 1)$  polynomials  $t_{i,0}, \dots, t_{i,d} \in R^k[x]$ .
  - (c): sends  $E_{pk}(p * t_{i,0} + F_1 * p' * t_{i,1} + \cdots + F_d * p^{(d)} * t_{i,d})$  to all the other players.
3. **Group Decryption:** All the players perform group decryption to obtain  $\Phi = F_d * p^{(d)} * (\sum_{i=1}^{c+1} t_{i,d}) + \cdots + F_1 * p' * (\sum_{i=1}^{c+1} t_{i,1}) + p * (\sum_{i=1}^{c+1} t_{i,0})$ .
4. **Recovering Set:** Each player  $i = 1, \dots, n$  determines the resulting set depending on the type of set operation.

Because the difference between the over-threshold set-union protocol proposed in [7, 8] and our protocol is only in the polynomial representation method of the element reduction of a multiset; it does not affect the security of the protocol. Thus, our protocol has the same security as that of [7, 8] in the set intersection and set union cases when we follow their set operation protocol.

## 5. Conclusion

A privacy-preserving element reduction method can be an important tool to identify internet users with anomalous behavior, while it preserves the privacy for normal users. In Crypto 2005, Kissner and Song introduced a polynomial representation of an element in a multiset for element reductions and proposed an over-threshold set-union protocol using the polynomial representation and homomorphic public-key encryption scheme. In this paper, we have shown that their polynomial representation is not correct and its impact to their protocol can be somewhat critical for privacy-preserving techniques. We present a correction for the polynomial representation of element reduction of a multiset. We also modify their over-threshold set-union protocol and propose an over-threshold set operation protocol on the basis of the corrected polynomial representation. Our over-threshold set-operation protocol can be combined with a privacy-preserving set operation, and it outputs the elements that are greater than a predetermined threshold number in the multiset resulting from the set operation.

## References

- [1] R. Agrawal, A. Evmimievski, and R. Srikant, *Information sharing across private databases*, In SIGMOD 2003, Proceedings of the 2003 ACM SIGMOD international conference on Management of data, pp. 86–97, ACM Press, 2003.
- [2] D. Boneh, E.-J. Goh, and K. Nissim, *Evaluating 2-DNF formulas on ciphertexts*, In TCC 2005, Lecture Notes in Comput. Sci. Vol. 3378, pp. 325–341. Springer-Verlag, 2005.
- [3] P.-A. Fouque and D. Pointcheval, *Threshold cryptosystems secure against chosenciphertext attacks*, In Advances in Cryptology—Asiacrypt 2000, Lecture Notes in Comput. Sci. Vol. 1976, pp. 573–584, Springer-Verlag, 2000.
- [4] M. Freedman, K. Nissim, and B. Pinkas, *Efficient private matching and set intersection*, Advances in cryptology—EUROCRYPT 2004, Lecture Notes in Comput. Sci. Vol. 3027, pp. 1–19, Springer, Berlin, 2004.
- [5] O. Goldreich, *The Foundations of Cryptography - Vol. 2*, Cambridge University Press, 2004.
- [6] L. Kissner, *Privacy-preserving distributed information sharing*, Ph. D. Thesis, 2006, <http://www.cs.cmu.edu/~leak/papers/thesis.pdf>
- [7] L. Kissner and D. Song, *Privacy-preserving set operations*, Advances in cryptology—CRYPTO 2005, Lecture Notes in Comput. Sci. Vol. 3621, pp. 241–257, Springer, Berlin, 2005.
- [8] ———, *Private and threshold set-intersection*, Technical Report CMU-CS-05-113, Carnegie Mellon University, 2005.
- [9] P. Pallier, *Public-key cryptosystems based on composite degree residuosity classes*, In Advances in Cryptology—Eurocrypt 1999, Lecture Notes in Comput. Sci. Vol. 1592, pp. 223–238, Springer-Verlag, 1999.
- [10] J. Seo and H. Yoon, *Analysis of privacy-preserving element reduction of a multiset*, Memoir of the 2nd cryptology paper contest, arranged by a government organization, pp. 77–91, 2006.

JAE HONG SEO  
 DEPARTMENT OF MATHEMATICAL SCIENCES AND ISAC-RIM  
 SEOUL NATIONAL UNIVERSITY  
 SEOUL 151-747, KOREA  
*E-mail address:* jhsbhs@gmail.com

HYOJIN YOON  
 DEPARTMENT OF MATHEMATICAL SCIENCES AND ISAC-RIM  
 SEOUL NATIONAL UNIVERSITY  
 SEOUL 151-747, KOREA  
*E-mail address:* jin25@math.snu.ac.kr

SEONGAN LIM  
 DEPARTMENT OF MATHEMATICS  
 INHA UNIVERSITY  
 INCHEON 402-751, KOREA  
*E-mail address:* seongannym@inha.ac.kr

JUNG HEE CHEON  
 DEPARTMENT OF MATHEMATICAL SCIENCES AND ISAC-RIM  
 SEOUL NATIONAL UNIVERSITY  
 SEOUL 151-747, KOREA  
*E-mail address:* jhcheon@snu.ac.kr

DOWON HONG  
ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE  
TAEJON 305-700, KOREA  
*E-mail address:* `dwhong@etri.re.kr`