2002년 한국정보처리학회 춘계학술발표논문집 제 9 권 제 1 호

속임수를 가지는 비밀분산 방식에 대한 고찰

김문정*, 오수현**, 원동호**, 채영도*,
*성균관대학교 수학과
**성균관대학교 전기전자 및 컴퓨터공학부
e-mail : kmj@math.skku.ac.kr

A Study on Secret Sharing Scheme with Cheaters

Moon-Jeong Kim*, Soo-Hyun Oh**, Dong-Ho Won ** , Young-Do Chai*,
*Department of Mathematics, Sungkyunkwan Unversity
** School of Electrical & Computer Eng., Sungkyunkwan Unversity

요 약

A. Shamir 는 비밀 정보 D를 n개의 조각으로 나눈 후, $k(k \le D)$개의 조각으로는 D를 복원할 수 있으나, $(k-1)$개 이하의 조각으로는 복원할 수 없는 비밀 분산 방식을 처음으로 소개 하였다. 이 방식은 대수학에서의 체(field) 이론에서 중요한 위치를 차지하고 있는 다항식에 대한 Lagrange interpolation 에 기반하고 있으며, 키 복구 시스템 등에 실제로 사용되고 있다. 그 후에, Tompa 등은 Shamir 의 방식이 어떤 형태의 속임수에 대하여 안전하지 않음을 보이고 이 방식을 약간 변형하여 속임수에 대해 안전한 방식을 제안하였다. Tompa 등이 제안한 방식은 안전성이 어떤 증명 되지않은 가정에 의존하지 않는다는 Shamir 방식의 특성을 보존한다. 본 고에서는 위의 방법들을 수학적으로 세밀히 분석하여 보다 자세한 증명들을 제시하고, 키 복구 시스템에 적용하는 경우에 있어 각 방식의 유용성과 단점들을 비교한다.

1. 서론

비밀분산(Secret sharing) 방식이란 비밀정보 D에 대한 부분 정보를 n 명의 참가자들에게 분배하고 필요한 경우에 허가된 참가자들의 부분 집합에 의해 본래의 비밀을 복원할 수 있도록 하는 암호학적 기술을 말한다. 이는 비밀정보의 관리뿐만 아니라 다자간 프로토콜이나 그룹 암호 방식, 키 복구 시스템[4] 등에 적용될 수 있다. 이러한 비밀분산 방식은 Blakely[1]와 Shamir[7]에 의해 각각 1979 년에 처음으로 제안되었으며, Shamir 다항식 보간법을 이용하는 (k,n) 역치(Threshold) 방식을 제안하였다. Shamir 의 방식은 각 참가자들이 보관해야 하는 부분 정보의 크기가 본래의 비밀의 크기와 거의 같고 허가되지 않은 참가자들이 비밀 D에 대해 아무런 정보도 얻을 수 없는 무조건적으로 안전한 비밀 분산 방식이라는 장점이 있다.

그 후에, Tompa 등[8]은 Shamir 의 방식이 특정 형태의 속임수(cheating)에 대하여 안전하지 않음을 보이고, 비밀을 분산하는 과정에 조건을 추가하여 원래 방식의 안전성과 효율성은 유지하면서, 그들이 제안한 속임수에 대해 안전한 방식을 제안하였다. 또한, 그들이 제안한 방식의 안전성은 RSA[6]나 Elgamal 암호 방식[3]과 같이 해결하기 어려운 정수론적 문제들의 안전성에 의존하지 않는다는 원래 방식의 특성을 보존한다. 본 논문에서는 Shamir 가 제안한 방식이 Lagrange interpolation 에 어떻게 기초를 두어서 해결되는지를 자세히 분석하고, 실제의 예를 들어 암호키를 복구하는 과정을 설명한다. 또한, Shamir 방식의 유용한 특성들을 분석하고 실제 키 복구 시스템에 적용하는 경우에 발생하는 단점들을 분석한다. 한편, 본 논문에서는 일반적인 비밀 방식 분산에서 (k,n) -역치 방식의 속임수에 대하여 거의 언급 되지 않고 있음에 주목하여, Tompa 등이 언급한 속임수에 대해 좀 더 자세한 수학적 계산을 통해 확률 값들을 분석한다.

2. 비밀 분산 방식

2.1 (k,n) - 역치 방식(Threshold Scheme)

비밀 분산 방식이란 비밀 데이터 D를 다음 조건을 만족 하는 n개의 조각으로 나누는 암호학적 기술을 말한다.

(1) $k(k \le D)$개 이상의 Share D_i 조각을 알면 쉽게 D를 구할 수 있다.
(2) $(k-1)$개 이하의 D_i 조각으로는 D를 구할 수 없다.

Lagrange interpolation 에 기반한 비밀 분산 방식을 구성하기 위해 먼저 서로 다른 x_i 값$(1 \le i \le n)$을 선택한다. 2 차 평면상의 점 $(x_1, y_1), \ldots, (x_k, y_k)$에 대해, $q(x_i) = y_i$를 만족하는 $(k-1)$ 차의 다항식 $q(x)$가 단 한 개만 존재한다. 랜덤한 $(n-1)$ 차의 다항식을 $q(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{k-1} x^{k-1}$ 라 할 때, $a_0 = D$가 되도록 한다. 이때, k개의 D_i 값들이 주어지면, Lagrange interpolation 에 의해 $q(x)$의 계수들을 얻을 수 있고, 비밀 정보 $D = q(0)$ 을 구할 수 있다 (그림 1 참조). 반면에, $(k-1)$ 개의 값으로는 유일한 D를 구할 수 없게 된다.

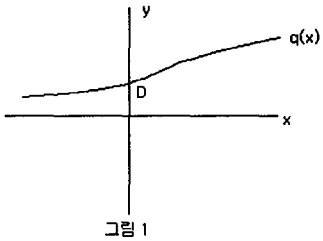

그림 1

2.2 비밀 분산 과정

Lagrange interpolation 에 대한 조건들을 만족 시키기 위해서 소수 p 상에서의 모듈라 연산을 이용한다. 여기서 $p > n$이고, Z_p는 Lagrange interpolation 이 가능한 하나의 체를 형성한다. 계수 $(a_1, \cdots, a_{n-1})$ 은 $[0, p-1]$ 의 정수 중 uniform 한 분포로 선택되고, $D_1, \cdots D_n$ 의 값들은 mod p 상에서 계산한다. 본 절에서는 Shamir 의 비밀 분산 방식의 구체적인 절차를 알아보고, 논문에서는 제시되지 않은 실제적 계산방법을 소개하고 이를 이용하여 암호키를 계산하는 과정을 설명한다.

비밀정보 D 에 대한 Share($D_1, \cdots, D_n$)을 분배 받은 n명의 참가자들을 $(p_1, p_2, \cdots, p_n)$ 라 하고, Share 를 분배하는 사람을 딜러(dealer)라 하자.

(1) 딜러는 Z_p원소 중 n개의 서로 다른 0 이 아닌 원소 x_i 를 선택한다. 단, $1 \le i \le n$ 이고 $p \ge n+1$ 인 소수이다. 딜러는 x_i 값을 p_i 에게 분배하고, x_i 값들을 공개한다.
(2) 딜러는 Z_p의 $(k-1)$ 개의 원소들을 랜덤하게 비밀스럽게 선택하여, $y_i = q(x_i)$ 를 계산한다. 단, $q(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{k-1} x^{k-1}$, mod p 이다.
(3) 딜러는 분산 정보 y_i 를 각 참가자 p_i 에게 분배한다$(1 \le i \le n)$.

$q(x)$는 최대 $(k-1)$ 차의 다항식 이므로 k 개의 독립된 선형 방정식이 주어지면 미지수 $a_0, \cdots a_{k-1}$을 유일하게 풀어 낼 수 있고, 이때 $a_0 = D$ 이므로 비밀 정보를 복구할 수 있게 된다.

2.3 비밀 복원 과정

모든 k 개의 (x_i, y_i) 들을 알고 있다고 하더라도, 주어진 연립 일차방정식을 풀어서 키를 구하는 것은 번거롭고 복잡하다. 그러나 Lagrange interpolation 식에 의하여 다음의 식을 얻을 수 있다.

$$q(x) = \sum_{j=1}^{k} y_{i_j} \prod_{\substack{1 \le l \le k \\ l \ne j}} x - x_{i_l} / x_{i_j} - x_{i_l}$$

$q(0) = D$ 이므로 $D = \sum_{j=1}^{k} b_j y_{i_j}$ 이고, 다음 식이 성립한다.

$$b_j = \prod_{1 \le l \le k} x_{i_l} / x_{i_l} - x_{i_j}$$

(예제)
$p = 17$, $k = 3$, $n = 5$이고 $x_i = i, 1 \le i \le 5$라 하자. 참가자 p_1, p_3, p_5가 각각 비밀 D에 대한 분산 정보로 8, 10, 11 로 분배 받은 경우에 비밀키 D를 복구하는 과정은 다음과 같다.

(풀이)
앞에서 설명한 식에 의해, b_1, b_2, b_3를 구할 수 있다.

$$b_1 = x_3 x_5 / (x_1 - x_3)(x_1 - x_5) \mod 17$$
$$= 3 \cdot 5 / 2 \cdot 4 = 4 \mod 17$$

이와 같은 방법으로 $b_2 = 3$, $b_3 = 11$를 구할 수 있고, 주어진 8,10,11 을 공식에 대입하면 다음과 같이 D를 구할 수 있다.

$$D = 4 \times 8 + 3 \times 10 + 11 \times 11 = 13 \mod 17$$

2.1 (k,n)-역치 방식의 특징

앞 절에서 설명한 (k,n)-역치 방식은 다음과 같은 장점을 갖는다.

(1) 각 D_i의 크기는 D의 크기를 넘지않는다.
(2) K가 고정되어 있을 경우, 다른 D_i에 영향을 주지않고, D_i를 자유롭게 첨삭할 수 있다.
(3) 새로운 다항식 $q(x)$를 이용하여 D를 변경하지 않고도 D_i조각을 변경할 수 있다. 만일 이전의 분산 정보들이 노출되더라도 D를 구하는데 도움이 되지 않으므로 빈번하게 변경할수록 안전성을 향상시킬 수 있다.
(4) 비밀 분산에 참여하는 참가자에 따라 서로 다른 권한을 갖는 계층적 방식을 만들 수 있다.

그러나 실제로 키 복구 시스템에 이 방식을 적용할 경우에는 다음과 같은 문제점들이 발생하게 된다.

(1) 딜러가 의도적으로 각 참가자들에게 분배하는 정보를 정당하지 않게 생성하는 경우에도 참가자들은 분배 정보의 정당성을 확인할 수 없게 된다. 따라서, 필요한 경우에도 본래의 정보를 복원할 수 없게 된다.
(2) 각 사용자의 암호키에 대해 하나의 분배 정보를 보관해야 하므로 기관이 관리해야 하는 Share 의 수가 사용자의 수에 비례하여 증가하게 된다.
(3) 본래의 비밀정보가 한번 복원된 이후에는 Share 를 재사용할 수 없다.

최근 들어, 비밀 분산 방식을 실제 키복구 시스템에 적용하기 위해 위의 단점들을 해결할 수 있는 여러 방식들이 제안되고 있다.

3. 속임수를 가지는 비밀 분산 방식

3.1 비밀 분산 방식에서의 속임수

Shamir 의 방식에 대해 살펴보면, 각 참가자는 n,k와 비밀키가 가능한 집합을 알고 있다.

[Inputs]
- 음이 아닌 정수들 n,s와 $k \le n$
- 비밀키 $D \in \{0,1,\cdots,s-1\}$

[Problem]
D를 다음 조건을 만족하는 $D_1,\cdots,D_n$으로 나눈다.
(a) 어떤 k개 이상의 조각을 알면 D를 재구성할 수 있다.
(b) 어떤 $(k-1)$개 이하의 조각을 알면 D에 대한 어떤 정보도 알 수 없다.

다항식에 대한 Lagrange interpolation 을 기초로 하여 비밀 데이터 D를 찾아내는 풀이는 앞에서 자세히 언급 하였다. 본 절에서는 이 방식에 대해 가능한 속임수에 대해 설명한다.

본 절에서 설명하는 비밀 분상 방식에 대한 속임수는 최대 $(k-1)$명의 참가자가 k번째 사람이 올바른 D를 복원하는 것을 방해하고 k번째 사람은 이를 검출하지 못하는 것을 말한다. 따라서 Shamir 의 방식은 신뢰할 수 없는 참가자들에 대해서도 고려해야 하므로 위에서 설명한 (a), (b) 외에 다음과 같은 특징 (c)가 추가되어야 한다.

(c) $k-1$명의 참가자 $i_1,\cdots,i_{k-1}$가 참가자 i_k를 속일 수 있는 $D_{i_1}',D_{i_2}',\cdots,D_{i_{k-1}}'$들을 위조할 수 있는 확률$(\varepsilon > 0)$은 작아야 한다.

여기서 i_k를 속인다는 의미는 $D_{i_1}',D_{i_2}',\cdots,D_{i_{k-1}}'$와 D_{i_k}로부터 다시 만들어진 비밀 D'가 합법적$(D' \in \{0,1,\cdots,s-1\})$이지만 본래의 D와 다른 옳지 않은 값이라는 것이다. D_{i_k} 한 점에서만 만나면서 $D' \in \{0,1,\cdots,s-1\}$인 다항식이 항상 존재하므로, (c)에서 $\varepsilon = 0$인 경우는 불가능하다. (그림 2 참조).

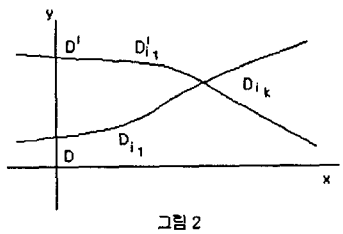

그림 2

3.1.1. Shamir 방식에 대한 속임수

Shamir 의 방식에서 소수 p $(D_i \in Z_p)$에 대해, $p = s$이면 모든 D'들은 합법적이므로 i_k는 어떠한 경우에도 속임수를 감지할 수 없게 된다. 따라서, s보다 p를 훨씬 크게 잡아야 한다.

또한, 다음의 방법을 이용하여 한 명의 참가자는 높은 확률로 $(k-1)$명을 속일 수 있다. 다른 참가자들을 속이려는 참가자 i_1(cheater)은 Lagrange interpolation을 이용하여 $\Delta(0) = -1$, $\Delta(i_2) = \cdots = \Delta(i_k) = 0$ 인 $(k-1)$차의 다항식 $\Delta(x)$를 생성한다. 그리고 비밀 복원과정에서 분배받은 Share D_{i_1} 대신에 $D_{i_1} + \Delta(i_1)$을 다른 참가자들에게 알려준다. Lagrange interpolation에 의해, 다른 참가자들은 $q(x) + \Delta(x)$를 재구성하게 되고, $q(0) + \Delta(0) = D - 1$을 구할 수 있다. $D \neq 0$이면 속임수를 감지할 수 없다. 만약 $D = 0$이면 D'은

-1 이므로 합법적이지 않다. $i_2,\cdots,i_k$ 는 $D-1$을 올바른 값으로 알게 되고, i_1 만이 $\Delta(i_1)$ 을 연산하여 올바른 D를 알게 된다.

위와 같은 속임수에 대한 한가지 해결책은 딜러가 각각의 D_i 조각에 위조할 수 없는 서명을 하는 것이다. 그러나 Tompa 등이 제안한 방식은 다음의 두 가지 장점을 갖는다.

(1) 안전성이 소인수분해나 이산대수의 어려움과 같은 증명되지 않은 가정에 의존하지 않는다. 따라서, Cheater 가 무한한 계산 능력을 갖고 있는 경우에도 안전하다.
(2) 별도의 디지털 서명 방식을 추가할 필요가 없다.

3.2 Tompa 등이 제안한 개선 방식

본 절에서는 Shamir 의 방식을 앞에서 설명한 속임수에 대해 안전하도록 변형한 Tompa 등의 방식에 대해 설명한다.

(1) 소수 $p>\max((s-1)(k-1)/\varepsilon+k,n)$를 선택한다.
(2) Z_p 상에서 랜덤하게 $a_1,\cdots,a_{k-1}$를 선택한다.
(3) 다항식 $q(x)=D+a_1x+\cdots+a_{k-1}x^{k-1}$를 구성한다.
(4) $\{1,2,\cdots,p-1\}$ 로부터 n 개의 서로 다른 원소 $(x_1,\cdots,x_n)$ 을 랜덤하게 선택하고, $d_i=q(x_i)$ 를 계산한다. $D_i=(x_i,d_i)$라 한다.

개선된 방식에서 다른 참가자들을 속일 확률에 대해서 살펴보자. 참가자 $i_1,\cdots,i_{k-1}$ 이 각각 (x_{i_1}',d_{i_1}'), $\cdots$, $(x_{i_{k-1}}',d_{i_{k-1}}')$ 을 위조하여 참가자 i_k 에게 전송하고 $D'(\in\{0,1,\cdots,s-1\})$을 만든다. 이것으로 $(0,D')$과 위조된 $(k-1)$ 개의 점들로 이루어진 $q_{D'}(x)$를 구성한다. i_k 에 대한 cheating 이 성공하기 위해서는 i_k 가 자신이 잘못된 비밀 값을 복원했다는 사실을 감지할 수 없도록 해야 하므로 $q_{D'}(x_{i_k})=q(x_{i_k})$이고 $D'\neq D$를 만족해야 한다. $D'\neq D$라면 $q_{D'}(x)$와 $q(x)$는 최대 $(k-1)$ 개의 점들에서 만날 것이고, 만약 k 개의 점들에서 만난다면 $D'=D$가 된다.

x_{i_k} 는 $\{1,\cdots,p-1\}-\{x_{i_1},\cdots,x_{i_{k-1}}\}$에서 선택된 랜덤 수이므로 x_{i_k} 의 수는 $(p-1)-(k-1)=(p-k)$ 이다. 또한, 두 식 $q_{D'}(x)$ 와 $q(x)$ 가 만나는 점의 수는 최대 $(k-1)$ 개이므로 $D'\neq D$인 각각의 다항식 $q_{D'}(x)$에 대하여, $q_{D'}(x_{i_k})=q(x_{i_k})$ 일 확률은 최대 $(k-1)/(p-k)$이다.

또한, 합법적이지만 옳지않은 비밀 값이 될 D' 의 가지수는 $(s-1)$ 개이므로 위조된 값에 상응하는 다항식이 가능한 가지수는 $(s-1)$ 개이다. 이러한 다항식들 중에서 어느 다항식도 $(k-1)/(p-k)$ 의 확률로 i_k 를 속일 수 있으므로 확률은 참가자 i_k 를 속일수 있는 확률은 $(s-1)(k-1)/(p-k)$가 된다. 따라서, (c)에 의해 $(s-1)(k-1)/(p-k)<\varepsilon$ 이어야 하므로 $p>(s-1)(k-1)/\varepsilon+k$ 인 소수를 선택해야 한다.

이 방식에서는 cheating 이 발생하는 경우 참가자 i_k 를 속일 확률은 매우 작지만 이미 참가자들의 Share 가 공개된다는 단점이 있다. Tompa 등은 논문에서 이러한 문제점을 해결할 수 있는 또 다른 방식에 대해서도 언급하고 있다. 그러나, 그 방식은 실제로 사용하는데 있어 각 참가자들이 유지해야 하는 비밀 정보의 양이나 계산량 면에서 비 효율적이므로 본 논문에서는 생략하기로 한다.

4. 결론 및 향후 연구 계획

본 논문에서는 Shamir 가 제안한 비밀 분산 방식의 비밀 분산 및 복원 과정에 대해 자세히 살펴보고, Shamir 의 방식이 갖는 장점과 이를 실제 키 복구 시스템에 적용하였을 경우에 발생할 수 있는 문제점에 대해 분석하였다. 또한, Tompa 등이 언급한 Shamir 의 방식에 대한 특정 형태의 속임수에 대해 살펴보고 그들의 개선 방식을 수학적 계산을 통해 분석해 보았다. 이러한 비밀 분산 방식을 실제 키 복구 시스템에 적용하기 위해서는 앞 절에서 언급한 몇 가지 문제점을 해결할 수 있는 비밀 분산 방식에 대한 연구가 계속되어야 할 것이다.

참고문헌

[1] G. R. Blakely, "Safeguarding cryptographic key", In proceeding of AFIPS National Computer Conference, pp. 313-317, 1979
[2] W. Diffie and M. Hellman, "New Directions in Cryptography", IEEE Trransaction on Information Theory IT-22, pp. 644-654, 1976.
[3] T. ElGamal, "A Public key cryptosystem and a signature scheme based on discrete logarithms", IEEE Transactions on Information Theory IT-31, pp. 469-472, 1985
[4] FIPS PUB-185 Escrowed Encryption Standard, 1993
[5] W. Hungerford, Algebra,
[6] R. Rivest, A. Shamir, and L. Adleman, "A method for Obtaining Digital signature and Public key Cryptosystems", Communication of the ACM, pp. 120-128, FEB. 1978.
[7] A. Shamir, How to Share a secret, Comm. ACM, 22(1979), 612-613
[8] M. Tompa, H. Woll, How to Share a secret with cheaters, J. Cryp. 1 (1988), 133-138