情報保護學會論文誌
第12券 第3號, 2002. 6

확장된 Interactive Hashing 프로토콜

홍 도 원*, 장 구 영*, 류 희 수*

Extended Interactive Hashing Protocol

Dowon Hong*, Ku-Young Chang*, Heuisu Ryu*

요 약

Interactive hashing은 Naor, Ostrovsky, Venkatesan, Yung[1]에 의해 소개된 프로토콜로 주어진 스트링 크기 t비트에 대해 $t-1$번의 라운드 복잡도(round complexity)와 t^2-1 비트의 전송 복잡도(communication complexity)를 가진다. 본 논문은 t를 나누는 m에 대해서 $t/m-1$번의 라운드 복잡도와 $t^2/m-m$ 비트의 전송 복잡도를 갖는 NOVY 프로토콜보다 효율적으로 확장된 Interactive hashing 프로토콜을 제안하고 그 안전성을 증명한다.

ABSTRACT

Interactive hashing is a protocol introduced by Naor, Ostrovsk, Venkatesan, Yung[1] with $t-1$ round complexity and t^2-1 bits communication complexity for given t bits string. In this paper, we propose more efficiently extended interactive hashing protocol with $t/m-1$ round complexity and $t^2/m-m$ bits communication complexity than NOVY protocol when m is a divisor of t, and prove the security of this.

Keyword : *Interactive hashing protocol, Universal hash family*

Ⅰ. 서 론

Interactive hashing은 Naor, Ostrovsky, Venkatesan, Yung[1]에 의해 소개된 프로토콜로 information-theoretic security를 가진 Commitment, Oblivious Transfer, Zero-knowledge 프로토콜에 유용하게 사용되고 있다.[1,2,3,4,5] Cachin, Crépeau, Marcil[6]은 interactive hashing 기법의 새롭고 유용한 분석을 내어놓았으며, 이 기법은 bounded storage 모델 하에서 Oblivious Transfer 프로토톨의 unconditional security를 얻기위하여 [7]에서도 최근 사용되고 있다.

Interactive hashing은 입력 값을 가지지 않는 수신자 Alice와 입력 스트링 χ를 가진 전송자 Bob 사이의 프로토콜로 두 개의 스트링을 공유하는 방법을 제공한다. 한 개의 스트링은 Bob에 의해 선택된 입력값 χ이고, 다른 스트링은 Bob에 의해 영향을 받지 않는 임의의 랜덤값으로 선택되어야 한다. 하지만, Alice는 두 개 중 어느 것이 χ인지 알 수 없어야 한다. 현재까지 사용된 모든 interactive hashing 프로토콜은 매우 큰 라운드 복잡도를 가진 NOVY 프로코콜[1]을 근본적으로 이용해 왔다. NOVY 프로토콜은 전달되는 스트링 크기 t 비트에 대해 $t-1$번의 라운드 복잡도와 t^2-1 비트의 전송 복잡도를 가지기 때문에, 이 기법에 기반한 프로토콜은 결과적으로 과도한 전송 복잡도를 가질 수 밖에 없다.

이 논문에서는 NOVY 프로토콜에 비해 효율적으로 확장된 interactive hashing 프로토콜을 제안하고 그 안전성을 증명한다. 즉, t를 나누는 m에 대해 NOVY 프로토콜보다 더 작은 $\frac{t}{m}-1$번의 라

* 한국전자통신연구원 정보보호연구본부 ({dwhong, jang1090, hsryu}@etri.re.kr)

운드 복잡도와 $\frac{t^2}{m}-m$ 비트의 전송 복잡도를 가지는 프로토콜로 확장할 수 있으며, 게다가 전송자와 수신자 사이에 전송자의 입력 스트링을 포함한 두 개 또는 그 이상의 스트링을 공유하는 방법을 제공한다.

프로토콜을 구성하기 위하여 NOVY 프로토콜에서 사용된 universal hash family 대신 새로운 universal hash family를 도입한다. Universal hash family는 interactive hashing 프로토콜을 구성하는데 가장 근본적인 역할을 담당한다.

II. Universal hash family

Universal hasing 기법은 1979년 Carter와 Wegman[9]에 의해 소개된 이후 전산학과 암호학의 많은 응용 분야에서 사용되고 있다.[8,10]

[정의 1]

H hash family를 $X\to Y$로의 함수 집합 $\mathcal{T}$에서 $|H|$ 개의 부분집합이라 하자. H hash family가 다음 조건을 만족하면 universal이라 한다: 모든 서로 다른 $x_1, x_2\in X$에 대해, 기껏해야 $\frac{|H|}{|Y|}$ 개의 $h\in H$ 만이 $h(x_1)=h(x_2)$를 만족한다.

이제 자연수 t, m에 대해 m이 t를 나누는 경우에 $GF(2)^t\to GF(2)^m$로의 universal hash family를 정의하자.

$f(x)$의 차수(degree)는 m이고, $GF(2)$위에서 기약(irreducible)이라고 하자. 그러면, 유한체 $GF(2^m)=GF(2)[x]/(f(x))$는 $\{\sum_{i=0}^{m-1}a_ix^i \mid a_i\in GF(2)\}$으로 표현할 수 있다. $GF(2)^m$에서 $GF(2^m)$로의 전단사함수 $\phi: (a_{m-1},\cdots,a_0)\mapsto a_{m-1}x^{m-1}+\cdots+a_0$를 고려하자. 임의의 t에 대해 $t=lm$, $m<t$라고 하자. 그러면, $GF(2)^t=(GF(2)^m)^l=\{(A_{l-1},\cdots,A_1,A_0)\mid A_i\in GF(2)^m, 0\le i\le l-1\}$로 표현될 수 있다. 이제 $S=(GF(2)^m)^l$라고 하고 $GF(2)^t$를 S로 간주하자. S에서 $GF(2)^m$으로의 universal hash family를 정의하기 위해, 위에서 정의한 함수 ϕ를 이용하여 임의의 $\zeta=(\zeta_{l-1},\cdots,\zeta_1,\zeta_0)\in S$에 대하여 다음과 같이 hash 함수를 정의한다.

$$h_\zeta : S\to GF(2)^m$$
$$(A_{l-1},\cdots,A_0)\mapsto \phi^{-1}\Big(\sum_{i=0}^{l-1}\phi(A_i)\cdot\phi(\zeta_i)\Big) \quad (1)$$

이러한 S에서 $GF(2)^m$로의 hash 함수들의 집합

$$H=\{h_\zeta: \zeta=(\zeta_{l-1},\cdots,\zeta_1,\zeta_0)\in S\}$$

을 고려하자.

[보조정리 1]

H는 universal hash family이다.

(증명)

S에 속해 있는 서로 다른 두 원소 $x=(x_{l-1},\cdots,x_0)$, $y=(y_{l-1},\cdots,y_0)$에 대해, 우리는 $h_\zeta(x)=h_\zeta(y)$를 만족하는 $\zeta=(\zeta_{l-1},\cdots,\zeta_0)\in S$의 개수를 셀 필요가 있다. $x\neq y$이기 때문에 $x_{i_0}\neq y_{i_0}\in GF(2)^m$를 만족하는 어떤 $i_0\in\{0,\cdots,l-1\}$가 존재한다. 그러면 임의의 $\zeta\in S$에 대해

$$\begin{aligned}
&h_\zeta(x)=h_\zeta(y)\\
&\Leftrightarrow \phi(\zeta_{i_0})(\phi(y_{i_0})-\phi(x_{i_0}))\\
&\quad+\sum_{i\neq i_0}\phi(\zeta_i)(\phi(y_i)-\phi(x_i))=0\\
&\Leftrightarrow \phi(\zeta_{i_0})(\phi(y_{i_0})-\phi(x_{i_0}))\\
&\quad=\sum_{i\neq i_0}\phi(\zeta_i)(\phi(y_i)-\phi(x_i))\in GF(2^m)\\
&\Leftrightarrow \phi(\zeta_{i_0})=\sum_{i\neq i_0}\phi(\zeta_i)(\phi(y_i)-\phi(x_i))\\
&\quad\cdot(\phi(y_{i_0})-\phi(x_{i_0}))^{-1} \qquad (2)
\end{aligned}$$

가 성립한다. ϕ는 전단사함수이기 때문에, 각각의 $i\neq i_0$를 만족하는 ζ_i의 선택에 따라 식 (2)는 유일한 하나의 해 ζ_{i_0}를 가진다. $i\neq i_0$인 ζ_i의 개수는 $l-1$개이고, 각각이 $\zeta_i\in GF(2)^m$이기 때문에 $h_\zeta(x)=h_\zeta(y)$를 만족하는 해쉬함수 $h_\zeta\in H$의 수는 정확히 $2^{m(l-1)}=|H|/2^m$이다. 그러므로, H는 universal hash family이다. ■

위에서 정의한 universal hash family H는 다음과 같은 성질들을 가진다.

[보조정리 2]

H는 위에서 정의한 universal hash family라고 하자. 임의의 0이 아닌 두 원소 $x\neq y\in S$와 임의의 $b\in GF(2)^m$에 대하여

$$T_b=\{h\in H\mid h(x)=b, h(y)=b\}$$

라고 하자. 그러면, $|T_b| = |H|/2^{2m}$ 이다.

(증명)

임의의 0이 아닌 두 원소 $x=(x_{l-1},\cdots,x_0)$, $y=(y_{l-1},\cdots,y_0)\in S$에 대해 $x\neq y$라고 하자. 정의에 의해 $T_b=\{\zeta\in S \mid h_\zeta(x)=b, h_\zeta(y)=b\}$이다. $x\neq y$이기 때문에 $x_j\neq 0, y_k\neq 0\in GF(2)^m$을 만족하는 서로 다른 두 index $j,k\in\{0,\cdots,l-1\}$가 존재한다. 그러면 임의의 $\zeta=(\zeta_{l-1},\cdots,\zeta_0)\in S$에 대해

$$
\begin{aligned}
&h_\zeta(x)=b\\
&\quad\Leftrightarrow \phi(x_j)\phi(\zeta_j)+\sum_{i\neq j}\phi(x_i)\phi(\zeta_i)=\phi(b)\\
&\quad\Leftrightarrow \phi(\zeta_j)=\Big(\sum_{i\neq j}\phi(x_i)\phi(\zeta_i)+\phi(b)\Big)\\
&\qquad\cdot\phi(x_j)^{-1},\\
&h_\zeta(y)=b\\
&\quad\Leftrightarrow \phi(\zeta_k)=\Big(\sum_{i\neq k}\phi(y_i)\phi(\zeta_i)+\phi(b)\Big)\\
&\qquad\cdot\phi(y_k)^{-1}
\end{aligned}
$$

가 성립한다. 따라서 보조정리 1의 증명과 비슷한 방법에 의해 $|T_b|=2^{m(l-2)}=|H|/2^{2m}$이다. ■

[보조정리 3]

H는 위에서 정의한 universal hash family라고 하자. 그러면, 임의의 0이 아닌 $s\in S$와 임의의 $b\in GF(2)^m$에 대하여

$$|\{h\in H \mid h(s)=b\}|=|H|/2^m$$

이다.

(증명)

보조정리 2와 유사한 방법으로 증명할 수 있다. ■

Ⅲ. Interactive Hashing 프로토콜

Naor, Ostrovsky, Venkatesan, Yung[1]에 의해 소개된 interactive hasing은 두 당사자 Alice와 Bob 사이의 프로토콜이다. 전송자 Bob은 t비트 스트링 $\chi\in T\subset GF(2)^t$를 가지고 있고, χ와 T는 수신자 Alice에게 비밀이다. 여기에서 $|T|\leq 2^{t-k}$이고 k는 security 변수이다. 프로토콜이 종료되면 Alice와 Bob은 χ를 포함한 두 개의 스트링을 서로 공유하게 되지만, Alice는 두 개 중 어느 것이 χ인지 알 수 없다. 또한 Bob도 무시할 만큼 작은 확률 $\nu(k)$를 제외하고는 두 스트링 모두 T의 원소가 되게 프로토콜을 구성할 수는 없다.

다음의 interactive hashing 프로토콜은 [1]에서 제안된 것이다.

[NOVY 프로토콜]

Alice는 $t-1$개의 일차 독립(linearly independent)인 벡터들 $a_1,\cdots,a_{t-1}\in GF(2)^t$를 임의로 선택한다. 프로토콜은 $t-1$ 라운드 수행되며, 각 라운드 i는 다음 단계들로 구성된다.

1. Alice는 a_i를 Bob에게 전송한다.
2. Bob은 $b_i=a_i\cdot\chi$를 계산하고, b_i를 Alice에게 보낸다.

프로토콜 종료 후 Alice와 Bob은 $GF(2)$상에서 같은 일차 방정식들 $a_i\cdot\chi=b_i$, $i=1,\cdots,t-1$를 가지게 된다. 벡터들 $a_i\in GF(2)$, $i=1,\cdots,t-1$은 일차독립이므로 시스템은 정확하게 두 개의 t비트 스트링 χ_0, χ_1을 해로 가지며, 둘 중 하나는 χ이다. Alice는 어떤 해가 χ인지 모르는 것은 information-theoretic 관점에서 분명하다. 또한 두 해 모두 T에 속하도록 Bob이 Alice를 속일 수 없다는 것이 [6]에서 증명되었다.

NOVY 프로토콜은 전달되는 스트링 크기 t 비트에 대해 $t-1$번의 라운드 복잡도와 t^2-1 비트의 전송 복잡도를 가지기 때문에, 이 기법에 기반한 프로토콜은 결과적으로 과도한 전송 복잡도를 가질 수 밖에 없다.

Ⅳ. 확장된 Interactive Hashing 프로토콜

이 절에서 우리는 입력 값을 가지지 않는 수신자 Alice와 입력 스트링 χ를 가진 전송자 Bob 사이의 프로토콜로 두 개 이상의 스트링을 공유하는 방법을 제안한다. 전송자 Bob은 t비트 스트링 $\chi\in T\subset GF(2)^t$를 가지고 있고, χ와 T는 수신자 Alice에게 비밀이다. 여기에서 $|T|\leq 2^{t-k}$이다. t를 어떤 자연수 l, m에 대해 $t=lm$, $m<t$로 표현하자. 프로토콜은 다음과 같은 조건들을 만족해야 한다:

1. Bob은 비밀 입력 값인 t비트 스트링 χ를 포함한 2^m개의 t비트 스트링들을 Alice가 구할 수 있도록 χ를 전송하지만, Alice는 2^m개 중 어떤 것이 χ인지 알 수 없어야 한다.
2. Bob은 무시할 만큼 작은 확률 $\nu(k)$를 제외하고는 2^m개의 스트링 중에서 서로 다른 어떤 두 개도 모두 T의 원소가 되게 할 수는 없다.

$t = lm$인 경우 $GF(2)^t$를 $(GF(2)^m)^l$로 간주하고, $S = (GF(2)^m)^l$라고 하자. Bob은 t비트 비밀 스트링 $\chi \in S$를 선택한 뒤에, $\chi = (\chi_{l-1}, \cdots, \chi_1, \chi_0)$로 표현한다. 여기에서 $\chi_i \in GF(2)^m$, $0 \le i \le l-1$이다. II절에서 정의한 S에서 $GF(2)^m$로의 hash 함수들의 universal family $H = \{h_\zeta : \zeta = (\zeta_{l-1}, \cdots, \zeta_1, \zeta_0) \in S\}$를 고려하자. 여기에서 h_ζ는 식 (1)과 같이 정의된다.

우리가 제안하는 프로토콜은 다음과 같다.

[Protocol]

프로토콜은 $t/m-1$ 라운드를 진행한다. 각각의 $i = 1, \cdots, t/m-1$에 대하여, 라운드 i에서는 다음과 같은 과정을 실행한다.

1. Alice는 균일한 분포(uniform distribution)를 가진 H에서 hash 함수 h_i를 선택한다. $a_i \in GF(2)^t$를 $h_i = h_{a_i}$되는 벡터라고 하자. 만약 a_i가 $a_1, \cdots, a_{i-1}$과 일차 종속이면 Alice는 일차 독립인 a_i가 선택될 때까지 이 단계를 반복한다. Bob에게 a_i를 전송한다.
2. $a_i = (a_i^{(t/m-1)}, \cdots, a_i^{(0)}) \in S$, $a_i^{(j)} \in GF(2)^m$, $0 \le j \le t/m-1$라고 하자. Bob은 받은 a_i를 가지고 m 비트값 $b_i = h_{a_i}(\chi) = \phi^{-1}(\sum_{j=0}^{t/m-1} \phi(a_i^{(j)}) \cdot \phi(\chi_j))$를 계산하여 Alice에게 보낸다.

프로토콜이 종료된 $t/m-1$ 라운드 후에, Alice와 Bob은 모두 $GF(2)^m$ 상에서 χ에 의해 만족되는 $t/m-1$개의 일차 방정식을 가진다. 시스템은 정확하게 2^m개의 t비트 스트링 $\chi_0, \cdots, \chi_{2^m-1}$을 해로 가지며, 이 해들 중 하나는 χ이다. 우리는 이 기법을 "확장된 interactive hashing"이라 부른다. $m=1$인 경우 우리의 프로토콜은 interactive hashing과 같다.

(참고)

단계 2가 종료된 후에 Bob이 계산된 2^m개의 t비트 스트링의 해 집합에서 χ를 포함하여 임의로 2개를 선택하여 Alice에게 보내는 단계를 수행하면 확장된 interactive hashing 프로토콜은 interactive hasing 프로토콜을 수행할 수 있다.

Information-theoretic 관점에서 Alice는 어떤 해가 χ인지 알 수 없으므로 확장된 interactive hashing의 조건 1은 만족된다. 이제 정직하지 못한 전송자 Bob에 대한 안전성 조건 2가 만족하는 지를 살펴보자. 제안된 프로토콜에서 Bob이 고정된 집합 T의 서로 다른 두 원소 s_1, s_2에 대해 같은 방법으로 Alice의 query들에 응답할 수 있다면, Bob은 Alice를 속일 수 있다. 정리 1에서 우리는 $m < (k-2)/6$이 만족하면 T의 크기가 $|GF(2)^t| = 2^t$와 거의 같을 때에만 Bob이 속일 수 있다는 것을 보일 것이다. 이를 위해서는 다음과 같은 몇 개의 보조 정리가 필요하다.

다음 보조 정리는 프로토콜의 각 라운드에 대해 T가 어느 정도 큰 경우, 앞의 라운드에 비해 T의 크기가 약 2^{-m}배 줄어든다는 것을 보인다. 이런 식의 증명은 [6]에서 interactive hashing 프로토콜의 안전성을 보이기 위해 처음으로 사용되었다.

[보조정리 4]

T는 $GF(2)^t$의 부분집합이고 $0 < \alpha < 1$에 대해 $|T| = 2^{\alpha t}$라 하고, p는 $p \le \alpha t/3$을 만족하는 자연수라 하자. m은 t를 나누는 자연수라 하자. H를 위에서 정의한 $GF(2)^t$에서 $GF(2)^m$으로 가는 hash 함수들의 universal family라 하자. U를 H 위에서 정의된 균일 분포(uniform distribution)를 가지는 확률 변수(random variable)라 하자. 그러면 임의의 $b \in GF(2)^m$에 대하여, U는 적어도 $1-2^{-p}$의 확률로 다음을 만족하는 hash 함수 h를 가진다:

$$|\{s \in T | h(s) = b\}| < \left(\frac{1}{2^m} + \frac{1}{2^{p+m/2}} + \frac{1}{2^{3p}} \right) |T| \quad (3)$$

(증명)

임의의 $s \in T$와 $b \in GF(2)^m$에 대하여 다음과 같은 확률 변수들

$$X_{(b,s)} = \begin{cases} 1 & \text{if } U(s) = b \\ 0 & \textit{otherwise} \end{cases}$$

와 그들의 합 $X_b = \sum_{s\in T} X_{(b,s)} = |\{s\in T: U(s)=b\}|$ 를 고려하자. 우리는 이 보조 정리를 증명하기 위해서 임의의 $b\in GF(2)^m$에 대하여, X_b가 적어도 $1-2^{-p}$의 확률로

$$x < \left(\frac{1}{2^m} + \frac{1}{2^{p+m/2}} + \frac{1}{2^{3p}}\right)|T|$$

를 만족하는 값 x를 가짐을 보이면 된다.

Case 1 : $b\neq 0\in GF(2)^m$

$s\neq 0$이면 보조정리 3에 의해 확률 변수들 $X_{(b,s)}$와 $X^2_{(b,s)}$의 기댓값들은

$$E[X_{(b,s)}] = E[X^2_{(b,s)}] = 1\cdot\frac{|\{h\in H|\ h(s)=b\}|}{|H|} = \frac{1}{2^m}$$

이고, hash 함수의 정의에 의해 $X_{(b,0)} = X^2_{(b,0)} = 0$이다. 따라서 $E[X_b] = \frac{|T|-1}{2^m}$ 이다. 확률 변수 X_b의 정의에 의하면

$E[X_b^2] = E[\sum_{s\in T} X^2_{(b,s)} + 2\sum_{s_i<s_j\in T} X_{(b,s_i)}X_{(b,s_j)}]$

이다. $b\neq 0$이므로 $X_{(b,0)}=0$이다. 이 사실과 보조정리 2를 이용하여 다음 식을 얻는다.

$$\begin{aligned} &E[\sum_{s_i<s_j\in T} X_{(b,s_i)}X_{(b,s_j)}] \\ &= \sum_{0<s_i<s_j\in T} E[X_{(b,s_i)}X_{(b,s_j)}] \\ &= \sum_{0<s_i<s_j\in T} \frac{\{h\in H:\ h(s_i)=h(s_j)=b\}}{|H|} \\ &< \frac{(|T|-1)^2}{2}\cdot\left(\frac{1}{2^m}\right)^2 \end{aligned}$$

그러므로 $E[X_b^2] < \frac{|T|-1}{2^m} + \left(\frac{|T|-1}{2^m}\right)^2$이고, X_b의 분산은 다음 수식을 만족한다.

$$Var[X_b] = E[X_b^2] - (E[X_b])^2 < \frac{|T|-1}{2^m}$$

Chebychev 부등식에 의해, 임의의 $b\neq 0\in GF(2)^m$와 $\sigma>0$에 대해 다음 식을 얻는다.

$$\Pr\left[\left|X_b - \frac{|T|-1}{2^m}\right| \geq \sigma\right] < \frac{|T|-1}{2^m\sigma^2}$$

$\sigma=\sqrt{2^p(|T|-1)/2^m}$으로 두면, 다음 식을 얻는다.

$$\Pr\left[\left|X_b - \frac{|T|-1}{2^m}\right| \geq 2^{\frac{p+at-m}{2}}\right] < 2^{-p}$$

따라서, $p\leq at/3$이면, 적어도 $1-2^{-p}$의 확률로

$$\begin{aligned} X_b &< \left(\frac{1}{2^m} + 2^{\frac{p+at-m}{2}-at}\right)|T| \\ &< \left(\frac{1}{2^m} + \frac{1}{2^{p+m/2}}\right)|T| \end{aligned}$$

성립하고, 식 (3)을 만족한다.

Case 2 : $b=0\in GF(2)^m$

$X_{(0,0)}=1$이라는 사실과 보조정리 2와 3을 이용하여 Case 1와 비슷한 방법으로 $E[X_0] = \frac{|T|-1}{2^m}+1$과 $E[X_0^2] < \frac{3(|T|-1)}{2^m}+1+\left(\frac{|T|-1}{2^m}\right)^2$를 구한다. 그러면 $Var[X_0] < \frac{|T|-1}{2^m}$이고, Chebychev 부등식에 의해, 임의의 $\sigma>0$에 대해

$$\Pr\left[\left|X_0 - \left(\frac{|T|-1}{2^m}+1\right)\right| \geq \sigma\right] < \frac{|T|-1}{2^m\sigma^2}$$

이 성립한다. $\sigma=\sqrt{2^p(|T|-1)/2^m}$으로 두면, $p\leq at/3$임으로 적어도 $1-2^{-p}$의 확률로

$$X_0 < \left(\frac{1}{2^m} + \frac{1}{2^{p+m/2}} + \frac{1}{2^{3p}}\right)|T|$$

가 성립하고, 보조정리가 증명된다. ■

다음 보조정리는 [6]에서 증명되었다.

[보조정리 5][6]

T는 $GF(2)^t$의 부분집합이고 $0<\alpha<1$에 대해 $|T|=2^{\alpha t}$이다. p, q는 $2\alpha t < mq-p$를 만족하는 자연수이고 p, $mq\leq t$라 하자. 여기서 m은 t를 나누는 자연수이다. H를 $GF(2)^t$에서 $GF(2)^{mq}$으로 가는 hash 함수들의 universal family라 하자. U를 H 위에서 정의된 균일 분포를 가진 확률 변수라 하자. 그러면 서로 다른 값들 $s_1, s_2\in T$에 대해,

$$\Pr[U(s_1)=U(s_2)] \leq 2^{-p}$$

가 성립한다.

[보조 정리 6]

Alice와 Bob은 위에서 정의한 t비트 스트링의 확장된 interactive hashing 프로토콜을 수행한다. T는 $GF(2)^t$의 부분집합이고 $0<\alpha<1$에 대해 $|T|=2^{\alpha t}$라 하고, r은 자연수로 $\log_2 t \le r \le \alpha t/6$를 만족한다. m은 t를 나누는 자연수로 $m \le 2r$이다. 만약 $\alpha < 1-\frac{8r+2m+2}{t}$ 라면, Bob이 Alice의 query들에 대해 T에 속하는 서로 다른 두 원소 s_1, s_2을 가지고 같은 대답을 할 확률은 기껏해야 $\frac{1}{m2^r}$ 이다.

(증명)

$T_0=T$라 두고, $i=1,\cdots,t/m-1$과 $b_i \in GF(2)^m$에 대해 $T_i=\{s \in T_{i-1} \mid h_i(s)=b_i\}$라 정의하자. $p=2r$로 두자. $r \le \alpha t/6$ 이므로 $\alpha t \ge 3p$ 이고, $\alpha < 1-\frac{8r+2m+2}{t}$ 이므로 $\frac{\alpha t-3p}{m}+1 < \frac{t}{m}-1$ 이다. 그러므로 다음 식을 만족하는 자연수 i_j ($1 \le i_j < t/m-1$)가 존재한다:

$$\frac{\alpha t-3p}{m} < i_j \le \frac{\alpha t-3p}{m}+1 \qquad (4)$$

$i=1$ 부터 i_j-1 까지 보조정리 4를 적용하면 기껏해야 $i\cdot 2^{-p}$의 확률을 제외하고 다음 식이 만족한다.

$$|T_i| < \left(\frac{1}{2^m}+\frac{1}{2^{p+m/2}}+\frac{1}{2^{3p}}\right)^i |T|$$

따라서, $|T_{i_j}| < 2^{\alpha t-mi_j}(1+2^{m/2-p}+2^{-3p+m})^{i_j}$이고, 조건들 $t \le 2^r$, $m \le 2r$, $p=2r$을 이용하면 $i_j \log_2(1+2^{m/2-p}+2^{-3p+m}) < \frac{t}{m}\cdot(2^{m/2-p}+2^{-3p+m}) < 1$를 얻을 수 있다. 따라서 식 (4)에 의해

$$\log_2|T_{i_j}| < (\alpha t-mi_j) + i_j\log_2(1+2^{m/2-p}+2^{-3p+m}) < 3p+1 \qquad (5)$$

을 얻는다.

이제 i_j 라운드부터 $t/m-1$ 라운드까지를 선택해 보조정리 5를 적용하자. $\alpha < 1-\frac{4p+2m+2}{t}$ 이므로 $4p < t-\alpha t-2m-2$ 가 성립하고, 식 (5)으로부터

$$2\log_2|T_{i_j}| < 6p+2 < 2p+t-\alpha t-2m$$

이 성립한다. (4)를 이용하면

$$2p+t-\alpha t-2m = t-m\left(\frac{\alpha t-3p}{m}+2\right)-p < t+m(-i_j-1)-p \qquad (6)$$

이 성립한다. 따라서 식 (6)에 의해 $2\log_2|T_{i_j}| < m\left(\frac{t}{m}-1-i_j\right)-p$가 만족한다. 그러므로 보조정리 5를 적용할 수 있고, 전체 실패할 확률은 기껏해야 $(i_j+1)2^{-p} < \frac{t}{m}\cdot 2^{-p} < \frac{1}{m2^r}$ 이므로 정리가 성립한다. ■

다음 정리는 확장된 interactive hashing 프로토콜의 조건 2가 만족함을 보인다.

(정리 1)

Alice와 Bob은 위에서 정의한 t비트 스트링의 확장된 interactive hashing 프로토콜을 수행하고, 적당한 자연수 l, m에 대해 $t=lm$, $m<t$을 만족한다. T는 $GF(2)^t$의 부분집합이고 $|T| \le 2^{t-k}$라 하고, k는 $\log_2 t \le k \le 2t/3$를 만족하는 변수이다. 만약 $m < \frac{k-2}{6}$ 가 성립하면, Bob이 Alice의 query들에 대해 T에 속하는 서로 다른 두 스트링을 가지고 같은 대답을 할 확률은 기껏해야 $\frac{2^{-O(k)}}{m}$ 이다.

(증명)

$\log_2 t \le r \le (t-2)/18$ 을 만족하는 임의의 r에 대해 $k=12r+2$라 두면 $r \le (t-k)/6$ 이고, $m<(k-2)/6$ 이라는 조건때문에 $m<2r$이 성립한다. 따라서 정리 1은 보조정리 6으로부터 분명하다. ■

V. 복잡도(Complexity)

이 절에서는 제안한 프로토콜의 복잡도를 기존의 NOVY 프로토콜과 비교 분석한다.

t비트 스트링의 NOVY 프로토콜은 $t-1$의 라운드 복잡도와 $(t-1)\cdot(t+1)=t^2-1$ 비트의 전송 복잡도를 가진다. 반면에 제안한 프로토콜은 t를 나누는 양의 정수 m에 대하여 $\frac{t}{m}-1$의 라운드 복잡도와 $(\frac{t}{m}-1)\cdot(t+m)=\frac{t^2}{m}-m$ 비트의 전송 복잡도를 가진다. [표 1]은 제안된 프로토콜과 NOVY

〔표 1〕 프로토콜 복잡도 비교 (m은 t를 나누는 정수이다.)

	라운드 복잡도	전송 복잡도
NOVY 프로토콜	$t-1$	t^2-1
제안된 프로토콜	$\frac{t}{m}-1$	$\frac{t^2}{m}-m$

프로토콜에 대한 복잡도 비교를 나타낸다.

1은 항상 t를 나누기 때문에 $m=1$인 경우, 제안한 프로토콜은 기존의 NOVY 프로토콜과 같음을 알 수 있다. 1보다 큰 m이 존재하는 경우에는 〔표 1〕에 보듯이 약 m배 정도 효율적인 프로토콜로의 확장이 가능하다. 따라서, 제안한 프로토콜은 NOVY 프로토콜의 확장이라고 볼 수 있다. 제안된 프로토콜은 t를 나누는 m의 크기가 클수록 복잡도가 좋아지지만, 프로토콜의 안전성 관점에서 m에 대한 경계값(bound)를 살펴볼 필요가 있다. 정리 1에 의해 security 변수 k에 대해 m은 $1 \le m < (k-2)/6$을 만족해야 한다. 따라서 m은 t를 나누면서, $(k-2)/6$보다 작은 수들 중에서 가장 큰 정수를 택한다면 가장 좋은 복잡도를 가지는 m을 선택할 수 있을 것이다.

Ⅵ. 결 론

NOVY 프로토콜은 주어진 스트링 크기 t비트에 대해 $t-1$번의 라운드 복잡도와 t^2-1 비트의 전송 복잡도를 가지는 interactive hashing 프로토콜이다. 본 논문에서 제안된 프로토콜은 t를 나누는 m에 대해서 $t/m-1$번의 라운드 복잡도와 $t^2/m-m$ 비트의 전송 복잡도를 가짐으로 NOVY 프로토콜에 비해 효율적임을 알 수 있다. 게다가 전송자와 수신자 사이에서 전송자의 입력 스트링을 포함한 두 개 또는 그 이상의 스트링을 공유하는 방법을 제공한다.

본 논문에서는 안전한 프로토콜을 구성하기 위하여 m은 t를 나누면서 $(k-2)/6$보다 작아야 한다는 조건이 필요하나, 향후 이러한 조건을 향상시킨 프로토콜에 대한 연구가 필요하다. 또한 bounded-storage 모델에서 oblivious transfer 프로토콜과 같이 제안된 interactive hashing 프로토콜을 이용할 수 있는 응용들에 대한 연구가 계속 진행되어야 할 것으로 생각된다.

참 고 문 헌

〔1〕 M. Naor, R. Ostrovsky, R. Venkatesan, and M. Yung, "Perfect sero-knowledge arguments for NP using any one-way function", Journal of Cryptology, 11(2), pp. 87~108, 1998. Preliminary version presented at CRYPTO '92.

〔2〕 I. Damgård, "Interactive hashing can simplify zero-knowledge protocol design without computational assumtions", In Advances in Cryptology - CRYPTO '93, pp. 100~109, 1993.

〔3〕 I. Damgård and R. Cramer, "On monotone function closure of perfect and statistical zero-knowledge", CWI technical report, CS-R9618, May 1996.

〔4〕 R. Ostrovsky, R. Venkatesan, and M. Yung, "Fair games against an all-powerful adversary", SEQUENCES '91, Positano, 1991.

〔5〕 R. Ostrovsky, R. Venkatesan, and M. Yung, "Interactive hashing simplifies zero-knowledge protocol design", In Advances in Cryptology - EUROCYRPTO '93, pp. 267~273, 1993.

〔6〕 C. Cachin, C. Crépeau, and J. Marcil, "Oblivious tranfser with a memory-bounded receiver", In Proc. 39th IEEE Symposium in Foundations of Computer Science, pp. 493~502, 1998.

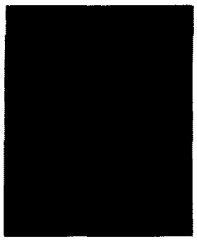
〔7〕 Y. Z. Ding, "Oblivious Transfer in the Bounded Storage Model", In Advances in Cryptology - CRYPTO 2001, pp. 155~170, 2001.

〔8〕 D. R. Stinson, "Universal hash families and the leftover hash lemma, and applications to cryptography and computing", 2002.

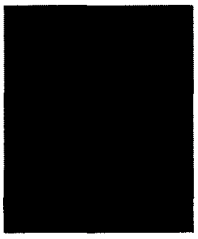
〔9〕 J. L. Carter and M. N. Wegman, "Universal classes of hash functions", Journal of Computer and System Sciences 18, pp. 143~154, 1979.

〔10〕 D. R. Stinson, "On the connections between universal hasing, combinatorial designs and error-correcting codes", Congressus Numerantium 114, pp. y~27, 1996.

〈著者紹介〉



홍 도 원 (Dowon Hong)
1994년 2월 : 고려대학교 이과대학 수학과(학사)
1996년 2월 : 고려대학교 수학과(석사)
2000년 2월 : 고려대학교 수학과(박사)
2000년 4월~현재 : 한국전자통신연구원 선임연구원
〈관심분야〉 정보보호 이론, 이동통신 정보보호



장 구 영 (Ku-Young Chang)
1995년 2월 : 고려대학교 이과대학 수학과(학사)
1997년 2월 : 고려대학교 수학과(석사)
2000년 8월 : 고려대학교 수학과(박사)
2000년 12월~현재 : 한국전자통신연구원 선임연구원
〈관심분야〉 정보보호 이론, 유한체 이론

류 희 수 (Heuisu Ryu)
1990년 2월 : 고려대학교 이과대학 수학과(학사)
1992년 2월 : 고려대학교 수학과(석사)
1999년 5월 : Johns Hopkins University 수학과(박사)
2000년 7월~현재 : 한국전자통신연구원 선임연구원
〈관심분야〉 정보보호 이론, 이동통신 정보보호