CASCADE 프로토콜의 Binary Search 알고리즘에 대한 확률적 분석

김광제[2], 염용진[1),2)], 강주성[1),2)]*
국민대학교 정보보안암호수학과[1)] / 금융정보보안학과[2)]
{rhkdwp0105, salt, jskang*}@kookmin.ac.kr

Probabilistic Analysis on the Binary Search Algorithm of CASCADE Protocol

Gwangjae Kim[2)], Yongjin Yeom[1),2)], Ju-Sung Kang[1),2)]*
Dept. of Information Security, Cryptology, and Mathematics[1)] /
Financial information security[2)], Kookmin Univ.

요 약

양자키분배 프로토콜은 양자 물리학의 기본 법칙을 기반으로 비밀키를 안전하게 공유하는 프로토콜이다. 이 프로토콜은 두 사용자가 양자 채널을 통해 키를 공유하는 과정에서 양자의 불안정성으로 인해 두 사용자의 키에 오류가 발생하게 되는데, 이를 효과적으로 정정하기 위해 오류 정정 프로토콜을 사용한다. 오류 정정 프로토콜은 두 사용자의 키 중 일치하지 않는 비트를 일반 채널에서의 통신을 이용하여 오류를 정정한다. 본 논문에서는 대표적인 오류 정정 프로토콜인 CASCADE 프로토콜의 오류 정정 과정을 분석하고 각 단계에서의 오류 정정 기댓값을 제시한다. 또한 시뮬레이션 결과를 바탕으로 통계적 가설 검정을 실시하여 제시한 오류 정정 기댓값의 타당성을 검증한다.

Ⅰ. 서 론

양자키분배(Quantum Key Distribution, QKD) 프로토콜은 정당한 두 사용자가 양자를 이용하여 안전하게 비밀키를 생성하고 공유하기 위한 프로토콜이다. 기존 키분배 프로토콜이 계산 복잡도에 기반한 조건부 안전성을 제공하는 것과는 달리 QKD 프로토콜은 양자 역학적 특성을 이용하여 정보이론적 안전성을 보장한다[1]. QKD 프로토콜은 raw 키 생성, 걸러진 키 생성, 비밀키 생성 3단계로 구성되며, 비밀키 생성 단계는 오류 정정 단계와 비밀 증폭 단계로 나뉜다[2]. 이 중 오류 정정 단계는 두 사용자가 양자 채널을 통해 나누어 가진 걸러진 키의 오류를 일반(classical) 채널을 통해 정정함으로써 걸러진 키를 일치시키는 단계이며 대표적인 오류 정정 프로토콜로는 CASCADE 프로토콜이 있다.

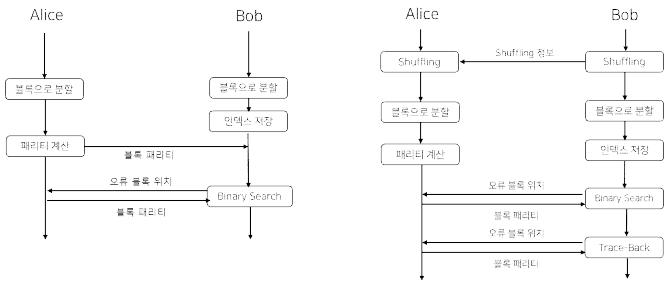
오류 정정을 위한 사용자 간 통신은 일부 키 정보의 노출로 이어지는데, 비밀 증폭 단계에서 이 노출 정보량에 해당하는 키 비트를 제거한다. 따라서 효율적인 키 분배를 위하여 노출 정보량을 최소화하는 오류 정정 과정의 최적화가 필요하고, 이에 따른 프로토콜의 이론적인 분석이 요구된다. 본 논문에서는 CASCADE 프로토콜을 구성하는 세 가지 핵심 알고리즘이 오류 정정에 미치는 영향을 분석하고, 각 단계에서 정정되는 오류 비트 수의 기댓값(이하 오류 정정 기댓값)을 제시한다. 이후 CASCADE 프로토콜 시뮬레이션 결과를 바탕으로 통계적 가설 검정을 통해 논문에서 제시한 오류 정정 기댓값을 검증한다.

Ⅱ. CASCADE 프로토콜

1993년 Gilles Brassard가 제안한 CASCADE 프로토콜은 두 사용자의 걸러진 키와 양자 채널에 의해 제공된 비트 오류율의 추정치를 입력으로 받아 패리티 정보를 비교하여 걸러진 키 사이의 오류를 식별하고 수정한다[3]. 프로토콜은 크게 Binary Search, Random Shuffling, Trace-back 세 가지 알고리즘의 반복 수행으로 이루어져 있으며, 각 반복과정을 패스라고 칭한다. CASCADE 프로토콜은 [그림 1]과 같이 첫 번째 패스에서는 Binary Search만을 수행하며, 이후 패스에서는 세 알고리즘을 모두 수행한다.

ⅰ. Binary Search

Binary Search는 수신자가 홀수 개의 오류를 가지고 있을 때 1비트의 오류를 수정하는 알고리즘으로, 패리티 정보 비교를 반복 수행하여 오류를 수정한다. 수신자는 송신자에게 비트 위치 정보를, 송신자는 수신자에게 위치 정보에 해당하는 패리티 정보를 전송하는 것을 반복하여 오류 비트를 찾는다. Binary Search 알고리즘은 한 번에 한 개의 오류만을 수정할 수 있으며, 두 사용자의 패리티 정보가 달라야 한다는 조건 하에 수행된다. 따라서 각 사용자는 적절한 길이의 블록으로 키를 분할하고, 각 블록의 패리티 정보를 비교하여 패리티가 다른 블록에 대해 Binary Search를 수행한다. 이 알고리즘은 각 블록에 오류가 홀수 개일 때 하나의 오류를 정정하기 때문에, 알고리즘 종료 후 각 블록에는 짝수 개의 오류가 남아있게 된다.



[그림 1] CASCADE 프로토콜 순서도 (좌: 첫 번째 패스, 우: 나머지 패스)

ⅱ. Random Shuffling

CASCADE 프로토콜의 Random Shuffling은 오류를 키 전반으로 분산시키기 위해 사용된다. 두 사용자는 프로토콜 진행 전, 혹은 프로토콜 진행 중 Shuffling 정보를 공유하여 같은 방식으로 Shuffling을 진행한다.

ⅲ. Trace-back

Trace-back 알고리즘은 이전 패스의 정보를 활용하여 기존 Binary Search 알고리즘의 사용자 간 통신을 줄이는 알고리즘이다. Binary Search 알고리즘으로 오류를 정정하기 위해서는 먼저 블록 내 오류가 홀수 개 존재하는지 확인해야 하므로, 두 사용자는 알고리즘 동작 전 패리티 정보를 교환해야 한다. Trace-back 알고리즘은 Binary Search 알고리즘이 종료된 후의 블록 내 오류 개수가 짝수임을 이용하여 패리티 정보 교환 과정을 생략한다. 예를 들어 두 번째 패스에서 오류가 한 개 수정된다면, 첫 번째 패스에서 해당 오류가 존재했던 블록의 패리티 정보가 달라지므로, 두 사용자는 해당 블록에 오류가 추가로 존재함을 알 수 있다. 따라서 두 사용자는 패리티 비교를 하지 않고도 해당 블록

에 Binary Search 알고리즘을 적용할 수 있다.

Ⅲ. Binary Search 오류 정정 기댓값 분석

CASCADE 프로토콜의 Binary Search 알고리즘을 분석하기 위하여 다음과 같은 가정을 한다.

1. 두 사용자 간 오류는 I.I.D(independent and identically distributed)이다.
2. Random Shuffling 알고리즘에 의해 각 비트가 각 자리에 위치할 확률은 모두 동일하다.

Binary Search는 두 사용자 간 패리티 정보가 일치하지 않는 블록에 대해 한 개의 오류를 정정한다. 즉 한 패스의 전체 키 비트에서 정정되는 오류의 개수는 두 사용자 간 오류가 홀수 개 발생한 블록의 개수와 같다. p를 키의 비트당 오류율이라고 하고, 각 블록의 크기를 k라고 하자. $X^{(k)}$를 k 길이 블록의 오류 개수를 나타내는 확률변수라고 하면,

$P(X^{(k)}=i)=\binom{k}{i}p^i(1-p)^{k-i}, i=0,1,\cdots,k$이고,

$p_{odd} := P(X^{(k)} is\ odd)=\sum_{i=1}^{\frac{k}{2}}\binom{k}{2i-1}p^{2i-1}(1-p)^{k-2i+1}$이므로,

$p_{odd}=\frac{1-(1-2p)^k}{2}$ 이다.

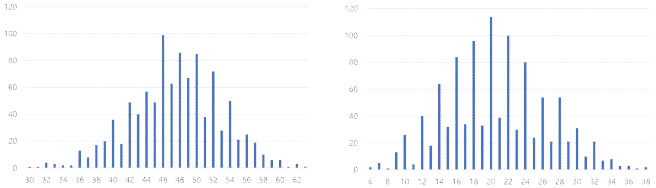
두 사용자의 전체 키의 크기가 N이라고 하면 두 사용자는 $\left\lfloor \frac{N}{k} \right\rfloor$ 개의 블록을 가지고, 확률변수 $Y^{(k)}$를 두 사용자의 k 길이 블록 중 오류가 홀수 개인 블록의 개수라고 하면 $Y^{(k)}$는 이항분포 $B(\frac{N}{k}, p_{odd})$를 따른다. 따라서 Binary Search 알고리즘을 통해 정정되는 오류의 개수는 $\left\lfloor \frac{N}{k} \right\rfloor \times p_{odd}=\left\lfloor \frac{N}{k} \right\rfloor \times \frac{1-(1-2p)^k}{2}$ 이다.

Ⅳ. 실험적 분석

실험적 분석을 위해 다음과 같이 실험 조건을 설정한다. 두 사용자의 전체 키의 크기를 10000bit, 비트 당 오류율 $p_1=0.01$, 첫 번째 패스의 블록 크기 $k_1=73$, 이후 패스가 진행될 때마다 블록의 크기를 2배씩 증가시킨다고 하자. 해당 조건으로 CASCADE 프로토콜을 1000번 시행한 후, 통계적 가설 검정을 통해 앞 절에서 제시한 정정되는 오류 개수의 기댓값과 실제 시뮬레이션 결과를 비교하여 기댓값을 검증한다.

3절에서 제시한 Binary Search 후 남은 오류 개수의 기댓값 E_1을 위 조건을 대입하여 계산하면, $E_1=N\cdot p-136\cdot p_{odd}=47.56$이다. 첫 번째 패스의 Binary Search 후 남은 실제 오류 개수의 평균을 μ_1이라고 하면, 시뮬레이션의 시행 횟수 n이 충분히 크므로 표준화된 표본평균 $z=\frac{\overline{x_1}-\mu_1}{\sigma_1/\sqrt{n}}$는 표준정규분포를 따른다. 귀무가설 $H_0:\mu_1=47.56$이고 대립가설 $H_1:\mu_1\neq 47.56$이라고 하면 유의수준 $\alpha=0.01$일 때, 기각역은 $\{z|z<-2.58 \text{ or } z>2.58\}$이다. Binary Search의 시뮬레이션 결과는 [그림 2]의 왼쪽 그래프와 같고, 이때의 표본평균 $\overline{x_1}=47.57$, 분산 $\sigma_1^2=28.919$이다. 시뮬레이션 결과를 토대로 검정통계량 z를 계산하면, $z=\frac{47.57-47.56}{5.3776/\sqrt{1000}}=0.059$이므로, 귀무가설을 채택한다.

두 번째 패스에서의 오류 정정 기댓값을 구하기 위해서는 두 번째 패스에서의 오류율이 필요하다. 앞선 가정에서 Random Shuffling 알고리즘에 의해 각 비트가 각 자리에 위치할 확률은 모두 동일하다고 가정했기 때문에, Shuffling 후에도 오류가 모두 독립적으로 발생했다고 가정할 수 있다. 따라서 두 번째 패스 시작 전 오류율 p_2는 첫 번째 패스 후 남은 오류 기댓값 E_1을 전체 키 비트 수로 나눈 값으로 계산한다. 그러므로 두 번째 패스에서의 오류 정정 기댓값은 $\left\lfloor \frac{N}{k_2} \right\rfloor \times \frac{1-(1-2p_2)^{k_2}}{2}=25.58$ 이고, Binary Search 후 남은 오류 기댓값 $E_2=21.98$이다. 두 번째 패스의 Binary Search 단계에서 정정되는 실제 오류의 평균을 μ_2라고 하면, 시뮬레이션의 시행 횟수 n이 충분히 크므로 표준화된 표본평균 $z=\frac{\overline{x_2}-\mu_2}{\sigma_2/\sqrt{n}}$는 표준정규분포를 따른다. 귀무가설 $H_0:\mu_2=21.98$이고, 대립가설 $H_1:\mu\neq 21.98$이라고 하자. 유의수준 $\alpha=0.01$이라고 했을 때, 기각역은 $\{z|z<-2.58 \text{ or } z>2.58\}$이다. Binary Search의 시뮬레이션 결과는 [그림 2]의 오른쪽 그래프와 같고, 이때의 표본평균 $\overline{x_2}=21.87$, 분산 $\sigma_2^2=29.34$이다. 시뮬레이션 결과를 토대로 검정통계량 z를 계산하면, $z=\frac{21.87-21.98}{5.4166/\sqrt{1000}}=-0.64$이므로, 귀무가설을 채택한다. 결과적으로, 실험적 입증에 의해 앞서 제시한 Binary Search 알고리즘의 오류 정정 기댓값은 타당하다.



[그림 2] Binary Search 시뮬레이션 결과 (좌: 첫 번째 패스, 우: 두 번째 패스)

Ⅴ. 결론

본 논문에서는 QKD 오류 정정 프로토콜 중 하나인 CASCADE 프로토콜의 세부 알고리즘을 분석하고, Binary Search 알고리즘이 오류 정정에 미치는 영향을 확률적으로 계산하여 기댓값을 제시하였다. 또한 이를 시뮬레이션 결과와 비교하여 제안한 기댓값의 타당성을 확인하였다. 향후 연구에서는 Trace-back 알고리즘의 오류 정정에 미치는 영향을 확률적으로 계산하고, 이를 바탕으로 CASCADE 프로토콜의 전체적인 오류 정정 기댓값을 확률적으로 분석할 예정이다. CASCADE 프로토콜의 확률적 분석은 프로토콜의 파라미터인 블록 크기와 오류 정정 확률의 직접적인 상관관계를 밝혀 사용자 간 더 효율적인 양자키분배를 가능하게 할 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2021-0-00046, 국가공공정보시스템 안전성 및 활용성 제고를 위한 차세대 암호체계 개발)

참 고 문 헌

[1] Xu, Feihu, et al. "Secure quantum key distribution with realistic devices." Reviews of modern physics 92.2, 2020.

[2] 김나영 외, “양자 키 분배 - 제2부: BB84 프로토콜,” TTAK.KO-12.0329-Part2, 2018.12.

[3] Brassard, G., Salvail, L., “Secret-Key Reconciliation by Public Discussion,” Advances in Cryptology, EUROCRYPT ’93, Lecture Notes in Computer Science, pp. 410-423, 1993.