J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

프라이버시를 보호하는 DNA 매칭 프로토콜

Privacy-Preserving DNA Matching Protocol

노 건 태[1*]
Geontae Noh

요 약

기술의 발전에 따라 유전 정보를 수월하게 얻을 수 있게 되었으며, 이것의 활용도 및 미래 가치는 매우 높다. 하지만, 유전 정보는 한 번 유출되면 변경할 수 없으며, 피해의 정도도 개인에만 국한되지 않고, 대용량 데이터이기 때문에 이를 고려한 처리 기술 또한 필요하다. 즉, 대용량에서도 프라이버시를 고려하며 유전 정보를 처리할 수 있는 기술의 개발이 필요하다.

본 논문에서는 Gentry 등의 준동형 암호 기법을 사용하여 먼저 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜을 제안하고, 이 프로토콜을 활용하여 효율적인 프라이버시를 보호하는 DNA 매칭 프로토콜을 제안한다. 우리가 제안하는 프라이버시를 보호하는 DNA 매칭 프로토콜은 효율적이며, 정확성, 기밀성, 프라이버시를 만족한다.

☞ 주제어 : 프라이버시 보호 내적 연산, DNA 매칭 프로토콜

ABSTRACT

Due to advances in DNA sequencing technologies, its medical value continues to grow. However, once genome data leaked, it cannot be revoked, and disclosure of personal genome information impacts a large group of individuals. Therefore, secure techniques for managing genomic big data should be developed.

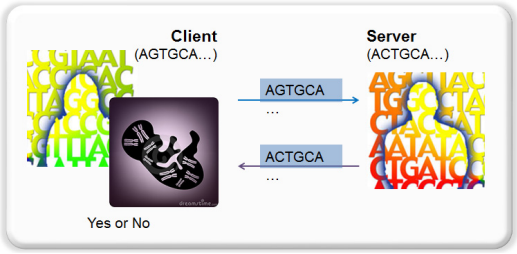
We first propose a privacy-preserving inner product protocol for large data sets using the homomorphic encryption of Gentry et al., and then we introduce an efficient privacy-preserving DNA matching protocol based on the proposed protocol. Our efficient protocol satisfies the requirements of correctness, confidentiality, and privacy.

☞ keyword : Privacy-Preserving Inner Product, DNA Matching Protocol

1. 서 론

최근 DNA 시퀀싱(Sequencing) 기술이 발달함에 따라 자신의 유전 정보를 저렴한 가격에 쉽게 얻을 수 있게 되었다. 이러한 유전 정보는 병에 걸릴 가능성을 예측한다던지, 친자 확인 등 여러 가지 중요한 응용 환경에 사용될 수 있다.

우리나라의 경우, 2016년 생명윤리 및 안전에 관한 법률 개정으로 인해 탈모, 피부 노화, 피부 탄력, 체질량 지수 등 건강 및 미용과 관련된 일부 유전자에 대해 개인이 직접 유전자 검사 서비스를 받을 수 있게 되었으며, 2017년부터는 생명윤리 및 안전에 관한 법률 시행령 개정안으로 인해 질병과 관련된 일부 유전자의 검사를 받을 수 있게 되었다.



(Figure 1) Paternity Test

하지만 유전 정보는 그 사람의 인종, 조상, 질병에 대한 가족력 등과 같은 민감한 개인 정보들을 포함하고 있기 때문에 노출되었을 경우 프라이버시 문제 등과 같은 심각한 문제들이 발생된다. 또한, 유전 정보는 한 번 유출되면 다른 개인정보들과는 달리 변경할 수 없으며, 특정 개인의 유전 정보로부터 가족의 유전 정보까지 파악할 수 있기 때문에 피해의 정도가 개인을 넘어서게 된다. 이렇기 때문에 유전 정보는 개인 정보를 암암리에 거래하는 암시장에서도 가장 높은 가치를 인정받고 있으며, 따

1 Dept. of Information Security, Seoul Cyber University, (Miadong) Solmaero 49 Gil 60, Gangbuk Gu, Seoul, Korea 01133
* Corresponding author (gnoh@iscu.ac.kr)
[Received 19 July 2017, Reviewed 26 December 2017, Accepted 12 January 2018]

라서 이러한 유전 정보의 활용에 있어서 유전 정보의 프라이버시는 반드시 보호되고 지켜져야만 한다.

또한 한 사람의 전체 유전 정보(fully sequenced genome)는 이를 저장하기 위해서 수 백 기가 바이트의 저장 용량이 필요할 만큼 대용량 데이터이다. 따라서 이러한 대용량 데이터를 프라이버시를 보호하면서도 효율적으로 활용할 수 있는 기법을 설계하기가 쉽지 않다. 최근 프라이버시를 보호하는 친자 확인 등 유전 정보 활용을 위해 프라이버시를 보호하는 기법[1]이 제안되었다(Figure 1). 하지만 인간의 전체 유전 정보를 활용하기에는 너무 비효율적이다. 두 사람이 친자 관계인지 살펴보기 위해 전체 유전 정보를 활용하는 경우 9일의 시간이 걸리고, 약 400GB 정도의 저장량이 필요하다(Table 1). 그리고 논문[1]에서는 효율성을 극대화하기 위해 모든 DNA를 전수 조사하는 방법 외에도, 전체 DNA의 약 1%만을 추출한 데이터로부터 비교 분석한 결과를 포함하였다. 이것은 전체 DNA의 약 1%만으로도 전체 DNA를 대표할 수 있다는 근거로부터 기인한 것이며, 단지 전체 DNA 대신에 1%만의 DNA 정보를 가지고 비교한 분석이다. 해당 방법을 사용하였을 때에는 약 4GB 정도의 저장량이 필요하다(Table 2).

(Table 1) Computation and Communication Costs [1]

	오프라인 시간	온라인	
		시간	저장 용량
클라이언트	4.5일	4.5일	358GB
서버	4.5일	4.5일	414GB

(Table 2) Computation and Communication Costs - 1% of the Human Genome [1]

	오프라인 시간	온라인	
		시간	저장 용량
클라이언트	67분	67분	3.57GB
서버	67분	67분	4.14GB

인간의 전체 유전 정보를 활용하는 것은 결과의 정확도를 위해 꼭 필요하지만, 대용량 데이터이기 때문에 실제로 안전하게 활용하는 데는 효율성이 급격이 떨어지는 단점이 있다. 따라서 대용량 데이터인 전체 유전 정보를 안전하게 효율적으로 활용할 수 있는 기법이 필요하다.

따라서 본 논문에서는 대용량 데이터에서 프라이버시를 보호하는 기법을 제안하고, 이 프로토콜을 활용하여 DNA 매칭 프로토콜을 제안한다. 제안하는 프로토콜은 사용자의 프라이버시를 보호할 뿐만 아니라 매우 효율적이다.

2. 배경 지식

2.1 Gentry 등의 준동형 암호

2009년, Gentry가 제안한 완전 준동형 암호 기법이 촉매제가 되어 래티스에서 수많은 암호 기법들이 설계되었으며[2], 공개키 암호 기법[3], ID기반 암호 기법[4, 5], 브로드캐스팅 암호 기법[6], 준동형 암호 기법[2, 7] 등 다양한 종류의 암호 기법들이 래티스 구조에서 제안되었다.

2009년 제안된 Gentry의 완전 준동형 암호 기법은 아이디얼 래티스에서 제안되었으며[2], 이후 2010년에 Gentry 등은 래티스에서 LWE(Learning with Errors) 문제의 어려움에 기반을 두고 새로운 준동형 암호 기법을 제안하였다[7]. 해당 기법은 암호문에서 여러 번의 덧셈과 한 번의 곱셈이 가능하다.

Gentry 등이 제안한 준동형 암호 기법 이후로도 준동형 암호 기법의 설계와 분석, 구현에 관한 래티스에서의 연구는 지속되고 있다. 2013년에 Gentry 등에 의해 제안된 기법[8]과 2014년에 Boneh 등에 의해 제안된 준동형 암호 기법[9]은 모두 속성 기반 암호 기법과 연관하여 준동형 암호의 새로운 매커니즘에 대해 연구하였고, 이러한 연구는 2016년에 Peikert와 Shiehian에 의해 발전되었다[10]. 이러한 연구들은 암호문에서 여러 번의 덧셈과 여러 번의 곱셈이 가능하다.

본 논문에서 우리는 2010년에 Gentry 등이 제안한 준동형 암호 기법을 사용한다[7]. 해당 기법은 최근의 연구들과는 달리, 여러 번의 덧셈과 단 한 번의 곱셈을 지원하는데, 우리가 본 논문에서 제안하고자 하는 프로토콜에서는 여러 번의 덧셈과 더불어 단 한 번의 곱셈만 필요하기 때문에 2010년에 Gentry 등이 제안한 준동형 암호 기법을 사용하는 것이 더 효율적이다.

Gentry 등의 준동형 암호 기법에 사용되는 파라미터는 다음과 같다:

- n은 시큐리티 파라미터
- $c = c(n) > 0$
- p와 $q = \omega\left(p^2 n^{3c+1} \log^5 n\right)$는 서로소
- $m = \lfloor 8n \log q \rfloor$
- $\beta = 1/\left(27 n^{1+(3c/2)} \log n \log q \sqrt{qm}\right)$

Gentry 등의 준동형 암호 기법은 다음과 같다:

- $GHV.Key(1^n,1^m,q)$: 이 알고리즘은 공개키 $pk=A\in Z_q^{n\times n}$ 과 비밀키 $sk=T\in Z^{n\times m}$ 을 생성한다. 여기서 T는 역행렬이 존재하고, $TA=0 \bmod q$ 를 만족하며, T의 원소들은 $O(n\log q)$에 바운드된다.
- $GHV.Enc(pk,B)$: 이 알고리즘은 평문 $B\in Z_p^{n\times m}$ 을 암호화하기 위해서 랜덤한 행렬 $S\in Z_q^{n\times m}$ 와 가우시안 에러 행렬 $X\in Z_q^{n\times m}$ 를 선택하여 암호문 $C=AS+pX+B \bmod q$를 생성한다.
- $GHV.Dec(sk.C)$: 이 알고리즘은 암호문 $C\in Z_q^{n\times m}$ 를 복호화하기 위해서 먼저 $E=TCT^t \bmod q$를 계산하고, 그 다음 평문 $B=T^{-1}E(T^t)^{-1} \bmod p$를 생성한다.

3. 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜

우리는 Gentry 등의 준동형 암호 기법을 사용하여 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜을 제안한다. 제안하는 프라이버시를 보호하는 내적 연산을 활용하여 본 논문에서 제안하는 프라이버시를 보호하는 DNA 매칭 프로토콜의 설계가 가능하다.

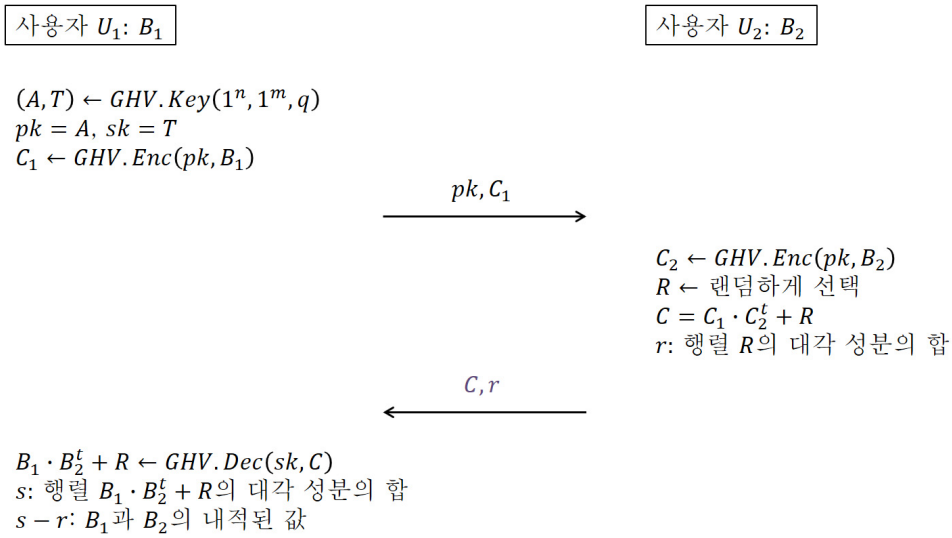
3.1 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜

본 프로토콜은 이자간에 이루어지며, 각각은 대용량의 이진(Binary) 형태의 정보를 가지고 있다. 본 프로토콜을 수행하여 각각이 가진 이진(Binary) 정보의 내적된 값만을 한 명만 알아낼 수 있도록 구성하며, 내적값을 제외한 다른 추가적인 정보는 어느 누구도 알 수 없다. 즉, 사용자 U_1과 U_2가 서로 프로토콜을 수행하며, 이진 정보의 내적된 값에 대한 정보는 U_1만 얻을 수 있으며, 다른 추가적인 정보는 U_1 조차도 알 수 없다.

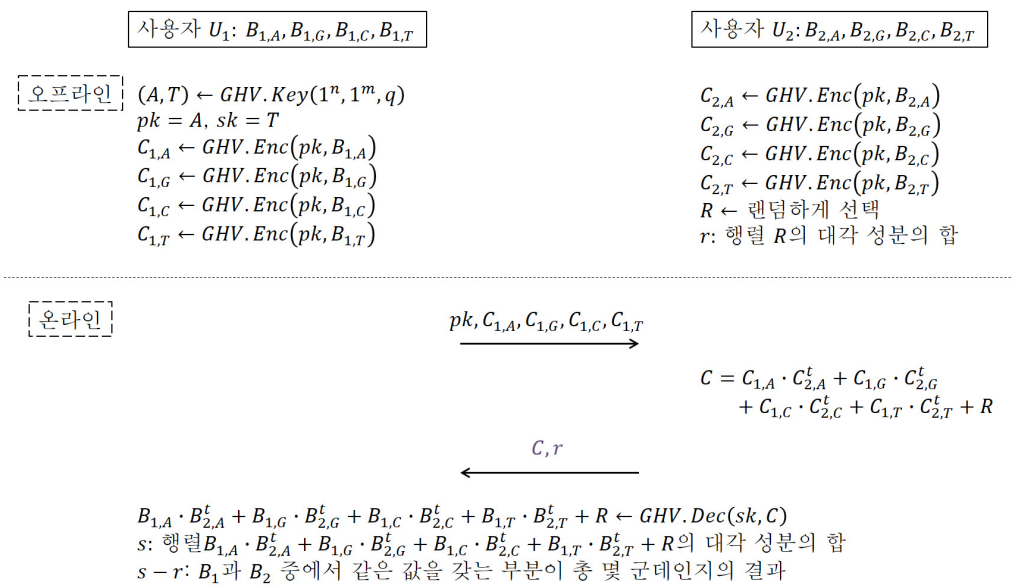
본 프로토콜에서 사용될 파라미터는 Gentry 등의 준동형 암호 기법에 사용되는 파라미터와 동일하며, 사용자 U_1의 대용량 이진 정보를 행렬로 구성하여 $B_1\in Z_2^{n\times m}$ 라 하고, 사용자 U_2의 대용량 이진 정보를 행렬로 구성하여 $B_2\in Z_2^{n\times m}$ 라 하자. 본 논문에서는 두 행렬 B_1, B_2의 내적값을 $B_1 \bullet B_2^t$의 대각성분의 합으로 정의한다.

그러면, 본 프로토콜의 기술은 다음과 같다:

- $U_1 \rightarrow U_2$: 사용자 U_1은 $GHV.Key(1^n,1^m,q)$ 알고리즘을 수행하여 $pk=A\in Z_q^{n\times n}$ 와 $sk=T\in Z^{n\times m}$ 을 생성하고, $GHV.Enc(pk,B_1)$ 알고리즘을 수행



(Figure 2) Privacy-Preserving Inner Product Protocol



(Figure 3) Privacy-Preserving DNA Matching Protocol

하여 B_1에 대한 암호문 C_1을 생성하고, 이것과 pk를 함께 사용자 U_2에 전송한다.

- $U_2 \rightarrow U_1$: 사용자 U_2는 $GHV.Enc(pk, B_2)$ 알고리즘을 수행하여 B_2에 대한 암호문 C_2를 생성하고, m차원 정사각 행렬 $R \in Z_p^{m \times m}$을 랜덤하게 선택한 다음, $C = C_1 \bullet C_2^t + R$을 계산한다. 마지막으로, 행렬 R의 대각 성분의 합 r을 계산하고, C와 r을 함께 사용자 U_1에 전송한다.
- 최종 단계: 사용자 U_1은 $GHV.Dec(sk.C)$ 알고리즘을 사용하여 C를 복호화하여 $B_1 \bullet B_2^t + R$을 계산하고, 계산된 행렬 $B_1 \bullet B_2^t + R$의 대각 성분의 합 s를 계산한다. 그러면 사용자 U_1은 B_1과 B_2의 내적된 값을 $s - r$과 같이 계산할 수 있다.

3.2 안전성 분석

우리의 프라이버시를 보호하는 내적 연산 프로토콜은 정확성, 기밀성, 그리고 프라이버시를 만족한다.

정리 1. [정확성] 우리의 프로토콜은 정확성을 만족한다.
증명. 우리가 제안하는 프로토콜은 Gentry 등의 준동형 암호 기법을 사용하며, 따라서 암호화 과정과 복호화 과정의 정확성은 Gentry 등의 준동형 암호 기법의 정확성을 따른다.

최종 단계에서 계산되는 $B_1 \bullet B_2^t + R$의 대각 성분의 합은 행렬 B_1과 B_2의 내적값에 행렬 R의 대각 성분의 합이 더해진 값이다. 따라서 행렬 $B_1 \bullet B_2^t + R$의 대각 성분의 합 s에서 R의 대각 성분의 합 r을 제거하면 행렬 B_1과 B_2의 내적값을 얻어낼 수 있다. □

$B_1 \bullet B_2^t$의 i번째 대각 성분은 행렬 B_1과 B_2의 i번째 행을 벡터로 보았을 때 두 벡터의 내적값이라고 할 수 있다. 행렬 B_1과 B_2가 이진 정보로 이루어져있기 때문에 i번째 대각 성분은 행렬 B_1과 B_2의 i번째 행에서 같은 위치에 둘 다 1이 있는 경우의 개수를 나타낸다.

정리 2. [기밀성] 우리의 프로토콜은 기밀성을 만족한다.
증명. 우리가 제안하는 프로토콜에서 사용자 U_1과 U_2가 주고받는 값은 공개키 pk, 암호문 C, 랜덤한 값 r이다. pk는 Gentry 등의 준동형 암호 기법의 공개키로, 애초에 공개되는 값이다. C는 Gentry 등의 준동형 암호 기법의 암호문이기 때문에 기밀성을 보장하며, r만 보고서는 어느 누구도 어떠한 정보를 얻어낼 수 없기 때문에 우리가 제안하는 프로토콜은 기밀성을 만족한다. □

정리 3. [프라이버시] 우리가 제안하는 프로토콜은 프라이버시를 만족한다.

증명. 우리가 제안하는 프로토콜에서 사용자 U_2는 비밀키 sk에 대한 정보가 없기 때문에 사용자 U_1으로부터 받은 공개키 pk와 암호문 C_1을 보더라도 얻을 수 있는 정보가 없다.

사용자 U_1은 $B_1 \bullet B_2^t + R$과 정사각 행렬 R의 대각 성분의 합 r을 얻을 수 있다. 행렬 R은 사용자 U_2가 랜덤하게 선택한 값이기 때문에 사용자 U_1은 행렬 R이 없다면 $B_1 \bullet B_2^t$의 원소들의 정보를 전혀 알아낼 수 없고, 단지 행렬 R의 대각 성분의 합 r을 이용하여 B_1과 B_2의 내적된 값만을 $s-r$과 같이 계산할 수 있다. □

4. 프라이버시를 보호하는 DNA 매칭 프로토콜

본 장에서 우리는 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜을 활용하여 프라이버시를 보호하는 DNA 매칭 프로토콜을 제안한다.

4.1 프라이버시를 보호하는 DNA 매칭 프로토콜

본 프로토콜은 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜과 마찬가지로 이자간에 이루어지며, 사용자들은 A, G, C, T의 연속으로 이루어진 대용량의 DNA 정보를 가지고 있고, 각각이 가진 대용량의 DNA 정보 중에서 같은 값을 갖는 곳이 총 몇 군데인지를 한 명만 알아낼 수 있도록 구성한다. 즉, 사용자 U_1(엄마)과 U_2(아빠)가 서로 정보를 주고받으며, 같은 DNA 값을 갖는 곳의 개수에 대한 정보(친자 유무)는 U_1(엄마)만 얻을 수 있고, 어느 부분에서 같은 값을 가졌는지에 대한 정보는 어느 누구도 알 수 없다.

본 프로토콜에서 사용될 파라미터는 Gentry 등의 준동형 암호 기법에 사용되는 파라미터와 동일하며, 사용자 U_1의 대용량 DNA 정보를 행렬로 구성하여 $B_1 \in Z_p^{n \times m}$라 하고, 사용자 U_2의 대용량 DNA 정보를 행렬로 구성하여 $B_2 \in Z_p^{n \times m}$라 하자. 그러면, 본 프로토콜의 기술은 다음과 같다:

- 초기 단계: 사용자 U_1과 U_2는 자신의 대용량 DNA 정보 중 A가 포함된 부분을 1로 남기고 나머지 값들로 이루어진 부분은 0으로 채운 새로운 이진 행렬 $B_{1,A} \in Z_2^{n \times m}$과 $B_{2,A} \in Z_2^{n \times m}$를 각각 생성한다. 이와 동일한 방식으로 사용자 U_1과 U_2는 각자 자신의 대용량 DNA 정보 중 G, C, T와 관련된 새로운 이진 행렬 $B_{1,G}, B_{1,C}, B_{1,T} \in Z_2^{n \times m}$와 $B_{2,G}, B_{2,C}, B_{2,T} \in Z_2^{n \times m}$를 각각 생성한다.
- $U_1 \rightarrow U_2$: 사용자 U_1은 $GHV.Key(1^n, 1^m, q)$ 알고리즘을 수행하여 $pk = A \in Z_q^{n \times n}$과 $sk = T \in Z^{m \times m}$을 생성한다. 그리고 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜에서처럼 $B_{1,A}, B_{1,G}, B_{1,C}, B_{1,T}$에 대한 암호문 $C_{1,A}, C_{1,G}, C_{1,C}, C_{1,T}$을 생성하고, 이것과 pk를 함께 사용자 U_2에 전송한다.
- $U_2 \rightarrow U_1$: 사용자 U_2는 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜에서처럼 자신이 가진 4개의 이진 행렬 $B_{2,A}, B_{2,G}, B_{2,C}, B_{2,T}$에 대한 암호문 $C_{2,A}, C_{2,G}, C_{2,C}, C_{2,T}$를 생성하고, m차원 정사각 행렬 $R \in Z_p^{m \times m}$을 랜덤하게 선택한 다음, $C = C_{1,A} \bullet C_{2,A}^t + C_{1,G} \bullet C_{2,G}^t + C_{1,C} \bullet C_{2,C}^t + C_{1,T} \bullet C_{2,T}^t + R$을 계산한다. 마지막으로, 행렬 R의 대각 성분의 합 r을 계산하고, C와 r을 함께 사용자 U_1에 전송한다.
- 최종 단계: 사용자 U_1은 $GHV.Dec(sk.C)$ 알고리즘을 사용하여 C를 복호화한 결과인 $B_{1,A} \bullet B_{2,A}^t + B_{1,G} \bullet B_{2,G}^t + B_{1,C} \bullet B_{2,C}^t + B_{1,T} \bullet B_{2,T}^t + R$을 계산하고, 계산된 이 행렬의 대각 성분의 합 s를 계산한다. 그러면 사용자 U_1은 B_1과 B_2 중에서 같은 값을 갖는 부분이 총 몇 군데인지를 $s-r$과 같이 계산할 수 있다.*

4.2 안전성 분석

우리의 프라이버시를 보호하는 DNA 매칭 프로토콜은 정확성, 기밀성, 그리고 프라이버시를 만족한다.

정리 4. [정확성] 우리의 프로토콜은 정확성을 만족한다.

증명. 우리의 프라이버시를 보호하는 DNA 매칭 프로토콜의 정확성은 대용량에서 프라이버시를 보호하는 내적

* 99.5% 이상이 일치하면 친자 관계이다[1].

연산 프로토콜의 정확성에 기반을 두고 있으며, 정리 1에 따라 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜이 정확성을 가지기 때문에 우리의 프라이버시를 보호하는 DNA 매칭 프로토콜도 정확성을 만족한다. □

정리 5. [기밀성] 우리의 프로토콜은 기밀성을 만족한다.
증명. 우리의 프라이버시를 보호하는 DNA 매칭 프로토콜의 기밀성은 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜의 기밀성에 기반을 두고 있으며, 정리 2에 따라 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜이 기밀성을 가지기 때문에 우리의 프라이버시를 보호하는 DNA 매칭 프로토콜도 기밀성을 만족한다. □

정리 6. [프라이버시] 우리가 제안하는 프로토콜은 프라이버시를 만족한다.
증명. 우리의 프라이버시를 보호하는 DNA 매칭 프로토콜의 프라이버시는 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜의 프라이버시에 기반을 두고 있으며, 정리 3에 따라 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜이 프라이버시를 만족하기 때문에 우리의 프라이버시를 보호하는 DNA 매칭 프로토콜도 프라이버시를 만족한다. □

4.3 효율성 분석 - 전수 조사

우리의 프로토콜은 아래 (Table 3)과 같은 예제 파라미터들을 가진다. 우리의 프로토콜에서의 암호문은 약 50GB인데 비해, 논문 [1]의 암호문은 약 400GB로 우리의 기법이 저장량 측면에서 훨씬 효율적이다(Table 1, Table 3). 또한 논문 [1]의 경우 지수승 연산으로 인해 엄청난 시간이 소요되는데 반해, 우리의 프로토콜은 상대적으로 효율적인 행렬의 곱셈, 덧셈만 수행하면 되기 때문에 빠른 시간 내에 DNA 매칭이 가능하다. 연산 효율성 분석을 위해서 우리는 대문자-O 표기법(big-O notation)을 사용하

(Table 3) Example Parameters

	예제 1	예제 2
n	192	120
q	6.8×10^{19}	6.6×10^{18}
p	151,795	90,034
m	101,196	60,022
C	79GB	26GB

(Table 4) Comparison of Related Work

논문	오프라인시간 복잡도	온라인	
		시간 복잡도	저장 용량
[1]	$O(m^2)$	$O(m^2)$	약 400GB
본 논문	$O(m^{3/2})$	$O(m^{3/2})$	약 50GB

였으며, 분석을 위해 Baldi 등의 논문[1], Gentry 등의 논문[7], 그리고 Cristofaro 등의 논문[11] 등을 참고하였다. 대문자-O 표기법을 사용하여 논문 [1]과 연산 효율성 및 저장 용량 측면에서 비교 분석한 것은 아래 (Table 4)과 같다. (Table 4)에서 보는 것과 같이 우리의 기법은 연산 효율성 측면과 저장 용량 측면에서 모두 기존의 연구와 비교하여 상대적으로 뛰어나다.

4.4 효율성 분석 - 1% 추출 조사

논문 [1]에서는 효율성을 극대화하기 위해 모든 DNA를 전수 조사하는 방법 외에도, 전체 DNA의 약 1%만을 추출한 데이터로부터 비교 분석한 결과를 포함하였다. 이것은 전체 DNA의 약 1%만으로도 전체 DNA를 대표할 수 있다는 근거로부터 기인한 것이며, 단지 전체 DNA 대신에 1%만의 DNA 정보를 가지고 비교한 분석이다. 해당 방법을 사용한다면, 우리의 프로토콜을 사용하였을 때의 암호문은 약 0.5GB로 채 1GB가 되지 않는다. 하지만 논문 [1]에서의 암호문은 1% 추출 조사 방법을 사용한다고 하더라도 여전히 4GB 정도가 됨을 알 수 있다.

5. 결 론

본 논문에서는 프라이버시를 보호하는 DNA 매칭 프로토콜을 제안하였다. 제안하는 프로토콜은 대용량에서 프라이버시를 보호하는 내적 연산 프로토콜을 이용하며 정확성, 기밀성, 프라이버시를 만족하는 안전하면서도 효율적인 프로토콜이다.

참고문헌(Reference)

[1] P. Baldi, R. Baronio, E. Cristofaro, P. Gasti, and G. Tsudik, "Countering GATTACA: Efficient and Secure Testing of Fully-Sequenced Human Genomes", CCS '11, pp. 691-702, Oct. 2011.

https://doi.org/10.1145/2046707.2046785

[2] C. Gentry, "Fully Homomorphic Encryption using Ideal Lattices," STOC '09, pp. 169-178, May 2009. https://doi.org/10.1145/1536414.1536440

[3] D. Micciancio and C. Peikert, "Trapdoors for Lattices: Simpler, Tighter, Faster, Smaller," Advances in Cryptology, EUROCRYPT '12, LNCS 7237, pp. 700-718, Apr. 2012. https://doi.org/10.1007/978-3-642-29011-4_41

[4] D. Cash, D. Hofheinz, E. kiltz, and C. Peikert, "Bonsai Trees, or How to Delegate a Lattice Basis," Advances in Crytology, EUROCRYPT '10, LNCS 6110, pp. 523-552, May 2010. https://doi.org/10.1007/s00145-011-9105-2

[5] S. Yamada, "Adaptively Secure Identity-Based Encryption from Lattices with Asymptotically Shorter Public Parameters," Advances in Cryptology, EUROCRYPT '16, LNCS 9666, pp. 32-62, May 2016. https://doi.org/10.1007/978-3-662-49896-5_2

[6] G. Noh, D. Hong, J.O. Kwon, and I.R. Jeong, "A Strong Binding Encryption Scheme from Lattices for Secret Broadcast," IEEE Communications Letters, Vol 16, No. 2, pp. 781-784, Jun. 2012. https://doi.org/10.1109/LCOMM.2012.041112.112495

[7] C. Gentry, S. Halevi, and V. Vaikuntanathan, "A Simple BGN-Type Cryptosystem from LWE," Advances in Cryptology, EUROCRYPT '10, LNCS 6110, pp. 506-522, May 2010. https://doi.org/10.1007/978-3-642-13190-5_26

[8] C. Gentry, A. Sahai, and B. Waters, "Homomorphic Encryption from Learning with Errors: Conceptually-Simpler, Asymptotically-Faster, Attribute-Based," Advances in Cryptology, CRYPTO '13, LNCS 8042, pp. 75-92, Aug. 2013. https://doi.org/10.1007/978-3-642-40041-4_5

[9] Dan Boneh, Craig Gentry, Sergey Gorbunov,, Shai Halevi, Valeria Nikolaenko, Gil Segev,, Vinod Vaikuntanathan, and Dhinakaran Vinayagamurthy, "Fully Key-Homomorphic Encryption, Arithmetic Circuit ABE and Compact Garbled Circuits," Advances in Cryptology, EUROCRYPT '14, LNCS 8441, pp. 533-556, May 2014. https://doi.org/10.1007/978-3-642-55220-5_30

[10] Chris Peikert and Sina Shiehian, "Multi-key FHE from LWE, Revisited," TCC '14-B, LNCS 9986, pp. 217-238, Nov. 2016. https://doi.org/10.1007/978-3-662-53644-5_9

[11] Emiliano De Cristofaro, Paolo Gasti, and Gene Tsudik, "Fast and Private Computation of Cardinality of Set Intersection and Union," CANS '12, LNCS 7712, pp. 218-231, Dec. 2012. https://doi.org/10.1007/978-3-642-35404-5_17

◐ 저 자 소 개 ◑

노 건 태(Geontae Noh)
2008년 고려대학교 산업시스템정보공학과(공학사)
2010년 고려대학교 정보경영공학과(공학석사)
2014년 고려대학교 정보보호학과(공학박사)
2014년~2017년 고려대학교 정보보호연구원 연구교수
2017년~현재 서울사이버대학교 정보보호학과 조교수
관심분야 : 암호 이론, 프라이버시 향상 기술, 데이터베이스 보안
E-mail : gnoh@iscu.ac.kr