Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.29, NO.2, Apr. 2019 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2019.29.2.299

영지식 증명을 활용한 블록체인 기반 개인정보 관리 기법*

이 정 혁,[1†] 황 정 연,[2] 오 현 옥,[1] 김 지 혜[3‡]
[1]한양대학교, [2]한국전자통신연구원, [3]국민대학교

Personal Information Management System with Blockchain Using zk-SNARK*

Jeong-hyuk Lee,[1†] Jung Yeon Hwang,[2] Hyun-ok Oh,[1] Ji-hye Kim[3‡]
[1]Hanyang University,
[2]Electronics and Telecommunications Research Institute,[3]Kookmin University

요 약

현재 개인정보의 활용가치가 높아짐에 따라, 개인정보를 제공하는 방법에 대한 논의가 활발하게 이루어지고 있다. 현재 가장 일반적인 개인정보 제공방법 중 하나로 개인정보를 활용하는 집단이 개인에게 동의를 얻어 개인정보를 사용하는 방법이 있다. 그러나 위의 방법은 두가지 문제점을 갖고 있는데, 첫째 개인정보 활용을 위해 기관에서 필요로 하는 정보 이상의 정보가 노출되고 있는점, 둘째 기업에서 개인정보를 요청 할 때마다 신뢰기관(Trusted Party)이 해당 정보에 대한 인증정보를 기업에 제공해야 하는 문제점이 있다. 위의 문제점을 해결하기 위해 본 논문에서는 zk-SNARK (zero-knowledge Succinct Non-interactive ARgument of Knowledge)기법과 블록체인을 사용하여 프라이버시 보호형 개인정보 관리기법을 제안한다. 프라이버시 보호형 개인정보 관리기법은 zk-SNARK를 통해 개인정보를 제공할 때 프라이버시를 보장하면서도 정보의 신뢰성을 보장할 수 있다. 또한 블록체인을 통해 데이터의 무결성을 보장하면서도 개인정보 데이터를 관리할 수 있으며 개인정보 공유를 기존의 인증방식보다 용이하게 수행할 수 있다.

ABSTRACT

As the utilization value of personal information becomes higher, discussions about providing personal information are being conducted actively. One of the most common methods of providing personal information is that a group obtains a personal information with a consent of individual. However, the above method has 2 problems. First, more information is exposed than the information required by organization for utilization of personal information. Second, trusted party should provide organization with an authentication of personal information whenever they require personal information. To solve these problems, we propose a personal information management system with blockchain using zk-SNARK(zero-knowledge

Received(12. 27 2018), Modified(02. 07. 2019), Accepted(03. 19. 2019)
* 본 연구는 과학기술정보통신부(No.2016-6-00599, 함수 서명 설계기법 및 응용기술 연구, No.2017-0-00661,영상 정보 프라이버시 침해 방지 및 인증기법, No.2018-0-01369, O2O 서비스를 위한 무자각 증강인증 및 프라이버시가 보장되는 블록체인 ID 관리 기술 개발)의 재원으로 정보통신기획평가원(Institute for Information & communications Technology Promotion)의 지원과 교육부(2017R1A2B4009903,2016R1D1A1B03934545)의 재원으로 한국연구재단(National Research Foundation of Korea) 이공분야기초연구사업의 지원과 과학기술정보통신부(2017R1A4A1015498)의 재원으로 한국연구재단(National Research Foundation of Korea) 이공분야기초연구사업의 지원을 받아 수행된 연구임.
† 주저자, ahoo791@hanyang.ac.kr
‡ 교신저자, jihyek@kookmin.ac.kr(Corresponding author)

Succinct Non-interactive ARgument of Knowledge) for privacy. Our proposal enables individuals to guarantee reliability of their information and protect their privacy concurrently using zk-SNARK when they provid organization with their personal information. In addition, it is possible to manage the personal information data while ensuring the integrity of the data using blockchain and it is possible to share the personal information more conveniently than existing systems.

Keywords: zk-SNARK, blockchain, personal information management, NIZK

I. 서 론

기업에서 소비자에게 제공하는 재화가 다양해짐에 따라, 소비자의 재화별 잠재수요를 파악하는 것은 필수적 요소가 되었다. 기업은 소비자에게 개인정보를 요청하여 개인정보를 수집하고 이를 분석하는 것을 통해 소비자의 잠재수요를 파악할 수 있다. 그러나 소비자는 기업이 개인정보를 수집할 때, 자신의 개인정보가 어느 정도까지 수집되는지 파악하는 것이 어려운 측면이 있으며 설령 기업에서 수집하는 개인정보 내역을 공개할지라도 기업에서 제공하는 정보 수집내역을 신뢰 하는 것이 어렵다 [1]. 즉, 기업에서 필요 이상의 개인정보를 소비자에게 요구하더라도 소비자 입장에서 이를 인지하는 것이 매우 어렵기 때문에 소비자의 개인정보가 과도하게 노출되는 문제가 발생할 수 있다.

위와 같은 과도한 정보 노출 문제를 방지하는 이상적인 정보제공 방법으로 기업에서 필요로 하는 정보만 개인이 산출하여 인증된 형태로 제공하는 것을 생각해 볼 수 있다. 예를 들어 기업에서 개인의 성인 여부 정보가 필요하다면 개인은 정확한 나이 정보를 제공할 필요 없이 성인 여부 정보만 인증된 형태로 제공하면 되는 것이다. 하지만 신뢰성 있는 정보의 제공을 위해서 정보에 대한 인증을 사용하는 경우 정보의 인증을 위해 제3의 신뢰기관이 필요하다. 그러나 개인정보를 활용할 때마다 신뢰기관에서 인증정보를 발급하는 것은 비효율적이다. 따라서 신뢰기관에 과도한 오버헤드가 부여되지 않도록 하는 새로운 개인정보 관리 방법이 필요하다.

본 논문에서는 개인이 기업에 개인정보를 제공할 때 기업이 필요한 정보만 제공하면서 신뢰기관에 부여되는 오버헤드가 많지 않도록 하는 새로운 개인정보 관리기법을 제안한다. 개인이 자신의 개인정보를 기업에게 제공할 때 개인정보 자체를 기업에 제공함과 동시에 해당정보에 대한 정확성을 신뢰기관으로부터 받아서 개인정보에 대한 신뢰성을 확보하는 방법과 달리 본 제안에서는 미리 인증된 개인정보에 대해서 개인이 실제로 개인정보를 활용할 때 기업에 필요에 맞게 정보를 가공하여 제공하는 것을 제안한다. 정보를 가공하여 제공하는 경우에도 해당 정보에 대한 정확성을 제공해야 하는데 이를 영지식 증명을 사용하여 해결하는 방법을 제안 한다 [2][3][4][5].

문헌 [2][3][4][5]등의 영지식 증명기술은 증명에 필요한 어떠한 정보도 드러내지 않고 Statement에 대한 정합성을 증명하는 기술이다. 특히, [2][3][4][5]의 경우 영지식 증명을 위해 상호작용(Interaction)이 필요 없이 간결한 증명(Succinct) 생성이 가능하기 때문에 소비자가 효율적으로 가공데이터에 대한 증명을 생성하는 것이 가능하다. 즉, 위의 영지식 증명을 통해 소비자는 제3의 신뢰기관을 거치지 않은 채 자신의 개인정보를 드러내지 않고 개인정보를 가공한 데이터를 기업에게 신뢰성 있게 제공할 수 있다.

인증이 필요한 개인정보를 관리하는 방법으로 인증기관에서 개인정보를 서버에 각각 저장하고 개인은 위 정보를 요청해 사용하는 방법을 생각 할 수 있는데, 개인정보를 단일 기관의 서버에 저장하는 것은 서버가 악의적인 공격자에 의해 공격당해 관리권한이 탈취되면 서버 내의 정보가 조작될 수 있는 위험성이 있다. 또한 서버 내에서 개인정보를 관리할 때 데이터 공유를 위해서는 서버 전체를 다시 구축하여 정보를 복제하거나, 서버 관리자에게 데이터를 요청하여 사용해야하는데 서버 전체를 복제하는 것은 비용적 측면에서 효율적이지 않고, 현행과 같이 서버 관리자에게 데이터를 요청하여 사용하는 것은 서버관리자를 완전히 신뢰할 때 가능하다는 점에서 한계점이 있다. 따라서 가공데이터를 생성하기 위한 개인정보를 관리하기 위한 방법 중 하나로 서버가 아닌 블록체인을 활용할 수 있다.

본 논문에서는 개인정보의 조작을 방지하고 정보 공유의 유용성을 위해 블록체인을 사용하여 개인정보를 공유하는 방법을 제안한다. 개인정보를 인증하는 신뢰기관은 개인정보를 해시한 값과 개인정보에 대한 암호문을 블록체인에 기록한다. 이후 기업이 개인의 개인정보를 활용할 필요가 있을 때 개인은 암호화된 개인정보를 복호화 한 후 자신의 개인정보를 기업에

서 필요로 하는 정보로 가공한다. 이후 개인은 가공한 정보에 대해 증명값을 추가한다. 개인은 가공된 정보와 해당 정보에 대한 증명값을 기업에 제공하는 것을 통해 가공정보에 대한 신뢰성을 기업에게 제공할 수 있다.

본 논문은 다음과 같은 구성을 갖는다. 2장에서는 배경에 대하여 서술하고, 3장에서는 본 연구와 관련된 관련연구, 4장에서는 제안기법의 정의에 대해 서술하며 5장에서는 구체적인 제안기법에 대해 서술한다. 6장에서는 제안기법의 안전성증명에 대하여 서술하며 끝으로 7장에서는 결론을 서술한다.

II. 배 경

2.1 zk-SNARK

2.1.1 zk-SNARK 개요

zk-SNARK(zero-knowledege Succinct Non-interactive ARgument of Knowledge)는 Statement의 올바름에 대하여 어떠한 비밀정보도 드러내지 않고 증명할 수 있는 기술을 뜻한다[2][3][4][5]. zk-SNARK는 영지식 증명이 가능하며 특히 생성되는 증명의 크기가 증명 수식(N)의 $\log(N)$이하의 크기를 가진다. 또한 증명자(Prover)와 검증자 (Verifier) 간 상호작용 (Interaction)이 존재하지 않으며, 지식(Knowledge)를 가진사람만이 증명을 생성할 수 있는 특성을 가진다. zk-SNARK는 특정한 함수를 하나의 곱셈과 다수의 덧셈으로 구성되는 회로 (Circuit)로 구성하는 것을 전제로 두고 있으며, 회로데이터를 R1CS(Rank-1 Constraint System)[2][3]로 구성한 뒤 R1CS 데이터를 QAP(Qaurdratic Arithmetic Program) 또는 SAP(Square Arithmetic Program)의 형태로 만들어 하나의 함수에 대한 공용변수(CRS)를 생성한다.

2.1.2 NIZK (Non-Interactive Zero-Knowledge Arguments of Knowledge) 정의

R을 Relation Generator라고 할 때, 4개의 Non-Interactive Zero-Knowledge Arguments of Knowledge 알고리즘 $Setup$, $Prove$, $Verify$, $SimProve$은 Perfect Completeness, Computational Soundness, Zero-Knowledge를 만족한다. 먼저 각 알고리즘의 정의는 다음과 같다.

- $(crs,\tau) \leftarrow Setup(R)$: $Setup$알고리즘은 Relation R을 입력으로 공용변수 crs와 시뮬레이션 트랩도어(Simulation Trapdoor) τ를 출력한다.
- $\pi \leftarrow Prove(crs, \phi; w)$: $Prove$알고리즘은 공용변수 relation R에 대한 crs와 $(\phi,w) \in R$을 입력으로 증명 π를 출력한다.
- $0/1 \leftarrow Verify(crs,\phi,\pi)$: $Verify$ 알고리즘은 공용변수crs, 인스턴스 ϕ, 증명 π를 입력으로 증명이 맞으면 1, 올바르지 않으면 0을 출력한다.
- $\pi \leftarrow SimProve(crs,\tau,\phi)$: 시뮬레이터 $SimProve$는 공용변수 crs, 시뮬레이션 트랩도어(Simulation Trapdoor) τ, 인스턴스 ϕ를 입력으로 π를 출력한다.

R : Relation
crs : 공용변수(Common Reference String)
τ : 트랩도어(Trapdoor)
π : 증명(Proof)
ϕ : 인스턴스 (Instance)
w : 위트니스(Witness)

각 알고리즘이 만족하는 특성은 다음과 같다.

- Perfect Completeness : Perfect Completeness는 올바른 Statement가 주어졌을 때, 위트니스(Witness)를 알고있는 증명자(Prover)는 검증식(Verify)을 항상 통과할 수 있음을 의미한다. Perfect Completeness는 다음의 확률을 만족한다.

$$\Pr\left[\begin{matrix}(crs,\tau) \leftarrow Setup(R); \pi \leftarrow Prove(R,crs,\phi,w): \\ Verify(R,crs,\phi,\pi)=1\end{matrix}\right]=1$$

이때, 트랩도어 τ값은 증명자와 검증자 모두에게 드러나지 않고, 아래에 서술할 증명 시뮬레이션을 위해 사용된다.

- Perfect Zero-Knowledge : 제대로 된 인스턴스와 위트니스를 알고 $Prove$함수를 통해 만들어진 π와 위트니스를 모르고 $SimProve$로 만든 π를 공격자 A가 구분할 수 없다. Perfect Zero-

Knowledge는 다음을 만족한다.

$$\Pr[(crs,\tau)\leftarrow Setup(R);\pi \leftarrow Prove(R,crs,\phi,w):\ A(R,crs,\tau,\pi)=1] = \Pr[(crs,\tau)\leftarrow Setup(R);\pi \leftarrow SimProve(R,\tau,\phi):\ A(R,crs,\tau,\pi)=1]$$

- Computational Soundness :

L_R을 R에 속하는 위트니스에 부합하는 Statement로 구성된 Language라고 할 때 다음의 확률을 만족하고 이때 ε이 negligible이면 Computationally Sound하다고 정의한다.

$$\Pr\left[\begin{matrix}(crs,\tau)\leftarrow Setup(R);(\phi,\pi)\leftarrow A(R,crs):\\ \phi\notin L_R \text{ and } Verify(R,crs,\phi,\pi)=1\end{matrix}\right]<\varepsilon$$

2.2 Blockchain

블록체인[6][7][8][9]은 거래정보를 기록한 원장을 특정기관의 중앙서버가 아닌 P2P네트워크에 분산하여 거래참가자가 공동으로 관리하는 기술이다. 블록체인은 거래정보가 담긴 각각의 블록을 해시함수를 사용하여 연결한다. 블록체인 내의 블록이란 블록체인의 원소 개념으로 각각의 블록은 블록헤더와 거래정보 및 부가정보로 구성된다. 블록체인은 읽기 및 쓰기 권한을 제한하는 정도에 따라 공개블록체인과 사설블록체인으로 분류될 수 있는데, 본 논문에서 사용하는 블록체인은 읽기권한이 모두에게 주어져 있지만 쓰기 권한은 개인정보를 인증하는 제3의 신뢰기관에게 주어지는 제한적 공개 블록체인을 가정한다.

2.3 Encryption

본 논문에서는 사용자가 필요한 시점에 자신의 정보를 도출할 수 있도록 자신의 개인정보를 암호화하여 블록체인 내에 보관한다. 암호화 알고리즘은 다음과 같은 알고리즘으로 구성된다.

$(pk, sk)\leftarrow Setup(k)$: 안전변수 k를 입력으로 비밀키 sk와 공개키 pk를 출력한다. 본 논문에서는 각 개인별로 공개키와 비밀키를 갖는 것을 가정한다.

$CT\leftarrow Encrypt(pk,M)$: 공개키 pk, 메시지 M을 입력으로 암호문 CT를 출력한다.

$M \leftarrow Decrypt(sk, CT)$: 비밀키 sk, 암호문 CT를 입력으로 암호문에 대한 복호문 M을 출력한다.

본 논문에서는 개인별로 다른 암호화키 및 복호화키를 가지는 것을 가정하며 이에 따라 ID에 대한 공개키와 비밀키를 pk_{ID} , sk_{ID}로 정의한다.

또한 본 논문에서 사용하는 암호화 알고리즘 ($Setup$, $Encrypt$, $Decrypt$)는 Semantic Security를 만족함을 가정한다. 본 논문에서 정의하는 Semantic Security는 다음과 같다.

- Semantic Security

A가 다항시간의 알고리즘일 때 다음의 확률이 negligible의 확률을 갖는다. 즉, 공격자가 Ciphertext로부터 메세지를 구분할 수 있는 확률은 negligible 이다.

$$\left|\Pr\left[\begin{matrix}(pk,sk)\leftarrow Setup(k);(m_0,m_1)\leftarrow A(pk)\\ b\leftarrow\{0,1\};C\leftarrow Encrypt(pk,m_b);b'\leftarrow A(pk,C):b=b'\end{matrix}\right]-\frac{1}{2}\right|<\varepsilon$$

III. 관련연구

3.1 Zerocash

제로캐시(Zerocash)는 zk-SNARK를 사용한 응용 중 가장 먼저 상용화된 응용이다[8]. 제로캐시는 가장 널리 사용되고 있는 비트코인(Bitcoin)의 프라이버시 문제를 보완하기 위해 제안되었다. 비트코인의 모든 거래는 공개된 탈중앙화 블록체인에 기록되기 때문에, 사용자의 모든 거래정보가 모두에게 공개된다. 따라서 비트코인은 구조상 프라이버시 문제를 벗어날 수 없다. 하지만 제로캐시는 zk-SNARK를 기반으로한 DAP 기법(Decentralized Anonymous Payment Scheme)을 통해 사용자들이 정보를 노출하지 않고도 직접적인 거래를 할 수 있도록 하였다. 제로캐시는 사용자의 비밀정보에 대해 일방향연산을 수행하고 연산결과에 대하여 증명값을 추가함으로써 익명블록체인을 구현하였다. 즉 제로캐시의 익명거래는 zk-SNARK를 사용함으로써 사용자의 비밀정보를 드러내지 않고 각 사용자의 거래가 유효하다는 것을 증명할 수 있기 때문에 사용자들은 각 거래의 자금출처, 목적지, 거래금액을 숨긴 채 거래할 수 있게 되었다.

3.2 Hawk

이더리움 등 수많은 탈중앙화 암호화폐기반의 스마트 계약 시스템들은 제 3의 신뢰기관 없이 사용자들이 안전한 거래를 할 수 있도록 하였다[7]. 이는 블록체인을 활용한 암호화폐가 단순 화폐의 역할 뿐만 아니라 더욱 다양한 역할을 가질 수 있게 하였다. 하지만 모든 거래내역이 블록체인에 공개되어 있어 거래 프라이버시가 결여된 스마트 계약 시스템이라는 한계점을 가진다. 따라서 이러한 방법에서는 사용자의 개인정보가 블록체인을 통해 전파되고, 블록체인에 기록되기 때문에 개인정보를 전혀 보호할 수 없다. 이를 해결하기 위해 Hawk는 스마트 계약의 과정을 공개하지 않는 방법을 제안하였다[10]. Hawk에서는 계약에 필요한 연산이 공개되지 않지만 계약과정을 증명할 수 있는 zk-SNARK를 통해 거래의 타당성을 증명 할 수 있다. 이를 통해 Hawk에서는 사용자의 비밀정보가 블록체인에 노출되지 않도록 스마트 계약 연산을 수행할 수 있다.

3.3 Accountable privacy for decentralized anonymous payments

Accountable privacy for decentralized anonymous payments는 기존의 블록체인을 활용한 탈중앙화 암호화폐와 탈중앙화 암호화폐 기반의 스마트 계약 시스템들의 문제점이었던 사용자의 프라이버시를 보호하면서, 자금세탁과 같은 금융범죄에 암호화폐가 악용되는 것을 막기 위한 연구이다[11]. 제로캐시는 zk-SNARK를 기반으로 사용자들이 비밀적으로 직접거래를 가능하게 하였지만, 이는 정부기관 및 수사기관 역시 추적 및 조회를 할 수 없기 때문에 자본세탁 등의 범죄에 악용될 수 있다. 이는 결국 현금과 다를바가 없기 때문에 Accountable privacy for decentralized anonymous payments는 새로운 DAP(Decentralized Anonymous Payment)을 설계하여 네트워크 참여자의 프라이버시를 보장하면서 거래정책(Transaction Policy)을 강제하였다. 새로운 DAP system은 zk-SNARK 기법을 활용하여 모든 거래의 사용자 프라이버시를 보장하고 모든 거래가 거래정책을 준수하도록 강제하였다. 이를통해 Accountable privacy for decentralized anonymous payments는 탈중앙화 스마트계약 시스템의 네트워크에 참여하는 모든 사용자들이 특정한 거래정책을 준수하도록 강제함으로써 사용자들의 프라이버시를 지킴과 동시에 익명 블록체인이 범죄에 악용되는 것을 막을 수 있게 하였다.

IV. 개인정보 증명 블록체인

본 장에서는 제안하는 알고리즘을 정의하고, 알고리즘이 만족하는 안전성을 정의한다. 알고리즘은 크게 3개의 개체에 의해 수행된다.

제안 알고리즘에는 3개의 개체(Entity)가 존재함을 가정한다. 첫 번째 개체는 개인정보를 블록체인에 기록하는 신뢰기관(Trusted Party), 두 번째 개체는 개인정보를 기업의 필요에 맞게 가공하고 증명을 생성하는 개인(Individual Entity), 마지막으로 개인으로부터 받은 가공데이터를 검증하는 기업(Corporation)이 있다. 신뢰기관은 개인 혹은 공공기관이 될 수 있으며, 여러 신뢰기관에서 한 개인에 대한 정보를 암호화하여 인증된 형태로 블록체인에 기록하는 상황을 가정한다. 이후 개인은 블록체인에 기록된 자신의 개인정보를 기업이 요구하는 정보로 가공하여 기업에게 제공한다. 개인은 기업에게 정보를 제공하는 과정에서 zk-SNARK를 사용하며, 사용하는 zk-SNARK를 VC(Verifiable Computing)로 표기한다. 또한 본 논문의 제안기법에서 사용하는 블록체인은 읽기권한은 모두에게 주어져 있지만 쓰기 권한은 다수의 신뢰기관에게 제한되어 있는 블록체인을 가정한다.

4.1 개인정보 증명 블록체인 정의

개인정보 증명 블록체인은 다음과 같이 구성된다.

$ChainSetup(k)$: 신뢰기관(Trusted Party)에 의해 수행되는 알고리즘이며 안전변수 k를 입력으로 블록체인의 초기 블록(Genesis)을 생성하며, 블록체인의 접근권한 및 블록 등록권한을 설정한다.

$Register(ID,info,pk_{ID})$: 신뢰기관이 ID와 ID에 대한 개인정보 $info$, 암호화를 위한 암호키 pk_{ID}를 입력으로 $info$를 pk_{ID}로 암호화한 CT, $info$를 해시연산한 h와 ID를 출력하고 블록체인에 기록한다. 이때 기록되는 $CT, info, h$를 tx로 정의한다.

$Setup(f)$: 신뢰기관이 개인정보를 활용하는 함수 f를 입력으로 증명 및 검증에 사용되는 ek_f와 vk_f를 생성하여 출력한다.

$ProvePI(tx, sk_{ID}, ek_f, f)$: 개인(Individual Entity)에 의해 수행되는 알고리즘으로 ID, CT, h로 구성된 tx, 암호문을 복호화하기 위한 sk_{ID}, 함수 f 에 대한 증명 키 ek_f와 함수 f를 입력으로 함수 f에 대한 출력값 t와 수 f를 올바르게 수행했다는 것을 증명하는 π를 출력한다.

$VerifyPI(vk_f, tx, t, \pi)$: 기업(Corporation)에서 수행하는 알고리즘이며 $Prove$에서 출력한 π와 검증키 vk_f, 개인정보 해시 h, 함수결과 값 t를 입력으로 π를 검증한다. 검증결과가 맞으면 1, 아니면 0을 출력한다.

4.2 개인정보 증명 블록체인 안전성 정의

본 논문에서 제안하는 개인정보 증명 블록체인 기법은 다음 3가지 안전성을 만족한다.

1) 완전성(Completeness): 블록체인에 기록된 개인정보를 입력으로 올바른 정보 가공 과정을 통과한 출력에 대한 증명은 항상 검증식을 통과한다.

2) 무결성(Soundness) : 블록체인에 기록되지 않은 값으로 증명을 생성하거나(Invalid Input Case), 올바르지 않은 정보 가공과정을 통해 생성된 출력으로 생성한 증명(Invalid Output Case)이 검증식을 통과할 확률이 매우 낮다.

3) 데이터 프라이버시(Data Privacy)

블록체인에 기록된 암호화된 정보로부터 출력되는 가공된 정보와 증명값으로 추가적인 정보를 얻을 수 없다.

V. 제안 기법

제안기법은 Fig.1과 같은 과정으로 수행된다. 초기에 개인 혹은 개인정보 인증기관이 개인정보 $info$와 해시 계산에 필요한 r을 암호화하여 블록체인에 기록하며 이후 개인정보 활용 시 개인정보 증명을 위해 개인정보를 해시한 $H(ID\|info\|r)$를 암호문 CT와 함께 블록체인에 기록한다. 여기서 r은 해시의 안전성을 위해 추가한 랜덤변수를 의미한다. 또한 개인별로 다른 r이 부여된다. 이후 개인정보를 활용하는

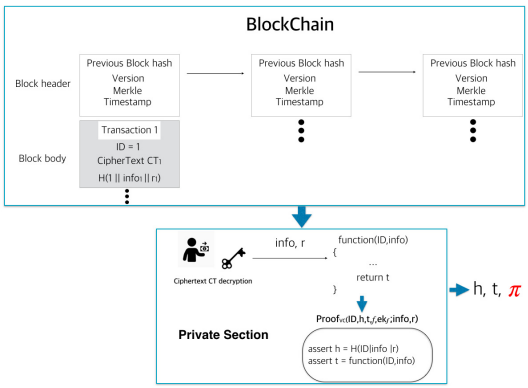


Fig. 1. Construction of our proposal

시점이 되면, 개인은 개인 수행 부분에서 암호문을 복호화한 후 얻은 개인정보 $info$와 ID를 입력으로 개인정보 활용을 위한 함수f를 수행한다. 그리고 함수 수행결과에 대한 증명을 위해 블록체인에서 개인정보의 해시데이터를 가져와서 증명의 입력으로 사용한다. 이때 증명함수에서는 두 가지를 증명하는데 첫 번째는 입력으로 사용하는 $H(ID\|info\|r)$값이 실제 $info$값을 통해 만든 값인지 증명하고, 두 번째로는 함수결과 t에 대하여 $info$값을 통해 만든 값인지 증명한다. 요약하면 증명함수에서는 함수 수행결과 t와 개인정보 해시데이터 $H(ID\|info\|r)$를 입력으로 수행결과에 대한 증명을 출력한다. 증명생성이 완료되면 개인은 개인정보 활용 기관에 자신의 증명과 함수 수행결과를 제출하여 이에 대한 신뢰성을 보장받을 수

k : Security parameter
ID : Personal ID
$info$: Personal information
pk_{ID} : Individual encrypt key to encrypt $info$, r
r : Randomization variable
f : Function
R : Relation
crs : Common Reference String
τ : Trapdoor
π : Proof
ϕ : Instance
ek_f : Evaluation key
vk_f : Verify key

Fig. 2-1. Variable definition

있다. 자세한 알고리즘은 Fig. 2-2.와 같다.

*ChainSetup*에서는 신뢰기관(Trusted Party)이 개인정보를 등록하기 위한 블록체인의 초기블록을 생성하는 과정이며, 안전변수 k를 입력으로 블록체인의 Genesis Block을 출력한다. 이후 모든 개인정보는 *ChainSetup*으로부터 생성된 블록체인에 기록된다.

*Register*에서는 개인의 *ID*와 개인정보 *info*, 암호문을 생성하기 위한 공개키 pk_{ID}를 입력으로 개인의 *ID*, 개인정보를 암호화한 암호문 *CT*, 개인정보를 해시 연산한 h를 블록체인에 기록한다. 이때 해시 h에는 선택공격 안전성을 위한 랜덤 r이 입력으로 사용된다. 또한 ID, CT, h를 tx라 정의한다.

개인이 새로운 정보를 등록할 때에는 새롭게 정보를 암호화하고 또 다른 해시를 생성하여 등록하며 개인정보를 가져오는 것은 최신의 블록에서 가져오는 것을 가정한다.

*Setup*에서는 함수f를 입력으로 Relation R을 정의하고 NIZK의 $Setup_{VC}$를 호출하여 개인정보를 활용하기 위한 증명을 생성하기 위한 ek_f와 증명을 검증하기 위한 vk_f를 생성한다. *Setup*은 개인이 아닌 신뢰기관에서 생성하는 것을 가정한다.

*ProvePI*는 개인이 수행하는 함수인데, tx, sk_{ID}, ek_f, f를 입력으로 하며 증명생성 이전에 블록체인에서 가져온 개인정보 암호문 *CT*를 복호화 하여 개인정보 *info*와 랜덤변수 r을 얻는다. 이후 개인은 *ID*와 *info*를 입력으로 함수 f에 대한 결과 값 t를 얻는다. 결과 값 t를 얻고 나면 개인은 ek_f, f 를 crs로 두고 tx, t를 인스턴스 ϕ로 zk-SNARK의 Prove함수 $Prove_{VC}$ 를 수행하여 증명 π와 함수 결과 값 t를 출력한다. 이때 $Prove_{VC}$에서는 $t = f(ID, info)$, $h = H(ID \| info \| r)$ 을 확인하며 이에 대한 증명을 생성한다.

*VerifyPI*는 기업이 수행하는 함수이며, 증명키 vk_f, ID, CT, h의 집합인 tx, 함수f에 대한 출력 t, 증명π 를 입력으로 vk_f, f를 crs로 두고 tx, t를 인스턴스 ϕ로 $Verify_{VC}$를 수행하여 $Prove_{VC}$에서 생성한 증명 π를 검증하고 검증결과가 1이면 검증값을 return하고 검증결과가 0이면 $\perp$을 return한다.

ChainSetup(k) :
Trusted party generates the initial block(Genesis block) of the blockchain by security parameter k

Register$(ID, info, pk_{ID})$:
$r \xleftarrow{\$} Z_p$
$M \leftarrow (info \| r)$
$CT \leftarrow Encrypt(pk_{ID}, M)$
$h \leftarrow H(ID \| info \| r)$
$tx \leftarrow (ID, CT, h)$
record tx to the blockchain

Setup(f) :

$$R = \left\{ \begin{array}{l} (\phi, w), \\ \phi = (tx, t), w = (info, r), tx = (ID, CT, h), \\ t = f(ID \| info), h = H(ID \| info \| r) \end{array} \right\}$$

$(ek_f, vk_f) \leftarrow Setup_{VC}(R)$
return ek_f, vk_f

ProvePI(tx, sk_{ID}, ek_f, f):
$M \leftarrow Decrypt(sk_{ID}, CT)$
$(info \| r) \leftarrow M$
$t \leftarrow f(ID, info)$
$crs \leftarrow (f, ek_f)$, $\phi \leftarrow (tx, t)$, $w \leftarrow (info, r)$
$\pi \leftarrow Prove_{VC}(crs, \phi; w)$
return t, π

VerifyPI(vk_f, tx, t, π) :
$crs \leftarrow (f, vk_f)$ $\phi \leftarrow (tx, t)$
$b \leftarrow Verify_{VC}(crs, \phi, \pi)$
if $b \equiv 1$ return b
if $b \equiv 0$ return $\perp$

Fig. 2-2. Our specific construction

VI. 안전성 증명

1) 완전성(Completeness) : *ProvePI*함수 내의 $Prove_{VC}$의 입력 h는 해시 함수를 통해 생성되어 같은 입력에 대해 항상 같은 출력을 얻기 때문에 올바른 *info*에 대하여 항상 같은 h를 얻는다.

또한 함수 f에 대하여 입력이 h이고 출력이 t임을 증명하는 π는 완전성을 만족하는 zk-SNARK 알고리즘 $Prove_{VC}$로부터 생성되어 완전성을 보장할 수 있고 위의 π가 *ProvePI*의 출력으로 사용되기 때문에 *ProvePI* 알고리즘 역시 완전성을 만족할 수 있다.

2) 무결성(Soundness) :증명 π의 조작이 불가함은 제안 알고리즘에서 사용하는 zk-SNARK의 무결성에 기반하며, π의 입력으로 사용되는 h의 조작 불가능성은 기반하고 있는 블록체인의 무결성에 기반한다. 즉, 함수 f에 대하여 입력이 h이고 출력이 O임을 증명하는 π는 계산적 무결성(Computational Soundness)을 만족하는 zk-SNARK 알고리즘 $Prove_{VC}$로부터 생성되어 계산적 무결성을 보장할 수 있고 위의 π가 *ProvePI*의 출력으로 사용되기 때문에 *ProvePI* 알고리즘 역시 계산적 무결성을 만족한다.

3)데이터 프라이버시(Data Privacy) : 블록체인에 기록된 암호문 *CT*의 비밀성은 암호문을 생성하는 암호화 기법의 선택 평문공격 안전성(CPA-security)에 기반하기 때문에 암호문 *CT*로부터 다른 비밀정보 $info_1$과 $info_2$를 구분할 수 있는 확률이 매우 낮으며(negligible probability) 또한 블록체인에 기록된 해시 값으로부터 개인정보 *info*값을 도출하는 것은 사용하는 해시의 일방향성(Onewayness)에 기인하여 매우 낮은 확률을 가진다. 또한 소비자가 생성한 증명 π로부터 개인정보 *info*값이 드러날 확률은 사용하는 증명 알고리즘의 영지식성(Zero-Knowledge)에 기반 하여 매우 낮은 확률을 지닌다. 따라서 위의 세 가지 경우를 고려했을 때, 암호문과 가공데이터, 증명으로부터 추가적인 개인정보 값이 드러날 확률이 매우 낮다. 이를 구체화하면 다음의 식과 같다.

A_{enc}를 두 개의 암호문으로부터 메시지를 구분하는 공격자, A_{hash}를 해시 값으로부터 해시의 입력값을 도출하는 공격자, A_{zk}를 실제 프로토콜을 통해 생성한 증명과 시뮬레이션을 통해 생성한 증명을 구분하는 공격자라고 가정했을 때, adv_{CPA} 를 A_{enc}가 공격을 성공할 확률,adv_{hash}를 A_{hash}가 공격을 성공할 확률, adv_{zk}를 A_{zk}가 공격을 성공할 확률이라고 가정하면 adv_{CPA}의 확률은 본 논문에서 제안하는 암호화 기법의 Semantic Security에 의해 negligible이며, adv_{hash}의 확률은 해시 함수에서 제공하는 일방향성에 의해 negligble이다. 또한 adv_{zk}의 확률은 A_{zk} 공격자가 a와 b의 확률을 구분할 확률에 수렴한다.

$$\text{a) } \Pr[(crs,\tau)\leftarrow Setup_{VC}(R);\ \pi\leftarrow Prove_{VC}(R,crs,\phi,w):\ A(R,crs,\tau,\pi)=1]$$

$$\text{b) } \Pr[(crs,\tau)\leftarrow Setup_{VC}(R);\ \pi\leftarrow SimProve_{VC}(R,\tau,\phi):\ A(R,crs,\tau,\pi)=1]$$

즉, 실제 비밀정보 $Prove_{VC}$에서 위트니스 w를 가지고 생성되는 증명 π와 시뮬레이터 $SimProve_{VC}$를 통해 생성되는 증명 π가 같은 분포(Distribution)를 가질 때 두 확률이 같은 확률을 지니는데, 제안기법에서 사용하는 *VC*(Verifiable Computation) 증명함수는 Perfect Zero-Knowledge를 만족하고 $Prove_{VC}$를 통해 생성된 증명 π는 랜덤화 되어 있는 값이고 $SimProve_{VC}$를 통해 생성되는 증명 역시 랜덤 변수이기 때문에 같은 분포를 가져 실제 증명값과 시뮬레이터를 통해 생성된 증명을 구분할 확률은 0이다. 따라서 $adv_{CPA}, adv_{hash}, adv_{zk}$의 확률을 모두 더한 확률은 negligible의 확률을 지닌다. 즉, *ProvePI*함수의 출력 t,π로부터 개인정보 *info*에 대하여 얻어내는 공격자 A_{PI}가 공격을 성공할 확률 adv_{PI}는 다음의 확률을 가진다.

$$adv_{PI} \leq adv_{CPA} + adv_{hash} + adv_{zk}$$

따라서 본 논문에서 제안하는 알고리즘의 데이터 프라이버시의 공격확률은 negligible의 확률을 갖는다.

VII. 결 론

본 제안에서는 zk-SNARK를 통해 개인이 개인정보를 활용함에 있어 자신의 프라이버시를 보호함과 동시에 개인정보에 대한 신뢰성을 얻을 수 있음을 보였다.또한 개인정보를 블록체인에 암호화 하여 기록하기 때문에 데이터 관리에 있어 무결성을 보장할 수 있다.

또한 여러기관이 하나의 블록체인에 개인정보를 암호화하여 기록하고 개인이 필요한 시점에 블록체인에서 인증된 데이터를 받아 가공하여 기업에 제공하고, 기업은 블록체인의 데이터와 개인으로부터 받은 증명을 통해 가공데이터의 신뢰성을 확보할 수 있다. 블록체인에 데이터를 기록하기 때문에 신뢰기관, 개인, 기업 간 데이터 공유가 용이한 장점을 지니고 있다.

본 제안의 한계점으로 현재 개인이 개인정보 데이터를 갱신하면, 데이터의 갱신 여부가 블록체인에 기

록될 수 밖에 없는데 이는 개인정보가 자주 바뀌는 사람의 프라이버시를 침해할 수 있는 가능성이 존재한다. 따라서 데이터 갱신 여부를 드러내지 않고 개인정보를 제공할 수 있는 방법에 대해 향후연구가 필요하다고 사료된다.

References

[1] Jones William, "Personal information management," Annual review of information science and technology 41.1 , pp.453-504, 2007.

[2] Parno, Bryan, et al, "Pinocchio: Nearly practical verifiable computation," 2013 IEEE Symposium on Security and Privacy. IEEE, 2013.

[3] Groth Jens, "On the size of pairing-based non-interactive arguments," Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 2016.

[4] Groth Jens, Mary Maller, "Snarky signatures: Minimal signatures of knowledge from simulation-extractable snarks," Annual International Cryptology Conference. Springer, Cham, 2017.

[5] Groth Jens, et al, "Updatable and universal common reference strings with applications to zk-SNARKS," Annual International Cryptology Conference. Springer, Cham, 2018.

[6] Nakamoto Satoshi, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[7] Wood Gavin, "Ethereum: A secure decentralised generalised transaction ledger," Ethereum project yellow paper 151, pp. 1-32, 2014.

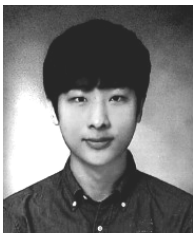
[8] Sasson, Eli Ben, et al, "Zerocash: Decentralized anonymous payments from bitcoin," 2014 IEEE Symposium on Security and Privacy (SP). IEEE, 2014.

[9] Zyskind Guy, Oz Nathan, "Decentralizing privacy: Using blockchain to protect personal data," Security and Privacy Workshops (SPW), 2015 IEEE. IEEE, 2015.

[10] Kosba Ahmed, et al, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," 2016 IEEE symposium on security and privacy (SP). IEEE, 2016.

[11] Garman, Christina, Matthew Green, and Ian Miers, "Accountable privacy for decentralized anonymous payments," International Conference on Financial Cryptography and Data Security, Springer, Berlin, Heidelberg, 2016.

〈저자소개〉



이 정 혁 (Jeong-hyuk Lee) 학생회원
2016년: 한양대학교 정보시스템학과 졸업 (학사)
2016년~현재: 한양대학교 정보시스템학과 석 박사통합과정
〈관심분야〉 암호이론, 블록체인

황 정 연 (Jung Yeon Hwang) 정회원
1999년 2월: 고려대학교 수학과 졸업
2003년 2월: 고려대학교 정보보호대학원 석사
2006년 8월: 고려대학교 정보보호대학원 박사
2009년 5월~현재: 한국전자통신연구원 책임연구원
〈관심분야〉 암호이론, 프라이버시 강화 암호 기법, 바이오 암호 시스템

오 현 옥 (Hyun-ok Oh) 정회원
1996년: 서울대학교 컴퓨터 공학과 학사
1998년: 서울대학교 컴퓨터 공학과 석사
2003년: 서울대학교 컴퓨터 공학과 박사
2005년: UC Irvine 연구원
2008년: ARM Inc. Staff S/W Engineer
2008년~현재: 한양대학교 공과대학 정보시스템 학과 부교수
〈관심분야〉 암호학, 비휘발성 메모리, 설계 자동화, 실시간 분석

김 지 혜 (Ji-hye Kim) 종신회원
1999년 2월: 서울대학교 컴퓨터공학부 졸업. 2003년 2월 서울대학교 전자컴퓨터공학부 석사
2008년 8월: UC Irvine Dept.of Information & Computer Science박사
2008년 9월~2008년 11월: UCIrvine Post Doc.
2008년 11월~2011년 8월: 서울대학교 수학연구소 Post Doc.
2011년 9월~현재: 국민대학교 전자공학부 부교수
〈관심분야〉 정보보호기술, 응용암호, 분산시스템