환자 익명성 보장을 위한 블록체인 알고리즘

조영복[1*] · 우성희[2]

[1]대전대학교 · [2]한국교통대학교

Block Chain Algorithm to Ensure Patient Anonymity

Young-bok Cho[1*] · Sung-hee Woo[2]

[1]Daejeon University, [2]Korea Transport National University

E-mail : ybcho@dju.ac.kr

요 약

본 논문에서는 블록체인 기반의 환자익명성 보장을 위한 알고리즘을 제안하였다. 환자의 익명성과 의사와 환자 사이 연계 불가성 지원을 위해서 신원증명용 주소와 별개로 스텔스 주소를 사용하였다. 또한 의무기록을 사용함에 있어 그 해쉬 값 등 사용 정보를 블록에 입력하여 의무기록의 무결성과 투명성을 보장 한다.

ABSTRACT

In this paper, we propose an algorthm for ensuring patient anonymity based on block chaining. For the anonymity of the patient and the inability to connect between the doctor and the patient, we used the stealth address separately from the identification address. Also in using the medical record, the use information such as the hash value is inputted into the block to guarantee the integrity and transparency of the medical record.

키워드

블록체인, 의료정보, 익명성, 프라이버시. 무결성

Ⅰ. 서 론

최근 의료정보의 전산화 및 네트워크화 다량의 개인정보가 수집, 보관되어 여러 관련 기관이 공유할 수 있게 됨에 따라 환자의 사생활 보호 측면에서 새로운 문제점이 야기되고 있다. 개인 정보는 개인의 신체, 신념, 사회적 지위, 신분 등과 같이 개인의 인격 주체성을 특정 짓는 사항으로서 그 개인의 동일성을 식별할 수 있게 하는 일체의 정보라고 할 수 있고, 반드시 개인의 내밀한 영역이나 사적인 영역에 속하는 정보에 국한되지 않고, 공적 생활에서 형성되었거나 이미 공개된 정보까지 포함하고 있다[1,2,3]. 또한, 의료 정보란 의료 제공의 필요성을 판단하기 위해서 또는 의료 행위를 통하여 수집된 자료 및 이 자료들을 기초로 하여 연구, 분석된 정보들을 포괄하는 것으로 진단과 치료행위, 치료 후의 관찰 등을 포함하여 의료 행위의 전 과정에서 수집된 환자의 건강상태 등에 관한 정보이다. 위 정의들을 기초로 판단해 볼 때, 건강과 관련된 개인 정보는 개인의 외면적인 부분뿐만 아니라 내면, 즉 심리적 정보까지 포함된 매우 민감한 정보이다. 또한 건강과 관련된 개인 정보는 개인의 외면적인 부분 뿐 아니라 내면, 즉 심리적 정보까지 포함된 매우 민감한 정보이다2,4]. 따라서 악의적인 내부 유출자 또는 해커는 이러한 업체들이 돈을 들여서라도 개인 의료기록을 구매하려한다는 점을 착안하여 유출 또는 탈취를 하여 금전적인 이익을 취하려 하고 있다. 본 논문에서는 의료정보 공유를 위협하는 위협원과 위협을 살펴보고, 안전한 의료정보공유를 위한 법제적 보안 요

* speaker

구사항과 기술적 요구사항을 분석하여 시스템이 갖추어야할 보안 요구사항을 도출하고 클라우드 기반의 환자의 프라이버시가 지켜지는 안전한 의료정보 공유 시스템을 제안한다.

II. 관련연구

2.1 블록체인

블록체인(Blockchain)은 Peer-to-Peer방식을 기반으로 관리하고자 하는 데이터를 블록(Block)이라는 분산 데이터베이스에 저장하고, 이를 각 블록들 간에 체인이 형성되어 있어서 누구도 임의로 수정할 수 없고, 누구나 변경의 결과를 열람할 수 있는 분산 컴퓨팅 기반의 데이터 위/변조 방지 기술이다.[5] 블록체인의 모든 트랜잭션은 네트워크의 모든 당사자가 있는 분산 원장에 저장된다. 블록체인 기술은 신뢰할 수 없는 트랜잭션을 도입하는 신뢰되는 제 3자 없이 분산 시스템을 구현한다. 이 때, 해당 트랜잭션의 소유권 및 무결성은 신뢰기관이 아닌 암호학적 기술을 통해서 입증이 된다. 따라서 블록체인의 특징은 분권화, 불변성, 그리고 비신뢰 환경에서의 합의 과정이라 할 수 있다. 블록체인의 종류는 크게 두 가지로 나눌 수 있다. 먼저 비트코인, 이더리움과 같은 누구나 네트워크에 참여할 수 있는 블록체인인 퍼블릭 블록체인(Public Blockchain)과 하나의 기관에서 독자적으로 사용하는 프라이빗 블록체인(Private Blockchain) 그리고, 허가된 여러 기관들이 컨소시움을 구성하여 참여하는 블록체인인 컨소시움 블록체인(Consortium Blockchain) 이 있다

2.2 블록체인을 이용한 프라이버시 보호

현재 Bitcoin 이나 Ethereum 같은 암호화폐 기반의 블록체인 기술에서는 사용자의 주소를 암호학적 해쉬의 결과로 사용하고 있다[4,5]. 이런 주소 값은 가명성(pseudonymity)에 기반을 둔다. 가명성은 사용자가 자신의 신원을 노출하지 않고 자원이나 서비스를 사용할 수 있지만, 그 사용에 대해서 책임을 추적할 수 있음을 보장한다. 이러한 가명성 기반의 주소체계에서는 연결성 순환 그래프(Directed Acyclic Graph) 형태로 분석하여 어떤 개인을 특정할 수 있다. 그러므로 가명성 기반의 주소체계는 환자의 프라이버시를 보호하는데 안전하지 않다. 이를 보완하는 대표적인 방법으로는 Monero의 Ring-CT 기술을 사용하는 방법과 ZCash의 영지식 기반의 ZKSNARK을 들 수 있다. 먼저 Ring-CT에서 사용되는 스텔스 주소나 링 서명같은 경우는 송수신자의 익명성(Anonymity) 및 비연결성(Unlinability)를 보장한다. 또한, ZK-SNARK은 거래 정보를 노출시키지 않고, 거래 내용의 무결성을 증명할 수 있는 방법이다

2.3 블록체인 기반의 의료 정보

의료 정보 보안을 위해 Bio Sensor에서 랜덤 키를 추출하여 대칭키 암호화 시스템을 사용하는 Biometric-based key distribution protocol을 제안하였다. 그러나 Bio Sensor는 작은 센서라서 충분한 랜덤성을 갖지 못한다는 문제점이 있다. 또한 블록체인 기반의 의료정보공유에 있어서 가장 중요한 점은 크게 프라이버시와 저장 공간의 두 가지가 대표적이다[6]. 먼저, 프라이버시 이슈는 앞서 클라우드 기반의 의료정보공유에서도 충분히 설명하였으므로 생략한다. 앞 서 언급하였듯이 의료정보공유에 있어서 블록체인을 이용할 경우 의료정보의 무결성과 투명성을 확보할 수 있다는 장점이 있다. 그러나 생성되는 의료정보는 몇 킬로바이트(KiloBytes)의 단순 텍스트 위주의 자료도 존재하지만, CT나 MRI 등 의료 영상의 크기가 수십 또는 수백 메가바이트(Mega Bytes)에 이르기도 하고, 유전체 정보의 경우에는 기가바이트(Giga Bytes)의 단위를 넘어서기도 한다. 따라서 환자의료정보에 관한 모든 내용을 블록체인에 넣는 것은 비효율적이며, 현실적으로 불가능한 일이다. 기존 이더리움이 지니는 컨센서스 확립 개념인 작업증명(Proof of Work) 방식을 적용한 의료계 블록체인을 에서는 기존의 금전적 지원 같은 보상 개념을 사용하기는 쉽지 않다

III. 제안방법

본 논문에서는 스는 환자의 익명성이 보장되는 클라우드 중심의 의료 시스템을 제안한다. 해당 시스템의 목적은 네트워크 환경에서 보내는 주소를 숨김으로써 환자의 익명성을 유지하는 것이다.

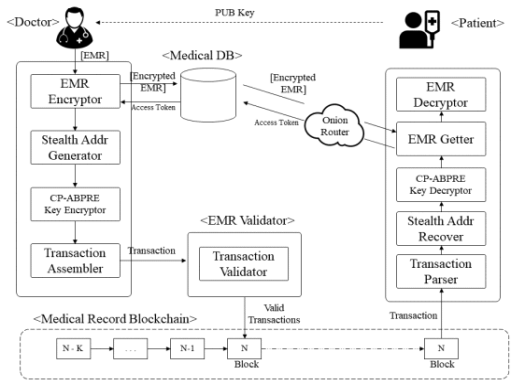

그림 1. 제안시스템

그림1은 환자 익명성 보장을 위한 의료정보공유 서비스의 네트워크 아키텍처로 는 주로 환자의 바이오센서(Bio Sensor), 싱크 노드(Sink Node), 클라우드 스토리지를 포함한 의료 서비스 플랫폼, 의료진의 장치 등 4개의 주요 개체로 구성된다. 제안

시스템에서 사용자 등록은 다음과 같이 KDC를 이용해 공개키와 개인키 쌍을 생성 하고 마스터 키는 다음과 같이 생성된다.

$$Y_a = \{y_a^1, y_a^2, ...y_a^n\}, X_a = \{x_a^1, x_a^2, ...x_a^n\} \quad (1)$$
$$S_a = \{s_a^1, s_a^2, ...s_a^n\},$$
$$(y_a^i, x_a^i, s_a^i \in GF(q),\ y_a^i = x_a^i \bullet P(1 \le i \le n)$$

환자 A 의 신분을 익명으로 유지하기 위해 (α, β) 익명성을 채택하였다. 먼저, 신체에 붙어 있는 바이오센서가 환자 A의 생체 데이터를 수집하고 데이터를 암호화한다. 다음에 암호화된 데이터를 환자 A 의 PA로 전송한다. PA는 메시지 m을 생성하고 PA는 신뢰할 수 있는 주변 PA를 주기적으로 검색하고 발견 시 전송한다. 전송 시마다 체인이 생성되고 생성된 체인을 이용해 최종적으로 병원에서 공개키를 이용해 암호화 한다. 이때 암호화되는 메시지 구조는 다음과 같다.

$$M = \{E(PU_{pr}, c_i), E(PU_{ci}, < c_a, \alpha >), E(PU_{ca}, < c_{di}, \beta, m)\} \quad (2)$$

환자 a는 전송할 장치에 M을 $1/\alpha$ 확률로 d_i가 발견되고, 발견될 때까지 M을 다른 장치로 연속적으로 전달한다. 마지막으로 d_i는 프록시 pr에 M을 전달한다. 다음으로 pr은 $E(PU_{pr}, c_i)$를 복호화 하고 다음 목적지 c_i를 검사한다. 그런 다음 pr은 M의 나머지 부분 M'을 c_i로 전송한다.

$$M' = \{E(PU_{ci}, < c_a, \alpha >), E(PU_{ca}, < c_d, \beta, m)\} \quad (3)$$

새로 생성된 체인 c_i는 M'에서 $E(PU_{ci}, < c_a, \alpha >)$을 복호화하고 c_i는 α개의 체인을 선택한다. 이때 구체적인 목적지를 숨기기 위해 c_a를 포함한 다른 $\alpha - 1$ 개의 클론에 식4와 같은 형태로 M''를 해독하고 m을 저장한다.

$$M'' = \{E(PU_{ca}, < c_d, \beta, m)\} \quad (4)$$

3.2 체인 간 통신

체인과 체인사이 다양한 의료정보를 전송할 수 있다. 이 단계에서는 각 체인간의 통신으로 c_a는 의사의 체인 c_d에 환자정보 m을 전송한다. c_a는 $\beta - anonymity$를 확보하기 위해 의사의 체인 c_d과 연결된 β개의 체인을 검색한다. 다음 c_a는 β개의 수신자 그룹내의 임의의 의사에서 생성된 체인 c_j를 선택한다. 이때 c_j가 선택된 이유는 악의적인 관리자가 존재하는 경우 실제 메시지가 어떤 목적지를 가지고 전달되는지 혼동을 주기 때문이다. c_a는 새로운 메시지 $\tilde{M}$을 식 (5)와 같이 c_i로 전송함으로 악의적인 관리자로 하여금 최종 목적지를 인식하기 어렵게 한다.

$$\tilde{M} = \{E(PU_{ci}, < c_j, E(PU_{c_j}, < c_d, \beta)\alpha >), E(PU_{cb}, m) >\} \quad (5)$$

Ⅳ. 실험 및 결과

본 논문에서는 환자의 의료정보 공유에 있어 제공되는 보안 요구사항은 다음과 같다.

- 의료정보에서 익명성(anonymity) 보장 및 연계불가성(Unlinkanility) : 저장되는 의무기록은 인가자를 제외하고는 누구의 것인지 확인 불가능해야 한다. 또한, 한 환자로부터 발생되었던 여러의무기록이 한 환자 것임이 알려지면 안 된다. 따라서 다른 사용자들이 의무기록 간의 연계를 할 수 없어야 한다.
- 의무기록 암호화 : 해커 및 데이터베이스 관리자를 통해서 불법적인 도청공격이 가능하므로, 반드시 의무기록은 암호화해서 저장이 되어야 한다. 사용자 신원인증은 사이비 진료 등 비인가자에 의한 의료행위 및 의무기록 공유를 방지하기 위해서 각 개체의 신원은 확인되어야 한다.
- 의무기록의 무결성 입증 : 저장, 공유된 의무기록은 불법적인 변경이 없었음이 확인되어야 하며, 시스템 내 누구나 확인 가능해야 한다.
- 응급상황 및 환자의무기록 공유 미동의 상황에서의 공유 : 응급상황이 발생하였을 때, 환자에게 동의를 받지 않았다 하더라도 최소한의 권한으로 제 3자와 의무기록공유가 가능해야 한다.

Algorithm 1 : 블록체인 set-up

Input: security parameter 1^λ
Output: public parameter pp, master key msk
1 choose random values $a, \alpha \in \mathbb{Z}_p^*$;
2 setup TCR hash functions
$H_1 : \{0,1\}^{2k} \to \mathbb{Z}_p^*$,
$H_2 : \mathbb{G}_T \to 0, 1^{2k}$,
$H_3 : \{0,1\}^* \to \mathbb{G}$,
$H_4 : \{0,1\}^* \to \mathbb{G}$,
$H_5 : \{0,1\}^k \to \mathbb{Z}_p^*$,
$H_6 : \{0,1\}^* \to \mathbb{G}$;
3 public parameter
$pp = (p, g, \mathbb{G}_T, \mathbb{G}, g_1, g^a, e(g,g)^\alpha, H_1, H_2, H_3, H_4, H_5, H_6)$;
4 master key g^α ;

본 논문은 프라이빗 블록체인 형태를 그대로 사용하므로 거래 장부를 분산한다는 점은 같으나, 작업증명과 채굴과정을 생략하고 대신에 감독기관이라는 관리주체가 거래의 승인 및 블록 생성권한을 보유한다.

해당하는 블록의 트랜잭션을 찾은 환자 A는 직접 해당 의무기록이 있는 데이터베이스에 접속을 시도한다. 이때 접속하기 전에 자신의 위치가 네트워크 주소가 노출되지 않게 하기 위해서 네트워크에서 익명성을 제공하는 네트워크를 사용하여 접속한다. 또한 환자 A는 해당 URL에 접속을 하여 데이터베이스에 암호화된 의무기록을 요청한다. 의료 데이터베이스는 저장되어 있던 CT_{EMR}를 반환하고, 환자 A는 Algorithm 2와 같이 복호화가 가능하다.

Algorithm 2 :환자A의 의료정보 복원

Input: a private key sk'_S, Encrypted Key $CK_{(\mathbb{M},\rho)}$
Input: Encrypted Record CT_{EMR}
Output: Electronic Medical Record M_A

1 Parse $CK_{M,\rho)} = (A_1, A_2, A_3, (B_1, C_1), \ldots, (B_l, C_l), (\mathbb{M}, \rho), D)$;
2 Check the original cipher text validity: $\leftarrow 1$ or $\perp$ iff
$e(A_2, g_1) = ?e(g, A_3)$,
$e(A_3, H_4(A_1, A_3, (B_1, C_1), \ldots, (B_l, C_l), (\mathbb{M}, \rho)) = ?e(g_1, D)$,
$e(\prod_{i\in I} B_i^{wi}, g) = ?e(A_2, g^a) \cdot \prod_{i\in I}(e(C_i^{-1}, H_3(\rho(i))^{wi}))$;
3 Compute $Z = e(A_2, K)/(\prod_{i\in I}(e(B_i, L) \cdot e(C_i, K_{\rho(i)}))^{wi})$
$= \frac{e(g^s, g^{a\cdot l} \cdot g^{\alpha})}{\prod_{i\in I} e(g^{u\cdot\lambda_i} \cdot H_3(\rho(i))^{-r_i}, g^t) \cdot e(g^{r_i}, H_3(\rho(i))^t))^{w_i}}$
$= \frac{e(g^s, g^{a\cdot t} \cdot g^{\alpha})}{e(g, g^{a\cdot t})^{\sum_{i\in I} \lambda_i \cdot w_i}}$
$= e(g^s, g^{\alpha})$;
4 Then $H_2(Z) \bigoplus A_1$
$= H_2(e(g^s, g^{\alpha})) \bigoplus (Key_{EMR}||\beta) \bigoplus H_2(e(g^s, g^{\alpha}))$
$= Key_{EMR}||\beta$;
5 Extract Key_{EMR} ;
6 Decrypt $EMR_A = S_{DEC}(CT_{EMR}, Key_{EMR})$;
7 Parse and Check the integrity EMR_A ;

본 논문에서는 환자의 익명성을 보장하고, 주치의와 환자 사이의 연계를 없애기 위해서 스텔스 주소를 만들어 사용한다. 의무기록을 공유하기 위해서는 환자는 자신의 주소를 블록체인으로 부터 찾기 위한 노력을 해야 한다. 그림1은 환자가 자신의 주소를 찾기 위한 스텔스 주소를 찾아내는 데 걸리는 시간을 의미한다. 본 논문에서는 스텔스 주소를 구현하여 복원 검증하기 위해 대표적인 NIST 타원 곡선인 Curve P-256 [52], Curve P-384, Curve P-512 를 중심으로 실험을 해 보았다. 범위는 블록 내에 찾아야 하는 스텔스 주소값을 1000개 범위를 지정하였다. 1000개의 주소를 복호화하고 검색 비교하기 위한 시간은 P-256의 곡선을 사용할 경우에는 약 4초의 시간이 걸리고, P-512 곡선을 사용하는 경우에는 약 9초의 시간이 걸린다. 환자는 블록이 전파될 때 마다 자신과 연계된 스텔스 주소를 찾는 노력을 한다면 그리 시간적으로 부담스럽지 않다

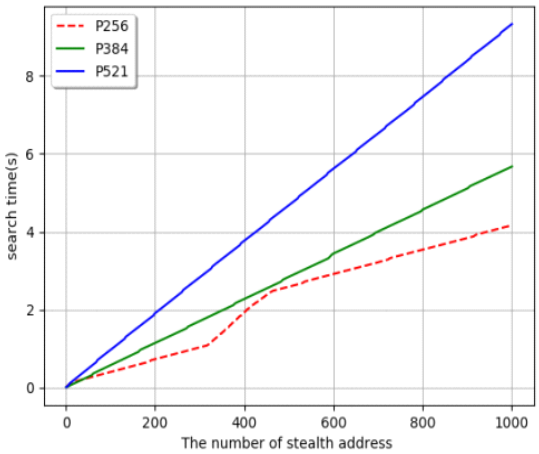

그림 2. 주소복원 시간

V. 결 론

본 논문에서는 블록체인 기반의 환자익명성 보장을 위한 알고리즘을 제안하였다. 환자의 익명성과 의사와 환자 사이 연계 불가성 지원을 위해서 신원증명용 주소와 별개로 스텔스 주소를 사용하였다. 또한 의무기록을 사용함에 있어 그 해쉬 값 등 사용 정보를 블록에 입력하여 의무기록의 무결성과 투명성을 보장하였으며, 사후 의무기록이 위변조 되는 것을 방지하였다

Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2018R1C1B 5083789).

References

[1] S.S Baek, S.H. Seo, and S.J Kim, “Preserving patient’s anonymity for mobile healthcare system in iot environment,” *The International Journal of Distributed Sensor Networks*, Vol. 12, No. 7, pp. 2171642, Jul. 2016.

[2] J. Kirk, “Premera, anthem data breaches linked by similar hacking tactics,” 2015

[3] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, “Medrec: Using blockchain for medical data access and permission management,” in *Proceeding of the 2016 2nd International Conference on Open and Big Data(OBD),* Vienna:Austria pp. 25 – 30, IEEE, 2016.

[4] X. Yue, H. Wang, D. Jin, M. Li, and W. Jiang, "Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control," *The Journal of medical systems*, Vol. 40, No. 10, pp. 218, Oct. 2016.
[5] Frost & Sullivan, Blockchain Technology in Global Healthcare, 2017-2025, 2017
[6] Gartner, Blockchain-Based Transformation: A Gartner Trend Insight Report[Internet]. Available : https://www.gartner.com/doc/3869696/blockchainbased-transformation-gartner-trend-insight