의료정보 공유 플랫폼을 위한 하이퍼레저 패브릭 응용 방안

이민기*, 민서현*, 지청민**, 홍만표*

*아주대학교 사이버보안학과, **아주대학교 정보컴퓨터공학과

Medical information sharing platform architecture based on Hyperledger Fabric

Minki Lee*, Seohyeon Min*, Cheongmin Ji**, and Manpyo Hong*

*Dept. of Cyber Security, Ajou University, **Dept. of Computer Engineering, Ajou University

요 약

정보통신 기술의 발전에 따라 데이터의 양이 많아지고 이를 공유하여 활용하고자 하는 연구가 활발히 진행되고 있다. 특히 의료분야에서는 의료정보를 활용하여 병원 간의 의료정보 공유 및 병원과 연구소 간의 의료정보 공유를 통해 환자의 진료비를 절감과 연구소의 의료정보를 공유를 통한 수월한 연구를 기대하고 있다. 하지만 의료정보 공유는 개인정보보호법, 의료법에 저촉되는 사안과 데이터 무결성, 외부로부터 들어올 수 있는 보안 취약점 때문에 쉽게 구현하기 어려운 실정이다[1]. 이를 블록체인을 통하여 해결하려고 할 때 데이터 무결성과 보안 취약점은 해결할 수 있으나 블록체인에 기록된 데이터는 삭제를 하지 못하기 때문에 개인정보 삭제권에 관한 법률이 저촉되는 문제점이 발생한다. 따라서 본 논문에서는 하이퍼레저 패브릭의 private data collection을 이용하여 분산된 데이터베이스, 데이터 무결성, 데이터 기밀성, 데이터 삭제권을 보장하는 플랫폼을 구상하였다.

I. 서론

현재 의료기관은 시스템 간 진료정보 교류가 제한돼 있다. 환자가 병원을 옮길 때마다 CD나 문서 등의 형태로 다시 제출해야 한다. 2014년 기준 CT, MRI 중복 촬영으로 월 평균 16억원이 낭비된다[1]. 이러한 문제를 해결하기 위해 의료정보를 공유하려고 하는 시도를 하고 있다. 예를 들어 메디블록과 같이 환자의 개인정보 주체를 보장하면서 환자, 의료공급자, 의료 연구자에게 자유로운 데이터 공유를 보장하기 위한 블록체인 플랫폼이 존재한다.

블록체인은 데이터 무결성을 위해 데이터를 분산화하고 정해진 알고리즘의 합의에 따라 유지하고 동기화하는 시스템을 말한다. 블록체인 네트워크 참여에 제약이 없는 시스템을 비허가형 또는 퍼블릭 블록체인이라 하고 블록체인 네트워크 참여 시 검증을 받는 과정이 있다면 허가형 또는 프라이빗 블록체인이라 한다. 비트코인, 이더리움과 같은 시스템은 비허가형 블록체인, 하이퍼레저 패브릭, R3가 있다. 금융의 탈중앙화를 위해 나온 것이 블록체인의 시초였지만 블록체인의 장점을 통해 다양한 시스템으로 파생되고 있다.

본 논문의 구성은 다음과 같다. Ⅱ장에서 의료정보의 공유 시 발생하는 개인정보보호법과 의료법에 대하여 살펴보고 이를 해결하는 방안의 예시인 블록체인 플랫폼 메디블록과 하이퍼레저 패브릭을 살펴본다. Ⅲ장에서 하이퍼레저 패브릭 기반 의료정보 공유 시스템을 설명하고 Ⅳ장에서 결론과 추후연구 방안을 서술한다.

II. 배경

2.1 의료정보 공유

정보통신의 기술이 발전함에 따라 많은 데이터가 생성되었고 이는 의료정보에도 영향을 끼쳤다. 한 환자에 대하여 여러 병원의 협진이 필요한 경우나 어떤 병원의 의료정보를 바탕으로 임상실험 및 연구를 유연하게 시행하기 위해 의료정보 데이터를 공유하려고 하는 움직임이 생겼다[3]. 이 때 중요한 것은 환자의 개인정보에 대한 권리를 최대한 보장하기 위해 의료정보의 기밀성과 무결성을 보장해야 하는 것과 유럽연합의 GDPR(General Data Protection Regulation)을 비롯한 개인정보보호법의 컴플라이언스 충족이다[4].

의료법 제21조 2항에 의하면 의료인, 의료기관의 장 및 의료기관 종사자는 환자가 아닌 다른 사람에게 환자에 관한 기록을 열람하게 하거나 그 사본을 내주는 등 내용을 확인할 수 있게 할 수 없다. 하지

만 개인정보보호법 제18조 2항 4호에 따르면 학술연구의 목적을 위해 개인정보를 이용하기 위해서는 특정 개인을 알아볼 수 없는 형태로 만들어 사용할 수 있다[5]. 이를 준수하기 위해 개인정보를 알아볼 수 없는 형태로 만드는 익명화 과정을 거치게 되면 데이터가 훼손되어 임상실험 및 연구에 적합하지 않게 된다[6]. 환자의 데이터를 익명화 없이 사용하기 위해서는 결국 환자에게 개인정보 활용을 동의를 받아야 한다.

본 논문에서는 환자에게 개인정보 활용 동의를 받았다는 가정 하에 각종 기관 간에 의료정보를 안전하게 공유하는 블록체인 플랫폼을 만드는 방법을 설명하고자 한다. 블록체인 기술을 도입하면 의료정보 기록의 블록체인 연결 구조로 의료 데이터의 무결성이 보장되고 분산된 데이터베이스 구조 덕분에 외부로부터 중앙 데이터베이스 공격 시 발생하는 문제점도 해결된다. 의료 정보의 이동 시 환자의 동의를 얻을 수 있도록 설계를 하면 개인정보의 주체성도 보장된다. 이러한 블록체인의 특성을 갖고 의료정보 공유 플랫폼을 만들려고 시도하였던 블록체인으로 메디블록이 있다.

2.2. 메디블록

메디블록은 환자의 PHR(Personal Health Record)를 위해 만들어진 블록체인 기반 의료정보 플랫폼이다[2]. PHR은 다양한 의료기관으로 부터 제공되는 개인의 진료정보와 개인 스스로 기록한 건강기록을 통합적이고 포괄적 관점에서 바라본 개인의 평생건강기록과 그 기록을 관리할 수 있는 도구를 말한다[3]. 메디블록은 네트워크 참여 계정을 일반 사용자, 의료 공급자, 의료 연구자로 계정으로 나눈다. 하지만 메디블록은 비허가형 블록체인을 사용하기 때문에 모든 피어들이 동등한 자격을 갖게 되는 문제가 있다. 이에 따라 메디블록은 신뢰할 수 있는 기관으로부터 인증을 받는 방식과 이미 네트워크 안에서 인증된 사람으로부터 P2P 인증을 받는 자격증명 시스템을 운영한다. 의료 공급자나 의료 연구자 계정은 일반 사용자 계정으로부터 정보를 얻으려면 당사자에게 허가를 받아야 한다. 이 때, 일반 사용자 계정에서 이를 허가해주면 일정 토큰을 받게 되어 본인의 의료정보에 대한 권리가 강화됨과 더불어 경제적 동기까지 주어서 의료정보 공유가 더 활성화 되도록 유도한다.

IPFS(Interplanetary File System)[7]는 메디블록에서 의료정보를 저장하는 분산 데이터베이스이다. IPFS는 P2P 방식으로 구현되어 있으며 IPFS에 데이터를 저장하게 되면 해시 값이 나오는데 이 때 해시 값을 가진 사람 누구나 데이터를 열람할 수 있는 방식이다. 메디블록은 IPFS에 의료정보를 저장하고 이 해시 값을 블록체인에 넣는 방법을 쓰고 있다. IPFS는 기본적으로 네트워크에서 데이터 삭제가 불가하여 보안 컴플라이언스 중 개인정보 삭제를 충족하기가 힘든 구조인데 메디블록에서 이를 어떻게 극복하였는지 명확히 서술되어 있지 않다.

2.3 하이퍼레저 패브릭

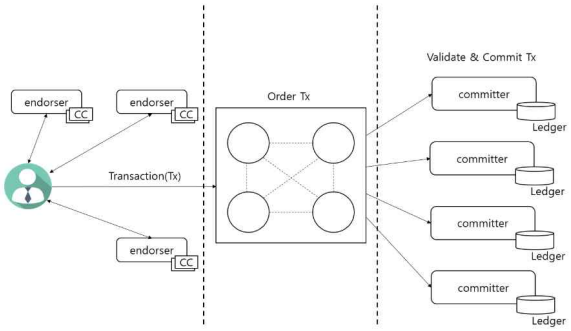

Fig 1. 하이퍼레저 패브릭 네트워크 거래처리

하이퍼레저는 리눅스 재단에서 주도한 B2B, B2C 사업을 위해 만들어진 오픈 소스 프로젝트이며 그 하위 프로젝트로 하이퍼레저 패브릭 프레임워크가 있다. 하이퍼레저 패브릭은 신원이 확인된 클라이언트를 네트워크에 참여시키는 허가형 블록체인이다. 하이퍼레저 패브릭의 네트워크는 여러 개의 채널로 이루어지며 각 채널은 다시 별도의 네트워크를 이룬다. 이 때 각 채널끼리의 통신 내용은 서로 독립되어 있으며, 이에 따라 기밀성이 보장된다. 네트워크 참여자인 피어는 여러 채널에 참여할 수 있으며 채널에 참여하면 MSP(Membership Service Provider)를 통하여 다른 피어에게 조직(Organization)을 할당해준다. 할당 받은 조직은 각 채널에서 규정한 채널 정책에서 권한을 준 조직이어야 참여할 수 있다.

Fig 1은 하이퍼레저 패브릭 네트워크에 들어온 피어의 거래가 어떻게 처리되는지 보여준다. 피어는 체인코드(스마트 컨트랙트)와 장부(Ledger)를 가지고 있으며 네트워크 밖의 클라이언트 애플리케이션과 연결되어 있다. 클라이언트 애플리케이션이 하이퍼레저 패브릭 채널내의 다른 피어와 거래를 하기 위해 보내는 신호인 트랜잭션을 호출하게 되면 호출을 받은 피어의 체인코드를 실행해야 하는데 이 때 실행결과가 전체 네트워크에 끼치는 악영향을 확인하기 위해 미리 체인코드를 실행하는 역할을 하는 보증(Endorsing) 노드에게 체인코드 실행을 의뢰한다. 보증 노드는 받은 체인코드를 실행하게 되며 트랜잭션과 실행 결과를 다시 피어에게 돌려준다. 피어는 다시 이를 순서화(Ordering) 서비스에게 보내며 순서화 서비스는 다른 피어들이 실행한 서로 다른 트랜잭션들의 순서를 결정하며 마지막으로 커미터(Commiter)에게 이 결과를 보낸다. 커미터는 장부에 데이터를 쓰는데 문제가 없는지 확인한 후 모든 피어가 가진 장부에 트랜잭션 결과를 배포하게 된다.

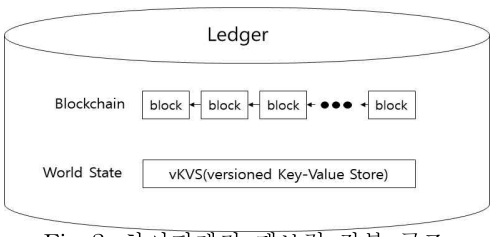

Fig 2. 하이퍼레저 패브릭 장부 구조

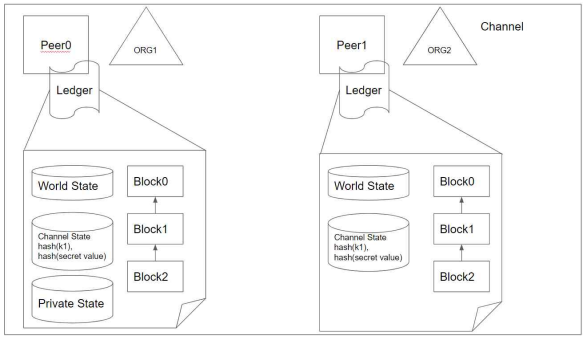

Fig 3. 피어 Ledger의 Channel State, Private State

Fig 2는 하이퍼레저 패브릭의 장부 구조를 보여준다. 피어의 장부는 블록체인과 상태(State) 데이터베이스로 이루어져 있다. 블록체인은 트랜잭션의 기록, 이전 블록의 해시 값이 저장된 블록이 서로 연결되어 있는 방식으로 이루어져 있다. 이는 생성된 모든 트랜잭션들의 무결성을 보장해준다. 상태 데이터베이스는 키-값-버전 형태로 저장되어 있다.

이러한 구조를 통해 하이퍼레저 패브릭은 특수 목적을 가진 조직들끼리 연합을 이루거나 하나의 데이터 공유 플랫폼을 만들기 적합하다. 특히 금융이나 의료정보와 같이 빠른 데이터 공유와 데이터의 기밀성이 보장되어야 하는 곳에 적합하다. 하지만 블록체인과 World State로 이루어진 장부를 사용하게 된다면 채널 안 모든 피어의 트랜잭션 기록으로 어떤 데이터가 오고 갔는지 참조할 수 있게 된다.

2.4 Private data collection

하이퍼레저 패브릭 1.2버전부터 private data collection 자료구조를 지원한다[8]. 기존 하이퍼레저 패브릭은 서로 다른 조직간 트랜잭션 내용의 기밀성을 보장하려 할 때마다 새로운 채널을 만들어야 했는데 이는 네트워크에 부담이 되었다. 이에 따라 같은 채널에 있으면서 데이터에 대한 접근 권한을 다르게 설정하여 데이터의 기밀을 보장하는 private data collection이 업데이트 되었다.

Fig 3의 채널에는 조직 1에 속하는 피어 0과 조직 2에 속하는 피어 1이 있다. 피어 0과 1의 장부 모두 World State, 블록체인, 채널 안 Private State의 자료에 접근할 키를 가진 Channel State을 가지고 있으나 Private State의 접근 권한은 피어 0에게 있기 때문에 Private State는 피어 1에 존재하지 않는다. 따라서 Private State을 읽거나 쓰는 작업은 피어 0만 가능하며 만약 피어 0이 속한 조직 1이 Private State에 피어 1이 속한 조직 2의 접근을 허용하면 피어 1의 장부에도 Private State가 복사된다. 조직 2도 마찬가지로 조직 2의 Private State와 접근 권한을 정의할 수 있다.

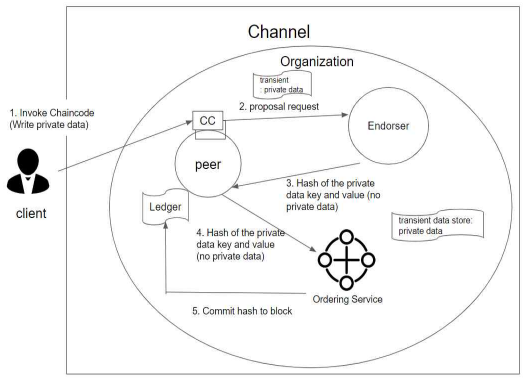

Fig 4. Private data 트랜잭션 흐름

1.2 이전 버전에서는 피어의 체인코드를 호출할 때마다 순서화 서비스에 트랜잭션이 전달되고 커미터가 이를 모든 피어에게 뿌려주는 방식인 트랜잭션 흐름을 가졌는데 private data collection을 사용하게 되면 순서화 서비스에는 private data의 해시 값이 전달되고 순서화 서비스는 이 해시한 키와 해시한 데이터에 관한 트랜잭션 기록의 순서를 결정하고 채널 안 모든 피어에게 전달하게 된다. 해시 값이 아닌 실제 private data의 키와 데이터는 보증 노드에서 임시적으로 관리하며 자격 증명이 된 피어의 트랜잭션 흐름이 끝나면 임시적으로 저장한 키와 데이터는 피어의 Channel State와 Private State에 저장된다.

III. 하이퍼레저 패브릭 응용 방안

Fig 5와 같이 의료정보 공유의 주체를 병원과 연구소라 할 때 병원 간 정보공유, 연구소 간 정보공유, 병원 연구소간 정보 공유를 위한 세 개의 채널을 만들 수 있다. 우선 이 하이퍼레저 패브릭 네트워크에 참여하려면 조직별로 만들어진 CA를 통해 인증받아야 하며 이는 MSP를 통해 이루어진다. MSP를 통하여 병원이나 연구소 조직을 할당 받았다면 각 채널 규정에 따라 접근 권한이 있는 채널에 참여가 가능하다. 채널 정책에는 조직 내 데이터를 어떻게 받고 처리할지 체인코드로 정의할 수 있어 효율적인 데이터 처리와 공유를 할 수 있게 된다. 합의 알고리즘은 병원과 연구소에서 원하는 방식대로 만들 수 있어 각 조직별 순서화 서비스가 나올 수 있으나 편의상 한 개로 가정한다.

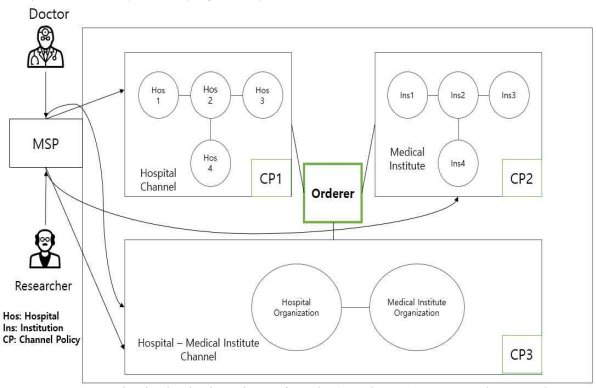

Fig 5. 하이퍼레저 패브릭 기반 의료정보 공유 플랫폼

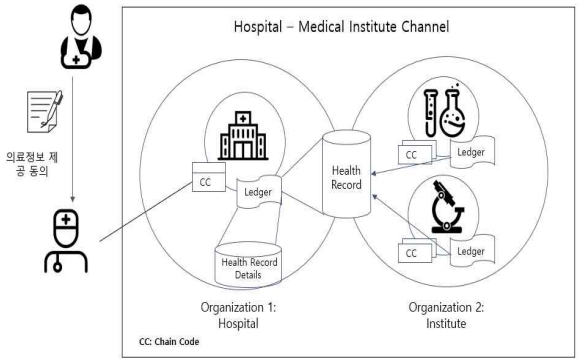

Fig 6. 채널 안 Private State를 이용한 거래 흐름

Fig 6은 병원 - 연구소 채널에서 일어나는 데이터 흐름을 보여준다. 병원-연구소의 거래 흐름은 다음과 같다. 우선 환자의 의료정보 제공 동의를 받은 의사가 하이퍼레저 패브릭의 병원 피어의 체인코드를 호출한다. 병원은 이 기록에 대하여 연구소에서 필요로 하는 의료 기록을 저장한 데이터베이스 Health Record와 환자의 개인 신상을 비롯한 모든 정보가 담긴 데이터베이스 Health Record Details를 만들어 같은 채널 내라도 조직별 접근권한을 다르게 한다. 조직 별 각 private data collection에 들어갈 JSON 형식은 체인코드로 정의할 수 있다. 본 논문은 병원에서 기록한 의료 기록은 이름, 나이, 진단, 처방이고 개인정보 보호를 위해 연구소에 제공할 데이터에는 환자의 이름을 빼야하는 상황을 가정한다.

Fig 7은 Golang으로 작성된 체인코드 내에 하이퍼레저 패브릭 private collection에 저장할 형식인 JSON을 선언하는 부분이다. “healthRecord” 구조체는 환자의 이름이 없으며 연구소만 열람할 수 있게 설정하며 “healthRecordDetails”는 환자의 이름을 포함하며 오직 병원만이 열람할 수 있도록 접근 권한을 정의할 수 있다.

```
// 체인코드 중 JSON 정의 부분
import ("encoding/json")

type healthRecord struct {
  ObjectType string `json:"docType"`
  Age        int    `json:"age"`
  Diagnosis  string `json:"diagnosis"`
  Prescription string `json:"prescription"`
}

type healthRecordDetails struct {
  ObjectType string `json:"docType"`
  Name       string `json:"name"`
  Age        int    `json:"age"`
  Diagnosis  string `json:"diagnosis"`
  Prescription string `json:"prescription"`
}
```
Fig 7. Private collection JSON (Golang)

```
// collections_config.json
[
 {
  "name": "healthRecord",
  "policy": "OR('HospitalMSP.member', 'Institute
            MSP.member')",
  "requiredPeerCount": 3,
  "maxPeerCount": 10,
  "blockToLive": 1000000
},
 {
  "name": "healthRecordDetails",
  "policy": "OR('HospitalMSP.member')",
  "requiredPeerCount": 4,
  "maxPeerCount": 8,
  "blockToLive": 1000000
 }
]
```
Fig 8. Collection 접근 권한 설정

Fig 8은 collection의 각 이름을 나타내는 “name”, 조직의 접근 권한 정책을 나타내는 “policy”, collection의 private data를 사용하며 이를 유지하는 최소 피어의 개수 “requiredPeerCount”, 최대 피어의 개수 “maxPeerCount”와 데이터가 기록된 시점을 기준으로 몇 번의 데이터 트랜잭션 후 자동으로 데이터가 삭제되어야 하는지 나타내는 “blockToLive”가 설정한 코드를 나타낸다. 병원은 “HospitalMSP” 연구소는 “InstituteMSP”이다. “policy”에서 각 private collection에 대한 접근 권한을 설정한다. 이에 따라 각 private collection을 읽고 쓰는 것에 대한 권한 설정을 할 수 있고 private collection의 기능에 따라 데이터 삭제도 가능하다. 나머지 옵션은 병원과 연구소의 상황에 맞게 설정할 수 있다.

Ⅳ. 결론

본 논문에서는 의료 정보 공유의 필요성과 관련된 법조항 그리고 이를 바탕으로 구현한 메디블록의 특징 살펴본 후 하이퍼레저의 private data collection의 특징을 바탕으로 데이터의 접근 통제와 기밀성 그리고 데이터 삭제를 가능한 의료 정보 공유 플랫폼 구현 방안을 설명하였다. 하이퍼레저 패브릭에 기반한 병원, 의료 연구기관들 사이의 데이터 공유 플랫폼은 데이터의 무결성과 기밀성을 보장함과 더불어 네트워크 접근통제까지 이루어져 안전한 데이터 공유 플랫폼이 될 것으로 기대된다.

하이퍼레저 패브릭은 키-값 쌍으로 데이터를 저장하는 JSON 방식의 자료구조를 이용하여 데이터베이스에 저장한다. 따라서 병원들의 의료기록을 공유하기 위해 사용하는 전자의무기록(EHR)을 하이퍼레저 패브릭의 데이터베이스에 저장하기 위한 자료구조 형식에 관한 협의 및 통일에 관한 연구가 필요하다.

[Acknowledgement]

본 연구는 보건복지부의 재원으로 한국보건산업진흥원의 보건의료기술연구개발사업 지원에 의하여 이루어진 것임(과제고유번호 : HI18C0316)

[참고문헌]

[1] CT·MRI 중복 촬영, 3년간 34.5% 증가…월 16억 낭비, <http://news.jtbc.joins.com/article/article.aspx?news_id=NB10579846>, (2018.11.01)
[2] MEDIBLOC WHITEPAPER, <https://medibloc-homepage.oss-us-west-1.aliyuncs.com/whitepaper/medibloc_whitepaper_kr.pdf> (2018.11.01.)
[3] Myung-Kyu-Yi. "A Study on Security Weakness and Threats in Personal Health Record Services", The Journal of The Institute of Internet, Broadcasting and Communication (IIBC) Vol. 15, No. 6, pp.163-171.
[4] Mostert, Menno, et al. "Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach." European Journal of Human Genetics 24.7 (2016): 956.
[5] 생명윤리법, <http://www.law.go.kr/> (2018.10.26.)
[6] 개인정보 비식별 조치 가이드라인, <https://www.privacy.go.kr/> (2018.10.26.)
[7] IPFS, <https://ipfs.io/#uses> (2018.11.01.)
[8] Hyperledger Fabric - Read the Docs, <https://hyperledger-fabric.readthedocs.io/en/release-1.3/> (2018.10.29.)