DOI:10.13878/j.cnki.jnuist.2019.05.011

杨明[1]　丁龙[1]　许艳[1]

基于区块链的医疗数据云存储共享方案

摘要

云服务可提供大量的存储空间,但是单一的云环境无法提供安全的数据存储和共享.而医疗数据的安全存储和共享是保证医疗数据安全的重要内容.公开审计技术保护存储在云中的数据不被篡改,从而实现医疗数据的安全存储.区块链的可公开访问及其上数据不可篡改,实现了医疗数据的安全共享.本文使用公开审计技术,提出基于区块链的医疗数据云存储共享方案,为医疗数据提供安全有效的存储和共享服务.

关键词

区块链;云存储;医疗数据;存储与共享

中图分类号 TP309
文献标志码 A

收稿日期 2019-07-02
资助项目 国家自然科学基金(61702005);安徽省自然科学基金(1708085QF136)

作者简介

杨明,男,硕士,实验师,研究方向为无线传感网安全. myang@mail.ustc.edu.cn

许艳(通信作者),女,博士,副教授,研究方向为信息安全,隐私保护. xuyan@ahu.edu.cn

---

1 安徽大学 计算机科学与技术学院,合肥,230601

0　引言

病人的医疗数据往往产生于不同的医院,由各个医院管理,很难做到以病人为中心来对数据进行管理使用.随着智慧医疗概念的提出,病人医疗数据的共享需求日益提高.电子健康记录(electronic medical records,EMRs),是个人健康信息的电子记录,可被拥有授权的医师或其他人员访问参考,能够解决医疗数据的共享问题.但 EMRs 在快速发展的同时,也面临一些安全问题,如隐私保护和完整性保护等.EMRs 涉及病人过多的隐私信息,这些隐私信息能否得到保护是 EMRs 得以推广的关键,因此病人的隐私必须优先得到保护.完整性保证了信息的可用性,EMRs 包含病人就医信息,用户就诊时医生结合可病人之前的 EMRs.如果 EMRs 的完整性遭到破坏,医生只能依靠当前诊断结果,甚至出现误诊,重复用药等更为严重的后果.

区块链可以被看成是去中心化的分布式数据库,数据区块通过共识机制后按照时间顺序进行连接,区块链中的数据可追踪且不可篡改.将 EMRs 存储在区块链中可以保证医疗数据的完整性不被破坏,已有学者提出基于区块链的医疗信息管理系统.Ekblaw 等[1]提出了利用区块链技术来处理 EMRs 的分散记录管理系统,利用区块链的不可篡改的特点保证 EMRs 的准确性.但该方案针对数据访问并未设置访问策略,导致医疗数据有可能被泄露.Shae 等[2]将新型的区块链架构应用于医疗领域,利用系统架构组件来保证医疗数据的完整性,集成医疗数据实现并行计算,而且能够保护访问者的身份隐私.Al Omar 等[3]提出了以病人为中心的医疗数据管理系统,并借助区块链来存储病人的医疗数据.Guo 等[4]提出了基于属性的多权限的签名方案.由多个授权机构来管理用户属性,但是该方案难以抵抗合谋攻击.Siyal 等[5]指出了区块链在医疗领域所面临的挑战和发展:区块链的公开可验证性使得 EMRs 不需要第三方验证,但是 EMRs 数据安全和隐私保护难以实现.Azaria 等[6]提出借助区块链实现对医疗数据的管理和访问.Roehrs 等[7]使用区块链和开放的电子医疗记录建立一个以病人为中心的医疗架构模型.但该模型只是将分布在不同医疗机构的医疗数据集成到一个视图中,仍将数据存储在区块链中.

但是受到节点的存储空间和网络的限制,区块链无法存储大量的 EMRs 数据.云存储能够为用户提供大量的存储空间,Hua 等[8]提出将医疗数据加密后外包存储在云中,在保护病人隐私的同时利用

云中存储的数据提供精准的医疗服务.Biswas 等[9]提出了三层架构的医疗云,用于存储病人的医疗数据,但是无法保证数据在云中不会被篡改.Liang 等[10]利用移动设备来收集用户的健康数据,并基于树的数据处理和批处理方式来处理移动端收集的数据,随后将其链接到区块链中.但该方案并未涉及对用户隐私信息保护.Xia 等[11]结合区块链和云存储的优点,提出了一个基于区块链的数据共享框架,可实现云中医疗数据的访问控制.但是用户共享数据的过程较为复杂.Esposito 等[12]指出医疗数据涉及病人的隐私,可以使用区块链来保护病人的隐私.然而将数据存储在区块会降低系统的效率.Cao 等[13]提出了基于区块链的云辅助的电子健康系统.将病例做成区块链,利用区块链的不可篡改性来保证病例不会被更改,实现数据共享.徐健等[14]指出将区块链和密码技术相结合来保证数据的安全性,通过智能合约来实现数据的上传和共享.董黛莹等[15]使用拜占庭容错机制提出基于区块链的电子医疗记录共享模型.该模型以角色和访问目结合的方式来实现数据共享,但效率较低.Liu 等[16]基于区块链技术,提出隐私保护的数据共享方案 BPDS,该方案将医疗数据加密后存储在云中,随后将数据在云中的索引值写入区块链.访问者通过获取数据在云中的索引值和密钥,实现数据共享.但是访问者一旦获取数据的索引值和密钥,就可以自由访问数据.如果数据泄露,该方案将难以追踪访问者.此外,BPDS 难以检测出存储在云上的医疗数据是否被篡改.针对文献[16]存在的问题,本文提出一个新的基于区块链的医疗云数据存储共享方案.该方案对用户每次访问数据的过程均加以控制,只有通过授权的用户才能够访问医疗数据.一旦数据泄密,授权机构可以追踪到恶意实体.此外,云存储数据的使用权和控制权分离,病人无法确定数据是否完整地存储在云服务器中.为了防止存储在云中的医疗数据被篡改,该方案使用公开审计的方式来确保存储医疗云数据的完整性.

1　预备知识

1.1　系统模型

基于区块链的医疗数据云存储共享方案,系统模型如图 1 所示.系统模型包含 6 种参与者:病人、数据用户、管理员、注册机构(PKG)、第三方审计(TPA)和云服务器(Cloud Server).

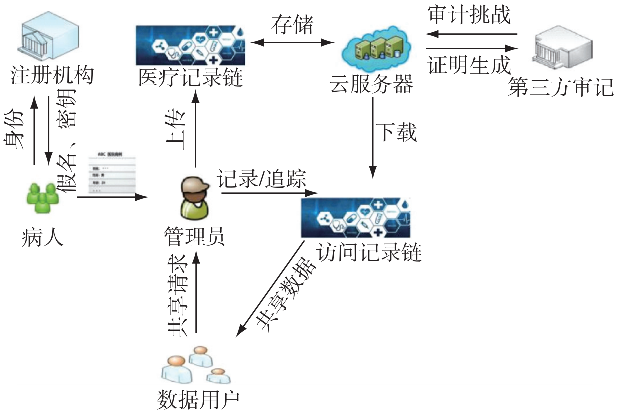

图 1　基于区块链的医疗数据云存储共享系统模型

Fig. 1　System model of cloud storage sharing for medical data based on blockchain

1)病人:医疗数据的拥有者,将其医疗数据存储在云服务器中.

2) 数据用户:可以是其他医疗机构的工作人员.拥有授权的数据用户可以查看存储在云服务器中的医疗数据,并对医疗数据进行分析.

3) 管理员:可以是医院等医疗机构,将病人的医疗数据加密后上传至云服务器,同时生成数据块(即病人的 EMRs).随后,将 EMRs 追加到病人的区块链上.此外,管理员还负责处理数据用户访问数据的请求.

4) 注册机构(PKG):可信第三方,负责生成系统参数、为病人选择假名.必要时,注册机构还要对用户的真实身份进行追踪.

5) 第三方审计(TPA):可信第三方,代替病人对存储在云中的数据进行完整性审计.

6) 云服务器(Cloud Server):拥有大量的存储空间,用来存储病人的医疗数据及对应的审计标签.

1.2　安全目标

本文使用公开审计技术,提出基于区块链的医疗数据云存储共享方案,为医疗数据提供安全有效的存储和共享服务.方案要实现的安全目标如下:

1) 身份隐私:协议必须保护病人的真实身份不被云服务器和攻击者获取,云服务器在接收到病人的医疗数据后,无法计算出病人的真实身份.

2) 完整性保护:TPA 能够代替病人对存储在云服务器中的医疗数据进行完整性审计.

3) 可追踪:只有管理员能够追踪病人的真实身份.若病人的医疗数据出现泄密,则管理员可以通过访问记录链中的访问记录追踪到泄密者.

1.3 区块链

2008 年,区块链的概念首次被提出[17].区块链可以被看成是去中心化的分布式数据库,数据区块通过共识机制的验证后按照时间顺序进行连接.除非有用户能够拥有整个网络 51%的计算能力,否则无法篡改存储在区块链中的数据,保证了数据的可靠性.此外,由于区块链具有公开性,即存储在区块链中的数据对于网络中的节点是公开的,便于网络中的节点访问存储在区块链中的数据.按照区块链公开程度划分,将区块链分为公有链,联盟链和私有链.本文将智能合约部署在联盟链中,一旦有访问请求的区块追加到区块链上,将会触动智能合约执行相应操作,实现数据的共享.

1.3.1 智能合约

智能合约是区块链的核心,是部署在区块链中由事件驱动的计算机程序.一旦预定义条件被激活,智能合约的脚本就可以自动执行.智能合约已成功的应用于以太坊(Ethereum)[18].本方案由管理员设置智能合约的执行条件.当访问记录链中增加新的访问记录时,智能合约会被触发执行.智能合约依据医疗数据在云中存储的索引下载医疗数据,并验证该医疗数据是否完整.如果医疗数据完整可用,智能合约将利用对称密钥解密密文.随后使用数据用户的公钥加密医疗数据,并将密文发送给数据用户.如果医疗数据不完整,则返回失败.智能合约具体过程如下:

算法 1:医疗数据共享合约

```
Input:存储在云中的医疗数据下标 $index$;密文的哈希值 $H_3(c)$;对称密钥 $k$;数据用户公钥 $pk_d$
Output:密文 $file$.
begin
  if(当区块链上增加新的访问区块)
  {利用 index 从云中下载医疗数据密文记为 $c_1$;
      if($H_3(c)==H_3(c_1)$)
      {$file_1=Dec_k(c_1)$;
        $file=Enc_{pk_d}(file_1)$;}
      else
          {return failure;}}
end
```

1.3.2 共识机制

区块链是一种分布式网络,所有的网络节点都会将数据备份,利用共识机制来保证所有节点的数据一致.目前的共识机制主要有工作量证明机制(POW),权益证明机制(POS)和权益授权证明机制(DPOS),本文使用 DPOS 来确保数据的一致性.系统初始时,我们依据医院的等级评分选择前 50 名的医院作为验证节点.数据产生节点我们采取医院等级在三甲以上的医院才可以产生数据块.所有参与这个系统的节点都会受到奖励,奖励采用积分制,正确执行操作的节点会获得相应的积分.如果没有正确执行操作则会扣除相应的积分,一旦积分低于一定的值后,则该节点将会失去相应操作的资格,失去的资格的节点将会由后来达到积分要求的节点来代替执行相应的操作.

2 基于区块链的医疗数据云存储共享方案

2.1 技术解释

本方案使用假名代替医疗数据中病人的真实身份,生成保护病人身份隐私的医疗数据.随后,医疗数据由管理员加密并上传到云服务器中进行存储,云服务器将存储医疗数据的索引值返回给管理员.管理员使用医疗数据的索引值、由医疗数据生成的消息摘要和医疗数据的简单描述生成相应的数据块,即病人的 EMRs.管理员随后将通过共识机制验证的 EMRs 追加到相应病人的区块链中.当有数据用户需要访问病人的医疗数据时,他通过区块中记录的信息向管理员发送请求.管理员接收到请求后,将数据用户的身份和病人在医疗记录链中的信息形成访问记录.随后将该访问记录添加到病人的访问记录链中.一旦病人的访问记录区块链添加了新的记录,智能合约将云中存储的医疗数据解密并使用数据用户的公钥加密后发送给该用户.当病人需要确定云中数据是否完整时,将委托 TPA 进行完整性审计.

2.2 具体方案

本方案利用云服务器对加密的数据进行存储,将消息摘要,医疗数据的简单描述和存储在云中的索引值写入数据块并追加到病人的医疗记录链中,以共享存储在云服务器中的医疗数据.拥有授权的用户利用这些信息可以对数据进行访问.同时管理员将访问记录写入到访问记录链中,如果数据泄露,管理员通过访问记录可以追踪到恶意用户.该方案分为系统初始化,数据存储,完整性审计和数据共享等 4 个阶段.

1)系统初始化:管理员生成系统参数.

(a) 注册机构 ID_R 生成 q 阶循环群 G_1,随机选择 $g,g_1\in G_1$,g 是群 G_1 的生成元,哈希函数 $H_1:\{0,1\}^*\rightarrow\{0,1\}^n$,$H_2:\{0,1\}^*\rightarrow G_1$,$H_3:\{0,1\}^*\rightarrow Z_q^*$.$ID_R$ 获

取系统公开参数 $params=\{q,g,g_1,H_1,H_2,H_3\}$.

（b）病人将真实身份 ID_p 提交给 ID_R,注册机构为病人生成假名 $PID_p=ID_p\oplus H_1(E_{ID_{Rpk_R}}(ID_p))$. 随后,$ID_R$ 随机选取 $a\in Z_q^*$,令病人私钥 $sk_p=a$ 并计算 $pk_p=g^{sk_p}$,并将 $\{PID_p,pk_p,sk_p\}$ 通过安全信道发送给病人.

2)数据存储:管理员将病人的医疗数据加密后存储到云服务器中,并形成数据区块(EMRs)追加到病人的医疗记录链中.

(a)病人使用假名代替其医疗数据中的真实身份,得到消息 m.随后将 m 提交给管理员.

（b）管理员 ID_m 对 m 进行验证,验证通过后将 m 分块为 $m=\{m_1,m_2,\cdots,m_n\}$.ID_m 随机选取 $k,r_{PID_p}\in Z_q^*$,用 k 对 m 进行对称加密得到密文 $c=\{m'_1,m'_2,\cdots,m'_n\}$;并计算 $R_{PID_p}=g^{r_{PID_p}}$,将 R_{PID_p} 公开.随后对 m_i 进行签名 $\delta_{m'_i}=g_1^{sk_p}\cdot(H_2(ID_m)\cdot R_{PID_p}{}^{m'_i})^{r_{PID_p}}$.最后,$ID_m$ 将 $\{c,\delta_{i(1\leqslant i\leqslant n)}\}$ 存储在云中并获取其在云中存储的索引 $index$.

（c）管理员 ID_m 对 m 进行简单描述,得到不涉及用户隐私信息的消息 M,并将 $info=\{H_3(c),index,M,t\}$ 做成区块.该区块通过共识机制的验证后将被追加到病人的医疗记录链.

3)完整性审计:当病人需要确定云服务器中存储的医疗数据是否完整时,将委托 TPA 对云中的数据进行完整性审计,过程如下:

(a)TPA 随机选择随机数 $i\in I(I\subseteq[1,n])$ 和 $v_i\in Z_q^*$,组成挑战信息 $chal=\{i,v_i\}$,并将其发送给云服务器.

（b）云服务器接收到 $chal$ 后,计算 $\lambda=\sum_{i\in I}v_i\cdot m'_i$ 和 $\sigma=\prod_{i\in I}\delta_i^{v_i}$,并将 $p=\{\lambda,\sigma\}$ 作为数据持有性证明发送给 TPA,进行完整性验证.

(c)TPA 接收到数据持有性证明 p 之后,验证（1）式是否成立,

$$e(\sigma,g)=e\left(g_1^{\sum_{i\in I}v_i},pk_p\right)\cdot e\left(H_2(ID_m)^{\sum_{i\in I}v_i}\cdot R_{PID_p}{}^{\lambda},R_{PID_p}\right),\tag{1}$$

如果成立,则表明医疗数据完整地存储在云中.否则表明医疗数据被破坏.

4）数据访问:数据用户想要访问某个医疗数据 m 时,执行下列步骤.

（a）当数据用户想要访问某个医疗数据 m 时,通过查找医疗记录链找到对应的记录 M.数据用户生成请求信息 M',M' 包含访问数据的 $index$,数据访问的目的,数据用户的名称,数据用户的公钥 pk_d 等信息.随后数据用户将请求信息 M' 发送给 ID_m.

(b)ID_m 将数据用户的请求信息 M' 附上时间戳形成数据块,广播到区块链中进行验证.数据块通过验证后,被链接到病人访问记录链上,用于追踪.

（c）如果访问记录链中增加了新的区块,智能合约将被触发执行算法 1,使用密钥 k 解密病人存储在云服务器中的医疗数据,使用数据用户的 pk_d 对解密后的医疗数据进行加密,并发送给用户.

3 安全性分析

本节从隐私保护,完整性保护、可追踪和数据共享等方面对方案进行分析.

1)隐私保护:本方案涉及的隐私部分包含用户的身份隐私,数据隐私两个部分.其中注册机构为病人生成假名,病人在医疗数据 m 中使用假名代替其真实身份,数据用户即使能够获取医疗数据,也无法通过医疗数据来推测出病人的真实身份.在完整性审计的过程中,病人使用假名与 TPA 进行交互.因此,本方案能够保护病人的身份隐私.此外,管理员对病人的医疗数据进行加密,并将密文存储到云服务器.除了拥有密钥的数据用户可以获取医疗数据外,其他任意第三方均无法获取有效的医疗数据.

2)完整性保护:在本方案中,存储在云服务器中的医疗数据可以通过等式(1)进行完整性检测.对于存储在区块链中的数据,由于每个数据块都包含一个当前时间戳和前一个块的散列值,因此按时间顺序嵌套的数据块保证记录的数据不能被更改,除非有人能够同时接管整个网络 51%的计算能力,否则无法更改区块链中的数据.此外,每个访问请求活动都被记录在区块链中,对数据的任何更改都可以审计和跟踪.因此,本方案能够保证数据不被篡改.式(1)的正确性分析如下:

$$e(\sigma,g)=e\left(\prod_{i\in I}\delta_i^{v_i},g\right)=$$

$$e\left(\prod_{i\in I}(g_1^{sk_p}\cdot(H_2(ID_m)\cdot R_{PID_p}{}^{m'_i})^{r_{PID_p}})^{v_i},g\right)=$$

$$e\left(\prod_{i\in I}g_1^{v_i},g^{sk_p}\right)\cdot e\left(\prod_{i\in I}((H_2(ID_m)\cdot\right.$$

$$\left.R_{PID_p}{}^{m'_i})^{r_{PID_p}})^{v_i},g\right)=$$

$$e\left(g_1^{\sum_{i\in I}v_i},g^{sk_p}\right)e\left(H_2(ID_m)^{\sum_{i\in I}v_i}\cdot\right.$$

$$\left.\prod_{i\in I}R_{PID_p}{}^{v_i\cdot m'_i},g^{r_{PID_p}}\right)=$$

$$e\left(g_1^{\sum_{i\in I'} v_i}, pk_p\right) e\left(H_2(ID_m)^{\sum_{i\in I'} v_i} \cdot R_{PID_p}{}^{\lambda}, R_{PID_p}\right).$$

3)可追踪:本方案借助区块链来存储数据用户对医疗数据的访问记录.当数据用户需要访问存储在云中的医疗数据时,会向管理员发送请求消息.管理员将请求消息生成数据块,并将通过区块链验证的数据块追加到访问记录连中.利用区块链的防篡改性保证访问记录链中的记录真实有效.当医疗数据出现泄密时,只有管理员通过访问记录链中的记录就可以追踪恶意的数据用户.

4 结束语

本文提出了基于区块链的医疗云数据存储共享方案.在保护数据隐私和用户的身份隐私的同时实现数据共享.通过安全性分析,证明该方案是安全可行的.随着智慧医疗的发展,数据的共享需求将会越来越高.我们将会进一步研究如何在保证数据安全性的同时提高数据的共享性.

参考文献
References

[1] Ekblaw A,Azaria A,Halamka J D,et al.A case study for blockchain in healthcare:“MedRec” prototype for electronic health records and medical research data[C]//Proceedings of IEEE open & big data conference,2016,13:13

[2] Shae Z,Tsai J J P.On the design of a blockchain platform for clinical trial and precision medicine [C] //Proceedings of the 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS).IEEE,2017:1972-1980

[3] Al Omar A, Bhuiyan M Z A, Basu A, et al. Privacy-friendly platform for healthcare data in cloud based on blockchain environment[J].Future Generation Computer Systems,2019,95:511-521

[4] Guo R,Shi H,Zhao Q,et al.Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems [J]. IEEE Access,2018,6:11676-11686

[5] Siyal A,Junejo A,Zawish M,et al.Applications of blockchain technology in medicine and healthcare:challenges and future perspectives[J].Cryptography,2019,3(1):3

[6] Azaria A,Ekblaw A,Vieira T,et al.Medrec:using blockchain for medical data access and permission management[C]//Proceedings of the 2016 2nd International Conference on Open and Big Data (OBD),IEEE,2016:25-30

[7] Roehrs A,Costa C A D,Rosa Righi R D,et al.Analyzing the performance of a blockchain-based personal health record implementation[J].Journal of Biomedical Informatics,2019,92:103140

[8] Hua J,Shi G,Zhu H,et al.CAMPS:efficient and privacy-preserving medical primary diagnosis over outsourced cloud[J].Information Sciences,2018

[9] Biswas S,Akhter T,Kaiser M S,et al.Cloud based healthcare application architecture and electronic medical record mining:an integrated approach to improve healthcare system[C]//Proceedings of the 2014 17th International Conference on Computer and Information Technology (ICCIT),IEEE,2014:286-291

[10] Liang X,Zhao J,Shetty S,et al.Integrating blockchain for data sharing and collaboration in mobile healthcare applications[C]//Proceedings of the 2017 IEEE 28th Annual International Symposium on Personal,Indoor,and Mobile Radio Communications (PIMRC),IEEE,2017:1-5

[11] Xia Q,Sifah E,Smahi A,et al.BBDS:blockchain-based data sharing for electronic medical records in cloud environments[J].Information,2017,8(2):44

[12] Esposito C,De Santis A,Tortora G,et al.Blockchain:a panacea for healthcare cloud-based data security and privacy? [J].IEEE Cloud Computing,2018,5(1):31-37

[13] Cao S, Zhang G, Liu P, et al. Cloud-assisted secure eHealth systems for tamper-proofing EHR via blockchain [J].Information Sciences,2019,485:427-440

[14] 徐健,陈志德,龚平,等.基于区块链网络的医疗记录安全储存访问方案[J].计算机应用,2019,39(5):1500-1506
XU Jian,CHEN Zhide,GONG Ping,et al.Secure storage and access scheme for medical records based on blockchain[J].Journal of Computer Applications,2019,39(5):1500-1506

[15] 董黛莹,汪学明.基于区块链的电子医疗记录共享研究[J].计算机技术与发展,2019,29(5):121-125
DONG Daiying,WANG Xueming.Research on electronic medical record sharing model based on blockchain[J].Computer Technology and Development,2019,29(5):121-125

[16] Liu J, Li X, Ye L, et al. BPDS: a blockchain based privacy-preserving data sharing for electronic medical records[C]//Proceedings of the 2018 IEEE Global Communications Conference (GLOBECOM),IEEE,2018:1-6

[17] Nakamoto S. Bitcoin: a peer-to-peer electronic cash system[EB/OL].[2019-05-01].https://bitcoin.org/bitcoin.pdf

[18] Wood G. Ethereum: a secure decentralised generalised transaction ledger [J]. Ethereum project yellow paper,2014,151(2014):1-32

A cloud storage and sharing scheme for medical data based on blockchain

YANG Ming[1] DING Long[1] XU Yan[1]
1 School of Computer Science and Technology,Anhui University,Hefei 230601

Abstract Cloud services can provide storage space,but a single cloud environment cannot provide secure data storage and sharing. However, the secure storage and sharing is an important content to ensure the security of medical data.Public audit technology protects cloud data from tampering,thus ensure the secure storage of medical data.The public access of the tamper-resistant blockchain can realize the secure sharing of medical data.The paper uses public audit technology to propose a blockchain-based cloud storage and sharing scheme for medical data,and analysis shows that the scheme is secure and effective.

Key words blockchain;cloud storage;medical data;storage and sharing