一种基于区块链的中医电子病历共享方法

常源恒,丁有伟,胡孔法

(南京中医药大学 人工智能与信息技术学院,江苏 南京 210046)

摘 要:随着大数据和人工智能等技术的发展与应用,基于中医电子病历的各类智慧医疗应用越来越多,这些应用需要对各个医院的中医电子病历进行共享与综合分析。当前中医电子病历共享主要采用云平台进行数据存储,并使用集中式的密钥分配与权限认证方法进行访问控制,容易因服务提供商的不可信导致数据泄漏及非法访问等安全问题。因此,提出一种基于区块链的中医电子病历安全共享方法,利用区块链的去中心化、可追溯、不可篡改等特性保证中医电子病历的安全存储、合法访问与访问存证,同时利用云平台超强的存储与计算能力构建弱中心化的区块链结构,以提高区块链的数据处理性能。通过理论分析与实验证明,在安全保障方面,该方案相比传统方案具有更强的安全性;在数据分享效率方面,每10个节点带来的延迟为200ms,具有一定可行性。

关键词:中医电子病历;区块链;数据共享

DOI:10.11907/rjdk.212103 开放科学(资源服务)标识码(OSID):

中图分类号:TP319 **文献标识码**:A **文章编号**:1672-7800(2021)011-0163-05

A Blockchain-based Method for Data Sharing of Traditional Chinese Medicine Electronic Medical Records

CHANG Yuan-heng, DING You-wei, HU Kong-fa

(*School of Artificial Intelligence and Information Technology*, *Nanjing University of Chinese Medicine*, *Nanjing* 210046, *China*)

Abstract: With the development and application of technologies such as big data and artificial intelligence, there are more and more smart medical applications based on electronic medical records of Chinese medicine. These applications need to share and comprehensively analyze the electronic medical records of Chinese medicine in various hospitals. At present, the sharing of electronic medical records of traditional Chinese medicine mainly uses cloud platforms for data storage, and uses centralized key distribution and authorization for access control. It is easy to cause security problems such as data leakage and illegal access due to the untrustworthiness of service providers. Propose a blockchain-based method for the secure sharing of traditional Chinese medicine(TCM) electronic medical records, which use blockchain's features such as decentralization, traceability, and non-tampering to ensure the secure storage, legal access and access certificate of TCM electronic medical records, and at the same time use the cloud platform's super strong storage and computing capabilities to build a weakly centralized blockchain structure to improve the data processing performance of the blockchain. Experimental results and theoretical analysis show that the security of the proposed scheme has stronger performance against traditional schemes, and the data sharing efficiency has a delay of 200ms for every 10 nodes, which shows ceveain feasibility.

Key Words: electronic medical records of TCM; blockchain; data sharing

收稿日期:2021-08-19

基金项目:国家重点研发计划项目重点专项(2017YFC1703500);国家自然科学基金项目(82004499,82074580);江苏省高等学校自然科学研究面上项目(19KJB520012)

作者简介:常源恒(1996-),男,南京中医药大学人工智能与信息技术学院硕士研究生,研究方向为中医药数据安全;丁有伟(1987-),男,博士,南京中医药大学人工智能与信息技术学院讲师,研究方向为中医药数据分析、中医药数据安全;胡孔法(1970-),男,博士,南京中医药大学人工智能与信息技术学院教授,研究方向为中医药数据分析、人工智能、知识图谱、数据安全等。本文通讯作者:丁有伟。

0 引言

中医药是中华文明的瑰宝，近年来随着大数据技术的发展，中医药大数据的管理与应用备受关注[1-2]。中医药大数据中包含许多敏感信息，如中医电子病历中包含病人私人信息、中医药方等敏感数据，一旦发生泄漏，后果不堪设想。当前在中医药数据共享方面一般采用中心化的方案，由数据中心集中分配密钥与权限并存储所有数据，但这种中心化的数据共享方法存在一定安全隐患，如因数据中心抗攻击能力不足导致外部攻击者窃取数据，或因数据中心本身不可信造成数据泄漏等。因此，安全、高效的中医电子病历共享方法是中医药大数据应用中亟待解决的关键问题。

现有的中医电子病历大多使用集中化的共享方法[3-4]，各级医院将自己的电子病历保存在本地数据库中，但随着数据规模的增长，对存储设备、机房空间和管理人员的需求也随之增长。同时，中心化的共享方法存在某些固有的安全漏洞，如数据库被破解、内部人员监守自盗等均可能会导致隐私信息泄露，极大地阻碍了中医药大数据共享。另外，由于当前中医药电子病历标准尚未统一，各医疗机构的数据格式及属性存在较大差异，很难使用传统的数据共享方案。因此，需要设计一种针对非标准化中医药电子病历的数据安全共享方案。

区块链技术是目前的热门技术之一，作为网络安全方面的佼佼者，区块链具有去中心化、不可篡改等特性[5]，对于交易双方又具有非对称加密、多方存证防篡改等安全机制，数据安全性可得到很好的保障[6]。区块链技术目前已成功应用于金融[7-8]、保险、物流[9]、电子票据、医学[10-12]等领域，但在中医药领域的应用较少。目前区块链在中医药领域的应用主要是探索其可行性，尚未考虑实际性能问题[13]。如肖丽等[14]提出基于区块链的中医电子病历系统，其需要各级医院配合使用相同电子病历，但目前各医院的中医电子病历标准并不相同，实际上难以实现。其他相关研究使用完全的去中心化方案[15]，如生慧等[16]提出基于联盟链的中医药海量异构数据共享方案，使用户可安全地分享自己的医疗数据，但这种方案并不利于医疗数据监管。从应用角度看，中医电子病历需要一种既能保证数据传输过程中的安全性，又能满足卫生部门监管要求的共享方案。

针对上述问题，本文提出一种基于区块链的中医电子病历共享方案。该方案结合区块链技术与传统数据共享的集中管理及中心化方法，将中医电子病历用区块链形式存储在中医药大数据中心，并搭建电子病历共享联盟链以保存共享过程中的请求答复信息。同时，为提高区块链传输速度，使用一条额外的区块链侧链进行数据传输，使中医药大数据能更好地实现安全共享。

1 基于弱中心化区块链的中医电子病历安全共享方案

在传统区块链中，多个陌生节点在区块链中通过共识机制达成信任，每个区块拥有上个区块的哈希值用于追溯，是一种去中心化结构，而传统数据库是中心化结构。本文将两者结合，使用中心化的数据中心作为区块链存储中心，并利用区块链可追溯、防篡改的特性，提出一种中医电子病历安全共享方法，即弱中心化方法。该方法包括主侧两条区块链，主链的用户有研究人员、数据中心和各级医院，用户加入区块链后，由系统分发唯一标识 uID 和密钥对（uPk，uSk），其中 uPk、uSk 分别为用户的公钥和私钥。医院作为数据所有者，也有唯一的身份标识符 hID 以及公钥 hPk 与私钥 hSk。由于区块链的整体效率较低，因此使用一条由数据中心监管的侧链以提升数据传输效率。

1.1 网络结构

区块链主链存储请求数据信息及对于请求的答复信息。在主链上，研究人员与各级医院具有访问数据中心目录、发起交易的权限，数据中心具有签发证书与存储数据的权限。侧链用于临时存储交易中请求者所请求的数据。数据中心具有完全的权限，包括发送区块、删除区块，而研究人员和各级医院只有读取区块的权限。整体网络结构如图 1 所示。

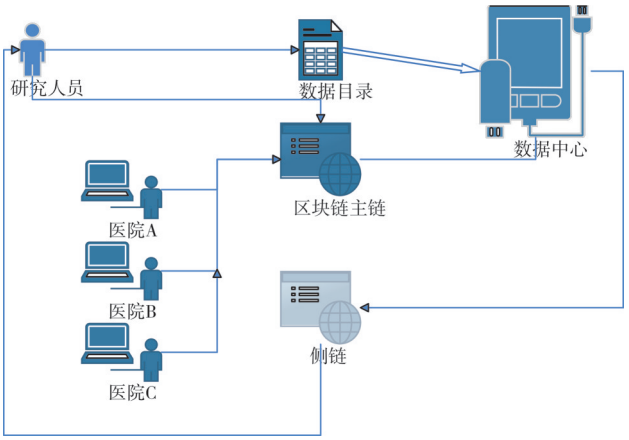

Fig. 1 Overall network structure

图 1 整体网络结构

在整体框架中，数据中心的各服务器负责完成去中心化的区块打包排序、分发区块链认证证书等任务，但数据中心对于整体而言依然是网络结构的中心，负责所有数据的存储与分发，既具有去中心化的高安全性，又保留了中心化的高性能特征。

1.2 数据格式

假设中医电子病历包含总编号、标准化临床表现、标准化病机、膏方、剂量、备注等属性，数据中心以区块链形式存储所有医院的数据，每 30 条数据打包为一个区块。每个数据区块的区块头包含该区块内所有电子病历的 Merkle

根与上个区块的哈希，区块体包含数据内容及其拥有者ID。

由于所有数据都存储在一条链上，为便于查询每个区块的大致内容，且数据中心能快速遍历整条链为数据创建目录，数据目录格式为列表，用户可访问目录表查询需要的数据。数据目录表如表1所示，其中hID为医院ID，number为数据所在区块高度。

Table1　Data catalog table

表1　数据目录表

hID	number，包含内容	number，包含内容	……
HS001	1，肺癌病例	3，肺癌病例	……
HS002	2，皮肤病例	9，肝癌病例	……
HS003	4，肺癌病例	5，肺癌病例	……

在主链上，每次用户的电子病历数据访问请求为一个交易，请求内容包括用户ID、请求的数据区块及被请求的医院ID。如图2中的主链请求区块体部分，其中number为请求数据所在区块高度。医院审核后，向区块链广播一条带有认证人信息的答复信息，并对该信息进行签名。答复格式如图2中的主链答复区块体部分，其中，uID为用户ID，verifier为认证人，answer为该次请求的结果。数据中心判断审核是否通过，如果通过，则数据中心将数据请求人的公钥加密后发送到侧链上。

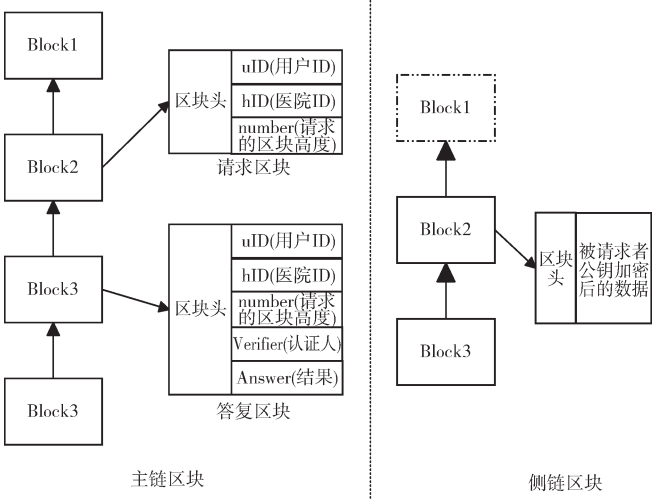

Fig. 2　Main chain and side chain blocks

图2　主链与侧链区块

对于非法请求，如出现所请求的数据与医院不匹配的情况，除医院对该请求的答复设置为非法（即将answer属性值设为n）外，区块链会将该请求永久存储，未来该请求者再次请求其他数据时，医院可查询该请求者非法请求的数量，酌情考虑对该节点的许可。如果非法数据连续超过5条，则该请求节点会被禁止发送请求。

如图2所示，主链上的区块均为用户请求区块与医院对该请求的答复区块，区块头包含上个区块的哈希，区块体则是交易具体信息。

若医院对请求者的认证通过，则数据中心使用请求者的公钥对授权区块加密后发送到侧链上，即侧链区块，侧链区块经过一定时间后被自动销毁。

1.3　共享流程

如图3所示，用户通过在数据中心存储的数据目录查寻自己所需数据，数据所属医院对用户访问请求进行审核。如果审核通过，数据中心通过侧链将数据提供给用户。

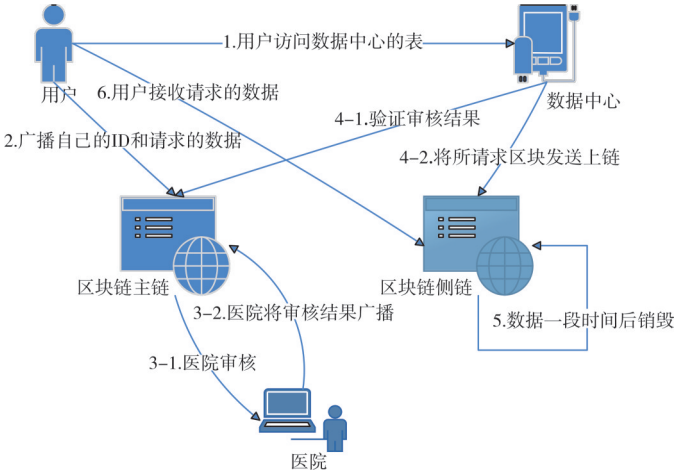

Fig. 3　Flow of safe sharing of TCM electronic medical records

图3　中医电子病历安全共享流程

数据共享具体流程如下：

（1）用户加入区块链，区块链向用户分发证书与密钥对，若需电子病历数据，查看数据中心的目录查询当前已存储数据。

（2）查询到所需数据的区块高度，用户将区块高度number、数据所属医院hID及自身uID打包为一个请求消息，使用自身的私钥加密后形成一个区块，广播到主链上。

（3）医院接收到该区块后使用用户公钥解密，并对用户身份进行审核，之后向主链发送对该请求的答复信息。答复信息中应包含所请求的数据高度number、请求用户uID、审核人verifier及审核结果answer，医院打包答复信息并使用自身的私钥加密形成区块，广播到主链上。

（4）数据中心使用医院公钥解密并检查审核结果，对于审核结果answer为y的答复消息中提到的数据高度，遍历数据区块链找到处在该高度上的区块，复制区块内容并使用答复消息中用户的公钥加密，发送到侧链上。

（5）用户接收从侧链上发送的区块，使用自身私钥解密得到所请求的数据。

（6）经过一定时间后，侧链上的数据将被自动销毁，保证非法用户无法获得数据，流程结束。

2　实验及分析

2.1　实验配置

实验使用计算机1：Intel® Core™ i5-2400 CPU @ 3.10GHz × 4，Ubuntu 18.04.5 LTS系统，4GB内存；计算机2：Intel® Core™ i5-4200M CPU @2.50 GHz × 4，Windows7系统，并使用VMware WorkStation虚拟机Ubuntu 18.04.5 LTS，

均搭建 Hyperledger fabric 平台，版本为 1.4.4。数据集为脱敏后近两年的中医电子病历数据，内容包含数据编号、标准化临床表现、标准化病机、中医膏方 4 个属性，共有 2 665 条数据，抽取其中部分数据作为实验用例。

首先构建网络，该网络使用一个排序（order）节点模拟数据中心，具有 Org1、Org2 两个组织，Org1 代表数据中心的内部组织，Org2 代表公有的区块链组织。两个节点 peer0.Org2 和 peer1.Org2 模拟医院与请求数据的研究人员，两个通道（channel）模拟区块链主链与侧链。然后对每个节点进行存储私有数据智能合约和请求数据智能合约的安装与初始化，其中针对存储数据的背书策略，设定为只能在数据中心组织内部查看其私有内容。在 Org1 中存储中医电子病历数据，在 Org2 中的研究人员节点 peer0 请求该数据，请求处理后 order 节点将请求数据发送至 channel2 上。

2.2 实验及结果分析

实验按照不同的总体存储数据量及每个区块存储数据量进行对照实验，在安全性由区块链进行保障的情况下，主要对其性能进行测试。对于传统数据库 MySQL，通常研究人员请求某医院电子病历数据需要等待的请求及审核时间与使用本文方法相同，二者的时间差值主要体现在发送数据给请求者的传输时间，故实验对传输时间进行比较。每组实验传输 10 次数据，在实验过程中，首次对接收者所在组织的锚节点进行传输时会产生一定延迟，若对组织的首次传输不是该组织的锚节点，则不存在该延迟情况。由于实验存在偶然性，故每组实验去除最大值和最小值，对剩余几次传输时间取平均值，得出实验结果。

首先使用 4 个节点构建网络进行 3 组实验，每组实验使用不同的单个区块存储量，每个区块分别存储 1/50/60 条数据，总共存储 600~1 800 条数据在侧链进行传输，结果如图 4 所示。由此可见，在节点数不变的情况下，不论区块链中存储数据总量为多少，每个区块存储数据量大小对传输时间影响不大，均在一个较稳定的区间内。

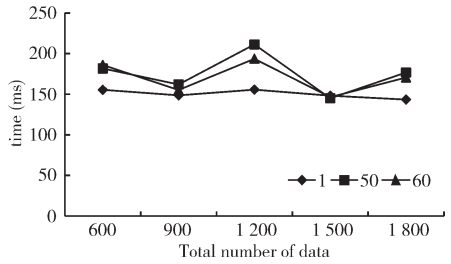

Fig. 4 Comparison of transmission speed of different total data

图 4 不同数据总量传输速度对比

同理，使用 4 个节点构建网络进行 3 组实验，控制数据存储总量分别为 600、1 200、1 800 条，每个区块存储数据量从 1~60 递增，对比数据传输耗时。如图 5 所示，实验结果表明，在数据总量相同的情况下，随着区块内存储量的增大，传输速度逐步增加，但在实际应用过程中，每个区块仅传输一条数据会使得每次传输过多区块，侧链变得冗杂，不利于接收与销毁。故在实际应用中，每个区块传输数据量应选择在 20~50 之间。

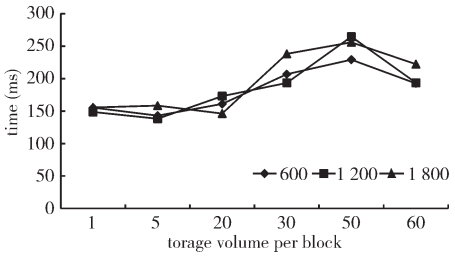

Fig. 5 Comparison of different transmission speeds for different numbers of data stored in each block

图 5 每个区块存储数据量不同时传输速度对比

最后，维持数据总量为 1 200 条进行两组实验，将节点数量由 2 个增加到 10 个，如图 6 所示。随着节点数的增多，数据传输效率有小幅增长，这是网络增大所带来的必然结果。由于实验环境限制，本实验未对大数据量的环境进行测试，因此在小数据量的情况下，本文方案的效果优于传统方案。

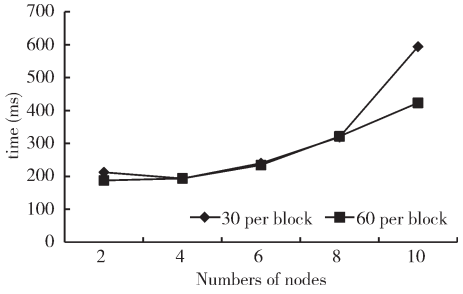

Fig. 6 Comparison of transmission speed with different number of nodes

图 6 不同节点数传输速度对比

2.3 安全性分析

本文方案的安全性主要通过区块链进行保障，使用 SHA-256 算法保证区块链的不可篡改性，对于一条任意长度的消息，SHA-256 都会生成一条 256 位的哈希值，称为消息指纹。当消息内容出现任何改变，就会使该指纹发生改变。目前对于 SHA-256 的破解方法主要是穷举法，这对于数据有 2^{256} 种可能的哈希值来说，破解几乎是不可能的。

对于数据加密，区块链数据加密过程中使用的公钥和私钥由加密算法 ECDSA（Elliptic Curve Digital Signature Algorithm，椭圆曲线数字签名算法）得出，该算法由私钥可得出公钥，但私钥无法由公钥逆推。破解椭圆算法除非解决离散对数难题，否则只能在一个巨大的空间内进行暴力破解。ECDSA 算法原理复杂，且具有严格的证明过程，本文不再赘述。

对于传统数据库结构，假设恶意节点想要窃取数据，恶意攻击者仅需攻破数据库本身（数据中心的防火墙）即可获得数据，且不会存在记录。而对于本文结构，攻击者首先需要在实名进入联盟链的情况下攻击数据中心防火墙，其次要破解使用 SHA-256 与 ECDSA 算法加密后的区块链。又假设恶意节点是已加入区块链的节点，想要篡改被医院驳回的请求信息，首先需要破解医院的私钥；对驳回信息所在区块 A 的内容进行更改。由于更改了区块内容，区块 A 的消息指纹发生了改变，恶意节点需要对区块 A

的区块头进行更改，生成新的消息指纹。又由于区块链中每个区块都包含上一个区块的消息指纹，对区块 A 的更改使得 A 的下一个区块发生变化。因此，恶意节点需要对下一个区块的消息指纹进行更改，之后所有区块才能成功进行攻击，这需要极大的算力，几乎不可能实现。

3　结语

针对现今中医药数据共享方案研究中不能有效保证安全性的问题，本文提出一种基于区块链的弱中心化中医电子病历分享方案。该方案使用区块链技术结合传统方案的中心化结构，可安全存储中医药数据，保留数据共享过程以便于追溯，并使用侧链提升区块链传输效率，在保证较好传输性能的同时，还能保证共享过程的安全性，最后通过实验证明本方案的可行性。但本文未对区块链本身的协议进行更改，未来将会对区块链协议进行针对性地修改，使其符合中医药电子病历数据特性，以期进一步提升方案的传输性能。

参考文献：

[1] HU W, HOU Z K, LIU F B, et al. Thoughts on clinical research of traditional Chinese medicine in the age of big data[J]. Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology, 2019,21(8):1656-1661.
胡文,侯政昆,刘凤斌,等.关于大数据时代的中医药临床研究的思考[J].世界科学技术-中医药现代化,2019,21(8):1656-1661.

[2] CUI M, LI H, HU X. Similarities between "big data" and traditional Chinese medicine information[J]. Journal of Traditional Chinese Medicine, 2014,34(4):518-522.

[3] LIU J M, PENG S L, LI K L, et al. Chinese herbal quality and safety management model based on blockchain[J]. Frontiers of Data&Computing,2020,2(5):65-75.
刘加梦,彭绍亮,李肯立,等.基于区块链的中草药质量安全管理模型[J].数据与计算发展前沿,2020,2(5):65-75.

[4] LIU H, XU L Q, CHEN S M. The construction and application of TCM electronic medical record cloud in Xiamen[J].Chinese Journal of Health Informatics and Management, 2019,16(6):690-693,712.
刘辉,徐乐勤,陈少玫.厦门市中医门诊电子病历云的建设与应用[J].中国卫生信息管理杂志,2019,16(6):690-693,712.

[5] LIU Y Z, LIU J W, ZHANG Z Y, et al. Overview on blockchain consensus mechanisms[J]. Journal of Cryptologic Research, 2019,6(4):395-432.
刘懿中,刘建伟,张宗洋,等.区块链共识机制研究综述[J].密码学报,2019,6(4):395-432.

[6] ZHU L H, GAO F, SHEN M, et al. Survey on privacy preserving techniques for blockchain technology[J]. Journal of Computer Research and Development, 2017,54(10):2170-2186.
祝烈煌,高峰,沈蒙,等.区块链隐私保护研究综述[J].计算机研究与发展,2017,54(10):2170-2186.

[7] LIU L C, CHENG H Y. Review on the progress of blockchain finance theory and application[J]. Modernization of Management, 2020,40(5):111-113.
刘凌晨,程宏宇.区块链金融理论与应用进展综述[J].管理现代化,2020,40(5):111-113.

[8] BA S S, QIAO R Y. Blockchain technology enables digital finance[J]. FinTech Time, 2021,29(7):14-18.
巴曙松,乔若羽.区块链技术赋能数字金融[J].金融科技时代,2021,29(7):14-18.

[9] TIAN Y, CHEN Z G, SONG X X, et al. Application of blockchain in supply chain management[J/OL].Computer Engineering and Applications, 2021-08-02. http://kns.cnki.net/kcms/detail/11.2127.TP.20210713.1333.008.html.
田阳,陈智罡,宋新霞,等.区块链在供应链管理中的应用综述[J/OL].计算机工程与应用,2021-08-02.http://kns.cnki.net/kcms/detail/11.2127.TP.20210713.1333.008.html.

[10] SOOKHAK M, JABBARPOUR M, SAFA N, et al. Blockchain and smart contract for access control in healthcare: a survey, issues and challenges, and open issues[J]. Journal of Network and Computer Applications,2021,178:102950.

[11] DONAWA A, ORUKARI I, BAKER C E. Scaling blockchains to support electronic health records for hospital systems[C].2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 2019: 550-556.

[12] ZHANG C, LI Q, CHEN Z H, et al. Medical Chain: alliance medical blockchain system[J]. Acta Automatica Sinica, 2019, 45(8): 1495-1510.
张超,李强,陈子豪,等.Medical Chain:联盟式医疗区块链系统[J].自动化学报,2019,45(8):1495-1510.

[13] JIN M L Y, LUO J T, HE L, et al. Exploration on the construction of Chinese medical electronic medical record standard conformity test platform[J].Asia-Pacific Traditional Medicine, 2020, 16(8): 207-210.
金木李由,罗迦腾,何黎,等.中医电子病历标准符合性测试平台建设探索[J].亚太传统医药,2020,16(8):207-210.

[14] XIAO L, LIN L, XIE P, et al. Study and application of blockchain in electronic medical record system of traditional Chinese medicine[J]. Lishizhen Medicine and Materia Medica Research, 2018, 29(12):3062-3064.
肖丽,林林,谢鹏,等.基于区块链的中医电子病历系统的应用研究[J].时珍国医国药, 2018,29(12):3062-3064.

[15] PAN H, PAN L, YAO Z Y, et al. A patient-controlled security access mechanism for electronic health records[J]. Journal of Applied Sciences, 2020,38(1):127-138.
潘恒,潘磊,姚中原,等.一种病人可控的电子病历安全访问方案[J].应用科学学报, 2020,38(1):127-138.

[16] SHENG H, ZHOU Y, MA J G, et al. A secure sharing solution for massive heterogeneous data of traditional Chinese medicine based on alliance chain[J].Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology,2019,21(8):1662-1669.
生慧,周扬,马金刚,等.一种基于联盟链的中医药海量异构数据安全共享解决方案[J].世界科学技术-中医药现代化,2019,21(8):1662-1669.

（责任编辑：黄　健）