计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2022,31(6):394−399 [doi: 10.15888/j.cnki.csa.008569]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于区块链对溯源数据的多方共享系统①

张　帅, 项　伟

(北京信息科技大学 自动化学院, 北京 100192)
通信作者: 项　伟, E-mail: shizsft@163.com

摘　要: 溯源可以辨别产品的真伪也可对流动人员的行动轨迹进行监控, 但由于数据存储时的安全问题很容易导致查询到虚假信息. 为了确保数据的真实性和可靠性, 我们提出了基于区块链对溯源数据的多方共享系统. 所提出的架构采用星际文件系统与区块链存储相结合的方式处理溯源数据量大的问题, 这样不仅可以大大缓解数据的存储压力还可以对链上数据施加另一层保护. 在安全问题上, 采用区块链和安全多方计算协议相结合的方法, 处理系统对外非法侵、对内横向渗透和隐私泄露的问题.

关键词: 区块链; 安全多方计算; 保密性; 隐私保护; 数据共享

引用格式: 张帅,项伟.基于区块链对溯源数据的多方共享系统.计算机系统应用,2022,31(6):394–399. http://www.c-s-a.org.cn/1003-3254/8569.html

Blockchain-based Multi-party Sharing System for Traceable Data

ZHANG Shuai, XIANG Wei

(School of Automation, Beijing Information Science and Technology University, Beijing 100192, China)

Abstract: Traceability enables the authenticity identification of a product and the monitoring of personnel mobility data. However, it is easy to find false data by query due to security issues during data storage. To ensure data authenticity and reliability, we propose a multi-party sharing system based on Blockchain for traceable data. The proposed architecture integrates the interplanetary file system and Blockchain storage to cope with a large amount of traceable data, which can significantly alleviate the pressure of data storage and impose another layer of protection on the chain. Using Blockchain and secure multi-party computation protocols can deal with the system's illegal intrusion into others, horizontal penetration, and privacy leakage in terms of security issues.

Key words: Blockchain; security multi-party computation; confidentiality; privacy-preserving; data sharing

溯源不仅可以获得流动人员的行动轨迹, 还能分辨产品的真伪. 政府、企业等机构应用了大数据、物联网以及云计算等技术竭力保证溯源数据的安全可靠. 然而, 数据仍面临着完整性和安全性等问题, 在此方面区块链技术有着很大的优势. 为了加强对系统数据的保护, 多数系统趋近于将数据转移到区块链上进行保存. 在应用区块链时如果出现数据被恶意篡改的情况, 其中运行节点会对比账本信息的一致性, 实施相应的措施并保障原有数据的安全. 区块链对数据带来的分布式、去中心化等特点如今已普遍存在于食品安全[1]、电力[2]、医疗[3] 等领域的系统中.

然而, 在中国信息通信研究院的区块链安全能力评估报告中指出当前基于区块链构建的系统存在如入侵检测能力不足、密钥保管不充分、对内存在横向渗透风险、智能合约代码审计不足、访问与监控能力不足、个人隐私易暴露、管理机制不完善等问题[4], 如何构建安全的区块链容器是当前研究的主要方向之一.

1　相关工作

区块链技术作为一个点对点的去中心化系统, 最

① 基金项目: 北京市自然科学基金 (L182032)
收稿时间: 2021-08-26; 修改时间: 2021-10-11, 2021-11-07; 采用时间: 2021-11-19; csa 在线出版时间: 2022-05-26

早出现在中本聪对比特币的介绍中[5]. 随着大众对其不断的深入研究, 区块链中加入了可编程智能合约这一概念. 智能合约[6] 就是在特定的条件下可以自动的执行交易流程或达成约定条件的合同. Vitalik Buterin 将智能合约这一概念引入到以太坊中[7], 随着概念的引入. 不同区块链平台中也出现相应的功能, 例如由 IBM 开发的企业级平台 Hyperledger Fabric 则将智能合约称之为链码[8].

区块链通过分布式网络保证数据的安全性, 但由于数据层、共识层等自身特性容易形成漏洞给黑客可乘之机. 文献 [9] 提出使用移动边缘、云计算的物联网和区块链架构, 利用去中心化实现数据的机密性、完整性和可用性, 但是对边缘设备的存储能力、计算能力等各方面性能都有一定的要求. 文献 [10] 针对于农业数据的追踪构建了双链系统, 通过与星际文件系统 (IPFS) 相结合对系统存储效率进行优化. 利用区块链实现数据的不可篡改性, 但是由于节点数据的公开性, 并没有考虑到恶意参与者对数据可以进行横向渗透从而造成数据的安全风险. 文献 [11] 设置了信任安全机制, 通过使用流量融合、聚合的方法减少恶意流量从而保护区块链节点. 由于所有的数据需通过网络服务器实现监控, 就有可能导致监控节点出现阻塞从而影响整个系统运行的情况. 文献 [12] 提出了一种压缩和私有数据共享框架为区块链上的产品数据提供了高效的私有数据管理. 但是脱链程序对产品数据进行压缩和加密的过程中存在数据处理间隙容易造成安全隐患以及在查阅数据时没有防护方法也容易出现信息泄露的情况.

安全多方计算可以处理没有可以信任的第三方的情况下, 多方共同安全地解决约定函数的问题[13]. 利用安全多方计算和区块链可以大幅提升区块链的安全等级. 其中文献 [14] 提出一种边缘智能电网的区块链双边隐私保护的多数据方案, 使用数据分段、加密、一次性地址和环签名等技术保证了至少有两个诚实节点的情况下不存在数据和身份的泄露. 文献 [15] 提出了一种基于区块链的鲁棒性安全多方计算方案, 通过奖励诚实节点惩罚恶意节点激励所有参与者合作. 用区块链维护账本的不可篡改性, 用安全多方计算维护各方隐私. 文献 [16] 利用信任矩阵构建可信评估机制形成基于安全多方的区块链可审计签名方案, 降低了恶意参与者带来的破坏并可以抵抗移动攻击. 文献 [17] 利用通信机制中的非阻塞方式支持安全多方计算中的相互通信, 将安全多方计算引入智能合约. 文献 [18] 利用链上和链下的两种方式存储数据并且使用代理重加密共享数据, 通过改进的共识算法获取多方的一致性决策并使用同态加密维护各方隐私数据. 文献 [19] 针对安全多方计算当中参与者不诚实的情况无法获取组织公平性作出研究, 结合区块链中智能合约构造惩罚机制提出更加安全的安全多方计算协议. 文献 [20] 基于 Pedersen 承诺与 Schnorr 协议的安全多方计算协议融入到区块链中, 能够充分保证节点各方的隐私.

本文通过使用区块链与 IPFS 对物联网形成的溯源数据进行保存增加数据的存储效率, 同时采用双链结构将安全多方计算协议与智能合约相结合实现数据对内隐私安全的提升并加大了区块链系统内信息的横向渗透的难度.

2 系统设计

2.1 系统整体架构

本文将区块链与安全多方计算的分布式架构相融合. 整个系统主要分为 3 部分即存储层、服务层、应用层. 安全多方计算协议思想贯穿于存储层、服务层当中, 从存储层开始对物联网所收集的数据压缩、加密; 服务层不仅通过对智能合约、SDK 以及脚本的运行实现安全多方计算中的约定函数, 还提供门限密钥为参与者提供子密钥的分发和信息确认等服务. 应用层就是打开系统应用的大门, 可以向使用人员提供便利的操作.

2.1.1 存储层

存储层将区块链的数据存储与 IPFS 相结合扩大系统本身的数据存储量. 通过智能合约保存由 IPFS 返回加密数据块的哈希值使得参与者所采集的溯源数据与区块链形成关联.

2.1.2 服务层

服务层在系统内实现了大部分的功能起到了一个承上启下的作用. SDK 与智能合约相交互通过操作逻辑向内映射完成对数据的存储、应用与查询, 进而实现安全多方计算. 更确切来说, 服务层对系统进行了更为精细的管理, 比如参与者节点的验证与授权、信息上传后加密密钥的存储、溯源链的形成.

2.1.3 应用层

应用层的实现可以将系统连接任意客户端如手机 APP、网页等, 与溯源参与者或查询人员进行交互. 通过客户端的调用连接服务层达到上传或下载相关数据

的目的.

2.2 基础性能的实现

2.2.1 双链的结构设计

目前溯源结构的实现有两种, 第一种是通过公有链或私有链将数据记录在一条链上; 第二种是使用联盟链构建主链和子链同时记录不同数据, 利用不同链间有加密措施的特性达到数据隔离的效果. 两种方式都可以对数据进行便捷查询, 也均存在区块链内部信息横向泄露的风险.

本文所构建的是区块链的双链式结构, 第一条链是保存原始加密信息的哈希值和哈希值之间的关联关系; 第二条链是保存形成溯源链后加密数据所代表的哈希值.

如图 1 所示, 第一条链主要记录不同节点各自溯源数据所代表的哈希值. 每个参与者的数据均被详细记录, 如名称、归属方、以及溯源数据哈希值, 其中哈希值的相互关联关系是指当初始节点代表的参与者与第二参与者产生交互, 根据初始参与者所提供的数据从而产生新的数据的相互关系.

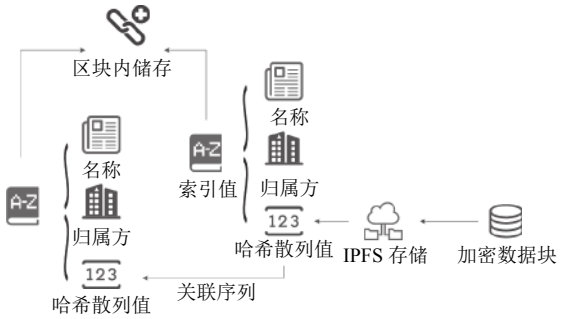


图 1 第一条链存储信息

第二条链是对第一条链数据的再加工, 通过对第一条链存储的数据由 SDK 查询、索引、组合等方式构建溯源链数据.

2.2.2 存储方式的介绍

第一条链的存储是整个系统的初始环节, 其中的溯源数据包含人为对数据的记录与说明、物联网根据传感器记录的数据. 为了极大的提升区块链的运行效率, 将采集而来呈现规律性和周期性数据进行压缩, 随后通过区块链与 IPFS 相结合扩大系统本身的数据存储量.

完整的溯源数据经过加密后保存到 IPFS 当中, 区块链数据保存其返回的哈希值. 其中, 溯源数据的加密是为了防止上传数据被不诚实节点所查看而导致数据的泄露, 同时也是为了满足安全多方计算将加密信息作为溯源链数据的组成部分的重要环节.

第二条链是当整个溯源环节已经完成, 对溯源链数据生成的存储. 溯源链数据通过参与者的交互环节进行搜集, 索引第一条链哈希散列值的相互关联关系构建溯源链. 在 SDK 中通过解密原有加密信息按照时间戳序列排列, 将溯源链数据组合完成后以同样的方式与 IPFS 相互关联.

2.3 安全多方计算约定函数

由于智能合约实现函数的复杂性有限, 为了尽量减少区块链的内存损耗, 安全多方计算中的约定函数通过链上智能合约与 SDK 编程相配合实现. 主要分为两部分: 智能合约实现信息存储并运行信息关联操作、通过 SDK 调用区块链信息组成函数.

智能合约是在区块链内部自动运行的业务逻辑. 通过智能合约可以将区块链系统内的底层架构与所运行的程序相结合, 从而实现具体的功能. 本文将自动化运行的逻辑用编程语言通过两个文件写入第一链条和第二链条中, 并使用 SDK 调用智能合约中的函数获取相应的返回值.

其中包含的智能合约与 SDK 接口所表达的功能分别如表 1 所示, 智能合约不仅支持参与者的数据初始化、上链、查询等功能, 还包含对溯源链数据的查询与检索; SDK 实现的函数包括对链上信息的上传和获取、将独立数据拼接成溯源链数据等.

表 1 函数说明

归属	函数名称	标识	函数描述
智能合约1	OnChainDataInit	f1	初始化上链数据
	OnChainData	f2	数据上链及哈希值关联
	QueryHashChangHistory	f3	查询哈希历史信息
	QueryOrgPeer	f4	查询参与者信息
智能合约2	OnChainData	f5	信息上链
	QueryNumber	f6	查询溯源链数据
SDK函数	OnChain	f7	连接智能合约传输溯源数据
	QueryData	f8	查询区块链所存储数据
	GeneratePkSk	f9	生成加密文件的公钥和私钥
	GetChainData	f10	拼接溯源链数据

至少含有i个组织的多方共享溯源系统, 根据由上述智能合约和 SDK 函数形成安全多方计算的约定函数. 参与方构建基础溯源数据是将原始数据x_i经过加密后传至 IPFS 获取相应的哈希值至链上, 如式 (1) 所示:

$$f_2\left\{IPFS\left[f_9\left(\begin{matrix}x_1\\ \vdots\\ x_3\end{matrix}\right)\right]\right\}=\begin{matrix}y_1\\ \vdots\\ y_3\end{matrix} \tag{1}$$

当参与方产生交互则溯源数据同时也会产生关联, 如式 (2) 所示:

$$f_2(y_i + y_j) \tag{2}$$

无数个关联数据构建溯源链, 构建完成的溯源链生成私钥即门限密钥共享体制的主密钥, 如式 (3) 所示:

$$f_9\{f_{10}[f_2(y_i + y_j) * n]\} = s_{k(private\ key)} \tag{3}$$

根据门限密钥共享体制存在任意t个属i中的参与者, 取大素数p构造多项式完成对主密钥的分割, 生成子密钥如式 (4) 所示:

$$f(x) = a_{t-1}x^{t-1} + a_{t-2}x^{t-2} + \cdots + a_1 x + s \tag{4}$$

且$a_1, \cdots, a_{t-1} \in$有限群$GF(p)$、$a_{t-1} \neq 0$随机选择满足$f(0) = s$的多项式并将$s_i = f(i)$分发给参与者$x_1, x_2, x_3, \cdots, x_i$手中, 此时所分发的多项式数据即为子密钥. 当子密钥创建完成后, 立即销毁主密钥.

当需要对溯源链数据进行查询时, 需在i中含k名的参与者提供子密钥. 通过k名参与者重构出主密钥, 如式 (5) 所示:

$$f(x) = \sum_{i=i}^{k} s_i \prod_{1 \leqslant i \leqslant k, i \neq j} \frac{x - x_j}{x_i - x_j} \tag{5}$$

请求溯源链数据的完整过程如图 2 所示, 当溯源查询者中的高权限查询申请者 y 发出查询请求时则会将请求消息传送至少 k 名组织参与者手中. 当至少 k 名组织参与者都同意数据的查看时, 高权限查询者将获取完整的溯源链数据. 若高权限查询者想要将溯源数据分享, 可以通过采用代理重加密技术将数据传送给普通查询者.

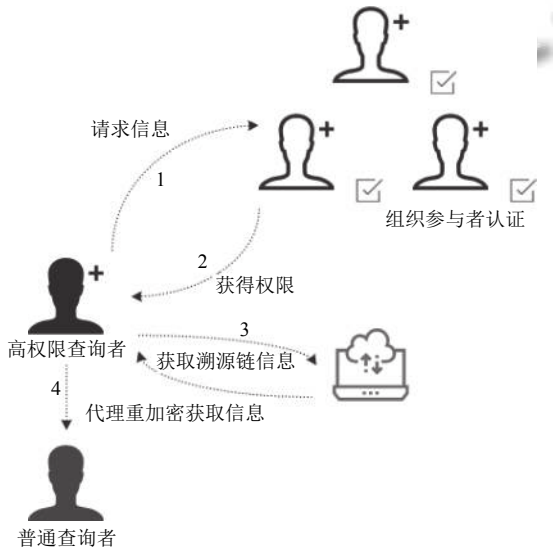


图 2 请求溯源链数据

3 系统性能分析

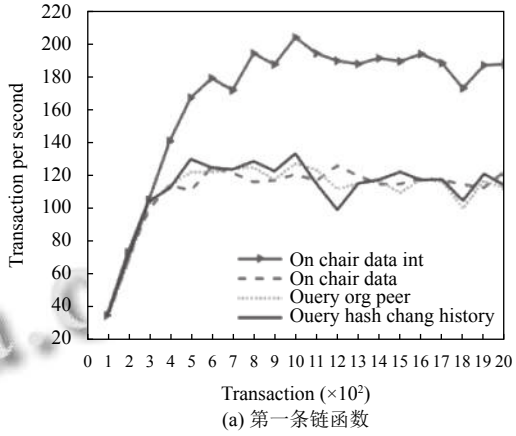
本文提出的一种基于区块链对溯源数据进行安全保护的系统, 可由不同区块链平台作为存储基础, 应用不同编程语言实现. 在测试中使用的区块链平台为企业级联盟链 Hyperledger Fabric, 并将此系统布置于一台 Intel(R)Core(TM)i7-7700HQ CPU@2.8 GHz 处理器、16 GB RAM 安装 Ubuntu 20.04 的计算机.

3.1 系统可用性测试

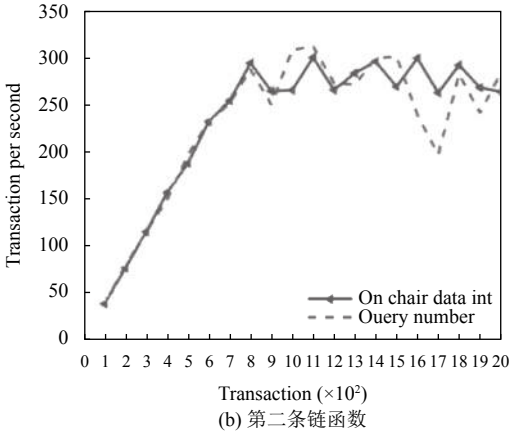
3.1.1 处理性能的分析

本文使用 Tape 对智能合约进行测试, 它是一个未使用 SDK 的轻量级测试框架, 仅通过 gRPC 向 Hyperledger Fabric 的节点发送请求, 可以准确测试出网络的真实性能[21]. 在测试环境中运行 4 个组织包含 8 个节点, 测试时将区块产生参数设置为定量, 传送总事件数作为变量, 根据传送总事件数评定智能合约函数的吞吐量性能.

如图 3 所示反映出发送待处理事件数对系统吞吐量的影响, 可以看出在大概不到三分之一处事件的处理效率随着事件数的增多而递增, 随后开始处于稳定状态.



(a) 第一条链函数



(b) 第二条链函数

图 3 智能合约中函数运行效率

根据上述智能合约运行效率分析可得出区块链系统对数据处理的速度稳定且运行效率良好, 基本可以满足企业级溯源系统的运行要求. 若对基础参数或硬件条件做出进一步的优化, 区块链网络处理事件的效率仍有很大的提升空间.

3.1.2 系统稳定性分析

通过区块链与安全多方计算构建的系统可能存在不稳定的状况. 使用 Go 语言自带的标准工具对整体数据收集、处理与分享的流程进行测试.

假设多方参与者加入系统并对数据进行存储, 数据从参与者上传至 IPFS 到区块链上的相互对应, 最后通过安全多方计算将溯源链数据分享给查询者为一个测试周期. 为了证明系统的稳定性进行了上千次测试, 随机选取其中的 500 个周期作为示例进行说明. 根据图 4 所示, 500 个测试周期中完成溯源数据的上传与分享的平均时间约 17.375 s, 这表明系统非常稳定. 并由图例可知测试周期中未出现崩溃或阻塞, 证明了区块链和安全多方计算搭建的系统具有良好的稳定性.

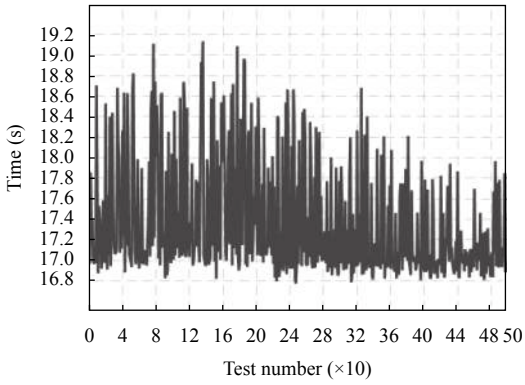


图 4 稳定性测试

3.2 系统对比分析

虽然许可链或私有链中加入系统的各节点均经过验证并且扰乱系统正常运行的可能性很低, 但会存在不诚实节点窃取系统内数据却不阻碍系统运行的情况. 其中受不诚实节点影响一些诚实节点很可能向不诚实节点转化, 由于查询操作不产生新的区块, 系统内没有有效的记录手段, 从而由区块链构建的系统无法保证对外的有效隔离.

分别对仅由区块链搭建的系统、仅由区块链和 IPFS 搭建的系统中内部节点从初始化到获取数据所需时间进行测试. 如图 5 所示获取数据共 120 次, 平均每 10 次取一次平均值, 可以看出内部节点获取数据的平均时间小于 2 s.

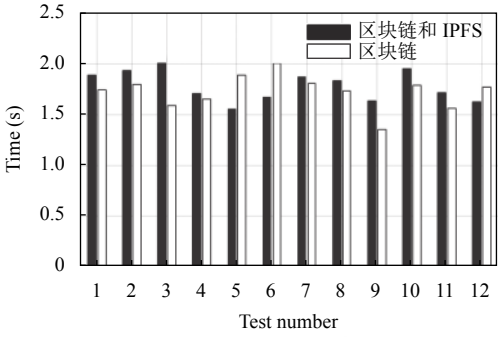


图 5 查询数据耗费时间

根据上述测试充分显示了没有保密措施时区块链内部系统容易面临着数据大量泄露的风险, 需要在区块链系统中加入一定的保护措施防止不诚实节点获取数据. 接下来对有无安全多方计算的系统在运行时进行基准测试并分析其 CPU 运行性能.

如图 6 所示, 仅由区块链和 IPFS 构建的系统在运行时 CPU 主要占比集中在对系统的运行处理上, 而加入安全多方计算后系统明显提升了对计算量、加密、数据读取的 CPU 占比, 从而可以直观的反应出本文所述方案对数据处理的复杂程度进一步提高, 在系统内数据的安全性得到了进一步的加强.

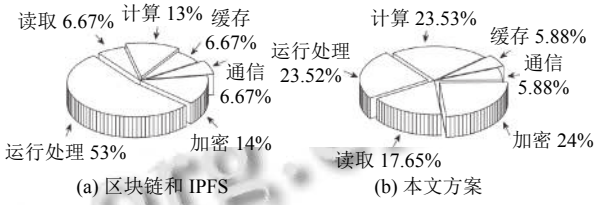


图 6 CPU 运行占比

3.3 应用效果分析

本文与其他区块链系统的方案的不同之处在于引入了安全多方计算, 并利用双链将数据采集与查询分开, 细化了各部分的职责更容易对数据的流向做出深度监控. 通过从数据保存和多方共享两个角度改善了以区块链为基础的系统对内数据渗透的安全问题和对外信息的保密程度.

具体应用时区块链仅保存 IPFS 返回的哈希序列. 这样不仅对区块链存储容量起到了扩展作用还在一定程度上避免了数据空间资源量大而导致的效率问题. 数据由外界显性记录与区块链隐性记录相结合充分保证了不可逆转和不可篡改的特性. 同时, 搭建安全多方计算协议为基础使各参与者数据得到保障, 最后在数据的共享阶段采用代理重加密技术将信息分享的路径透明化. 整个

系统不仅实现了对事件处理的多次核对还对上传数据进行了严密的安全保护, 这使得系统本身可以防御多种攻击手段如女巫攻击、分布式拒绝服务攻击等.

根据所述方案具体可应用于通过联盟链、私有链构建的多方数据交互系统中. 例如可以应用在如下会产生多方数据交互的场景中: 对产品监管的溯源链中, 在不泄露任何节点数据的基础上做到对产品上中下游整体流程的呈现; 疫情防控时各个节点对数据的采集, 保证形成个人行动轨迹数据的安全; 居家环境中对分布式智能家电收集数据的共享与查询.

4 结论与展望

本研究从实际溯源数据安全的角度出发提出了基于区块链对溯源数据的多方共享系统, 通过将安全多方计算协议应用到溯源数据的生成与查询中提高系统的安全性. 为基于区块链构建的系统提供了一种防止内部横向攻击的新思路.

目前随着区块链应用变得越来越广泛, 当中存在的问题也逐渐暴露. 对如何提升区块链应用的安全问题与运行效率还需做大量的研究工作.

参考文献

1 Baralla G, Pinna A, Corrias G. Ensure traceability in European food supply chain by using a Blockchain system. Proceedings of the 2019 IEEE/ACM 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain. Montreal: IEEE, 2019. 40–47. [doi: 10.1109/WETSEB.2019.00012]

2 Kim Y, Kim KH, Kim JH. Power trading Blockchain using hyperledger fabric. Proceedings of 2020 International Conference on Information Networking. Barcelona: IEEE, 2020. 821–824. [doi: 10.1109/ICOIN48656.2020.9016428]

3 Lamba R, Gupta Y, Kalra S, *et al*. Preventing waiting list manipulation and black marketing of donated organs through hyperledger fabric. Proceedings of 2019 International Conference on Computing, Communication, and Intelligent Systems. Greater Noida: IEEE, 2019. 280–285. [doi: 10.1109/ICCCIS48478.2019.8974526]

4 中国信通院. 区块链行业: 区块链安全能力测评与分析报告. http://www.caict.ac.cn/.

5 Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. Decentralized Business Review, 2008: 21260.

6 Butterin V. A next-generation smart contract and decentralized application platform. White Paper, 2014, 3(37).

7 Ethereum. https://ethereum.org.

8 Hyperledger Fabric. Advancing business Blockchain adoption through global open source collaboration. https://www.hyperledger.org/.

9 Pavithran D, Al-Karaki JN, Shaalan K. Edge-based Blockchain architecture for event-driven IoT using hierarchical identity based encryption. Information Processing & Management, 2021, 58(3): 102528. [doi: 10.1016/j.ipm.2021.102528]

10 Ren W, Wan XT, Gan PC. A double-blockchain solution for agricultural sampled data security in Internet of things network. Future Generation Computer Systems, 2021, 117: 453–461. [doi: 10.1016/j.future.2020.12.007]

11 Meng WZ, Li WJ, Zhou JY. Enhancing the security of Blockchain-based software defined networking through trust-based traffic fusion and filtration. Information Fusion, 2021, 70: 60–71. [doi: 10.1016/j.inffus.2020.12.006]

12 Qi SY, Lu YS, Zheng YQ, *et al*. Cpds: Enabling compressed and private data sharing for industrial internet of things over Blockchain. IEEE Transactions on Industrial Informatics, 2020, 17(4): 2376–2387. [doi: 10.1109/TII.2020.2998166]

13 Yao AC. Protocols for secure computations. Proceedings of the 23rd Annual Symposium on Foundations of Computer Science. Chicago: IEEE, 1982. 160–164.

14 Guan ZT, Zhou X, Liu P, *et al*. A Blockchain based dual side privacy preserving multi party computation scheme for edge enabled smart grid. IEEE Internet of Things Journal, 2021. [doi: 10.1109/JIOT.2021.3061107.]

15 Gao HM, Ma ZF, Luo SS, *et al*. BFR-MPC: A Blockchain-based fair and robust multi-party computation scheme. IEEE Access, 2019, 7: 110439–110450. [doi: 10.1109/ACCESS.2019.2934147]

16 王韫烨, 程亚歌, 贾志娟, 等. 基于安全多方的区块链可审计签名方案. 计算机应用, 2020, 40(9): 2639–2645.

17 朱岩, 宋晓旭, 薛显斌, 等. 基于安全多方计算的区块链智能合约执行系统. 密码学报, 2019, 6(2): 246–257. [doi: 10.13868/j.cnki.jcr.000299]

18 王童, 马文平, 罗维. 基于区块链的信息共享及安全多方计算模型. 计算机科学, 2019, 46(9): 162–168. [doi: 10.11896/j.issn.1002-137X.2019.09.023]

19 黄建华, 江亚慧, 李忠诚. 利用区块链构建公平的安全多方计算. 计算机应用研究, 2020, 37(1): 225–230, 244. [doi: 10.19734/j.issn.1001-3695.2018.07.0479]

20 刘峰, 杨杰, 李志斌, 等. 一种基于区块链的泛用型数据隐私保护的安全多方计算协议. 计算机研究与发展, 2021, 58(2): 281–290.

21 Tape. A light-weight tool to test performance of hyperledger fabric. https://github.com/Hyperledger-TWGC/tape.