区块链数字取证：技术及架构研究

范伟[1,2]，李海波[1,2]，张珠君[1,2]
（1.中国科学院信息工程研究所，北京 100093；2.中国科学院大学网络空间安全学院，北京 100049）

摘　要：针对传统数字取证存在场景适应性差、证据保全能力弱以及取证效率低的问题，分析将去中心化、不可篡改的区块链技术引入数字取证的可行性。首先，基于区块链取证技术的层次架构，提出了阶段化取证流程，并剖析了区块链技术在证据获取、保全和呈现阶段的研究进展。其次，通过分析现有研究的不足，结合区块链的分布式优势，设计了一套区块链全流程参与的数字取证架构，将证据信息融入链上数据结构并提出了配套的图分析算法，统一了各场景下的证据采集形式；利用链下分布式数据库实现了高效扩容存储；借助智能合约模板提升了同类型取证事务的合约复用性。最后，展望了区块链技术在未来取证应用中的研究方向。

关键词：数字取证；区块链；取证元数据；智能合约；图分析算法
中图分类号：TP309.2
文献标志码：A
DOI：10.11959/j.issn.1000–436x.2024204

Blockchain digital forensics: technology and architecture

FAN Wei[1,2], LI Haibo[1,2], ZHANG Zhujun[1,2]
1. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China
2. School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: Issues of limited scene adaptability, inadequate evidence preservation, and low efficiency in traditional digital forensics were addressed by analyzing the feasibility of incorporating decentralized, tamper-resistant blockchain technology into digital forensic practices. Initially, a phased forensic process was proposed based on a hierarchical architecture for blockchain forensic technology, examining the advancements of blockchain at each stage of evidence acquisition, preservation, and presentation. Subsequently, limitations in existing research were analyzed, and a digital forensic framework incorporating comprehensive blockchain involvement was designed by utilizing the distributed advantages of blockchain. This framework integrated evidence information into the on-chain data structure and introduced a complementary graph analysis algorithm to standardize evidence collection across various scenarios. An off-chain distributed database was employed to achieve scalable, efficient storage, while smart contract templates enhance the reusability of contracts for similar forensic transactions. Lastly, potential future directions for the application of blockchain technology in forensic science were explored.

Keywords: digital forensics, blockchain, forensic metadata, smart contracts, graph analysis algorithms

0　引言

Jordaan 等[1]将数字取证定义为“数字证据的识别、保存、检查和分析，经过科学认可的验证过程，最终在法庭上呈现该证据以回答法律问题”。现如今，数字取证技术已经涵盖了数字信息的各类场景，如物联网取证、社交媒体取证等。根据 Verizon 最新发布的 2024 年度数据泄露调查报告[2]（DBIR, data breach investigations report）显示，伴

收稿日期：2024–09–03；修回日期：2024–11–07
通信作者：张珠君，zhangzhujun@iie.ac.cn
基金项目：国家重点研发计划基金资助项目（No.2021YFB2700603）
Foundation Item：The National Key Research and Development Program of China (No.2021YFB2700603)

随数字化程度加深，新型网络犯罪层出不穷，涉及领域也愈发广泛，传统数字取证技术在取证场景适用性、数据安全、取证事务效率等诸多层面面临挑战。

区块链作为新兴技术手段，在2008年被提出后，凭借其不可篡改、可追溯、去中心化等特性，成为研究人员构建取证方案的重要技术支持。一方面，区块链的不可篡改特性为证据信息提供了安全保障；另一方面，其去中心化的分布式存储方式提升了证据保存的可靠性。

为梳理区块链技术在数字取证领域的应用现状，本文归纳了区块链取证技术层级架构及流程模型，在此基础上综述了区块链在各取证流程中的研究工作，并提出了一套区块链全流程参与的取证架构。最后基于现有研究的不足，展望了区块链技术在数字取证领域未来的研究方向。本文主要贡献如下。

1) 设计了区块链取证流程模型，涵盖取证、存证和呈现3个阶段，为后续研究提供了清晰的划分依据。

2) 全面剖析区块链技术在证据采集、保全和呈现中的应用，对应综述了在3个层面的研究进展、技术类型及各方案的优缺点。

3) 提出了一套区块链数字取证架构，配合3个阶段分别设计取证元数据结构和图分析算法、链下分布式存储和智能合约模板，实现区块链全流程参与的数字取证。

1　研究背景

1.1　传统数字取证

根据数字取证研究工作组（DFRWS, digital forensic research workshop）提出的取证模型，参照“取证过程”分类标准[3]，传统数字取证技术可以分为取证分析技术和证据管理技术两类，如表1所示。取证分析技术对潜在证据信息进行收集、处理、分析和解释，从而支持安全事件调查。证据管理技术对原始证据进行跟踪管理，实现证据全生命周期的数字化。

传统的取证分析技术在应对复杂证据链及烦琐取证步骤时，容易导致证据链断裂或混乱[3]。基于数据备份和克隆的证据管理技术在传输和应用过程中，证据完整性需要额外验证。面对新型取证场景和复杂网络环境，传统取证工具往往依赖于人工参与且效率低下。这些问题为数字取证领域的发展提出了严峻考验，而区块链技术的引入为解决这些问题提供了新的思路。

1.2　区块链技术特点

区块链作为分布式账本技术，具有不可篡改、去中心化、透明性、可追溯性等特性。

区块链的不可篡改特性依赖于自身的密码学机制，数据以链式结构进行加密存储，每个区块包含前一区块的哈希值，避免了链上信息被恶意篡改。区块链网络中数据分布式存储于多个节点，每个节点都拥有完整的账本副本，避免了单点故障问题，进一步提升了数据存储的安全性。区块链账本公开透明，所有链上操作均可通过链式结构记录并验证，从而维持数据的完整性与可追溯性。

1.3　数字取证与区块链的融合

1.3.1　区块中的证据形式

区块链由一系列按时间顺序连接的区块组成。基于区块的证据存储如图1所示，区块由区块头和区块体构成，区块头包含两类数据信息：一类表明区块的内容信息，包括前一区块哈希值、版本号、区块高度和Merkle树根哈希值；另一类指明交易属性，即挖矿难度目标和随机数等。在区块体中，证据信息以区块链形式交易，通过Merkle树结构进行逐级哈希，生成根值存入区块头，实现证据存储。

表1　传统数字取证技术

类别	取证技术	技术简介	取证工具
取证分析	磁盘分析	解析磁盘空间，提取被隐藏或删除的潜在信息	EnCase[4]、FTK[4]等
	内存取证	通过内存映像雕刻内核结构，逆向重建进程	Dumpit[5]等
	网络取证	记录分析网络流量，检测异常行为	Wireshark[6]等
证据管理	数据备份	创建副本以支持系统和数据恢复	Veeam[7]等
	数据克隆	制作数据的精确副本，确保数据完整性	Clonezilla[8]等

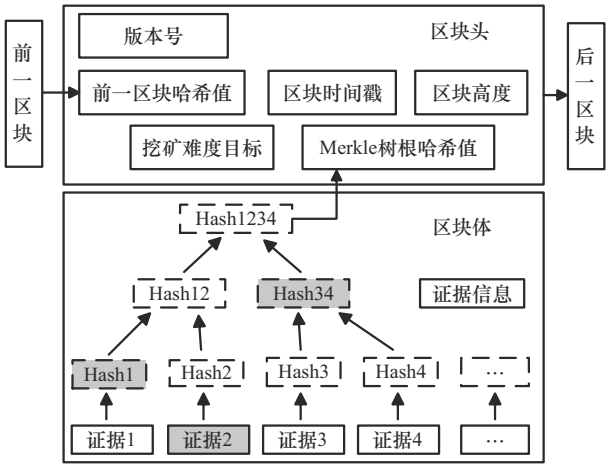

图1　基于区块的证据存储

Merkle树支持简易支付验证机制，区块中的证据信息可以通过核验根值路径上的部分节点完成验证。例如，图1中“证据2”仅通过灰色节点Hash1和Hash34完成验证。

1.3.2　基于区块链的证据保护

最高人民法院《关于互联网法院审理案件若干问题的规定》（2018）的第十一条指出，鼓励和引导当事人通过哈希摘要、可信时间戳、数字签名和区块链等技术固定、留存和收集证据。相关技术特点如表2所示。

表2　　数字证据处理形式

类型	主要优势	局限
哈希摘要	防篡改、不可逆、成本低	效果单一、摘要时间长、效率低
可信时间戳	提供时间追溯、成本低	效率低
数字签名	防篡改、不可抵赖	成本高、存在单点故障问题
区块链	防篡改、可追溯、自动化	扩展性受限

与传统数据保护技术相比，区块链技术覆盖面更广，且支持数据的去中心化存储。智能合约的应用也提升了数字证据处理过程的自动化程度。

1.3.3　自动取证的智能合约

智能合约最早由密码学家尼科萨伯于1994年提出。作为计算机程序化的协议，智能合约按照预定义的条款自动执行节点交互等行为。在数字取证领域，通过部署专门的智能合约来处理证据的采集、存储、验证和追溯等任务，可以在不需要第三方介入的情况下自动完成取证事务，在降低人工成本的同时，进一步提升了数字取证效率。图2展示了智能合约实现自动化取证过程。

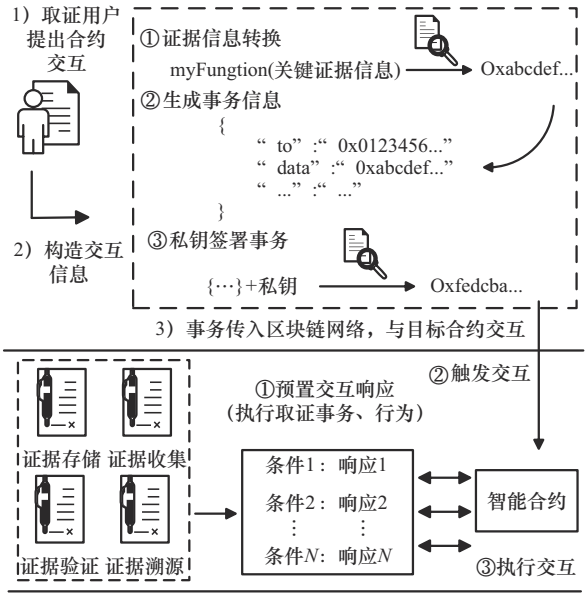

图2　智能合约实现自动化取证过程

2　区块链取证架构

本节通过分析数字取证场景下的区块链层次化架构，明确了取证过程中各层次的功能以及层间信息交互方式，进一步归纳了区块链取证的基础流程模型。层级架构提供的结构化方法将区块链技术融入取证过程中，确保取证程序系统化、安全且高效。通过定义各层的具体功能和信息交互方式，促进了取证流程模型的衔接和实施。

2.1　区块链取证层级

本节将区块链技术架构结合取证场景划分为5个层次，分别为数据层、网络层、共识层、取证合约层和取证应用层，如图3所示。

1) 数据层。作为区块链技术的底层结构，封装了数据区块的链式结构以及相关的非对称公私钥数据加密、时间戳等技术。此外，数据层还为网络层提供证据存储、检索及相关信息支持。

2) 网络层。负责节点间数据传输，实现取证区块传播及取证节点的维持与发现。同时，网络层负责维护网络安全，防止恶意攻击和数据窃取。

3) 共识层。负责达成与维护区块链网络的共识，确保所有节点对区块链状态的一致性判定，同时还负责取证区块和取证事务的有效性验证。共识消息借助数据层信息实现核验，通过网络层接收和传播，进一步与取证合约层交互执行与验证合约。

4) 取证合约层。负责执行和管理智能合约，实现预定义合约的自动化执行，并定义和管理取证

过程中的规则和流程。取证合约层与数据层紧密交互，证据信息作为参数触发合约执行，同时合约执行结果作为证据追溯信息。

5) 取证应用层。为用户和取证人员提供取证应用和接口，负责取证数据的可视化、分析和报告生成。通过接口获取智能合约的执行信息，实现与取证人员和其他用户的交互。取证应用层通过与取证合约层和数据层交互执行操作和获取数据。

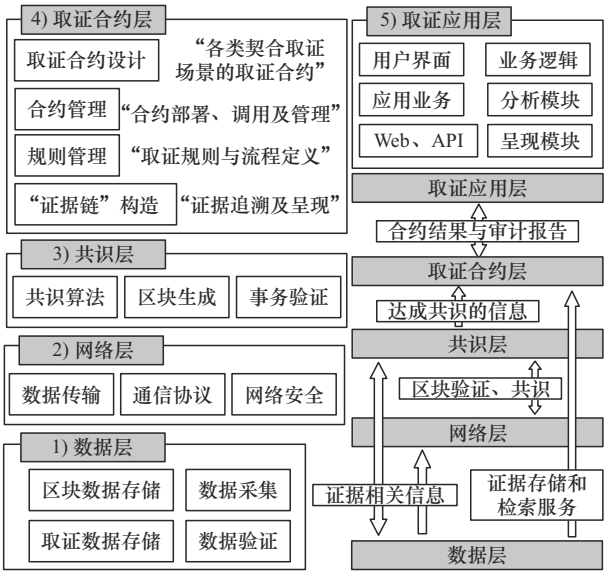

图3　区块链取证层级架构及交互方式

2.2　区块链取证基础流程模型

在区块链参与的数字取证中，证据信息的转化和区块链操作紧密耦合。对应数字取证的不同阶段，结合前文“取证过程”的思想，提出了区块链参与的三阶段数字取证流程基础模型，如图4所示，分为取证、存证和呈现阶段。

2.2.1　取证阶段:构造区块链证据

取证阶段通过数据层进行信息收集和处理，通过网络层实现信息的节点间传输，并由取证合约层实现标准化分析。该阶段重点在于逐步构造“区块链证据”，过程包括3个步骤：确定取证目标与收集、证据预处理和取证分析。

在这一阶段，取证系统需要准确识别并采集目标数据，在明确取证范围的基础上，将“潜在数据”转化为“数字证据”。进一步剔除主体信息缺失、来源验证错误等不规范的证据信息，挖掘证据间的关联关系，配合可信时间戳生成可追溯的链上“区块链证据”。

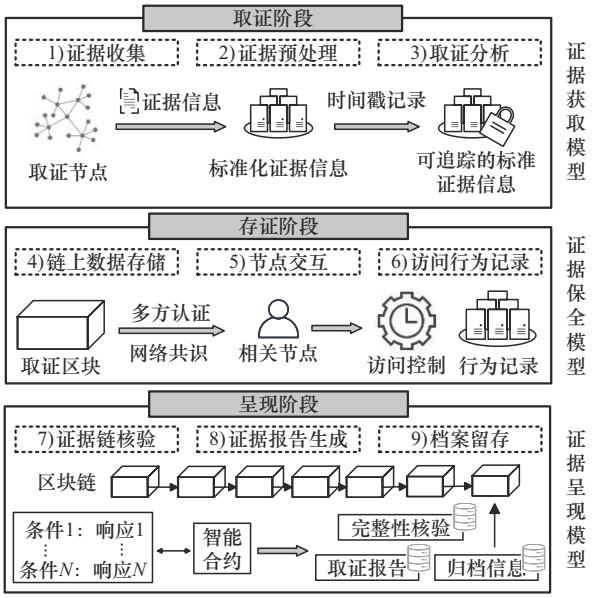

图4　区块链数字取证流程基础模型

2.2.2　存证阶段:证据上链存储

存证阶段通过数据层完成数据存储，并借助网络层实现节点交互，同时由共识层验证上链的过程。该阶段重点在于数字证据的安全存储，辅助链上数据管理以实现高效的存取过程，包含3个步骤：链上数据存储、节点交互和访问行为记录。

在这一阶段，将“区块链证据”存入区块，经过网络共识后配合不同的上链策略进行存储。取证节点与链上数据进行交互，实现许可下的区块信息索引查阅等行为，原始证据记录和链上访问信息作为溯源依据存入区块。

2.2.3　呈现阶段:智能合约溯源

呈现阶段利用取证合约层自动追溯“监管链”信息，配合取证应用层呈现和归档证据。该阶段包括3个步骤：证据链核验、证据报告生成和档案留存。

证据监管链（CoC, chain of custody）[9]是取证过程中的重要概念，代表对证据全生命周期的记录和呈现，其溯源方式和内容直接影响证据的可信度。传统取证技术面临大量数据核验，缺乏自动化分析形式[10]，而区块链取证在这一阶段利用智能合约自动执行的特性，事先规定从证据上链到呈现展示过程中的取证事务。进而将整个调查过程、证据监管链以及庭审意见进行归档留存，以便于配合后续审查和评阅。

3　区块链取证技术剖析

本节依照取证流程，综述证据获取、保全和呈

现3个阶段内区块链的相关研究进展、技术形式和改进思路。分析应用区块链技术实现“数据”到“可信数据”，生成“证据”并配合“链上信息”构造“证据监管链”，实现溯源和呈现的全过程。

3.1 证据获取阶段：数据采集与取证分析

区块链技术在该阶段辅助数据采集和取证分析，实现从采集“数据”到“证据”的转换过程，相关方案如表3所示。

3.1.1 区块链辅助的数据采集

物联网环境中存在大量设备充当证据来源[11]，研究者们探索了该场景下信息采集的高效性和准确性。文献[12]将智能合约与车辆中心系统集成，每当取证事件发生时，车辆中心系统配合相关车辆共同触发合约的采集行为。文献[13]采用类似做法，将智能合约集成在目标设备中，目标设备参数构建“指纹”信息，配合设备日志实现可信验证。智能合约还为每个取证事件生成唯一编号，并提供查询接口，为后续取证分析打下基础。

部分研究致力于在传统司法领域构建区块链取证平台，配合明确的取证流程规范化采集数据信息。文献[14]利用区块链在数字审计领域收集各系统分散的目标信息。针对被审计方不同系统间的兼容性问题，研究在客体系统上部署区块链审计节点，再通过主体链构建信息采集平台，实现客体链数据的统一和标准化采集。

以采集的信息为基础，研究还利用区块链共识进行可信验证，实现从“数据”到“可信数据”的转变。在文献[15]的区块链取证方案中，通过动态组建联盟节点对数据进行核验，并将验证结果一同打包作为取证数据上链。分布式共识提供的“事件证明”为数据采集增添了可靠性。

3.1.2 区块链辅助的取证分析

在取证分析环节，相关研究者尝试结合链上可信环境，辅助“可信数据”到“证据”的转换。

针对取证场景，文献[16]调整去除了区块原有的金融属性字段，提出了一种专用于取证分析的区块结构，额外记录信息来源和去向等信息。借助该结构将设备间的复杂通信统一处理为链上交易，避免了取证人员和设备之间的交互。文献[17]为了解决传统手机取证中，证据采集到取证分析过程中间隔过长且参与人员、机构复杂等问题，将灵活的内存分析技术与区块链结合，通过授权中心统一组织的可信环境，构建了全取证流程的实现方案并减少了对人员机构的依赖。

对于链上证据的取证分析而言，相关研究还探索构建区块链证据关联图提供便利。文献[18]设计了用于持续追踪、采集和更新区块的模糊智能合约（FCS, fuzzy based smart contract），通过配合采集信息的关键属性构建证据关联图（ECG, evidence correlation graph），为取证分析带来了辅助作用。采用类似的思路，文献[19]提出了基于图论算法的智能合约（GNNSC, graph neural network-based smart con-

表3 证据获取阶段的取证方案

事务类型	文献	技术特点	取证方案	方案评价
数据采集	文献[12]	合约验证	智能合约计算签名验证密钥，推断参与事故的相关车辆，进而采集各方数据	具有灵活性但存在中心化节点
	文献[13]	集成合约的取证设备	在EOSIO区块链中利用智能合约配合数字指纹不断采集设备的日志信息，确保取证信息来源可信	适用跨域场景的日志采集与溯源，仅理论分析
	文献[14]	双链审计	双链结构组成的混合审计模式，主客链配合实现数据采集和辅助评估	仅理论分析可行性，需要进一步实验验证
	文献[15]	动态节点选举	动态组建链上节点核验数据，为取证分析提供可靠“事件证明”	低时延、去中心化程度高，但有恶意节点问题
取证分析	文献[16]	取证事务区块	将物联网设备间的复杂通信统一处理，存入改进的取证区块，各个参与者均可验证信息完整性	简化了异构设备数据的存储过程，但链上存在隐私性问题
	文献[17]	实时数据取证	将区块链技术、自动取证工具和记忆取证技术融合，实现数据及时上链	实现数据的实时取证，但适用范围有限
	文献[18]	证据关联图分析	定义模糊智能合约（FCS）进行风险数据的针对性采集，生成逻辑证据图，提供更直观的取证分析形式	证据关联图依赖于人工构建，且无更深入的关联分析
	文献[19]	图论算法	利用基于图论算法的智能合约（GNNSC）追踪数据，简化取证分析过程	计算复杂度高且扩展性需要验证

tract），允许自动跟踪服务器上的数据访问情况，实现信息的全生命周期追溯，同样构建匹配日志属性的证据关联图用于辅助取证分析。上述研究利用合约采集的数据为取证分析提供图的分析形式，但依旧需要人为参与实现证据之间的关联。为进一步提升区块链在全取证流程中的参与程度，此处可以尝试引入图分析算法，以更直接的形式参与数字取证过程。

3.2　证据保全阶段：链上存储扩展与优化

在证据保全阶段，相关研究围绕链上存储的扩展与优化展开。过程中，对链上访问和操作等信息进行“二次取证”，实现“证据”到“证据+链上信息”的转换。

3.2.1　扩容的证据存储

随着取证系统的增长，区块链取证节点面临日益增加的存储成本和计算压力，链上扩容显得尤为紧迫。现有研究通过“减少链上数据”和“引入额外空间”实现扩容，相关方案如表 4 所示。

首先，研究人员尝试改变区块链的原有结构，仅存储部分区块或区块内的部分数据，实现空间扩容。在文献[20]提出的混合链结构中，多数解耦的区块节点仅存储块头信息，部分非解耦的区块节点参与区块验证，既减少了链上数据又提升了数据的一致性判定效率，但需要关注节点减少可能带来的安全问题。文献[21]将频繁访问的证据信息存储于“热”区块，而将不频繁或不变数据存储于“冷”区块，根据“冷热”程度实现以取证交互为主的热区块扩容。但在具体的“冷热”程度划分上，缺乏明确的标准。文献[22]设计的碎片化账本仅维护不同参与者的部分信息，并在共享账本中保留证明，以较小的存储和处理开销实现可追溯的取证分析，但同样面临数据完整性的问题。

“引入额外空间”主要通过两种策略——“链下存储”和“多链扩容”。“链下存储”在区块链外存储数据，在不影响原有区块结构的情况下扩展系统空间。“多链扩容”构造辅助链用于取证存储，在链结构和类型的选择上具有灵活性。

传统数据库首先被用于实现“链下存储”。文献[23]通过多个数据库将证据信息和取证系统分离，分离的取证事务为取证交互提供了便利。文献[24]采用私人托管数据库，区块链负责维护系统安全。文献[25]提出了一种链下证据协同管理模式，借助分组核心网中的数据管理功能，仅将提升溯源过程可信度的关键信息上链，但共识算法是阻碍该系统吞吐量进一步提升的关键因素。上述研究利用传统数据库实现了低成本的证据空间扩容，但中心化的存储方式需要额外的安全支持。

星际文件系统（IPFS, interplanetary file system）[26]具有内容寻址和分布式存储的特性，扩容的同时维持了取证系统的去中心化。相关研究通常将完整的证据信息存入 IPFS，并利用生成的内容标识符（CID, content identifier）配合证据哈希上

表 4　证据保全阶段的扩容方案

扩容思路	文献	技术特点	扩容方案	方案评价
链上缩减	文献[20]	区块解耦	混合链结构，区块头与区块体解耦，多数区块节点仅存储区块头信息	减少链上数据，但影响了原有区块或链的结构，需要额外注意数据完整性问题
	文献[21]	块分类存储	结合证据的“冷热”程度分区块存储，实现以交互为主的链扩容	
	文献[22]	轻量级框架	轻量级碎片化账本，降低存储和处理开销	
传统数据库	文献[23]	数据库备份	对数字证据信息通过链下多个数据库存储和备份	成本低、扩容效果好，但过于中心化，需要与去中心化的区块链技术进行协调
	文献[24]	私人数据库	利用 ISP 提供的托管数据库进行证据链下存储	
	文献[25]	链下协同	继承分组核心网特性的链下证据协同管理模式，提升整体效率	
IPFS	文献[26]	节点交互	多取证节点和 IPFS 交互，实现信息链下安全构建与存储	链下分布式扩容效果好，但需要避免过多交互造成的网络时延，同时需要处理证据内容安全问题
	文献[27]	链下防护	在 IPFS 基础上构建安全层，实现数据验证与追溯	
	文献[29]	节点交互	系统参与者均作为对等节点参与 IPFS 网络	
多链扩容	文献[32]	多链存储	多个成本较低区块链进行取证数据的辅助存储	延续了去中心化特性，但侧链、跨链等技术实施和互操作协议仍需发展完善
	文献[33]	双链结构	双链结构将取证事务分离，跨链实现信息交互	
	文献[34]	侧链扩容	Polygon 区块链平台的二层扩容方案	

链，此处称以上结合过程为“基础模式”，包括文献[26-30]。在“基础模式”上，各研究通过构造不同的交互方案提升对取证场景的适应性。

文献[29]将取证系统中的所有参与者均作为IPFS协议中的对等网络节点，通过证据标识检测存储过程中的非法访问和数据篡改，并记录到最终的取证报告中。文献[27]认为现有IPFS仍存在审计功能缺乏的问题，在提出的BlockIPFS中构建额外的安全层，清晰的审计轨迹为取证分析提供了高度可信的链上数据。Li等[31]认为，在涉及警务调查的数字取证中，应由警务处、法院或其他相关部门合作建立分布式存储系统，避免纯商业式的分布式存储。因此，文献[26]在取证系统中配合设计了警务处、法院等节点，各节点分别与IPFS进行交互，共同实现数据完整性验证。

在利用IPFS进行扩容的同时，需要注意该系统在取证场景下面临的问题，如缺乏证据审计，也需要注意复杂节点交互导致的网络时延等问题。此外，“链下存储”需要额外关注链外数据写入区块链过程的传输安全问题。

“多链扩容”策略利用分层区块链、侧链或多链等技术，分离取证事务或节点数据信息。文献[32]提出利用多个廉价链辅助进行数据存储，为主链增加存储空间。文献[33]将取证模型划分为事故认定链和证据保存链，事故信息存储在证据保存链中，通过双链运行模式，实现了取证事务的分离。此外，文献[34]借助支持跨链策略的区块链平台Polygon部署存证模型，利用以太坊二层网络为基础设施开发及扩容提供便利。

3.2.2 强安全性的证据存储

本节从链上数据加密和系统访问控制的角度，全面分析了区块链技术如何在取证场景中保障数据安全，相关方案如表5所示。

加密技术是区块链安全性的基石。一方面，研究人员针对取证场景不断改进区块链中的哈希方式和加密手段，提升取证的灵活性和安全性。另一方面，研究人员将一些新兴领域技术，如基于格的密码、零知识证明和人工智能算法等，配合区块链共同提升取证场景中的数据安全。

模糊哈希（FZ, fuzzy hashing）技术[35]，将目

表5 证据保全阶段的强安全性方案

技术类别	文献	关键技术	强安全性方案	方案评价
模糊哈希类	文献[37]	模糊哈希SSDEEP	SSDEEP模糊哈希构建Merkle树结构，以“区块”为单位进行相似性分析	灵活性高，但在适用性和准确性上需要提升
	文献[38]	模糊哈希mrsh-v2	采用模糊哈希mrsh-v2创建哈希值，结合分层布隆过滤器检测区块内的语义匹配程度	
数字签名类	文献[13]	设备数字指纹	将取证节点参数配合IP构建数字指纹，确保节点可信	取证节点的数据安全性、隐私性高，但需要注意计算开销、场景适应性和复杂度问题
	文献[31]	ElGamal	短随机签名结合ElGamal加密实现安全投票，将证据信息分片后安全传输给各取证参与方，确保法庭数据的隐私性	
	文献[39]	强鲁棒性水印	设计强鲁棒性水印嵌入数字证据，验证证据信息的完整性	
	文献[41]	改进Merkle签名	改进Merkle签名结合数字证书实现证据加密	
系统访问控制	文献[47]	阶段访问控制	配合取证进程的基于角色的分阶段访问控制	细化了取证节点的权限管理，利于证据监管链追溯，但为取证系统引入复杂程度
	文献[48]	基于角色和属性结合访问控制	综合角色和属性的访问控制，配合场景实现主体权限设置	
	文献[49]	分布式访问控制	分布式访问控制，利用智能合约将取证程序建模为有限状态机，确保节点间信任	
新兴领域技术	文献[18]	结合深度学习的椭圆曲线密码	深度学习结合椭圆曲线加密算法SA-DECC，根据证据的敏感程度进行加密	高灵活性，复杂程度高、计算资源消耗大
	文献[42]	格密码可编程哈希	基于格密码的可编程哈希函数对证据信息签密，实现低复杂度下的抗量子计算攻击，取证节点匿名化确保隐私性	抗量子攻击的安全程度，但复杂度高
	文献[44]	零知识证明	引入零知识证明验证取证节点的有效性	强隐私性和高效性，但计算资源消耗大
	文献[45]	人工智能	人工智能预取证场景并过滤证据信息，区块链技术配合实现取证节点数据安全传输	高适应性和自动化，存在认可度和隐私问题

标数据分块计算哈希值后，拼接得到关于整体的模糊哈希，进而借助近似匹配算法分析数据的相似度。已有研究将该技术应用于区块链取证领域，利用模糊哈希的灵活性作为对“不可篡改”的链上环境在应用局限性上的补充[36]。

在文献[37]的区块链取证方案中，先利用哈希函数 SHA-256 计算证据项用于数据完整性验证，而在上链过程中改用模糊哈希算法中的样本字符串相似性检测（SSDEEP, super string search deep）递归构建证据信息的模糊 Merkle 根值，以“区块”为单位进行相似性分析，推断可能存在修改的区块。类似地，文献[38]采用 mrsh-v2 模糊哈希算法，在图像取证领域提出了配合分层布隆过滤器的模糊哈希取证范式。该方案筛选匹配度高于 95% 的区块视为有效数据参与证据链的构建。但此类方案目前主要针对文本、图片等特定格式的数据进行处理，适用场景有限且现有研究以区块为单位的相似度比较，当在单个区块中存储多案件证据信息时，可能存在准确性上的问题。

一类研究将加密技术的改进，首先应用在区块链证据信息的处理上。文献[39]针对证据信息设计了水印嵌入与提取方案，该方案将公钥通过扩充、翻转等变换得到的水印图像，附着在取证信息上参与数据完整性验证，增强了取证过程的安全性和隐私性。类似于水印，文献[13]在先前研究 Chrono EOS[40]的基础上，利用取证设备参数信息生成独特的设备指纹，配合管理员私钥共同验证取证数据来源。

另一类研究则通过改进或结合不同的加密技术以适应区块链取证场景。文献[31]为庭审环节设计了安全投票方案，通过短随机签名验证用户身份，再利用 ElGamal 加密算法将分片后的投票信息进行加密并传给不同取证参与方，确保了庭审数据的隐私性。同样针对隐私问题，文献[41]允许各参与方利用一次性签名方案随机生成公钥对证据密签，进而与证据共同上链，实现可信标记。

研究人员还积极探索区块链与新兴技术在取证场景中的结合，寻求更安全、更智能的链上取证。文献[42]将格密码用于数字证据签密，以较低的复杂度实现节点身份匿名化，同时能够抵御可能的量子计算攻击。零知识证明[43]要求验证者在表明陈述是真实的同时不泄露关于该陈述的任何信息。文献[44]借助该技术配合区块链加密，确保了参与方无法利用伪造的私钥获取隐私数据。配合深度学习算法，文献[18]提出了深度椭圆曲线密码（SA-DECC, sensitivity aware deep elliptic curve cryptography）算法，实现了对数字证据信息的灵敏度感知功能。文献[45]同样将深度学习神经网络和区块链取证结合，构建的人工智能系统从以往取证事务中汲取经验，预识别潜在的取证场景，实现对特定场景数据的准确和可信采集。

设计取证场景下的访问控制是提升区块链取证系统安全性的另一种手段。基于角色的访问控制可能因不同取证阶段主体权限发生变化而引发复杂的状态变更[46]，因此文献[47]提出了分阶段的角色授权访问控制策略（RBAC-SA, role-based access control with staged authorization），集成了基于角色和规则的访问控制模块，动态赋予角色权限以适应复杂的取证场景。文献[48]提出了基于属性和角色的混合访问控制（ARBAC, attribute and role-based access control）架构，通过策略决策点（PDP, policy decision point）和取证事务的动态环境条件设定主体权限，以此实现日志取证系统中的安全访问。在另一类基于属性的访问控制中，如果证据信息在不同取证阶段存在不同属性，主体权限管理往往会变得复杂。文献[49]针对区块链背景，提出了分布式密钥策略属性基加密（DKP-ABEH, decentralized key-policy attribute-based encryption），配合区块链分布式特性与智能合约将取证流程建模为有限状态机，使各方在明确取证流程状态的基础上，以受信任的方式进行取证合作。

3.2.3　高可用性的证据存储

针对区块链取证在实际应用中的可行性，相关研究通过改进节点共识、调整链结构或降低取证成本等手段提升取证系统的效率，如表 6 所示。

一方面，研究者们针对取证系统的共识算法进行改进或融合，以实现证据信息的高效验证。文献[15]结合车联网取证的动态和自主性特点，提出快速领导节点选举算法，动态组建联邦共识节点以较小的时延解决突发取证环境下证据区块确认问题。文献[50]的作者认为，采用拜占庭容错的数字取证研究难以应对节点数增多带来的扩展问题。本文提出由代理权益证明（DPoS, delegated proof of stake）共识算法推选节点后，再利用实用拜占庭容错（PBFT, practical Byzantine fault tolerance）实现

表6 证据保全阶段的高可用性方案

类型	文献	技术特点	高可用性方案	方案评价
高效验证	文献[15]	动态联邦共识	改进领导节点选举算法，完成动态联邦共识中的节点选择	灵活性和场景适应性强，需要注意共识结合的安全性问题，平衡安全性、效率和灵活性之间的关系
	文献[50]	多共识结合	DPoS与PBFT在不同阶段结合，实现节点选取和高效共识	
	文献[51]	改进的DPoS	改进的DPoS满足车联网环境下实时取证需求	
快速索引	文献[52]	DAG图结构	利用DAG结构优化证据信息存储效率	高取证事务吞吐量，低时延与资源消耗，需要处理新结构下的共识问题
	文献[53]	改进默克尔树	构建Merkle DAG，并配合DASL索引结构实现高效查询	
	文献[55]	MQTT协议	配合轻量级协议MQTT，在资源受限下完成监控和数据收集	
	文献[56]	联邦学习	利用联邦学习训练设备数据模型参数，实现区块链聚合	

同步，既提升了单位时间的吞吐量，又适用于节点复杂的取证场景。文献[51]将联盟链上的取证节点分为主节点和其他节点，改进的DPoS共识算法允许节点批量共识，提升了系统实时取证的效率。

另一方面，研究者们探索改进传统链式结构以提升数据存储、查询或审计效率。文献[52]采用有向无环图（DAG, directed acyclic graph）组织证据区块，借助DAG自然拓扑结构的排序协议，优化链上确认时延和存储效率。文献[53]则改变区块内扁平化的Merkle树为Merkle图，在跳表索引的基础上设计确定性仅追加跳表（DASL, deterministic append-only skip list），配合实现高效证据索引。

降低成本是区块链取证系统长远发展的前提条件，研究围绕取证系统的资源成本展开。通常在物联网环境下数字取证面临计算、存储等方面的限制[54]，而采用轻量级取证框架是一种有效的手段。消息队列遥测传输（MQTT, message queuing telemetry transport）协议是物联网领域常用的轻量级消息传输协议，具有功耗低、传输数据包小等特点。文献[55]将MQTT协议与数字取证结合，利用MQTT协议的发布/订阅消息机制，实现轻量级数据监控和证据收集。系统中的证据分级模块能够动态调整传感器的监控和取证等级，进一步平衡了成本。文献[56]将物联网设备数据作为联邦学习模型参数，通过区块链聚合消除了共享原始证据信息的需要，从而实现整体架构的轻量化。

3.2.4 支持取证事务的证据存储管理

前文分析了区块链在各个层面为证据保全阶段提供的完善保障。还有研究注意到，链上存储过程能够针对可疑节点或恶意访问行为进行“二次取证”，采集的“链上证据”进一步提升了监管链的完整性。文献[30]提出了“可追溯取证”的概念，构建链上证据信息的存储映射，定位并判断可疑的授权及访问行为，相关信息记录为审计日志参与到监管链的构建。本文还基于B/S（brower/server）网络结构提出了结合IPFS的分布式B-DS（bowser/distributed server）结构，匿名、离散地存储非关键服务及数据，配合区块链实现去中心化的证据信息管理与隐私保护。

3.3 证据呈现阶段：“监管链”追溯

在该阶段，区块链技术负责将采集到的“原始证据”和二次取证的“链上证据”串联起来，构建“证据监管链”，实现证据追溯与呈现。

3.3.1 区块链证据“监管链”

在构建证据监管链的过程中，研究人员借助区块链对整体取证流程中的证据痕迹进行关联，本文称为“证据监管链管理”，完成“区块链证据”配合“链上信息”生成“证据报告”的过程。

文献[42]通过监管链CoC、证据链（EC, evidence chain）和案件链（CC, case chain）共同记录证据信息。区块链在处理单个案件的过程中衍生出证据链EC，EC伴随数字取证阶段不断扩展，得到案件链CC，关于证据的所有访问信息都保存在CoC，最终三者组合成CoC-EC-CC以呈现完整的证据信息。而文献[57]通过在取证起始阶段构建序列图，定义数据信息的处理流程、参与方的交互行为和业务逻辑。模块化架构将分布式核心系统与取证应用程序域解耦，实现证据监管链的透明呈现。

在文献[58]提出的证据监管链原型B-CoC中，将取证过程信息分为证据数据与证据日志两类。其中与构建CoC相关的信息存储与日志库中。在区块链基础设施上创建和部署智能合约，为取证用户提供前端接口实现证据链的追溯、验证与管理。文

献[59]中的Forensic-Chain模型定义了证据信息的数据结构，其中证据的创建、删除、转移和公示等行为，通过“转移链”结构被完整记录，利用区块链权限特性确保只有授权参与者访问证据，进一步增强了证据链的安全性和隐私性。

此外，文献[60]借助智能合约充当取证工具和区块链间的桥梁，将各个阶段取证工具和区块链的交互过程记录为不可篡改的链上交易，同时刻画原始证据图像、设备状态等信息封装在区块中，为证据监管链构建提供额外的数据支持。

3.3.2　集成智能合约的取证流程实现

智能合约基于预定义程序化条款，在取证操作中实现自动化执行和取证事务的验证，有效降低人为干预。在此背景下，文献[49]将取证流程通过智能合约建模为包含 8 个不同状态的有限状态机（FSM, finite state machine），分别为授权请求、授权、份额检索、数据收集、数据检查、数据分析、取证报告和完成。本文借助区块链“预言机”的概念，将外部信息反馈至区块链，动态调整延迟的取证状态，并由智能合约辅助实现取证过程核验。“验证后转发”的模式为区块链取证的整体流程实现提供了可审计的过程保障。文献[61]利用私有链协议和智能合约增强模型的自动化过程，强调数据和取证分析阶段之间的迭代相互作用，还明确了各取证阶段的数据合规性指标，以支持取证过程中的证据信息控制、转移、分析和保存监控。同样在私有链中，文献[62]通过定制的 4 种类型智能合约：代币合约、案件管理合约、访问控制合约和溯源合约，实现取证过程中的案件创建、管理和链上数据追溯。

如表 7 所示，当前研究充分展示了区块链技术在取证监管中的显著应用潜力，特别是通过智能合约实现的证据追溯、管理及全取证流程的自动化。

4　区块链全流程参与的取证架构

为进一步挖掘区块链技术的应用潜力，本节提出了基于区块链技术的数字取证系统架构。对应现有研究中的不足，该架构通过三方面的设计：1) 融合证据特点的链上数据结构及图论分析；2) 节点并行的高效链下分布式存储；3) 合约模板参与的规范化“证据监管链”溯源，促进了区块链技术与取证场景的进一步融合，提升了链下存储的效率和同类型取证事务合约的可复用性。

4.1　取证系统架构总述

本节对区块链数字取证系统架构进行描述，如图 5 所示。取证系统涉及用户节点User、取证节点（FN, forensic node）和庭审节点（TN, trial node），其中取证节点包括采集节点和分析节点，庭审节点涉及法院、检察院等节点。数字取证步骤如下。

表 7　证据呈现阶段的区块链方案

事务类型	文献	技术特点	呈现方案	方案评价
“监管链”溯源	文献[42]	监管链设计	多链层次结构，伴随取证进程提供的过程信息逐步扩展证据链EC，直至获得透明、可追溯的案件链CC	各研究中的“证据监管链”在内容和形式上存在差异
	文献[57]	智能合约交互	智能合约规定取证和信息处理规则，用序列图来展示智能合约的实现方式及交互，分布式核心系统与取证应用程序域解耦	
	文献[58]	溯源数据结构	利用证据ID、描述、创建者身份等信息作为链上日志，智能合约支持行为记录	
	文献[59]	监管链设计	定义转移链结构记录链上行为信息，配合权限区块链特性实现证据链的安全性和隐私性	
	文献[60]	智能合约交互	智能合约记录取证工具与区块链的交互过程，刻画取证图像等信息，为证据监管链构建提供数据支持	
取证流程自动化	文献[49]	取证行为建模	将取证过程通过智能合约建模为一个有限状态机，引入区块链“预言机”辅助取证流程数据转移	辅助智能合约实现的取证事务自动化，如何设计合约逻辑是关键
	文献[61]	溯源指标设计	利用智能合约规定各阶段指标，自动执行证据的全生命周期管理，支持取证过程中的证据信息控制、转移、分析和保存监控	
	文献[62]	定制取证合约	定制 4 种类型智能合约实现不同取证阶段的特定功能并简化数据管理，关注证据来源验证	

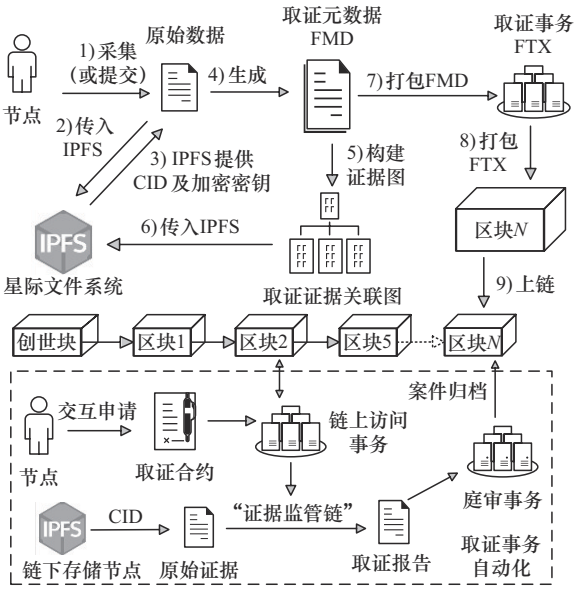

图5　区块链数字取证系统架构

1) 证据采集。该环节采集原始数据并生成证据关联图。节点在获取相应权限基础上，结合取证设备或取证工具，对文本、图片、音视频或其他格式的原始数据进行采集或提交。按照元数据结构提取关键信息，通过IPFS存储并生成唯一标识CID及加密密钥，参与生成“取证元数据”（FMD, forensic metadata）。

2) 证据分析。在获取各项取证元数据的基础上，利用区块链FMD选取关键属性构建证据关联图，结合图论算法中的社区检测算法和中心性分析，进一步组织提炼相关信息，为取证人员提供辅助。

3) 证据保存。该过程涉及原始数据和证据关联图的存储。提交或采集的原始数据存入IPFS后，IPFS返回每份文件的标识CID及加密后的对称密钥，配合构建FMD。基于FMD生成的证据关联图及分析内容同样存储于IPFS。

4) 证据访问。相关信息上链后，允许授权用户和取证节点访问存储在链上的证据信息。节点提出访问事务申请，依然借助FMD结构构建事务信息，将访问涉及的证据标识、CID标识、区块哈希等信息作为触发相应智能合约的条件值，利用合约触发与区块链的交互，进一步通过匹配链下IPFS中的原始数据实现访问操作。

5) 证据监管链追溯。该过程包括两方面内容：原始证据及“二次取证”的链上信息。在庭审相关节点触发合约审查“证据监管链”时，两方面内容在“追溯合约”的作用下关联生成取证报告，展示完整证据链，并参与庭审过程。

6) 取证归档。案件处理结束后，将审理结果和相关数据配合详细的审计日志归档存储。需要取证审查或后续核验时，通过访问接口确保审查人员能快速访问证据。

以上内容详细描述了本文所提架构中区块链与数字取证各阶段的技术融合。区块链的全程参与和多样化应用，确保了数据从采集到证据链构建的完整性和可靠性。因此，该架构适用于需要关联性分析、证据链构建和复杂溯源过程的取证场景，如金融交易与审计取证、知识产权保护与侵权纠纷、司法取证与法律诉讼以及供应链溯源等。

在存在复杂信息流传递的金融审计场景中，本文所提架构通过取证元数据结构对金融交易进行标准化采集，并通过图论算法对不同交易之间的关联性进行分析，揭示隐藏的交易网络，识别潜在的金融犯罪行为。当涉及知识产权保护及相应纠纷时，该架构能借助存储的证据链追溯侵权行为的来源，快速定位侵权行为的起源和责任方，缩短纠纷解决时间。在存在复杂流程的司法诉讼场景中，智能合约配合结构化数据采集以减少人为干预带来的误差，同时结合图论算法分析不同证据之间的关联性，辅助挖掘更多取证线索。对于供应链产品溯源，各环节流转信息能够被溯源字段完整记录，促使消费者、企业或监管机构借助区块链实现溯源验证。因此，本文所提架构在以上取证场景中的适用性充分体现了区块链全流程参与带来的取证优势。

4.2　融合证据信息的链上数据结构及图论分析

在数据采集阶段，为增强对各类取证场景下证据采集的适应性，本文为不同场景设计了一种专用数据结构“取证元数据（FMD）”，包括采集时间、来源、采集者身份等关键信息，如图6所示。该结构有三部分：1) 证据标识符（EI, evidence identifier）用于明确元数据归属，包含证据标识、归属标识和存储标识；2) 证据可追溯信息（ETI, evidence traceability information）可以详细记录取证的过程信息，包含证据类型（如文档、图片或视频等）、证据来源（如设备编号或提交证据的用户编号等）、证据采集时间戳等；3) 扩展元字段（EMF, extended metadata field）可以根据实际需求，实现额外的信息记录和扩展功能。

证据标识符EI			证据可追溯信息ETI					扩展元字段EMF
证据标识	案件标识	存储标识	证据类型	证据来源	时间戳	内容哈希	人员信息	文件大小、格式

图6　取证元数据结构

根据定义的数据结构，FMD可表示为

FMD=[EI,ETI,EMF]

FMD中的证据标识符（EI）与证据追溯信息（ETI）设计为分离部分，不仅确保证据的唯一标识，还可通过证据关联图的边建立证据之间的关联，便于溯源和验证。相比仅记录哈希值以关联证据的方式，该设计提升了证据溯源的清晰度和一致性。文献[12]、文献[13]和文献[16]等研究同样在区块中扩展辅助信息，并额外记录了证据来源作为“可信证明”。而FMD不仅包含以上信息，还记录了证据类型、访问时间戳、访问人员信息等，结构化记录了更全面的内容信息。在证据关联分析过程中，能够基于整体信息辅助完成证据间的关联性分析。为处理不同取证场景下的兼容问题，文献[14]针对不同系统设计多个客体链的做法过于复杂，且链间交互的安全性需要额外考虑。相比之下，FMD结构使系统可以在不改变基础结构的情况下，灵活地添加和调整信息字段。FMD允许用户或系统根据场景自定义需要记录的取证元数据，这种定制化功能使FMD结构能够应对从简单的文档证据到复杂的多媒体证据的广泛取证场景。例如，在处理音频证据时，添加音频采样率和时长等信息，而在处理图像证据时，记录分辨率和拍摄设备信息。取证元数据结构通过预留扩展元字段（EMF）的做法，为面临多种场景下的信息采集带来了灵活性。

在构建证据关联图ECG的过程中，以FMD包含的数据属性为基础，参与图节点的构建。表8展示了数据样本及属性，其中每个案件可以对应多条证据，案件标识用于明确归属，证据来源于用户User或由取证节点链上采集，由采集时间T、证据内容哈希hash和证据所处的区块哈希Hash共同辅助实现信息校验。在此基础上，按照以下步骤构建证据关联图。

1) 每条取证元数据作为一个节点。

2) 结合证据相关性构造边。

3) 结合时间戳及属性排列相关数据，构建ECG。

4) 引入社区检测Louvain算法辅助证据分析。

表8　　数据样本及属性

证据节点	案件标识	证据来源	采集时间	内容哈希	区块哈希
evidence_id_1	case_id_1	User A	T_{s1}	hash1	Hash1
evidence_id_2	case_id_1	FN A	T_{s2}	hash2	Hash1
evidence_id_3	case_id_2	FN A	T_{s1}	hash3	Hash2
evidence_id_4	case_id_3	User B	T_{s4}	hash4	Hash2
evidence_id_5	case_id_3	FN B	T_{s5}	hash5	Hash2
evidence_id_6	case_id_3	FN B	T_{s6}	hash6	Hash2

图7是根据样本构建的证据关联图。圆形顶点代表证据节点，连接顶点的边代表证据的相关关系，虚线代表节点属于相同案件，实线代表信息具有相同来源。

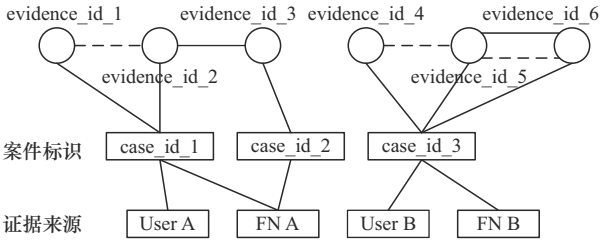

图7　证据关联图

以建立的ECG为目标，进一步引入Louvain算法[63]识别图中证据节点的“社区”信息，此处“社区”概念指具有某种相关属性的全部信息集合。以FMD中需要进行关联分析的属性为节点的社区标签，遍历相邻节点选择模块度增益最大（即贪婪思想）的社区加入，直至所有节点不能通过移位来增加模块度时，便成功构建证据节点“社区”信息。

算法1　基于Louvain的证据社区划分算法

输入　证据关联图，目标属性{**ECG**、targetAttribute}

输出　目标属性的证据社区集合community

1) 初始化证据节点v、单一社区$\boldsymbol{c}$、社区集合community

2) function Louvain(ECG, targetAttribute)

3)　每个证据节点处理为单一社区:$\boldsymbol{c}[v]=\{v\}$，初始化权重

4)　for every c in community do

5)　　计算内、外权重$W_\text{in}(\boldsymbol{c})$、$W\_\text{out}(\boldsymbol{c})$

6)　for every v in **ECG** do

7)　　移动节点至相邻社区，记录模块度变化

ΔQ 最大的节点
8) if $\Delta Q(v, c') > 0$ then
9) 移动 v 目标节点所在社区 c'
10) 更新 $W\_in(c)$、$W\_out(c)$
11) end if
12) end for
13) 压缩社区内所有节点为新节点，计算新节点内外边权重
14) 重复步骤 6)~步骤 12)的循环语句至整个图的模块度不再变化
15) return 证据社区集合 community
16) end function

算法 1 展示了图关联分析的具体实现，将在某属性上具有关联的所有证据组织起来，以“属性社区”的形式呈现给取证分析节点，清晰的图示结构避免了冗余信息干扰，辅助提升对证据信息的关联分析效果。

基于元数据结构的图分析算法，在涉及复杂信息采集的场景中具有独特优势。以跨平台、跨系统的取证场景为例，不同平台可能使用不同的文件系统、加密标准和存储格式，单一角色的区块链取证工具容易出现证据提取不完全、文件格式不兼容等问题。而 FMD 的优势体现在兼容性和信息透明性上，证据标识符 EI 可以确保不同系统之间证据归属的统一性，避免因系统差异导致的证据丢失或混淆。证据可追溯信息 ETI 针对性地记录各平台间证据的流转路径，确保证据从源头到提取过程中的全面记录；扩展元字段 EMF 使不同取证场景之间能够进行交互与集成，例如，EMF 可以根据不同需求添加相应的扩展元字段，包括特定证据的采集字段、采集工具的合法性记录、合作机构的审核记录或国际组织的标准化证明等，为取证分析提供完整信息并在证据链呈现过程中提供对取证行为的合法性证明等。

该结构同样适用于复杂场景中的信息追踪，如金融审计中对资金流动和合同履行情况的严格监督与取证，以及供应链管理中各环节的产品认证和溯源取证等。

4.3 节点并行交互的高效链下分布式扩容

此架构采用 IPFS 作为证据链下扩容的分布式数据库，与文献[23]和文献[25]等借助传统数据库的研究相比，维持了去中心化的特点，减少了单点故障风险，有效支持了取证流程中针对链上信息的“二次取证”以及后续的审查行为。

但基于点对点（P2P, peer-to-peer）网络的 IPFS，数据检索速度取决于网络中其他节点的响应速度和稳定性。文献[26]和文献[29]为取证系统的各种身份设计了多类节点，复杂的场景交互可能对取证事务的效率产生影响。同时，IPFS 作为纯商业的分布式存储系统[31]，为适应警务相关的取证调查，需要将相应节点纳入系统安全设计。文献[27]基于 IPFS 设计了额外的安全防护层用于提升数据的可信程度，但同样需要关注引入附加机制后的取证事务效率。

因此，为充分发挥 IPFS 的扩容性能，并确保取证行为的法律认可，提出了一个包括庭审节点在内的并行交互方案。该方案简化了整个区块链取证系统的参与方，设计为三类节点，分别为用户节点、取证节点和庭审节点。用户节点仅参与证据上传、查询和验证环节，作为取证系统的常规用户参与交互过程。取证节点负责实现数据采集和取证分析，并将两类数据与 IPFS 进行交互。庭审节点代表警务、法庭等可信节点，负责核验证据监管链及取证案件的归档存储。为了实现高效的 IPFS 证据存储，设计以下并行机制。

1) 并行上传。用户节点和数据采集节点可以并行上传证据到 IPFS 网络。用户和采集节点将证据文件分割成多个小文件，并利用多线程/并发上传技术上传到 IPFS。每个小文件上传时会返回一个 CID，用户节点将所有这些 CID 及其相关的元数据同时提交到区块链。两类节点同时将多个证据文件上传到 IPFS，充分利用网络带宽和存储资源。

2) 并行检索。取证和庭审两类节点可以并行检索存储在 IPFS 上的证据。在取证分析节点和庭审节点中，通过设置并行检索机制，多个节点可以同时获取不同证据的 CID 标识，提高取证和审查的效率。

3) 任务调度和负载均衡机制。针对与 IPFS 进行交互事务，创建多个任务队列（如上传、检索和审查队列），将待处理的任务放入相应的任务队列中。各类型节点提出交互申请后，每个任务队列可以按照任务类型及节点优先级进行组织，并分配空闲 IPFS 节点进行交互处理，避免节点过载。

通过上述方案，IPFS 能够并行处理不同节点的证据存储、索引和验证事务，在实现链下扩容的

同时，进一步提升了取证事务的效率。简化的各类节点与IPFS建立了明确的并行处理模式，一定程度上缓解了网络时延问题。

4.4　基于合约模板的规范化“监管链”

“监管链”是整个取证流程中的重要环节，其溯源方式及内容直接影响证据的可信度。然而，现有研究对CoC内容和追溯过程存在不同理解和描述，导致其在实际应用中的统一和规范困难。而在取证流程中，信息经历以下演变过程。从采集时获得的“数据”到初步共识验证后的“可信数据”，再到取证分析后的“证据”，上链存储后附加链上访问信息形成“证据+链上信息”，最终由建立监管链呈现于取证报告。在每个演变环节，区块链技术均为构建“监管链”提供支持，为了明确各环节中的区块链内容，提出了如下的规范化CoC构造步骤。

1) 以标准化采集的信息作为溯源起始。确保按照元数据结构采集信息，并打包取证事务上链存储。

2) 将链下存储的数据信息作为“存在性证明”，链外安全存储的证据与链上取证元数据形成关联。

3) 将链上访问记录作为“过程性证明”，记录相关访问操作并存储。

4) 将审计记录作为“可信证明”。借助智能合约在各操作中核验数据哈希，并记录核验结果。

上述步骤明确了监管链构建过程中需要的“溯源起始数据”“存在性证明”“过程性证明”和“可信证明”，规范了每个过程的监管内容。引入智能合约实现了监管链构造环节的自动化。本文基于模块化设计的思想，引入智能合约模板（SCT, smart contract templates）的概念[64]。通过将数字取证事务涉及的共性步骤（如数据哈希化、获取时间戳、数据验证、节点身份核验等）固定在相应合约模块中提高合约的可复用性。相比于文献[49]和文献[62]中将智能合约固化为几种固定类型的方式，SCT针对不同的数字取证事务预留了灵活性，允许其根据具体需求对合约进行演化和调整。作为通用智能合约，SCT可以通过节点传递不同参数来调用基础模块，既提升了合约的可复用性，又提供了一定的灵活性。此外，与文献[61]中智能合约实现的证据生命周期管理不同，本文方案基于区块链全流程参与，制定了适用于不同取证事务的合约演化形式，覆盖了整个数字取证环节，进一步增强了系统的适应性和扩展性。

将证据采集、存储、验证和“证据链”溯源等共性操作设计为SCT中的基础模块进行固定，配合不同取证事务实现扩展。图8展示了基于SCT的取证合约构建方法。SCT由模板ID、模块代码、合约参数字段和取证内容字段组成。模块代码封装了通用的取证操作，包括身份验证、数据哈希化及核验等。合约参数字段的基本信息包括节点身份信息、目标证据信息和取证行为参数。取证内容字段作为可变字段记录额外信息。依托传递不同取证类型的参数，首先在逻辑上实现各种取证合约的演化，进而实际部署在区块链上实现调用。

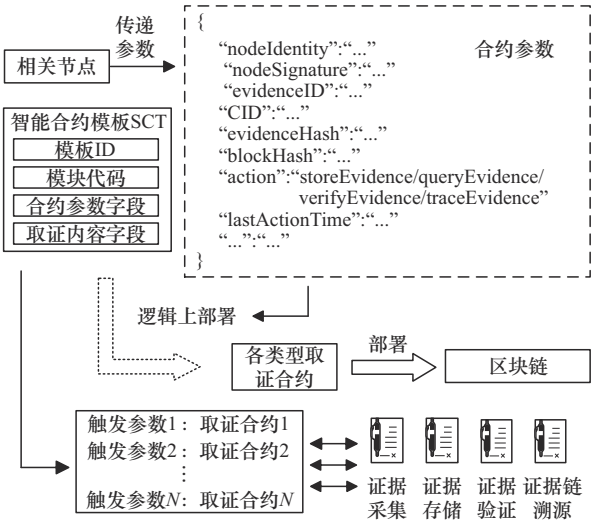

图8　基于SCT的取证合约构建方法

算法2　证据采集合约算法

输入　原始信息，授权主体信息，事务参数{evidenceData、authorizedEntity、forensicAction}

输出　取证元数据forensicMetadata

1) 分析取证事务类型，定向初始化合约模板
2) 初始化证据标识evidenceId，创建IPFS存储映射 **evidences**
3) function forensic_acquire(evidenceData, authorizedEntity)
4)　if authorizedEntity ==authorizedEntity' then
5)　　记录证据所属案件信息
6)　　记录时间戳timestamp←block.time-stamp
7)　　记录证据信息evidenceId←evidenceId + 1; evidenceHash←hash(evidenceData)

8) end if
9) 原始数据存入IPFS并返回CID标识与密钥信息ε
10) 提取取证元数据forensicMetadata←{evidenceId, caseId, CID, ε, evidenceType, authorizedEntity, timestamp}
11) 建立取证元数据存储映射**evidences**[evidenceId]←forensicMetadata
12) return forensicMetadata
13) end function

以构建证据采集合约为例，如算法2所示，取证节点向合约模板传递证据采集的相关参数，定向初始化合约。取证事务中的共性操作，如计算哈希、核验哈希以及获取时间戳等，将通过模板中的固定模块实现。引入的合约模板作为各节点共同调用的公共合约，简化了取证事务的复杂度，同时提升了取证合约的可复用性。

5 区块链取证技术展望

基于文献中整理的研究成果，区块链技术能够为数字取证提供新的技术支持。但在各个取证阶段中，区块链技术的应用仍存在一定的局限性。本节分别对取证流程的3个阶段给出当前区块链在应用中面临的挑战和可行的研究方向。

5.1 链上证据获取面临的挑战与研究方向

证据获取过程重点关注两方面内容：链上证据的存在形式及其在取证分析中的支持作用。部分研究在已有证据信息的基础上，将证据哈希处理后直接上链存储，或配合验证信息一并上链。虽然相较于传统取证技术，区块链技术为数字证据带来了额外的安全保障，但上述方式未能充分挖掘区块链的应用潜力。同时，区块链能够确保数据上链后不被篡改，但对上链前原始证据的真实性和准确性缺乏有效评估。尝试利用链上证据构建证据关联图或在可信的链上环境中进行取证分析的做法，在实施应用过程中需要充分考虑区块链和取证相融合的特点。

从单纯链上哈希固证到辅助参与证据分析，揭示了区块链与取证过程结合的不断深入。因此，进一步挖掘区块链技术如何以更直接的方式参与数字取证是下一步的研究方向。此外，在证据采集过程中借助区块链共识实现可信验证的形式，需要分析共识算法的适用性，研究取证共识中的节点和步骤设置是下一步的研究方向。

5.2 链上证据存储面临的挑战与研究方向

区块链的扩容能力将成为未来数字取证领域研究的重点。当前主流的扩容思想分为“减少链上数据”和“引入额外空间”。减少链上数据意味着对原有链结构进行调整，无论是通过区块解耦实现单一区块层面的削减，还是仅存储部分关键块数据实现空间扩容等方式，都不可避免地面临着数据完整性的问题。减少的链上节点或验证信息在一定程度上削弱了区块链的去中心化特性，需要结合实际应用场景综合考虑取证行为的安全性。因此，研究优化链上证据结构实现空间扩展是下一步的研究方向。

另一思路是通过引入额外空间实现扩容，此处的额外空间指链下空间和多链实现的扩展。引入链下空间首先要考虑证据信息的一致性问题，由于传统数据库面临中心化等问题，在区块链数字取证研究中逐步被分布式系统所替代。其中IPFS作为相关研究应用最多的链下扩展系统，其自身的商业属性、存在的网络堵塞问题在数字取证环境下如何处理，促进有效的链下扩容是下一步的研究方向。多链扩容的研究中辅助链的类型和结构具有较高灵活性，既延续了去中心化特性又适用于细分的取证场景。相应支持多链策略的区块链平台，如闪电网络、BitXHub等，也在不断涌现与发展。但取证场景下链间交互协议的复杂性、信息传递的安全性等问题是下一步的研究方向。

在提升证据存储的安全性和可用性层面，现有研究关注改进区块链上的加密方式和共识算法并与新兴领域相结合。模糊哈希被应用到分析链上证据块的相似程度，但在目前适用的数据类型范围和准确性上存在提升空间，进一步探索链上内容相似度的匹配方式，辅助判定证据的完整性是下一步的研究方向。根据取证环境特点，改进或提出新的共识算法是下一步的研究方向。此外，研究加密算法的改进或将其与其他技术领域（如人工智能、格密码等）结合，用以提升取证信息的安全性或可用性，但需要关注引入的资源消耗。

5.3 链上证据链呈现面临的挑战与研究方向

数字取证工作中创建和维护可靠的证据监管链是证据被认可的关键环节。一方面，现有研究关于

区块链取证中“证据监管链”的理解存在差异。虽然各研究提到的监管链基本建立在原始证据和链上访问信息的基础上，但在呈现形式和具体内容上不尽相同。另一方面，不同国家和地区对于区块链取证的流程和系统设计具有不同的标准。上述差异使得不同系统之间存在互操作性问题，并且在评估相关研究时缺乏统一的指标。因此，结合领域内权威的法律标准，探讨构造规范化的监管链是下一步的研究方向。

此外，区块链环境下的各类取证事务均需要借助智能合约实现。智能合约区别于常规代码执行，合约一旦部署将难以修改和调整，需要额外进行安全层面的考虑。在智能合约的应用方式上，取证事务相对固定且涉及诸多共性操作，因此，研究针对取证行为的模块化合约设计、标准化合约模板等提升复用性、安全性是下一步的研究方向。

6　结束语

本文深入探讨了区块链技术如何应用于数字取证领域，着重分析了区块链技术在各个取证阶段的作用与优势。在现有研究的基础上，提出了一套区块链数字取证架构，促进了区块链技术与取证场景的进一步融合，实现了高效的链下分布式存储，并借助智能合约模板提升了取证合约的可复用性。本文还深入展望了该领域下一步的研究方向，为推动区块链技术在数字取证领域的应用做出了贡献。

参考文献：

[1] JORDAAN J, BRADSHAW K. The current state of digital forensic practitioners in south Africa[C]//Proceedings of the 2015 Information Security for South Africa (ISSA). Piscataway: IEEE Press, 2015: 1-9.

[2] Verizon Threat Research Advisory Center. Data breach investigations report[R]. 2024.

[3] 李炳龙, 王鲁, 陈性元. 数字取证技术及其发展趋势[J]. 信息网络安全, 2011, 11(1): 52-55.
LI B L, WANG L, CHEN X Y. Digital forensic technique review[J]. Netinfo Security, 2011, 11(1): 52-55.

[4] MOHITE M P, DESHMUKH J Y, GULVE P R. Qualitative and quantitative analysis of cloud based digital forensic tool[C]//Proceedings of the 2016 10th International Conference on Intelligent Systems and Control (ISCO). Piscataway: IEEE Press, 2016: 1-5.

[5] GRUHN M, FREILING F C. Evaluating atomicity, and integrity of correct memory acquisition methods[J]. Digital Investigation, 2016, 16: S1-S10.

[6] BOCK L. Learn wireshark: a definitive guide to expertly analyzing protocols and troubleshooting networks using wireshark[M]. Birmingham: Packt Publishing, 2022.

[7] CHILDERHOSE C. Mastering veeam backup & replication: secure backup with veeam 11 for defending your data and accelerating your data protection strategy[M]. Birmingham: Packt Publishing, 2022.

[8] SHIAU S J H, SUN C K, TSAI Y C, et al. The design and implementation of a novel open source massive deployment system[J]. Applied Sciences, 2018, 8(6): 965.

[9] COSIC J, COSIC Z, BACA M. An ontological approach to study and manage digital chain of custody of digital evidence[J]. Journal of Information and Organizational Sciences, 2011, 35: 1-13.

[10] CIARDHUA S O. An extended model of cybercrime investigations[J]. International Journal of Digital Evidence, 2004, 3(1): 1-22.

[11] ORIWOH E, JAZANI D, EPIPHANIOU G, et al. Internet of things forensics: challenges and approaches[C]//Proceedings of the 9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing. Piscataway: IEEE Press, 2013: 608-615.

[12] TAO Q, DING H W, JIANG T, et al. B-DSPA: a blockchain-based dynamically scalable privacy-preserving authentication scheme in vehicular ad hoc networks[J]. IEEE Internet of Things Journal, 2024, 11(1): 1385-1397.

[13] FERNÁNDEZ-CARRASCO J Á, ECHEBERRIA-BARRIO X, PARE DES-GARCÍA D, et al. ChronoEOS 2.0: device fingerprinting and EOSIO blockchain technology for on-running forensic analysis in an IoT environment[J]. Smart Cities, 2023, 6(2): 897-912.

[14] 房巧玲, 高思凡, 曹丽霞. 区块链驱动下基于双链架构的混合审计模式探索[J]. 审计研究, 2020(3): 12-19.
FANG Q L, GAO S F, CAO L X. Exploring the hybrid audit approach based on double-chain architecture in a blockchain environment[J]. Auditing Research, 2020(3): 12-19.

[15] GUO H, LI W X, NEJAD M, et al. Proof-of-event recording system for autonomous vehicles: a blockchain-based solution[J]. IEEE Access, 2020, 8: 182776-182786.

[16] RYU J H, SHARMA P K, JO J H, et al. A blockchain-based decentralized efficient investigation framework for IoT digital forensics[J]. The Journal of Supercomputing, 2019, 75(8): 4372-4387.

[17] HU S S, ZHANG S H, FU K L. TFChain: blockchain-based trusted forensics scheme for mobile phone data whole process[C]//Proceedings of the 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC). Piscataway: IEEE Press, 2022: 155-165.

[18] POURVAHAB M, EKBATANIFARD G. Digital forensics architecture for evidence collection and provenance preservation in IaaS cloud environment using SDN and blockchain technology[J]. IEEE Access, 2019, 7: 153349-153364.

[19] RATHORE N K, KHAN Y, KUMAR S, et al. An evolutionary algorithmic framework cloud based evidence collection architecture[J]. Multimedia Tools and Applications, 2023, 82(26): 39867-39895.

[20] TIAN Z H, LI M H, QIU M K, et al. Block-DEF: a secure digital evidence framework using blockchain[J]. Information Sciences, 2019, 491: 151-165.

[21] KIM D, IHM S Y, SON Y. Two-level blockchain system for digital crime evidence management[J]. Sensors, 2021, 21(9): 3051.

[22] CEBE M, ERDIN E, AKKAYA K, et al. Block4Forensic: an integrated lightweight blockchain framework for forensics applications of connected vehicles[J]. IEEE Communications Magazine, 2018, 56(10): 50-57.

[23] XIONG Y, DU J. Electronic evidence preservation model based on blockchain[C]//Proceedings of the 3rd International Conference on Cryptography, Security and Privacy. New York: ACM Press, 2019: 1-5.

[24] BROTSIS S, KOLOKOTRONIS N, LIMNIOTIS K, et al. Blockchain solutions for forensic evidence preservation in IoT environments[C]// Proceedings of the 2019 IEEE Conference on Network Softwarization (NetSoft). Piscataway: IEEE Press, 2019: 110-114.

[25] LIN Q J, WANG H Z, PEI X F, et al. Food safety traceability system based on blockchain and EPCIS[J]. IEEE Access, 2019, 7: 20698-20707.

[26] SHILPA C, SHANTHAKUMARA A H. An implementation of blockchain technology in combination with IPFS for crime evidence management system[C]//Proceedings of the 2023 International Conference on Computer Communication and Informatics (ICCCI). Piscataway: IEEE Press, 2023: 1-6.

[27] NYALETEY E, PARIZI R M, ZHANG Q, et al. BlockIPFS-blockchain-enabled interplanetary file system for forensic and trusted data traceability[C]//Proceedings of the 2019 IEEE International Conference on Blockchain. Piscataway: IEEE Press, 2019: 18-25.

[28] LIU C Y, WANG Z H, XIONG A, et al. Research on industrial Internet traceability technology based on blockchain[C]//Proceedings of the 2022 IEEE 14th International Conference on Advanced Infocomm Technology (ICAIT). Piscataway: IEEE Press, 2022: 286-291.

[29] RENO S, BHOWMIK S, AHMED M. Utilizing IPFS and private blockchain to secure forensic information[C]//Proceedings of the 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI). Piscataway: IEEE Press, 2021: 1-6.

[30] SHANG S Y, ZHOU A Y, TAN M, et al. Access control audit and traceability forensics technology based on blockchain[C]//Proceedings of the 2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC). Piscataway: IEEE Press, 2022: 932-937.

[31] LI M, LAL C, CONTI M, et al. LEChain: a blockchain-based lawful evidence management scheme for digital forensics[J]. Future Generation Computer Systems, 2021, 115: 406-420.

[32] MERCAN S, CEBE M, TEKINER E, et al. A cost-efficient IoT forensics framework with blockchain[C]//Proceedings of the 2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). Piscataway: IEEE Press, 2020: 1-5.

[33] YAO Q, LI T T, YAN C, et al. Accident responsibility identification model for Internet of vehicles based on lightweight blockchain[J]. Computational Intelligence, 2023, 39(1): 58-81.

[34] RANA S K, RANA A K, RANA S K, et al. Decentralized model to protect digital evidence via smart contracts using layer 2 polygon blockchain[J]. IEEE Access, 2023, 11: 83289-83300.

[35] KORNBLUM J. Identifying almost identical files using context triggered piecewise hashing[J]. Digital Investigation, 2006, 3: 91-97.

[36] 任艳丽, 徐丹婷, 张新鹏, 等. 可修改的区块链方案[J]. 软件学报, 2020, 31(12): 3909-3922.
REN Y L, XU D T, ZHANG X P, et al. Scheme of revisable blockchain[J]. Journal of Software, 2020, 31(12): 3909-3922.

[37] MAHROUS W A, FAROUK M, DARWISH S M. An enhanced blockchain-based IoT digital forensics architecture using fuzzy hash[J]. IEEE Access, 2021, 9: 151327-151336.

[38] ALI M, ISMAIL A, ELGOHARY H, et al. A procedure for tracing chain of custody in digital image forensics: a paradigm based on grey hash and blockchain[J]. Symmetry, 2022, 14(2): 334.

[39] LI M, SHEN Y Z, YE G X, et al. Anonymous, secure, traceable, and efficient decentralized digital forensics[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(5): 1874-1888.

[40] FERNANDEZ-CARRASCO J A, EGUES-ARREGUI T, ZOLA F, et al. ChronoEOS: configuration control system based on EOSIO blockchain for on-running forensic analysis[C]//Blockchain and Applications, 4th International Congress. Berlin: Springer, 2023: 37-47.

[41] LE D P, MENG H S, SU L, et al. BIFF: a blockchain-based IoT forensics framework with identity privacy[C]//Proceedings of the TENCON 2018-2018 IEEE Region 10 Conference. Piscataway: IEEE Press, 2018: 2372-2377.

[42] KUMAR G, SAHA R, LAL C, et al. Internet-of-forensic (IoF): a blockchain based digital forensics framework for IoT applications[J]. Future Generation Computer Systems, 2021, 120: 13-25.

[43] GOLDREICH O, MICALI S, WIGDERSON A. Proofs that yield nothing but their validity and a methodology of cryptographic protocol design[C]//Proceedings of the 27th Annual Symposium on Foundations of Computer Science. Piscataway: IEEE Press, 1986: 174-187.

[44] LI M, CHEN Y F, LAL C, et al. Eunomia: anonymous and secure vehicular digital forensics based on blockchain[J]. IEEE Transactions on Dependable and Secure Computing, 2023, 20(1): 225-241.

[45] TYAGI R, SHARMA S, MOHAN S. Blockchain enabled intelligent digital forensics system for autonomous connected vehicles[C]//Proceedings of the 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT). Piscataway: IEEE Press, 2022: 1-6.

[46] AKBARFAM A J, BARAZANDEH S, MALEKI H, et al. DLACB: deep learning based access control using blockchain[J]. arXiv Preprint, arXiv: 2303.14758, 2023.

[47] AKBARFAM A J, HEIDARIPOUR M, MALEKI H, et al. ForensiBlock: a provenance-driven blockchain framework for data forensics and auditability[C]//Proceedings of the 2023 5th IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA). Piscataway: IEEE Press, 2023: 136-145.

[48] ISLAM M E, ISLAM M R, CHETTY M, et al. User authentication and access control to blockchain-based forensic log data[J]. EURASIP Journal on Information Security, 2023, 2023(1): 7.

[49] LI M, WENG J, LIU J N, et al. Toward vehicular digital forensics from decentralized trust: an accountable, privacy-preserving, and secure realization[J]. IEEE Internet of Things Journal, 2022, 9(9): 7009-7024.

[50] 孙靖超. 基于区块链的可扩展电子取证模型研究[J]. 计算机应用研究, 2021, 38(3): 671-674, 679.
SUN J C. Research on scalable digital forensics model based on blockchain[J]. Application Research of Computers, 2021, 38(3): 671-674, 679.

[51] 陈葳葳, 曹利, 顾翔. 基于区块链的车联网电子取证模型[J]. 计算机应用, 2021, 41(7): 1989-1995.
CHEN W W, CAO L, GU X. E-forensics model for Internet of vehicles based on blockchain[J]. Journal of Computer Applications, 2021, 41(7): 1989-1995.

[52] MIAO Z K, YE C X, YANG P, et al. Blockchain-based electronic evidence storage and efficiency optimization[C]//Proceedings of the 2021 International Conference on Artificial Intelligence and Blockchain

Technology (AIBT). Piscataway: IEEE Press, 2021: 109-113.
[53] RUAN P C, DINH T T A, LIN Q, et al. LineageChain: a fine-grained, secure and efficient data provenance system for blockchains[J]. The VLDB Journal, 2021, 30(1): 3-24.
[54] TAO Q, CUI X H. B-FLACS: blockchain-based flexible lightweight access control scheme for data sharing in cloud[J]. Cluster Computing, 2023, 26(6): 3931-3941.
[55] WAN C, MEHMOOD A, CARSTEN M, et al. A blockchain based forensic system for IoT sensors using MQTT protocol[C]//Proceedings of the 2022 9th International Conference on Internet of Things: Systems, Management and Security (IOTSMS). Piscataway: IEEE Press, 2022: 1-8.
[56] ALMUTAIRI W, MOULAHI T. Joining federated learning to blockchain for digital forensics in IoT[J]. Computers, 2023, 12(8): 157.
[57] KHAN A A, UDDIN M, SHAIKH A A, et al. MF-ledger: blockchain hyperledger sawtooth-enabled novel and secure multimedia chain of custody forensic investigation architecture[J]. IEEE Access, 2021, 9: 103637-103650.
[58] BONOMI S, CASINI M, CICCOTELLI C. B-CoC: a blockchain-based chain of custody for evidences management in digital forensics[J]. arXiv Preprint, arXiv: 1807.10359, 2018.
[59] LONE A H, MIR R N. Forensic-chain: blockchain based digital forensics chain of custody with PoC in hyperledger composer[J]. Digital Investigation, 2019, 28: 44-55.
[60] ALQAHTANY S S, SYED T A. ForensicTransMonitor: a comprehensive blockchain approach to reinvent digital forensics and evidence management[J]. Information, 2024, 15(2): 109.
[61] ALRUWAILI F F. CustodyBlock: a distributed chain of custody evidence framework[J]. Information, 2021, 12(2): 88.
[62] SHARMA P K, CHEN M Y, PARK J H. A software defined fog node based distributed blockchain cloud architecture for IoT[J]. IEEE Access, 2018, 6: 115-124.
[63] BLONDEL V D, GUILLAUME J L, LAMBIOTTE R, et al. Fast unfolding of communities in large networks[J]. Journal of Statistical Mechanics: Theory and Experiment, 2008, 2008(10): P10008.
[64] CLACK C D, BAKSHI V A, BRAINE L. Smart contract templates: foundations, design landscape and research directions[J]. arXiv Preprint, arXiv: 1608.00771, 2016.

[作者简介]

范伟（1984– ），男，北京人，博士，中国科学院信息工程研究所高级工程师、硕士生导师，主要研究方向为移动通信安全、云计算安全、虚拟化安全等。

李海波（2001– ），男，河南南阳人，中国科学院信息工程研究所硕士生，主要研究方向为区块链安全、数字取证等。

张珠君（1987– ），女，河北南宫人，博士，中国科学院信息工程研究所工程师，主要研究方向为区块链安全、系统安全等。