一种总分双链的银行可信数据管理方案*

杨 城[1,2] 颜 河[3] 段 江[1,2,3] 康 立[1,2] 陈 智[1,2]

[1](西南财经大学 经济信息工程学院 四川成都 611130)

[2](西南财经大学 中国区块链研究中心 四川成都 611130)

[3](成都恒图科技有限责任公司 四川成都 610041)

通讯作者: 杨城, E-mail: yangcheng@swufe.edu.cn

摘 要: 针对传统银行系统的中心化数据存储模式高效便捷但不透明，而新兴的去中心化的应用系统公开透明但共识机制低效的问题，本文提出了一种面向银行系统的总分双链的弱中心化可信数据管理方案。该模式的核心思想是应用区块链技术，打造彼此交叉且相互印证的总分双链的数据存储结构，结合大量客户端基于密码学技术的分布式监督，实现数据的中心化可信存储与管理。它将中心化的数据存储与去中心化的数据验证相结合，从而兼有中心化管理的高效性和分布式机制的透明性，其本质上是一种民主监督下的中心模式，在满足银行独立完全掌控数据的同时，兼顾了隐私与监督之间的平衡。

关键词: 总分双链; 弱中心化; 区块链; Merkle 树; 可信银行

中图法分类号: TP309.2

A Trusted Data Management Schema for Digital Bank based on Binary-Chain

YANG Cheng[1,2] YAN He[3] DUAN Jiang[1,2,3] KANG Li[1,2] CHEN Zhi[1,2]

[1](School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China)

[2](Blockchain Research Center, Southwestern University of Finance and Economics, Chengdu 611130, China)

[3](Chengdu Everimaging Ltd, 610041, China)

Abstract: Given the fact that the centralized data storage schema of traditional banking system is efficient but non-transparent, while the emerging decentralized data management system is transparent but inefficient in the consensus mechanism, this paper suggests a new trusted data management schema with a cross-chain for general ledger & subsidiary ledger in semi-centralized architecture, called “Binary-Chain Schema”. The core idea of the new schema is to build a cross-chain for the transaction data with the help of blockchain technology, so as to realize centralized trusted storage and management of data by combining the distributed supervision of a large number of clients based on cryptography technology. Combining the centralized data storage mode and decentralized data validation method, it is both efficient and transparent. Essentially, the Binary-Chain schema is a centralized approach based on distributed democratic supervision. Achieving the balance between privacy and supervision, it satisfies the need for banks to independently and completely control over the transaction data.

Key words: Binary-Chain; Semi-centralized; Blockchain; Merkle Tree; Trusted Bank

1 引言

传统银行体系一直采用“中心化”管理模式，该模式具备高效可控和管理便捷等优势，但近年来随着一些内部人作案和外部攻击事件的披露，中心化模式的弊端日益突显：中心集中存储所有的账户信息和交易信息，很容易成为黑客攻击的目标，造成巨大损失。更为重要的是，中心化管理缺乏透明性和可监督性，储户只能被动的完全信任银行，难以主动实施监督。银行拥有每一笔交易的记账权，具备伪造客户信息、篡改交易记录的能力，在特定情况下，存在中心作弊侵害储户利益的风险。

近年来，基于区块链技术的“去中心化”管理模式逐渐受到关注和热捧。建立在多方共同记账原理下

*基金项目: 教育部人文社科项目(17YJCZH210); 中央高校科研业务费专项资助项目(JBK120505)

Foundation item: MOE (Ministry of Education in China) Project of Humanities and Social Sciences (17YJCZH210); Fundamental Research Funds for the Central Universities (JBK120505).

的区块链技术，具备高安全性、公开透明、数据防篡改可追溯等优势[1-4]。但该技术目前仍处于初期，存在高耗低效、隐私泄露和责任主体缺失等问题。[5-8]此外，若采用该模式，银行将丧失系统的中心地位，政府更难以实施监管，无法保证金融体系的安全性和可控性。目前，以“比特币”为代表的虚拟货币是区块链在金融应用中的典型代表，我国央行也准备发行自己的数字货币，但央行的数字货币以国家信用作为背书，仍然采用中心化模式，无法从本质上体现区块链多方信任的特点，进而无法充分发挥虚拟货币的业务优势。同时，它对现有银行体系的运作模式改动较大，建设成本和维护成本都将十分高昂。

有鉴于此，为了更好的满足银行业需求，提升银行公信力，我们结合现有银行系统的业务特点和区块链的技术优势，提出了一种新型的银行数据存储和管理方案，使其在保持集中高效、低成本运作的同时，兼备透明可监督的特点。

本文后续部分的结构如下：第二节整体性阐述该新型数据管理方案的设计理念和核心思想；第三节介绍新方案实施的非对称加密账户管理模式和交易流程；第四节是全文的核心，对总分双链模式的账务数据管理进行详细解读；第五节阐述个体储户如何进行分布式数据验证；最后是全文的总结，对新方案的优势和扩展应用进行归纳。

2 总分双链的弱中心化管理模式

如前所述，传统银行的高效便捷源于数据的集中管理和对银行的无条件信任，而去中心化模式的安全透明源于结合密码学技术的分布式存储。基于此本文结合现有银行系统的业务特点和区块链的技术优势，提出了一种新型的银行数据存储和管理方案——***总分双链的弱中心化可信数据管理方案***，简称“BC 方案”（Binary-Chain Schema）。它将中心化的数据管理与分布式的、轻客户端的数据审计相结合，从而兼有中心化的高效便捷性与去中心化的安全透明性，它在本质上是一种民主监督下的中心化数据管理模式。

BC 方案的核心词是“总分双链”和“弱中心化”。其中，前者是底层技术手段，后者是上层制度设计，二者共同维系整个体系的数据共识，以实现“可信”银行的目标。

“总分双链”是指在数据存储模式上，通过引入块链结构，应用密码学技术打造彼此交叉且相互对照的两条数据链来“锁住”交易数据——“总账链”记录对应时间片内每一笔交易的哈希摘要（对包括交易明细和储户、银行双方的数字签名等信息求哈希值），既保证了所有交易的真实性和不可篡改性，又有效保护了储户和银行的隐私性；而“分户账链”以连环哈希的方式记录个体储户每一笔交易的哈希摘要（对包括交易明细和余额等信息求哈希值），用于验证个人交易的完整性和可追溯性。

“弱中心化”是指在数据管理模式上，采用“写验分离”的数据管理方式——一方面保留银行在系统中的核心地位，它是所有交易的共同中介，也是系统唯一的记账人，集中存储全部的账务数据，是维持系统高效便捷运作的关键；另一方面所有储户都是体系的分布式监督人，他们手握基于区块链技术生成的验证数据，在查验自身交易数据的同时，间接实施对系统整体账务数据的审计，一定程度上实现了数据的公开透明性，从而有效避免了银行篡改数据、记假账的可能，使其成为真正可信任的数字银行。

3 账户管理与交易流程

在现有银行体系中，目前普遍采用对称加密技术来保护账户安全和进行身份识别。这种技术虽然简单易行，但必须基于对银行的绝对信任。显然这与本文所提的“可信”银行是相矛盾的，因为银行可以利用手中的信息优势，任意伪造或篡改储户的交易记录。

因此，在本文 BC 方案的数字银行体系中，系统采用非对称加密技术来识别身份和保障安全，即储户和银行都有各自的私钥和公钥，通过数字签名来进行身份认证和实现信息的不可抵赖性。例如，储户通过数字签名来确保交易请求的真实性，银行通过数字签名来确保系统消息和转账回执的真实性。同时，类似比特币等虚拟数字货币的注册模式，体系中没有所谓中心化的 CA 机构，储户的密钥对由储户自己生成，其公钥与账户地址的对应关系在其向银行进行账户注册时完成映射。**无 CA 机构的非对称加密账户管理技术是实现可信数字银行的基础**。

图 1 展示了一笔转账交易的完整流程。其中，交易请求 TX 中，V_A、V_B分别表示转出方（Alice）和转入

方（Bob）的账号，Amount 为转账金额，$Time_1$ 为请求时间，Memo 为备注信息，Sgn_A 为转出方对该请求的签名，以确认该请求的真实性；转账回执 RE 中，$Time_2$ 为记账时间，$Balance_A$ 为交易完成后转出方的账户余额，$Head\_PB_A$ 表示转出方分户账链的链头信息（参见后面公式 1），Sgn_C 为银行对该交易的签名，以确认该交易已完成的真实性；转账通知 Msg 中的参数与 RE 类似。

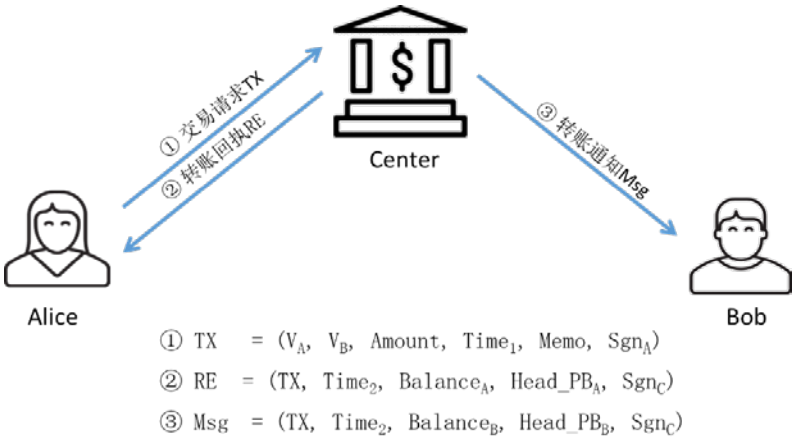

Fig. 1　A flow chart for transfer transaction

图 1　转账交易流程示意图

4 总分双链设计

将所有交易流水依据参与主体和时间维度交织成一张有向交易网，纵向为一个个的交易主体，横向为某个时间片内的交易关系，前者对应一个个储户的分户账信息，后者对应系统的总账信息，两条线分别从不同维度锁定交易记录，并通过发生额和余额的数值关系来相互印证。

如图 2 所示，我们将交易时间周期性的划分为等长的若干时间片（建议按天划定，设为 24 小时，以便与银行现行的“日记账”对应），单位时间片内的每一个储户视为一个节点 V_i，交易网内的每一个箭头都对应一笔转账交易，箭头始端为转出方，末端为转入方。横向实线代表不同节点间的资金流向，纵向虚线代表同一个节点内的余额流向。

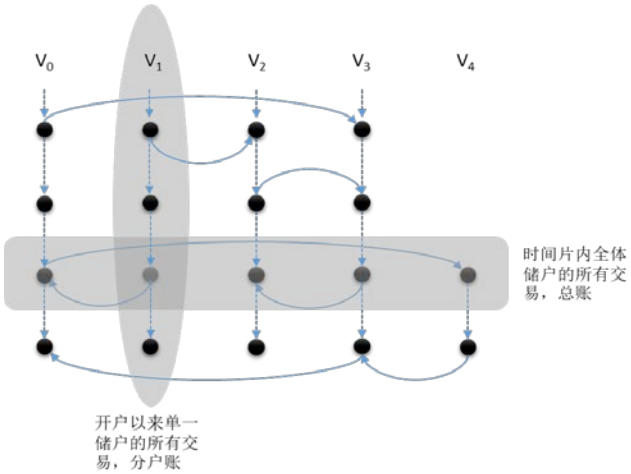

Fig. 2　A crossover network for transfer transaction

图 2　纵横交织的交易网

显然，这是一个不断生长的动态网络，并且不断有新的节点产生（新开户）。我们将银行本身视作一个特殊节点 V_0，初始状态全网只有 V_0 一个节点，其他储户节点 V_i 在后期逐渐并入网络。规定，所有节点的存款余额必须不小于零，V_0 的初始余额为一个足够大的正值 TB（Total-Balance），而其他节点 V_i 的初始余额为 0。由于所有交易都是交易网内节点间的转账，包括储户节点 V_i 对自身账户存取现金也可视为 V_i 与 V_0 之

间的转账[1]。因此，交易网内的资金流是一个闭环，总余额始终等于 TB。

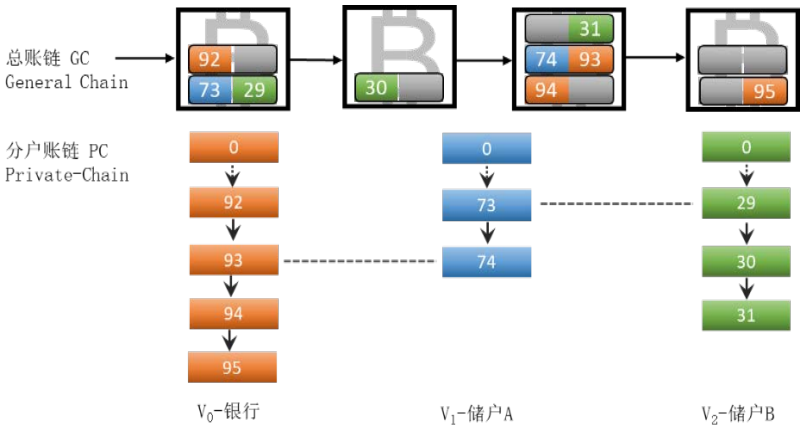

Fig. 3　The logical structure diagram of Binary-Chain

图 3　总分双链的逻辑结构示意图

如图 3 所示，我们将图 2 的有向交易网进一步抽象，分别从横竖两个维度进行切割，**横向为账务信息，纵向为账户信息，形成两条逻辑数据链：一条总账链 GC（General-Chain）和若干条分户账链 PC（Private-Chain）。**[9]前者按照块链结构存储信息，每个区块装载对应时间片内的所有交易，并利用哈希函数前后迭代，锁定历史交易；后者将个人的一条条交易记录按照时间先后顺序串联，并按公式 1 的计算模式组成连环哈希。其中，TX_i 表示第 i 笔交易的明细，$Balance_i$ 表示第 i 笔交易后的余额；末尾 H_n 称为该分户账链的链头 Head_PC，例如图 3 中的 31 号绿色方块就代表节点 V_2（储户 B）的 Head_PC。

$$H_0=hash(TX_0|Balance_0),\ H_1=hash(H_0|TX_1|Balance_1),\cdots,H_n=hash(H_{n-1}|TX_n|Balance_n) \qquad (公式\ 1)$$

由于整个系统只有一条总账链 GC，而每一位储户都有一条相对独立的分户账链 PC，因此在具体实施过程中，**我们以 GC 为主体，所有 PC 的链头信息同步并入 GC 的区块中，从而将两条逻辑链合并成一条物理链 BC（Binary-Chain）。**BC 主链的所有区块头信息对全体储户公开，通过对 BC“块头链”的分布式存储，实现全体储户的交易共识，进而达到对中心银行实施监督的目的。

图 4 展示了 BC 主链的详细结构图，其中上半部分为主链整体结构，左下部分为当期所有交易构成的总账树 TX_Merkle，右下部分为当期所有 PC 链头信息构成的分户账树 Head_PC_Merkle。（1）BC 主链由若干区块按时序前后衔接构成，每个区块包括区块头和区块体两部分。区块体储存当期的交易明细，区块头存储该区块的摘要信息，如时间戳、交易笔数、交易总额模 M 的值（M 是一个远大于单笔交易最大限额但小于单位时间片交易总额的整数值）和两颗默克尔树的根值，以及父区块头的哈希值。（2）TX_Merkle 是一颗扩展的默克尔树，每个叶节点对应当期的一笔交易 TX，存储的信息包括该交易的哈希值 h(TX)和发生额 Amount。[10][11]除了像传统默克尔树对哈希值层层迭代外，每个节点的 SumA 字段等于其最相近的两个子节点的发生额之和模 M。这样，根节点的 Root 字段锁定了当期的所有交易，同时 $SumA^*$字段等于当期的总发生额模 M。（3）类似的，Head_PC_Merkle 也是一颗扩展的默克尔树：每个叶节点对应一位个体储户的账务信息，储存的内容除各自 PC 的链头信息外，还包括该账户本期内的发生额之和∑Amount 和当期余额 Balance[2]。对应的，根节点的 Root 字段锁定了当期所有储户的分户账链 PC，同时 $SumA^{**}$字段等于当期总发生额的两倍值模 M（同一笔交易会同时记录在交易双方的分户账链上），SumB 字段等于系统总余额 X。**这样的设计既保留了可验证的数值特征，又有效的保护了银行的商业隐私。**上述数量关系可表述如下：

$$Head\_PC\_Merkle\_Root(SumA^{**}) = 2 * TX\_Merkle\_Root(SumA^{*})\ \%\ M \qquad (公式\ 2)$$

$$Head\_PC\_Merkle\_Root(SumB) = TB \qquad (公式\ 3)$$

---

1 普通储户 V_i 存入 X 的现金，等价于 V_0 向 V_i 转账 X，即 V_0 自身存款余额减少 X，而 V_i 的存款余额增加 X，同时银行的现金资产增加 X。类似的，普通储户 V_i 提取 Y 的现金，等价于 V_i 向 V_0 转账 Y，即 V_i 自身存款余额减少 Y，而 V_0 的存款余额增加 Y，同时银行现金资产减少 Y。V_0 的初始余额 TB 足够大意味着无论储户如何提现，V_0 的余额永远为正值。

2 这里∑Amount 表示当期该账户的所有发生额之和，发生额都是正值，转入和转出都需要计算。例如，如果按天划分时间片，∑Amount 就表示该账户当日的总交易额。

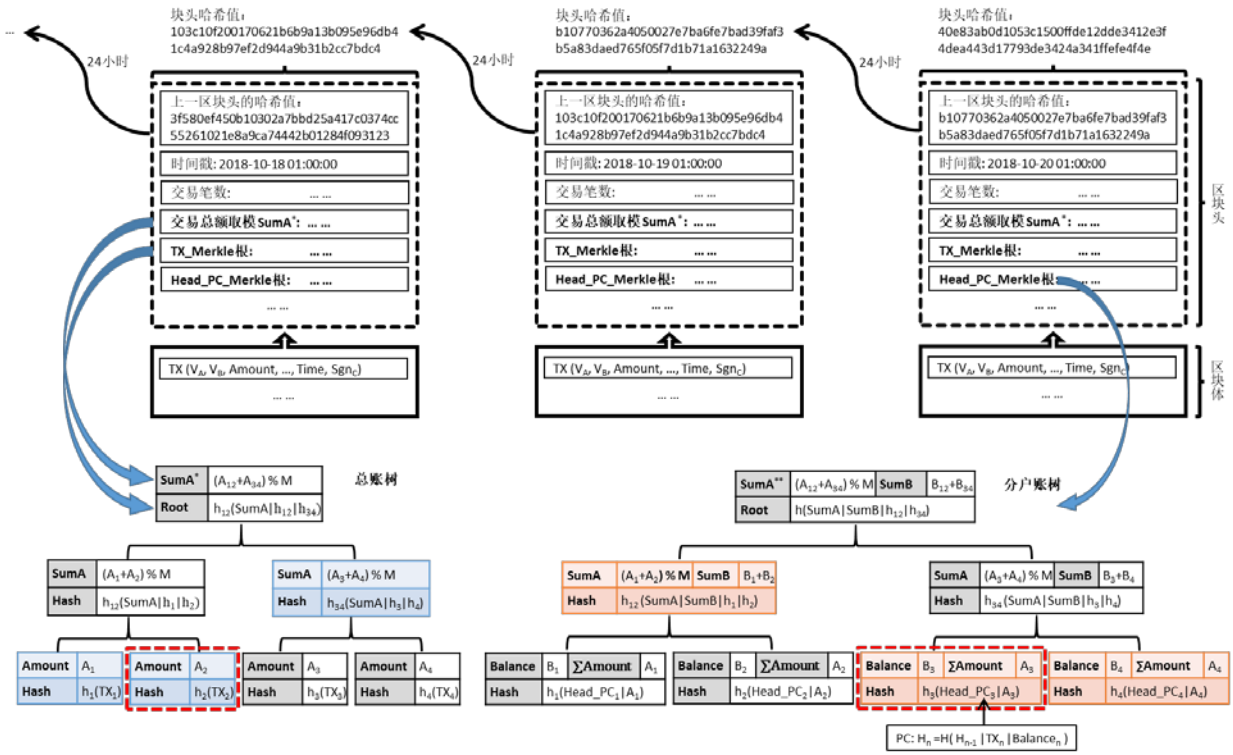

Fig. 4　The actual structure diagram of Binary-Chain

图 4 Binary-Chain 主链的详细结构图

如图 4 所示，***非对称加密技术保证了银行无法伪造虚假交易来直接侵害储户利益，TX_Merkle 锚定了所有的交易，Head_PC_Merkle 锚定了所有的分户账链 PC，而每条 PC 又是锚定每位储户个人交易的锚。***通过双链存证及相互对照，结合数值验证，可以有效“锁住”数据，预防银行作弊。如果只有总账链，没有分户账链，储户只能完成交易的存在性证明，但无法验证交易的完整性，即银行可能“漏掉”部分交易记录。相反，如果只有分户账链，储户虽然能够独立验证自身交易的完整性，但由于分户账链是彼此孤立无关联的，公众缺少总账数据无法充当交易见证人，银行可以利用信息不对称的优势，制造局部真实，整体虚假的局面，例如转账时记单边账，或利用内部账户伪造历史交易记录。

5 分布式验证

在数据的存储安排上，银行作为系统的中心，存储全部数据，包括完整的交易流水、分户账信息、BC 主链数据和每位储户的 PC 数据等。由于存储空间和计算能力的限制，以及对隐私安全的考虑，个体储户仅存储系统的 BC“块头链”（由所有 BC 主链区块的区块头构成的数据链）和自身的分户账链 PC。

基于上述 BC 方案的块链设计和存储安排，每一位储户都可以对自己的相关交易和账户信息进行独立审计，而不用依靠银行自身或第三方权威机构来验证。具体而言，储户的独立审计是通过其日常的交易查询来实施的，在查询过程中储户需进行两方面的验证来确保查询信息的真实性：**块头验证和交易验证。前者验证自身分户账链的链头信息 Head_PC 与 BC 主链最新区块的关系；后者验证具体交易与所在 BC 主链区块的关系。**

在实际的计算过程中，每一部分又细化为两类验证，**即对自身交易数据的验证和对系统整体交易的审核。前者是存在性验证**，通过银行返回对应区块分户账树的 Merkle 路径来验证自身 Head_PC 是否被如实的包含在 BC 主链的 Head_PC_Merkle 中，以及通过银行返回对应区块总账树的 Merkle 路径来验证自身分户账链上的某笔具体交易是否被如实的包含在 BC 主链的 TX_Merkle 中 [3][10]；**后者主要体现在对公式 2 和公式 3**

[3] Merkle 树是一种哈希二叉树，它是一种用作快速归纳和校验大规模数据完整性的数据结构。构造一棵完整的 Merkle 树需要递归的对数据节点进行哈希运算，并将新生成的哈希节点插入到 Merkle 树中，直到只剩下一个哈希节点，该节点就是 Merkle 根。

的数值验证，即从自身 PC 叶节点回溯到 Head_PC_Merkle 的根节点，查验总余额 SumB 是否恒定为 TB，总发生额 SumA**是否等于 BC 区块头中 SumA*的两倍值模 M，以及从自身 TX 叶节点回溯到 TX_Merkle 的根节点，查验总发生额 SumA*是否等于 BC 区块头中 SumA*的值[11]。

需要说明的是，储户在进行交易查询前，需要先进行数据同步，包括更新系统 BC 块头链和自身 PC，即下载自上次更新以来系统新增的区块头数据和自身 PC 新的连环哈希节点。注意，数据同步时需要对更新数据的正确性进行验证，即依据先前存储的数据逐步计算新数据的延续性，看是否能够顺利推导出 BC 块头链的最新“块头哈希值”和自身 PC 的最新 Head_PC。

从上述查验过程不难看出，大量分散的个体储户在对自身交易信息和账户数据进行直接查验的同时，通过十字双链对数据存在性、完整性和数值关系的交叉验证，间接实现了对银行整体交易信息和账务数据的审计，从而大大降低了银行篡改数据、做假账的可能，提升了银行的公信力。同时，BC 方案对于客户端而言是一种轻量级的、高效的、可信的数据应用体系。而对于银行而言，系统改造容易，实施成本低。银行无需额外建造新的公链或发行新的数字货币，仅仅通过引入非对称加密账户管理技术和对账务数据进行区块链式的存储改造，就能使得银行在保留系统中心地位的同时，实现众多虚拟货币的特色业务和功能。

6 结论

本文结合现有银行系统的中心化数据存储模式与基于区块链思想的总分双链数据验证机制，实现了一种面向数字银行的弱中心化可信数据管理方案的创新：***应用区块链技术，打造彼此交叉且相互印证的总分双链的数据存储结构，结合大量客户端基于密码学技术的分布式监督，实现数据的中心化可信存储与管理。***它在数据管理模式上为银行系统提供了一种区别于区块链全分布式管理和传统集中式管理之外的第三条路径——“弱中心化”模式，即通过总分双链结构，一方面保证数据公开透明，具有区块链的不可篡改性和可追溯性特征，同时数据集中存储，满足银行自己掌控数据，保护银行及储户账务隐私的需求。

相对于中心化的传统银行体系和以比特币为代表的去中心化虚拟货币体系，BC 方案的优势主要体现在以下两点：（1）兼顾数据透明与执行效率。通过数据的写验分离，实现了银行高效集中记账，储户分散民主监督的全新模式，通过总分双链监督机制“锁住”数据，大大提升银行公信力，将银行的角色由传统的管理者转变受监督的中间人。（2）兼顾隐私保护与资金监管。相比于比特币网络，无关人员无法获取其他储户的交易记录和账户信息，也无法获知银行的存款规模等整体性商业信息，同时数据存储中心化和不可篡改的特点也便于审计机构对银行实施监管和内部审计。

最后，本文的 BC 方案并不局限于银行系统的应用，而是一种全新的数据管理方案，具有广泛的应用前景，其上可存储电子凭证、财务数据、交易信息等众多的数字化信息，例如中央银行对数字货币的管理，商业银行对数字现金的管理，虚拟货币交易所对交易记录的管理等，用于打造一款应用广泛的、可信任的“数据银行”，成为未来数字普惠金融的重要基础设施。

Reference:

[1] Binanda S., Samiran B., Sushmita R., et al. Retricoin: Bitcoin based on compact proofs of retrievability[C]. Proceedings of the 17th International Conference on Distributed Computing and Networking. New York: ACM, 2016, 14: 1-10.

[2] Juels A., Kaliski B.S. PORs: Proofs of retrievability for large files[C]. Proceedings of ACM Conference on Computer and Communications Security. New York: ACM, 2007:584-597.

[3] Andrew M., Ari J., Elaine S., et al. Permacoin: Repurposing Bitcoin Work for Data Preservation[C]. Proceedings of IEEE Symposium on Security and Privacy. Washington D C IEEE, 2014: 475-490.

[4] Walsh K. 'Blockchain' Increases Online Trust: New Technology Could Bolster Digital Badges and E-Portfolios, among Other Innovations[J]. University Business, 2017, 20(2): 24.

[5] David S. Understanding the DAO Attack[EB/OL]. https://www.coindesk.com/understanding-dao-hack-journalists/, 2016.6.

[6] Pete R. Split or No Split? Bitcoin Miners See No Certainty in Segwit2x Fork[EB/OL]. https://www.coindesk.com/split-no-split-bitcoin-miners-see-no-certainty-segwit2x-fork/, 2017.11.

Merkle 树提供了一种快速验证数据存在性的方法。当 N 个数据元素经过哈希计算并插入 Merkle 树时，任意一个满节点回溯至多 $\log_2(N)$个路径节点就能到达 Merkle 根，即至多经过 $\log_2(N)+1$（包括自身哈希）次哈希计算就能检查出任意数据元素是否存在于该 Merkle 树，从而完成其存在性证明。

[7] Xiwei X., Cesare P., Liming Z., et al. The Blockchain as a Software Connector[J]. Software Architecture, 2016: 182-191.

[8] Roman B., Michel A., Matti R., et al. Blockchain Technology in Business and Information Systems Research[J]. Business & Information Systems Engineering, 2017, 59(6): 381-384.

[9] Ren Z., Cong K., Aerts T. V., et al. A Scale-out Blockchain for Value Transfer with Spontaneous Sharding. ARXIV PREPRINT ARXIV: 1801.02531, 2018.

[10] Andreas M.A. Mastering Bitcoin[M]. O'Reilly Media, 2014.12.

[11] Dagher G. G., Bonneau J., Clark J., et al. Provisions: Privacy-preserving Proofs of Solvency for Bitcoin Exchanges[C], 2015:720-731.