DAG 区块链中连通度极限值 CLV 研究

孙 晴[1] 李志淮[2] 贾志鹏[2] 李文祺[2]
[1](海军大连舰艇学院 辽宁 大连 116018)
[2](大连海事大学信息科学技术学院 辽宁 大连 116002)

摘 要 DAG(Directed Acyclic Graph)区块链技术在性能、确定性等方面较中本聪最长链显露出更多优势。DAG 区块链技术中的核心问题就是交易排序,其中引入了连通度极限值(Connectivity Limit Value, CLV)。可推定中本聪最长链的 CLV 取值为零,保证网络的安全达到极高的水平,但是交易验证的并行度差。为此对 DAG 区块链扩容方案进行泛化分析,DAG 区块链网络具有高并发性,但 CLV 取值非零,相应安全性受到影响。继而探索连通度极限值与网络的安全阈值、网络延迟等存在的关系,并针对在 DAG 区块链网络中要保障 CLV 取值在安全阈值与网络延迟之间的平衡问题,引用进化计算中的方法给出了平衡关系式,在安全性与延迟度之间提出合理优化。此外针对 DAG 区块链中沙漏区块的 CLV 取值进行分析并给出合理建议。

关键词 DAG 区块链 连通度极限值 中本聪最长链 沙漏区块 交易排序

中图分类号 TP3 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2024.07.037

GENERALIZATION ANALYSIS OF CONNECTIVITY LIMIT VALUE IN DAG BLOCKCHAIN

Sun Qing[1] Li Zhihuai[2] Jia Zhipeng[2] Li Wenqi[2]
[1](*Dalian Naval Academy*, *Dalian 116018*, *Liaoning*, *China*)
[2](*School of Information Science and Technology*, *Dalian Maritime University*, *Dalian 116002*, *Liaoning*, *China*)

Abstract Directed acyclic graph (DAG) blockchain technology has more advantages than Satoshi Nakamoto's longest chain in terms of performance and certainty. The core problem in DAG blockchain technology is transaction sequencing, which introduces the connectivity limit value (CLV). It can be inferred that the CLV of Satoshi Nakamoto's longest chain is zero, ensuring that the security of the network reaches an extremely high level, but the parallelism of transaction verification is poor. For this reason, this paper conducts a generalized analysis of the DAG blockchain expansion plan. The DAG blockchain network has high concurrency, but the value of CLV is non-zero, and the corresponding security is affected. This paper explored the relationship between the connectivity limit and the security threshold of the network, network delay, etc., and aimed to ensure the balance between the security threshold and network delay in the DAG blockchain network. The method gave a balance relationship and proposed a reasonable optimization between safety and delay. In addition, it analyzed the CLV value of the hourglass block in the DAG blockchain and gave reasonable suggestions.

Keywords DAG blockchain Connectivity limit value Satoshi Nakamoto's longest chain Hourglass block Transaction sorting

0 引 言

2008 年,随着中本聪(Satoshi Nakamoto)发表《比特币:一种点对点的电子现金系统》[1]一文,区块链技术应运而生。在 2019 年 10 月 24 日中央政治局第十八次集体学习时,习近平总书记更是强调要把区块链作为核心技术自主创新重要突破口,鼓励加快推动区块链技术和产业创新发展[2]。这一重要讲话对我国区块链技术的发展具有重大意义,确立了区块链核心技

收稿日期:2020-12-28。**孙晴**,硕士生,主研领域:区块链技术,作战仿真。**李志淮**,教授。**贾志鹏**,硕士生。**李文祺**,硕士生。

术创新在国家层面的战略地位。

目前区块链技术中仍存在很多亟待突破的问题，如规模扩展性（扩容）问题、区块验证缓慢问题、区块链隐私监管治理问题、存储问题等。其中扩容问题是区块链技术最急迫的核心问题之一[3]。针对扩容问题，许多项目提出解决方案，目前主要的扩容方案有：分片分层[4]、有向无环图 DAG 区块链技术[5]、状态通道[6-7]、闪电网络[7]、Rollup 方案等。其中 DAG 区块链技术作为重要的底层扩容方案，一直受到广泛关注。

2013 年，在 bitcointalk. org 论坛上用户首次提出“以有向无环图 DAG 作为区块链的底层数据结构，进而提高系统整体性能”，但此时提出的方案还停留在侧链思路上。2015 年 Lerner 在 bitslog[8] 网站的个人博客中发表文章《DaggerCoin：A cryptocurrency without blocks》，提出 Blockless DAG 的概念，首次提出 DAG 区块链技术把区块和交易均作为节点单元，可以直接对交易进行全网交易排序，达到无块交易 Blockless。随后 IOTA、Conflux、Hashgraph 等 DAG 区块链技术陆续出现。

DAG 区块链技术是 DAG 有向无环图的数据结构与区块链技术结合而形成的一项新的技术，在效率、确定性、去中心化等方面比以往的区块链显露更优的特性。目前面临的核心技术难点就是要实现安全、高效的 DAG 区块链技术以达到对交易全网排序并确保排序的唯一性和一致性。

连通度极限值（CLV）是在 DAG 区块链中能够有效反映出网络结构连通情况的一个参数，出现在 PHANTOM 协议[9] 中。该协议由以色列研究团队在 2020 年国际密码研究协会 IACR 首次提出，是 DAG 区块链的最新研究成果。PHANTOM 协议提出一种实现可伸缩性的在线方法，可以极大程度保障网络安全；协议中 CLV 度量值 k 用来限定网络的安全容忍度，同时可以保证并行出块。本文对 PHANTOM 协议提出的 CLV 值 k 进行多方面的分析研究。

在研究 DAG 区块链技术的交易排序的问题时注意到：连通度极限值 CLV 在中本聪最长链（即在区块共识时采用最长链原则进行共识的区块链）中的取值一定为 0。通过对比中本聪最长链与 DAG 的结构、安全性、性能、确定性上存在的差异，根据 CLV 在中本聪最长链中取值为 0 这一特殊值在 DAG 结构中，本文进行了泛化分析，得出 CLV 与区块链中的网络传播时延、区块链吞吐量等重要指标存在的关系，并进一步讨论三者之间的具体关联并提出合理的方案寻求其间的平衡。

本文主要贡献如下：

（1）对中本聪最长链的 CLV 在 DAG 区块链技术中进行泛化分析，给出几种特殊 DAG 区块链网络结构中 CLV 受 DAG 结构影响情况。

（2）分析 CLV 在通用 DAG 区块链中，与网络传播时延、区块链吞吐量具体存在的量化关系，并提出平衡三者关系的合理方案。

（3）发现 DAG 区块链中存在的一种特殊区块沙漏区块，分析沙漏区块的影响并推荐合理的 CLV 区间。

1 相关研究与技术分析

1.1 中本聪最长链

中本聪最长链即在区块共识时采用最长链原则进行共识的策略，用于解决 P2P 网络中时间序列的竞争与确定问题。为便于后续与 DAG 区块链技术作对比分析，本文中称经典的块链式区块链为中本聪最长链。

在中本聪最长链中，每个个体或组织作为区块链中的一个节点，这些节点之间进行交易的公共数据就记录在区块链这个分布式账本上，这就是区块链的核心。区块链涉及到技术、经济甚至政治等多个领域。从技术上来讲，区块链技术又涉及到四大核心技术[10]：数据结构、共识机制[11]、密码学、分布式存储。

单从词的表意可以了解到区块链的数据结构[12]由区块和“链”构成。中本聪最长链中每个区块由区块头和区块体构成，其中区块头存储着区块的头信息，包含上一个区块的哈希值（PreHash）、本区块体的哈希值（Hash）、时间戳信息（TimeStamp）等。区块体记录交易信息，包括区块验证的交易记录以及创建区块产生的其他信息等。每一个区块与其上一个区块由哈希指针链接形成链式结构，这样环环相扣就形成了一条“关系链”，这样的结构和内容就构成了区块链的数据结构。

区块链的块链式数据结构、去中心的分布式存储结构、区块链共识机制，保证节点能够主动并正确地“记账”。PoW[13] 是最初的一种共识机制，这种共识机制中，所有参与的节点通过争夺计算能力来竞争获得记账权，这也是目前公认最为公平和去中心化的一种共识机制。但是由于 PoW 这种共识机制存在浪费大量的算力资源和时间成本的弊病，后又出现了 PoS[14]、DPoS、PBFT[15] 及经典共识机制改进的多种共识机制[16]。

区块链背后的保障由密码学支撑，密码学中保障安全的一个重要内容是非对称加密技术。区块链中涉及到公钥和私钥经过哈希算法、椭圆曲线加密算法等

重要的算法进行加密解密,这也是保障区块链信息安全的一项重要技术。

在分布式存储技术中区块链一个最重要的思想就是去中心化,区块链中每一个区块中的数据信息都是由参与记账的节点竞争记录的,没有第三方的介入,保证交易数据信息的安全公开透明。

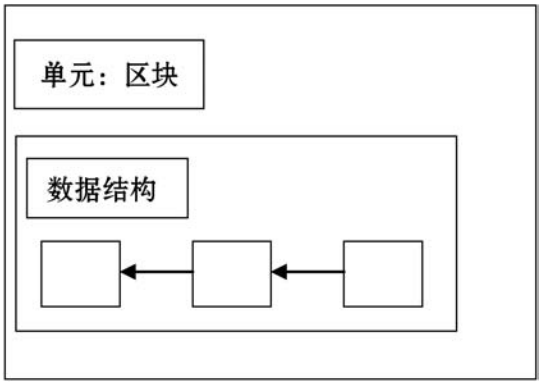
虽然区块链技术随着研究的深入不断完善和发展,但是仍面临一些问题,更存在许多需要突破的核心技术,其中一个重要的问题就是区块链交易验证的性能亟待提高。目前的区块链扩容方案有分片分层、DAG 区块链技术、状态通道、闪电网络等多层面的解决途径。本文选取 DAG 区块链技术进行相关工作。

1.2　DAG 区块链技术

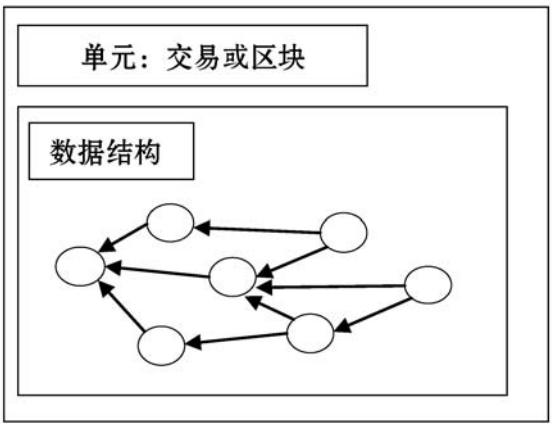
中本聪最长链运行至今被公认为最安全,但是被研究人员公认的弊端是吞吐量太低,这是因为它采用了最长链原则。最长链原则要求所有诚实节点能迅速接收到新创建的区块,所以要等待一个区块的消息同步到所有区块节点时才可以创建下一个区块。所以当前研究面临的一个难题是吞吐量与安全之间的权衡,DAG 区块链技术就是要做到在保障安全的同时实现并行。在这一理论背景下本文选取区块链中的核心技术问题扩容方案进行探索,并选取难点问题 DAG 区块链技术进行研究。

DAG 区块链技术实际上就是基于有向无环图的数据结构为基础发展的区块链技术,DAG 在计算机领域可以作为一种数据结构的类型。DAG 区块链技术明显区别于中本聪最长链主流区块链的特点有如下几个方面:单元是交易或区块;可以异步并发处理交易信息;每个单元记录单个用户交易。

DAG 区块链技术解决了中本聪最长链成块慢、缺乏并发处理机制等问题,将同步记账提升为异步记账解决了中本聪最长链的并发度问题,为区块链提供了有效的扩容方案。图 1 展示了中本聪最长链与 DAG 区块链在数据结构及交易单元上的差异。



(a) 中本聪最长链数据结构



(b) DAG 区块链数据结构

图 1　中本聪最长链与 DAG 区块链结构对比

DAG 区块链技术提高了中本聪最长链的并发度,解决了中本聪最长链孤块产生浪费资源的问题,为区块链的扩容提供了一个有效方案。但是技术的创新在其有显著优势的同时也会存在亟待解决的问题。DAG 区块链技术当前主要面临两个潜在问题:

(1) DAG 区块链中交易时间的不确定性。在中本聪最长链中当达成共识区块出块后,区块的交易时间相对可控,但是在 DAG 区块链中,由于其每个单元记录单笔交易的特点,每笔交易独立处理事务,所以没有能保证达到预期的机制确保交易确认时间。

(2) DAG 区块链技术中另外一个重要待突破的问题就是排序问题。在 DAG 区块链中交易节点的排序可以有效解决双花攻击,所有交易在 DAG 区块链技术中连通的状态下,如果出现双花交易,在路径上排序之后的先序交易视为有效交易,进而有效保证区块链的安全性。

另外,在一个涉及到加减乘除运算的较为复杂系统中,如果没有一种类似全局“账本”的机制保障每个节点最终所确认的操作顺序保持一致的话,很有可能出现节点间所记录的数据在运行一段时间后出现较大的偏差。总的来说,DAG 区块链技术中亟需从理论上解决的问题是交易时长不可控和不存在全局排序机制。

本文从 DAG 区块链技术解决排序问题切入,现有应用了 DAG 区块链技术的项目提出了许多排序方案。如文献[17]的 IOTA 项目中应用 Tangle 确认规则进行的节点排序,在 IOTA 中每个节点代表一笔交易,没有区块的概念,在 IOTA 中不存在挖矿这一步骤,这就使网络的吞吐量很高。

2018 年 12 月,图灵奖得主、清华大学交叉信息研究院院长姚期智的清华姚班创立区块链项目 Conflux[18] 从包括 Metastable、IMO Ventures、峰瑞资本、红衫中国

等投资机构,以及一些知名的科技公司获得了 3 500 万美金融资。Conflux 前期所结合的正是 DAG 区块链技术,在测试环境中,Conflux 实现了 3 000 ~6 000 TPS(TPS:每秒系统处理的数量)-公链比特币 7 TPS,以太坊 30 ~40 TPS,扩容性得到了极大的提升。但后期由于交易排序机制基础理论遇到瓶颈由 DAG 转为树图概念。

2020 年,以色列研究团队在国际密码研究协会 IACR(IACR 是非营利的国际性科研组织,致力于密码学及相关领域的前沿课题研究)上发表的文章[10]中提出新共识算法 PHANTOM,借助有向无环图 DAG 实现并行的同时保障安全性,采用 GHOST 进行交易排序,并引入了连通度极限值概念对 DAG 区块链中的节点进行分类并排序,限定了网络安全容忍度的同时保障了并行出块。

正如共识机制为中本聪最长链提供保障,是区块链一项核心技术,现存的种种排序机制也是支撑 DAG 区块链技术的重要一环。类似于中本聪最长链的共识机制有 PoW、PoS、DPoS 等,DAG 区块链的排序机制也有很多种,其中具有代表性的有 Witness、Tangle、GHOST 等。

2 CLV 泛化分析

2.1 基本思想

2.1.1 DAG 区块链中集合定义

不同于中本聪最长链具有单链的数据结构,DAG 区块链有着特殊的网状数据结构,这也决定了 DAG 数据结构区块链中每个节点都存在几组数据集。这些数据集在 CLV 值 k 选取的时候起到重要的作用。在 DAG 区块链中,每个块会确认它的矿工创建此块时知道的块。对于创世块为 Genesis 的 DAG 区块链 G 中的任意区块 X 而言,均有与之相关的几组数据集:

(1) $past(X)$:由 X 直接或间接引用的区块集合,一般在块 X 之前创建。

(2) $future(X)$:直接或间接引用块 X 的区块集合,一般在块 X 之后创建。

(3) $anticone(X)$:与块 X 之间顺序不明确的区块集合。

(4) $tips(X)$:DAG 区块链 X 中的叶子节点集合,这些块没有被其他区块所引用。

本文 CLV 值 k 的选取以上述数据集为基础,其中 $anticone(X)$ 对 CLV 选取尤为重要。

2.1.2 CLV 选取规则

在 PHANTOM 中对 k-集群的定义如下:

定义:给定一个 DAG, $G=(C,E)$, $\forall B \in S$: $|anticone(B) \cap S \leqslant k|$,则子集 $S \subseteq C$ 被称为 k-集群。

对于一个 DAG 区块链,CLV 值 k 的选取要通过 MSC_k 选取,MSC_k 就是要找到整个 DAG 网络中最大子图 S^* 子图,并且保证 S^* 子图中的任一区块 X 的 $anticone(X)$ 在 S^* 中的数量不超过 k。MSC_k 选取算法如下:

Maximum k-Cluster SubsetDAG(MSC_k)

输入:DAG G=(C,E)。

输出:A subset S^*。

$\subset$C of maximum size, s.t. $|anticone(X) \cap S^*| \leqslant k$ for all $B \in S^*$.

在区块链网络中,认为诚实区块的矿工群体拥有大部分算力,并且只会考虑最大的 CLV 值 k-集群,所以满足条件的区块代表合作节点正确挖出的区块。矿工对 CLV 值选取区分出的红色蓝色区块分别代表网络中的诚实区块与非诚实区块。所以在 CLV 值 k 的选取过程中,将网络中所有节点区分为蓝色区块和红色区块,除蓝色区块外所有区块即为红色区块。DAG 交易排序过程与红蓝色区块的判定密切相关,蓝色区块作为诚实区块排序在红色区块前。蓝色区块选择算法如算法 1 所示。

算法 1 蓝色区块选择算法

输入:G-a block DAG, k-the propagation parameter。

输出:$Blue_k(G)$-the dense-set of G。

```
1.  function Sele-Blue(G,k)
2.  if G == {genesis} then
3.  return {genesis}
4.  for B ∈ tips(G) do
5.  Blue_k(B) ← Sele-Blue(past(B),k)
6.  B_max arg max{ |Blue_k(B): |B ∈ tips(G)}
7.  Blue_k(G) ← Blue_k(B_max) ∪ {B_max}
8.  for B ∈ anticone(B_max) do in some topological ordering
9.  if |anticone(B) ∩ Blue_k(G)| ≤ k
10. then add B to Blue_k(G)
11.  return Blue_k(G)
```

根据 CLV 值选取方法,引出本文要探讨分析的问题:中本聪最长链及 DAG 区块链中的 CLV 值 k 存在什么样的关系;CLV 值 k 与网络安全容忍度和网络吞吐量又有何种联系。本文分别探讨连通度极限值 CLV 在中本聪最长链与特殊 DAG 区块链中的取值,探索 CLV 值在 DAG 区块链中的泛化规则。

2.2 中本聪最长链与 CLV

在比特币这种中本聪最长链的区块链网络中,比

特币的区块的创建间隔大约为 10 min,速率很低,但是安全性有极强的保障。中本聪最长链的结构如图 2 所示,其中黑色为最长链共识时所抛弃的区块,也被称为孤块。分析在中本聪最长链中 CLV 取值。

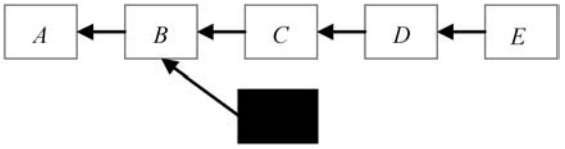


图 2　中本聪最长链结构

对此结构进行分析可知,所有区块的 *anticone* 均为空集,即:

$$anticone(A) = anticone(B) = \cdots = anticone(E) = \varnothing$$

所以 CLV 值 k 有且仅有一个值为 0,即全部区块均可视为蓝色区块。结果可以验证 k 的选取是合理的,因为在中本聪最长链算法中连接到网络中的所有区块(除被抛弃的孤块以外)均被验证为安全诚实的。中本聪区块链中这一确定性、特殊性的为 0 取值引发对 DAG 区块链中 CLV 值探索与关注,对 DAG 网络中的 CLV 值 k 进行泛化分析。

2.3　CLV 在 DAG 区块链中的泛化

PHANTOM 协议中首次引入连通度极限值,在此背景下,探究在中本聪最长链与 DAG 区块链中连通度极限值 CLV 的情况,计算并分析较为典型数据结构中值的情况,分析与网络吞吐量、延迟度存在的关系。DAG 区块链由于其结构多样,存在形式也千变万化,本文选取几种典型的 DAG 数据结构进行 CLV 值分析。

(1) 当区块链网络的数据结构为标准二叉树时,形式如图 3 所示,对 CLV 值 k 的取值进行分析。

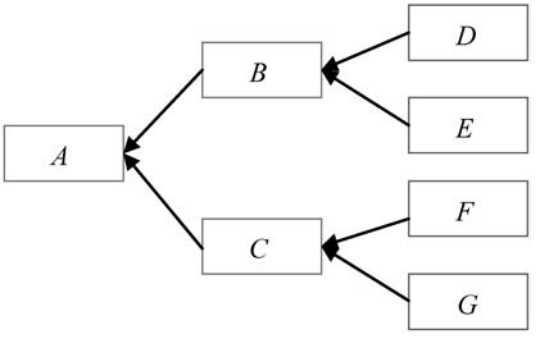


图 3　标准二叉树结构

此时求解所有区块的 *anticone*,结果如下:

$$anticone(A) = \varnothing$$

$$anticone(B) = \{C,F,G\}$$

$$anticone(C) = \{B,D,E\}$$

$$anticone(D) = \{E,C,F,G\}$$

$$anticone(E) = \{D,C,F,G\}$$

$$anticone(F) = \{G,B,D,E\}$$

$$anticone(G) = \{F,B,D,E\}$$

此时,分析 k 可取的值:

当 $k=1$ 时,只有区块 A 为蓝色区块,MSC_1 为由蓝色块 A 构成的子图。

当 $k=2$ 时,区块 A、B、C 为蓝色区块,MSC_2 为由蓝色块 A、B、C 构成的子图。

当 $k=3$ 时,两个分支分别取任意两个叶子节点为红色区块,其余所有节点均可构成蓝色块的最大 3-聚簇的子图 MSC_3。例如,可以取区块 B 分支下的 D 叶子节点及 C 分支下的 F 叶子节点为红色区块,蓝色块为块 A、B、C、E、G 构成的集合。

此数据结构的区块链网络模型属于 DAG 的图形区块链的一种特殊情况,分析此时的 k 取值。

此时 k 的设置为 1 显然不合理,因为此时只有创世区块为诚实节点,其他都被视为恶意节点。

k 值设置为 2 时,所有叶子节点在这种排序方案下均被视为恶意节点,此时根据排序规则所有叶子节点的顺序将排在最后,是合理的。

当 k 设置为 3 时,诚实区块与恶意节点的取值在这个少数区块链的网络就存在多种方案,有明显的不确定性,排序方案也必然存在多种,具有不确定性。此时虽然排序是可以进行的,但是显然不是最优的 k 取值。

综上所述,k 为 2 时是当前数据结构最为合理的 k 参数取值。验证此数据结构下的区块节点增多时,k 为 2 时同样合理。此时所有叶子节点均没有被验证,且也只验证唯一的父节点,将其视为恶意节点进而排序靠后是合理的。

但是这种标准二叉树的数据结构在 DAG 区块链项目中由于其节点数量指数型增长,不具有收敛性,很难达到最终一致性,所以实际 DAG 区块链项目中很少被应用,本文仅用于对 CLV 值分析。

根据上述分析,得知 CLV 在 DAG 区块链网络的数据结构为二叉树时取值为 2。而此时,根据网络的结构知网络的延迟会随区块链中的节点数量增多而增大,网络安全性也是随之降低,而网络节点数量的变化不会对 CLV 产生影响。

(2) Hashgraph 实际上是一种数据结构和共识算法,由以色列耶路撒冷希伯来大学学者提出。Hashgraph 能为分布式 App 提供高效、公平、安全的基础设施,高吞吐量和异步拜占庭容错的特点使得 Hashgraph 在公链和私有链领域都有潜在的使用价值,当区块链网络的数据结构为 Hashgraph 时,形式如图 4 所示,对 k 参数取值进行分析。

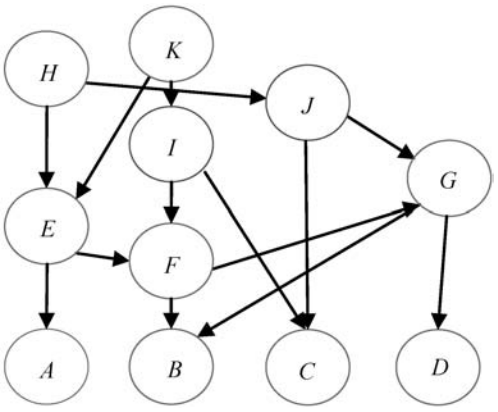


图 4 Hashgraph 区块链网络结构

求解此结构下的所有区块的 *anticone*,结果如下:

$anticone(A)=\varnothing$

$anticone(B)=\varnothing$

$anticone(C)=\varnothing$

$anticone(D)=\varnothing$

$anticone(E)=\{I,J\}$

$anticone(F)=\{J\}$

$anticone(G)=\{I\}$

$anticone(H)=\{K,I,F\}$

$anticone(I)=\{H,E,J,G\}$

$anticone(J)=\{K,I,F,E\}$

$anticone(K)=\{H,J\}$

由于 I、J 的 *anticone* 包含的区块数量较多,证明 I、J 在整个网络中的连通度较低,所以将 I、J 先判定为红色区块。当 I、J 为红色区块时,此时 k 取 2,MSC_2 为由蓝色块 A、B、C、D、E、F、G、H、K 构成的子图。网络数据结构为 Hashgraph 时,对 k 的取值分析较为复杂。创世节点并不是唯一节点,所以每个节点 X 的 $past(X)$ 与 $future(X)$ 会比其他 DAG 区块链多,$anticone(X)$ 集合较小,k 值相对而言取值均会偏小。但仍需根据具体情况分析,同时 Hashgraph 也有自己独立的一套排序方案。

根据上述分析,得知 CLV 在 DAG 区块链网络的数据结构为此种 Hashgraph 时取值为 2,而此时 CLV 会随网络节点数量的变化需视具体结构情况而判定,与网络中节点之间的连通性相关。但可以分析出,CLV 越大时,网络中连通度低的节点越多,网络安全性越低。

(3) 当区块链网络的数据结构为普通 DAG 结构时,形式如图 5 所示,这种 DAG 区块链网络应用最为广泛,因其他 DAG 区块链数据结构都基于此改进探索,因此也最具有研究意义。本文对其 CLV 值 k 参数进行分析需考虑节点构成网络的具体结构。

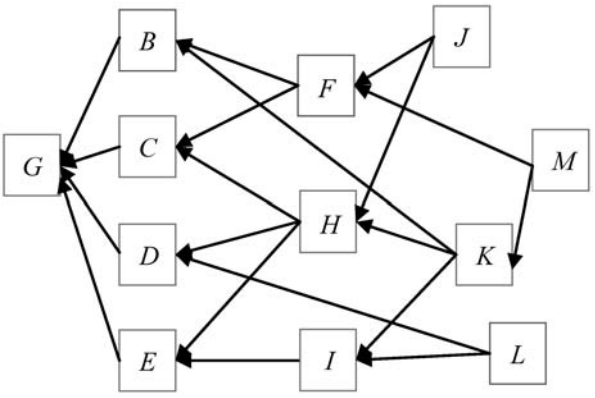


图 5 Hashgraph 区块链结构

求解此结构下的所有区块的 *anticone*,结果如下:

$anticone(G)=\varnothing$

$anticone(B)=\{C,D,H,E,I,L\}$

$anticone(C)=\{B,D,E,I,L\}$

$anticone(D)=\{B,F,C,E,I\}$

$anticone(E)=\{B,D,C,F\}$

$anticone(F)=\{D,H,K,E,I,L\}$

$anticone(H)=\{B,F,I,L\}$

$anticone(I)=\{B,F,J,C,D,H\}$

$anticone(J)=\{M,K,I,L\}$

$anticone(M)=\{J,L\}$

$anticone(L)=\{B,F,J,C,H,K,M\}$

$anticone(K)=\{F,J,L\}$

由于 B、F、L 的 *anticone* 包含的区块数量较多,证明 B、F、L 在整个网络中的连通度较低,所以将 B、F、L 先判定为红色区块。此时 k 可以判定为 3,MSC_3 为由蓝色块 G、B、C、D、E、F、G、H、K 构成的子图。

根据上述对 CLV 值 k 在 DAG 区块链中的取值泛化分析,我们可以分析得出相应结论:连通度取值远小于整个网络节点总数。CLV 值 k 取值的大小与网络的安全性紧密相关,根据具体案例分析可以推断 CLV 值 k 越大,网络安全稳定性会越差。综上所述,CLV 值 k 与网络吞吐量、网络延迟等因素存在关系。

3 优化方案

3.1 CLV 值优化选择

根据上述对 CLV 值 k 在 DAG 区块链网络中的泛化分析得知,CLV 的取值与 DAG 区块链网络中的一些因素有着密切联系。具体而言这些因素包括 DAG 区块链中的网络延迟、网络安全性等。由于标准二叉树结构和 Hashgraph 的 DAG 网络结构有明确的应用限制,所以本文针对普通 DAG 区块链结构进行优化。此种结构的 DAG 区块链是基础研究,更具有代表性,而

且应用更为广泛。

由于区块的创建频率服从泊松分布,因此对任意一个在时刻 t 创建的区块 X,在时间间隔$[t-D_{\max},t+D_{\max}]$中创建出的其他区块数量最多为 $k(D_{\max},\delta)$,且其概率至少为 $1-\delta$。已知泊松分布概率函数如式(1)所示。

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda} \tag{1}$$

根据上述分析,对 DAG 区块链网络中的 CLV 值 k 而言,整理关系如式(2)所示。

$$k(D_{\max},\delta):=\min\left\{\hat{k}\in N:\left(\sum_{j=\hat{k}+1}^{\infty}e^{-2\cdot D_{\max}\cdot\lambda}\cdot\frac{(2\cdot D_{\max}\cdot\lambda)^j}{j!}\right)<\delta\right\} \tag{2}$$

此关系式将 k、$D_{\max}$、δ 几个参数联系起来,其中 δ 用来设定描述安全阈值,关系式中通过选择较小的 δ 值收紧安全域。且须注意关系式中的 λ 值不能无限增大,会影响整个网络的吞吐量,以致出现网络拥堵,这里的目的是为同时创建的区块数量设计一个上限。

针对关系式,需要在大的安全边界和协议的快速收敛之间进行取舍。对 $D_{\max}$过高的估计会导致 k 的增大,使交易结算的等待时间变长。CLV 值 k 与安全阈值成反比,安全域越小,CLV 越大,为保证安全性,在 CLV 选取过程中要保证选取尽可能小的 k 值。

根据上述对 CLV 值 k 与安全阈值与网络延迟关系的分析,所以平衡三者之间的关系成了一个重要的问题。进化计算是基于进化论思想而产生的求解全局最优化问题的一种新型优化算法,可解决函数求极值问题,求出最佳参数。这种算法在诸多方面的应用较为广泛,比如模式识别、图像处理、人工智能等。结合进化计算的思想可构造函数f_{itness}如式(3)所示。

$$f_{\text{itness}}=v_{\text{alue}_1}\times\delta+v_{\text{alue}_2}\times D_{\max} \tag{3}$$

式中:v_{alue_1}和 v_{alue_2}分别为 δ 和 $D_{\max}$对应的权值。式(3)中的安全阈值 δ 与网络延迟 $D_{\max}$均与 CLV 值 k 存在上述公式的关系,通过调整相应权值找到f_{itness}的最小值,找寻到三者之间一个平衡,此时的 CLV 值 k 就认为在安全性与性能之间达到了平衡。

3.2　沙漏区块 CLV 及安全性分析

在 DAG 区块链结构中,存在一种特殊形式的区块。若某个区块 X 的 *anticone* 内没有蓝色区块,则其过去集中的所有蓝色区块在其过去集以外的蓝色区块之前,称之为沙漏特性。根据该引理,G 的蓝色区块一定在 B 的过去集、将来集、B 自身当中。因此所有的蓝色区块组成的形状就像一个沙漏,而 B 就是沙漏中间直径最小的两头相连的孔径。故将 B 称为沙漏区块。

可以推定,沙漏区块的过去集里的蓝色区块在排序时一定在沙漏区块过去集以外的区块之前。图 6 为以区块 K 为沙漏区块的 DAG 区块链网络。

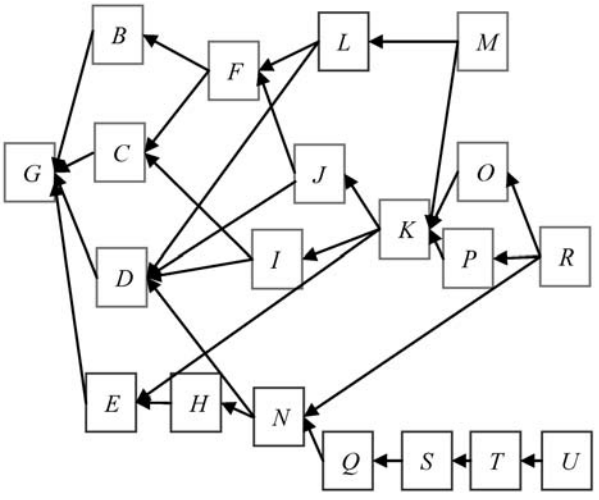


图 6　K 为沙漏区块的 DAG 区块链结构

求解此结构下的所有区块的 *anticone*,结果如下:

$anticone(G)=\varnothing$

$anticone(B)=\{C,D,H,E,I,L\}$

$anticone(C)=\{B,D,E,I,L\}$

$anticone(D)=\{B,F,C,E,I\}$

$anticone(E)=\{B,D,C,F\}$

$anticone(F)=\{D,H,K,E,I,L\}$

$anticone(H)=\{B,F,I,L\}$

$anticone(I)=\{B,F,J,C,D,H\}$

$anticone(J)=\{M,K,I,L\}$

$anticone(M)=\{J,L\}$

$anticone(L)=\{B,F,J,C,H,K,M\}$

$anticone(K)=\{F,J,L\}$

此 DAG 数据结构区块链网络中延迟参数为 $k=3$,其中 L、E、H、N、Q、S、T、U 区块为红色区块,其余所有区块的 *anticone* 均不超过延迟参数 $k=3$ 的蓝色区块,即剩余区块均为蓝色区块,可视为诚实区块。在这个区块链网络中我们应当注意的一个区块是区块 K,它的 *anticone* 中没有任何蓝色区块,则区块 K 为沙漏区块。根据引理,排序时 K 的过去集中的所有蓝色区块在 K 的过去集以外的蓝色区块之前,即 $Blue_{\text{past}}(K)=\{G,B,F,J,C,I,D,E\}$集合中的区块排序在 $Blue_{\text{other}}(K)=\{M,O,P,R\}$集合中的区块之前。

当网络中存在沙漏区块时,它所验证的诚实区块以及验证它的诚实区块,即与此区块有链接的区块相对较多,所以沙漏区块在网络中的安全性较强。

当出现沙漏区块时,整个 DAG 区块链网络受到影响。当出现沙漏区块情况时,整个网络的诚实区块出现集中化情况,大部分红色区块单独分离在外面一个链上,而所有蓝色区块集中围绕在区块 K 的周围。若

网络中的新区块 V 验证叶子节点中的 U 节点,使用 CLV 值 k 判定的方式认定新生成的区块 V 仍为红色区块。

所以,沙漏区块的出现将导致整个 DAG 区块链网络稳定性较差,在 DAG 区块链网络中设定沙漏区块的 CLV 值 k 区间时应该尽量避免新生成的区块出现“蜂拥而至”的情况,以免出现“孤链”影响整个网络的安全性。

4 结 语

DAG 区块链技术是当前区块链解决扩容问题的重要方案之一,本文确认了 CLV 在中本聪最长链中的值为 0,研究分析了 CLV 在 DAG 区块链中的泛化问题。CLV 为 0 时,区块链网络的并行度差,但是可以很好地保证网络的安全性;当网络为 DAG 区块链时,连通度取值非零,此时网络具有高并发性,但是安全上受到一定的影响。

本文给出了 CLV 与网络的安全阈值、网络延迟等具体存在的关系,得出网络安全阈值与连通极限值成反比,网络延迟与 CLV 成正比关系。在此三者之间关系的基础上,本文探讨了 DAG 区块链网络保障 CLV 取值在安全阈值与网络延迟之间的平衡问题,并引用进化计算中的方法给出了关系式,帮助安全性与延迟度之间的合理优化。

随着区块链中用户节点、交易积压度、计算复杂度等方面的增加甚至爆发,扩容成为区块链技术的核心技术问题。目前区块链的扩容方案主要从三个不同层(Layer)解决,DAG 区块链技术正是从第一层(Layer1)即改变区块链本身架构来解决扩容问题的重要方案。由于 DAG 是有向无环网状图的数据结构,而非普通区块链的块链式有序结构,所以在区块链中确定交易排序成为 DAG 区块链中的核心问题。连通度极限值在 DAG 区块链中能够有效反映出网络结构连通情况,探究 DAG 区块链中的连通度极限值与网络吞吐量、延迟度的相关关系,可以进一步规范确定交易排序方案,为解决区块链的扩容问题奠定基础。

DAG 区块链技术极大程度提高了区块链的并发度,有效地解决了采用中本聪最长链的区块链中存在孤块的问题,是当前一项重要的区块链技术。DAG 区块链技术虽然有效地提高了并发度,但安全性面临着挑战。CLV 在 DAG 区块链技术中的提出,对 DAG 区块链技术中的排序方案做了重要的理论补充。此外,CLV 值对 DAG 区块链网络中的交易确认时长、应用场景等其他方面存在的影响,值得进一步研究探索。

参 考 文 献

[1] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system [EB/OL]. (2008 - 08 - 22). http://bitcoin.org/bitcoin.pdf.

[2] 中国共产党新闻网. 习近平:把区块链作为核心技术自主创新重要突破口加快推动区块链技术和产业创新发展[EB/OL]. (2019 - 10 - 26). http://cpc.people.com.cn/n1/2019/1026/c64094-31421707.html.

[3] Javarone M, Wright C. From bitcoin to bitcoin-cash: A network analysis[EB]. arXiv:1804.02350,2018.

[4] Luu L, Narayanan V, Zheng C, et al. A secure sharding protocol for open blockchains[C]//2016 ACM SIGSAC Conference on Computer and Communications Security, 2016: 17 - 30.

[5] Chen J, Micali S. Algorand: A secure and efficient distributed ledger[J]. Theoretical Computer Science,2019,777(3): 155 - 183.

[6] Raiden foundation. Raiden network whitepaper[EB/OL]. (2018 - 05 - 11). http://raiden.network.

[7] Poon J, Dryja T. The bitcoin lighting network: Scalable off-chain instant payments[EB/OL]. (2016 - 01 - 14). https://lightning.network/lightning-network-paper.pdf.

[8] Lerner S. Dagcoin: A cryptocurrency without blocks[EB/OL]. (2015 - 10 - 15). https://bitslog.com/2015/09/11/dagcoin/.

[9] Sompolinsky Y, Wyborski S. PHANTOM GHOSTDAG: A scalable generalization of Nakamoto consensus[C]//3rd ACM Conference on Advances in Financial Technologies, 2021:57 - 70.

[10] 俞学劢. 区块链的 4 大核心技术[J]. 金卡工程,2016(10):9 - 14.

[11] 袁勇,倪晓春,曾帅,等. 区块链共识算法的发展现状与展望[J]. 自动化学报,2018,44(11):2011 - 2022.

[12] 赵艳杰. 基于区块链的物联网信息安全传输与存储研究[D]. 长沙:湖南师范大学,2018.

[13] Panagiotakos G. Proof of work[M]//Encyclopedia of Cryptography and Security. Springer,2011:1 - 3.

[14] Kiayias A, Russell A, David B. Ouroboros: A provably secure proof-of-stake blockchain protocol[C]//Annual International Cryptology Conference,2017:357 - 388.

[15] Lamport L, Shostak R, Marshall P. The byzantine generals problem[J]. ACM Transactions on Program Languages and Systems,1982,4(3):382 - 401.

[16] 韩璇,刘亚敏. 区块链技术中的共识机制研究[J]. 信息网络安全,2017(9):147 - 152.

[17] 高政风,郑继来,汤舒扬,等. 基于 DAG 的分布式账本共识机制研究[J]. 软件学报,2020,31(4):1124 - 1142.

[18] Li C, Li P, Zhou D, et al. Scaling Nakamoto consensus to thousands of transactions per second[EB]. arXiv: 1805.03870,2018.