

基于交易记录特征的自私挖矿检测方案

王贺立, 闫巧

(深圳大学计算机与软件学院, 广东 深圳 518060)

摘要: 自私挖矿攻击是一种挖矿策略, 存在基于工作量证明 (PoW, proof of work) 机制的区块链中。自私矿工通过延迟区块广播的时机来截取其他矿工的挖矿奖励, 以此获得比正常情况下更多的收益。这种攻击对工作量证明机制的激励相容性造成了破坏。近年来, 自私挖矿这一攻击行为被研究者从许多角度进行了研究和分析。然而这一挖矿策略利用了工作量证明机制区块链在网络延迟方面存在的缺陷, 使得采用这一挖矿策略的矿工在行为上与正常挖矿行为没有明显的区别, 导致当前缺少可以有效检测出自私挖矿行为的方案。因此, 提出了一种自私挖矿的检测方案, 该检测方案创新性地利用了区块链中区块的高度和区块中的交易记录特征, 可以做到对网络中的自私挖矿行为进行实时检测, 有一定的实用价值。所提检测方案为生成的新区块定义了一个状态值, 这个状态值与区块中包含的交易数量、支付给矿工的交易所需费用等数据有关。根据这些特征值间的数学关系, 可以判断出该区块是否来自自私矿工。通过仿真实验测试了攻击者在不同算力下, 该检测方案的检测情况。结果表明, 所提检测方案在判断区块是否来自自私挖矿时, 有86.02%以上的检测准确率, 可以有效地对自私挖矿产生的区块进行检测。

关键词: 区块链; 自私挖矿; 网络安全; 比特币

中图分类号: TP391

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2023025

Selfish mining detection scheme based on the characters of transactions

WANG Heli, YAN Qiao

College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

Abstract: Selfish mining is an attack strategy in Proof-of-Work based blockchains, where attackers withhold their mined blocks to intercept the awards of other honest miners, resulting in higher profitability. This attack undermines the incentive compatibility of Proof-of-Work. Although there are various studies from researchers in recent years, there is no effective detection scheme as the vulnerability of blockchain in propagation delay makes it challenging to

收稿日期: 2022-07-31; 修回日期: 2022-12-03

通信作者: 闫巧, yanq@szu.edu.cn

基金项目: 国家自然科学基金 (61976142); 深圳市科技计划项目 (JCYJ20210324093609025)

Foundation Items: The National Natural Science Foundation of China (61976142), Shenzhen Science and Technology Plan Project (JCYJ20210324093609025)

引用格式: 王贺立, 闫巧. 基于交易记录特征的自私挖矿检测方案[J]. 网络与信息安全学报, 2023, 9(2): 104-114.

Citation Format: WANG H L, YAN Q. Selfish mining detection scheme based on the characters of transactions[J]. Chinese Journal of Network and Information Security, 2023, 9(2): 104-114.

distinguish the attackers from honest miners. A scheme for selfish mining detection was introduced. In this scheme, a state value was assigned to the new block based on transaction amount and the transaction fee of the block. By analyzing the relationship between state value and transaction characteristics, we can determine if the block was mined by a selfish miner. The scheme is verified by an experiment with an accuracy of 86.02%.

Keywords: blockchain, selfish mining, network security, bitcoin

0 引言

工作量证明 (PoW, proof of work) 机制这种共识协议是激励相容的。激励相容是一个博弈论中的概念。若一个机制是激励相容的, 则说明当所有的参与者在按照机制诚实地公开任何被要求的信息时, 所有人都能趋向效用最大化, 即矿工的期望收益与其所拥有的算力成正比。在这样的前提下, 大的矿池不会为矿工带来收益上的优势, 矿工没有动力相互联合起来形成一个更大的矿池。因此, 一个正常规模的矿池对系统来说是无害的, 因为矿池不会自发地扩大。对于被广泛使用的 PoW 共识机制的区块链来说, 其安全性会受到来自拥有高算力比例的矿池或矿工的威胁。在最理想的情况下, 当矿工或矿池拥有的算力在全区块链网络算力中的占比为 σ 时, 该矿工或矿池获得的区块奖励在全区块链网络中的占比也为 σ 。

但是, 现有的研究指出, 自私矿工可以通过自私挖矿攻击^[1]的方式获得额外的挖矿收益。在自私挖矿攻击中, 自私矿工会在区块链中主动制造和维护一条属于自己的分叉链, 并根据公链长度和自己私链长度等因素选择性地延迟发布区块。由于 PoW 机制尚未对这种恶意延迟发布的区块进行检测, 当自私矿工发布的私链长度大于公链长度时, 自私矿工便能利用这一缺陷截取到其他正常矿工的区块收益。

1 相关研究

近年来, 自私挖矿这一攻击行为已经被研究者从多个角度进行了研究和分析。与广为人知的日食攻击^[2]相似, 自私挖矿也需要让自身的算力在整个网络中的占比达到某个阈值才可以营利。这个阈值在 Eyal 和 Sirer 最初提出的自私挖矿策略中是 33%^[1], 之后, 随着关于营利阈值和策略研究的进行, 这一阈值下降到了 25%左右^[3-6]。

目前也有关于自私挖矿过程分析的研究^[7-9], 但关于防御和检测自私挖矿的研究却相对较少。在自私挖矿的检测和防御相关的研究中, Heilman 等^[10]提出了被称为“新区块优先”的防御性策略, 这种策略可以将自私挖矿营利的算力要求从 25%提高到 32%。这种策略需要在区块中单独写入一个被称为“不可伪造时间戳”的字段, 矿工可以根据这个字段轻松判断出一个区块是否被恶意延迟广播, 然后选择接受那些具有较新时间戳的区块。这个策略虽然可以很好地防御自私挖矿, 但劣势在于需要额外增加一个“不可伪造时间戳”的管理单位, 同时矿工在接收到区块信息时需要做额外的验证工作, 这两个劣势对于区块链系统来说是不可接受的, 因此这一策略虽然提高了自私矿工的挖矿成本, 但并不能从根本上解决问题, 而且不能防御时间戳修改的情况。Solat 等^[11]提出了“零区块”策略来防御自私挖矿, 这个策略基于一个简单思路: 每个矿工在挖到新区块后, 必须在规定的时间内将其广播出去, 否则不会被接受。这个策略虽然在技术上是有效的, 但在实际应用中很难被矿工所接受。Chicarino 等^[12]提出了一种启发式的方案, 根据区块链网络的分叉率^[13]采用机器学习的方式来检测自私挖矿攻击, 这种方案在自私矿工拥有较低算力占比时拥有极高的准确率, 但随着自私矿工的算力占比上升, 检测的准确率则随之下降。此外, 这种方法不能被实时应用于区块链网络, 缺乏实用性。

为了弥补这些不足, 本文提出了一种基于区块交易记录特征的自私挖矿攻击检测方法, 使用这种方法, 可以有效地检测来自自私矿工的区块, 检测准确率达到 86.02%以上。

2 检测方案

2.1 自私挖矿原理

在文献[1]中, Eyal 和 Sirer 描述了自私挖矿

的最原始过程。自私矿工首先会从公链上分叉并维护一条与公链相对应的“私链”(简单来说,自私矿工接下来总是在这一条私链上进行挖矿,并将挖出的区块暂时不公开),然后根据自私挖矿策略来决定该区块的公开时机。初始状态下,公链和私链是完全相同的链,所有矿工开始在各自己的链上挖矿。随后,诚实矿工若先挖到一个区块,诚实矿工会将这个区块立即广播,那么自私矿工收到这一消息后,根据领先区块的数量和使用的策略,选择是否公开自己的私链;但是若自私矿工先挖到区块,这个区块则不会被立即广播公开,它会被其加入自私矿工的私链。若诚实矿工接下来挖到一个新的区块,此时公链的高度或者长度则追上了自私矿工的私链,那么自私矿工会选择立即公布自己的私链,企图让网络选择这一私链作为公链。假设自私矿工的算力占全网算力的比例为 α ,诚实矿工的算力比例则为 $1-\alpha$,进一步假设当区块链网络同时存在两个广播中的区块消息时,诚实矿工中会选择并接受自私矿工区块的矿工比例为 γ ,根据这一场景,图1给出了存在自私挖矿时,自私矿工领先区块数量的马尔可夫状态转移过程。图中的状态表示自私矿工私链领先公链的区块的数量,而当私链和公链的长度相同时,这一状态根据转移前的状态又分为状态 s_0 和状态 s'_0 ,对应的马尔可夫状态转移过程如下。

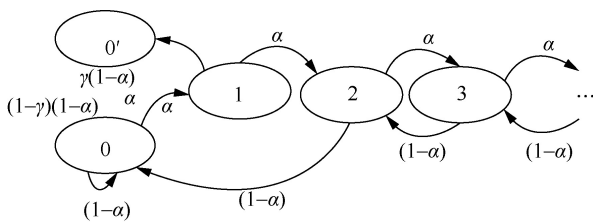


图1 领先区块数量的马尔可夫链过程
Figure 1 The Markov chain process for the amount of leading blocks

1) s_0 : 此时为初始状态,自私矿工在自己私链上挖矿,诚实矿工在公链上挖矿,此时自私矿工的私链和区块链网络公链的长度相同,即自私矿工领先于诚实矿工的区块数量为0。若自私矿工挖到一个区块,则马尔可夫状态转移到状态 s_1 (即领先公链1个区块),这一转移概率与算力比值相同,为 $1-\alpha$,更新后系统状态保

持不变。

2) s_1 : 这一状态表示自私矿工的私链的长度领先公链1个区块。若自私矿工选择不公开私链,并又挖到一个新区块,则状态会被转移为到状态 s_2 (即领先公链2个区块),这一转移概率同样为 α ;若此时诚实矿工挖到了1个区块,则自私矿工的状态会回到状态 s'_0 ,概率为 $1-\alpha$ 。

3) s'_0 : 这一状态表示自私矿工的私链长度从领先于公链1个区块转变为不领先的状态。当自私矿工的私链领先公链的区块数量为1个时,自私矿工在私链上挖矿,诚实矿工在公链挖矿。若诚实矿工挖到1个新区块并广播,自私矿工在接收到这一区块广播的消息后,会立即广播其私链上的区块,此时区块链网络中会同时存在两个长度相同的链,区块链网络需要根据全网其他节点对这两条链的确认情况选择其中一条。接下来会出现3种情况:① 自私矿工挖到下一个区块,私链会被接受为公链,自私矿工获得这两个区块的区块奖励,私链长度与公链相同,自私矿工的状态回到 s_0 ,这一概率为 α ;② 部分诚实矿工在自私矿工的私链上挖矿,然后挖到1个新区块,那么自私矿工和这部分矿工获得了这1个区块的奖励,自私矿工状态也会回到状态 s_0 ,这一概率为 $(1-\alpha)\gamma$;③ 另一部分诚实矿工继续在公链上挖矿并成功先挖到一个新区块,那么这些矿工将获得2个区块的奖励,自私矿工状态回到状态 s'_0 ,这一概率相对较低,概率为 $(1-\alpha)(1-\gamma)$ 。

4) s_2 : 这一状态表示自私矿工的私链领先于公链的区块数量为2。在这一状态下,如果自私矿工率先挖到1个新区块,状态会转移到状态 s_3 ,这一概率为 α ;但如果此时诚实矿工率先挖到一个区块并将其广播,自私矿工接收到这一消息后会选择立即广播其私链,由于其长度领先于公链,自私矿工的私链成为新的公链,自私矿工可以获得2个区块的区块奖励,这一概率为 $1-\alpha$,此时自私矿工会回到 s_0 的初始状态。

5) $s_n (n > 2)$: 这一状态表示自私矿工的私链领先于公链的区块数量为 n 。如果自私矿工率先挖到1个区块,状态转移会到状态 s_{n+1} ,这一概率为 α 。在 $n > 2$ 的情况下,如果是诚实矿工率先

挖到新的区块，那么自私矿工领先区块的数量减少为 $n-1$ ，所以状态相应回退到前一个状态 s_{n-1} ，这一概率为 $1-\alpha$ 。

2.2 自私挖矿的关键变量

在上述过程中，可以发现，自私矿池的风险只存在于状态 s_1 ，也就是领先区块数量只有 1 个时，诚实矿工可能挖到区块后与其竞争；而当领先区块数量超过 1 个后就不再存在收益损失的风险。因此，自私挖矿对矿池的算力阈值有一定的要求。在大多数自私挖矿文献中^[1,14-17]，矿池相对收益是重点研究的对象，其定义如下。

$$R_p = \frac{r_p}{r_p + r_0} \quad (1)$$

其中， R_p 表示成为研究对象的自私矿工的收益， r_0 代表除研究对象外的其他所有矿工的收益，包括其他自私矿工和诚实矿工。在自私挖矿的研究中，通常使用 γ 来表示在自私矿工进行一次区块的广播之后，正常矿工选择加入自私矿池在全网中所占比例；使用 α 表示矿工所拥有的算力占全网算力的比例。根据前面描述的自私挖矿过程可以知道， γ 的值会决定自私矿工能够营利所要求的算力比 α 。根据 Eyal 的研究，它们之间的关系如下。

$$\frac{1-\gamma}{3-2\gamma} < \alpha < 1/2 \quad (2)$$

在理想情况 $\gamma=100\%$ 时，即存在竞争时，所有诚实的矿工都跟随自私矿工的区块继续挖矿的情况下，无论自私矿工矿池的算力是多少总是能够营利（收益超过自私矿工诚实挖矿）。当 $\gamma=50\%$ ，自私矿工的算力需要达到 25% 才能通过自私挖矿获得额外收益。而当 $\gamma=0$ 时，即没有任何诚实矿工跟随自私矿工挖矿，若要获得额外收益，算力的要求则是至少大于 1/3。值得注意的是，相对收益曲线斜率会随着矿池规模的增加而增加，说明自私矿池中矿工的期望收益会随着更多矿工的加入而不断增加，这反过来会吸引更多矿工加入这个自私矿池。由于算力 α 超过 1/2 后，自私挖矿攻击的性质会变成 51%^[18] 攻击，本文研究的方法不再适合，因此 α 的上限被认为是 1/2。

2.3 区块中交易记录的数据结构

在基于工作量证明机制的区块链中，区块的交易信息通过默克尔树来存储。默克尔树是一种哈希二叉树，其结构如图 2 所示。其除叶子节点外的其他节点，会对其两个子节点再次进行哈希计算，而叶子节点的哈希值则对应其中一个交易。默克尔树的优势在于可以迅速完成数据查询和校验，通常被用于大型离散数据的完整性校验，如版本管理（Git）、分布式数据存储、可信计算以及区块链技术等。这种优势是比特币等采用 UTXO（unspent transaction output）交易记录存储模型的基础。在一个采用 UTXO 模型的区块链中，为了使用默克尔树，其交易记录的数据结构需要保留交易哈希值和时间戳等信息。

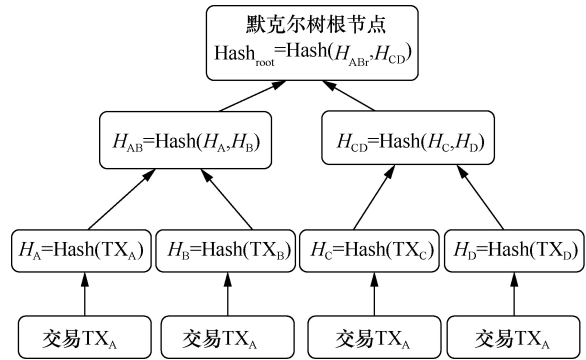


图2 默克尔树的结构
Figure 2 The structure of a Merkle tree

交易记录中的部分字段和区块的数据信息如表 1 所示。

表 1 交易记录中的部分字段和区块的数据信息
Table 1 Data information for the some fields and blocks in transaction records

字段	描述
区块哈希	本区块的哈希值
时间戳	生成本区块的时间戳
区块高度	本区块的区块高度
区块大小	本区块所占用的字节数
交易数量	本区块所包含的交易数量
交易哈希值	某条交易的哈希值
默克尔树根	本区块默克尔树根节点的哈希值
交易哈希值	某条交易的哈希值
交易所在区块	打包某条交易时的区块高度
交易记录大小	本交易所占用字节数
交易费用	完成本交易所支付给矿工的费用

2.4 交易的期望区块高度

所提检测方案中,需要对交易记录所出现的区块高度进行预测,本文将这一预测结果称为交易的期望区块高度,这里的高度指的是与当前区块链高度对比相对高度。

现有对交易的区块高度相关预测方法都与为矿工支付交易费用相关。当区块链中产生一次交易时,为了使得这个交易得以完成,即被矿工打包进区块,需要交易的发起者为矿工支付一定数量的费用,矿工将这个交易打包进区块并成功获得这个区块的记账权后,可以获得额外该区块中所有交易支付的交易费用。可以看出,若交易为矿工支付的费用越高,则该交易越有可能在短时间被矿工打包,完成交易的时间越短,在区块链上的表现则是该交易所出现的区块高度越低;反之若支付的费用越低,则完成的时间越长,出现在区块的高度则越高。对于交易的发起者来说,为了节省所需支付的费用,需要根据支付费用对完成交易对应的时间进行估计,即研究为交易支付的费用与交易所出现的区块高度间的关系。

在对交易记录对应区块高度预测的研究中,feesim方法达到了90%以上的预测成功率^[19],因此,本文方案也使用feesim方法来获得交易期望区块高度。在区块链网络中,有两种情况会导致低费用的交易被延迟打包。

1) 区块的大小不足。2021年7月~2022年7月,每个区块的平均交易数量约为1700笔^[20],如果有超过这一数量的交易等待确认,那么矿工将选择那些支付较高费用的交易,而其余的交易将不得不等待后续的区块。然而,这是由于区块大小的限制,而不是“交易数量”的限制,所以必须考虑到以字节为单位的交易大小。因此,矿工考虑的不是总支付费用,而是考虑“费率”,即费用除以交易数量。

2) 交易费率低于矿工能接受的最低费率。每个矿工通常会设定一个交易必须支付的最低费率;如果费率低于这个最低值,那么即使区块没有满,该交易也不会被打包。如果一笔交易的费率低于大多数矿工的最低费率,那么这笔交易可能在很长一段时期内都不会被确认。

feesim方法基于以下3个假设。

1) 区块的发现和交易的抵达是一个泊松过程。

2) 在考虑交易费率时,矿工对于交易费率(交易费与区块体积的比值)总是贪婪的,费率越高越好。

3) 交易费率受限于区块的最大体积,同时有一个最小值。这两个变量相互独立。

根据比特币网络的相关数据,如当前哈希率、最大区块分布和短期交易抵达率等,结合当前比特币交易池的状态信息,如交易记录大小、交易费用、交易所在区块高度和交易池中未被打包的交易记录的整体信息,对这些数据进行蒙特卡洛仿真,可以很好地预测短期网络的交易费率。此外,feesim方法通过记录在预测的区块数内被确认的交易比例来验证费用估算,结果表明feesim方法的准确率在90%以上。

3 自私挖矿检测方案

虽然已经有很多学者从多个角度研究了自私挖矿攻击,但在检测方面的研究存在缺失,本文提出一种启发式方法用于检测区分区块链中的区块是否产生于自私挖矿。由于自私挖矿等区块截留攻击^[21]是通过延迟发布区块来实现的,自私矿工和正常矿工在除开算力之外的差别并不明显,对于检测者来说,很难确定某一个矿工是否扣留了区块,也就导致很难直接检测出自私矿工是否存在。2018年,在Monacoin的自私挖矿攻击事件中^[22],自私矿工被发现的原因是其算力过高,最高达到了全网算力的57%。因此需要从新的角度来实现对自私挖矿的检测。

从区块的角度出发来检测这种攻击是一种可行的方案。对于一个公开的区块来说,无论是是否通过自私挖矿产生,其区块高度、时间戳和难度信息等区块元信息都是正常的,无法作为分辨是否合法的依据。然而,区块中的交易信息却可以作为一个有效的特征。

在使用非匿名地址的公开区块链中,所有交易信息可以被公开查询,因此表1中的所有数据都可以通过某种合法方式获取。根据这些

交易信息，本文提出了一个方案来检测和判断某个区块是否通过自私挖矿产生。为了更好地描述这一方案，表2列举了检测方案所使用的符号，这里的高度指的是与当前区块链高度对比相对高度。

表2 检测方案所使用的符号

Table 2 The notations used by the detection scheme

符号	符号说明
n	被检测链中领先于当前公链的总区块数
A_j	被检测链中某个交易的字节大小
A_n	被检测链中的总交易字节大小
p_i	第 i 个被检测区块中的总交易数
q	正常区块中的总交易数
H_h	正常区块的高度
H_{si}	被检测区块的相对高度
E_{Ts_j}	被检测区块交易 E_{Ts_j} 的期望区块高度
E_{Th_j}	正常区块中交易 E_{Th_j} 的期望区块高度
ST_h	正常链的状态值
ST_{si}	被检测链的状态值, $i=1,2,\dots,n$

根据前文关于自私挖矿攻击过程描述和关于区块中交易信息的讨论，对于产生于自私挖矿的区块来说，其中的交易信息存在以下特点。

1) 自私矿工在打包交易时会挑选支付费用较高的交易。

2) 交易数量较少。自私矿工为了使得自私区块被大多数节点确认，需要加快区块信息在网络中的传播速度。这只能通过减少区块大小的方式来实现，因此自私矿工会使区块中的交易数量尽可能减少（需要注意在实际上，交易数量至少为1，该交易用于记录对矿工的区块奖励）。

这两个特点会导致区块中交易的平均支付费用升高。因此，在存在自私挖矿的区块链中使用 feesim 方法会导致交易的期望区块高度降低。也就是说，在 feesim 方法的训练过程中，训练数据集中自私挖矿产生的区块会影响预测的结果，因此对用来训练的数据进行筛选，用正常数据得到的结果作为预测的参考标准。为了得到更多的可靠数据，根据需要在区块链中挑选出一个诚实矿工作为参考的对象（假设这个诚实矿工值得信赖，永不会使用自私挖矿策略）。这个诚实矿工并不难挑选。基于以上描述，本文为正常链 h 定义一个状态值。

$$ST_h = H_h - \frac{\sum_{j=1}^q E(Th_j)}{q} \quad (3)$$

可以看出，在正常情况下，根据 feesim 方法的预测结果，在高于 90% 的情况下， $ST_h \approx 0$ 。同样，为被检测链 s 也定义一个状态值。

$$ST_s = \sum_{i=1}^n H_{si} - \frac{\sum_{j=1}^p E(Ts_j)}{p}, i=1,2,\dots,n \quad (4)$$

在前文关于自私挖矿的原理中提到过，自私矿工的私链领先公链的区块数量可能存在于一个或多个，但从式(4)可以发现，自私矿工私链领先公链的长度越长，自私矿工的状态值越大。值得注意的是，若自私矿工选择不在区块中打包交易记录，那么区块中仅存在一条交易记录用于记录自私矿工的区块奖励，此时这条交易记录的交易费用为 0，feesim 方法将会认为这条交易记录不会被打包，因此预测区块高度为无穷大，导致被检测链的状态值小于 0。

基于以上分析，自私挖矿检测如算法 1 所示，总结来说，对于一个区块是否通过自私挖矿产生的检测过程和判断标准如下。

1) 在区块链网络中，只有在较短的时间间隔内检测到了两个或者两个以上数量的来自不同矿工的区块广播信息时，才触发检测机制。该时间间隔取决于区块链网络，如在与 PoW 相关的区块链网络中，一个广播平均需要约 10 s 才能在全网中传播^[23]，因此设置的检测间隔需要考虑这些因素。

2) 当被检测链的区块中存在交易记录时，如果 $ST_h < ST_s$ ，则认为这条链上的区块来自自私矿工。

3) 当被检测链的区块中不存在其他交易记录时， $ST_s < 0$ ，可以认为这条链的区块来自自私矿工。

4) 当检测出现 $ST_h = ST_s$ 的结果时，先计算出被检测链区块中的平均交易大小 A/n ，再通过比较该平均交易大小和 p 值的大小进行判断。若 $A/n < q$ ，则说明被检测链中的交易较小，有更大可能来自自私挖矿。反之则认为该区块正常。

算法 1 自私挖矿检测

- 1) Init
- 2) while not S and H do
- 3) Detect

```

4) end while
5) //当前网络中同时存在两个区块广播 S 和 H
6) if  $ST_h < ST_s$  then
7)     reject S
8) else
9)      $A = 0, n = \text{Amount}(S), q = \text{Size}(H)$ 
10)    for  $T$  in  $S$  do
11)         $A += \text{Size}(T)$ 
12)    end for
13)    if  $A/n < q$  then
14)        reject S
15)    end if
16) end if
17) goto Init

```

4 仿真实验

4.1 环境参数设置

为了验证所提方案,本节进行了仿真实验。实验时区块链系统的运行环境中,对区块验证采用的哈希算法使用 Bitcoin Core 的标准 SHA-256d 算法实现^[24]。SHA-256d 算法会对 SHA-256 哈希的结果额外进行一次哈希,其原理由式(5)表示^[25]。

$$\text{SHA}_{256d}(x) = \text{SHA}_{256}(\text{SHA}_{256}(x)) \quad (5)$$

SHA_{256d} 的双重哈希方式可以有效地避免长度扩展攻击^[26]。由于实验与交易有关,本文的区块链网络额外加入了对交易的支持。实验的挖矿平台使用的操作系统为 Ubuntu-20.04-server-amd64,内核版本为 5.11.7, CPU 型号为 i7-8600k,内存大小为 32 GB。挖矿使用的设备是 NVIDIA2070 显卡,实验过程中的驱动版本为 460.67。根据测试,使用 SHA-256d 算法挖矿时,该显卡的总算力稳定在 40 MH/s 左右,即每秒可以进行 40 000 000 次哈希计算。实验过程中,模拟了 1 000 个诚实挖矿节点,诚实矿工的总算力在节点间平均分配。根据 BTC.COM 对 2021 年比特币网络的算力分布统计^[27],包括独立矿工在内,共分为 17 个矿池,结果如图 3 所示。

根据图 3 的结果,将实验中的 1 000 个诚实节点随机自由组合形成 17 个矿池,最终各个矿池的算力占诚实矿工总算力的比例与图中矿池的情况

对应。本文实验采用的难度调整算法同样参考的是 Bitcoin Core。比特币中难度的变化由式(6)计算得到

$$\text{new_difficulty} = \frac{D_{\text{init}}}{\text{target}} \quad (6)$$

$$\text{target} = \text{current_target} \cdot \frac{\text{ActualTime}}{\text{ExpectedTime}} \quad (7)$$

其中, D_{init} 表示初始化阶段的默认难度,被写入创世区块中。出于对编码过程的考虑,在程序中通过调整 target 的值实现对难度的控制。在比特币中,每挖出 2 016 个区块后进行难度调整,每个区块的期望挖矿时间为 10 min,因此 ExpectedTime=20 160 min。在实验中,为了获得一个合理的难度值,难度调整的周期会加快,每挖出 18 个区块后进行一个难度调整,每个区块的期望挖矿时间同样为 10 min。根据测试,当区块链网络的难度设置为 0.3 时,区块的挖矿时间约为 10 min,满足实验的要求。由于区块链的难度调整不会对实验结果产生影响,因此在接下来的实验过程中,将难度值固定为 0.3,不再进行调整。

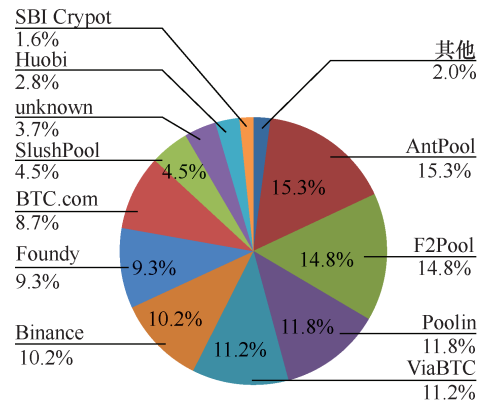


图 3 2021 年比特币网络的算力分布统计
Figure 3 The computational power distribution statistics of Bitcoin in 2021

4.2 实验设计

根据本文方案的算法,详细的检测流程如图 4 所示,具体实现设计如下。

1) 初始状态,在网络中创建一个观测者节点,这个节点不参与挖矿,同时定义一个消息计数的参数,这个参数用于记录在区块链网络中同时存在的新区块广播信息。每当观察节点收到一个新区块广播消息后,消息计数值需要加 1;若 10 s 内没有再收到新的区块信息,消息计数归 0。在实验中,将时间间隔设置为 10 s,因此在 10 s

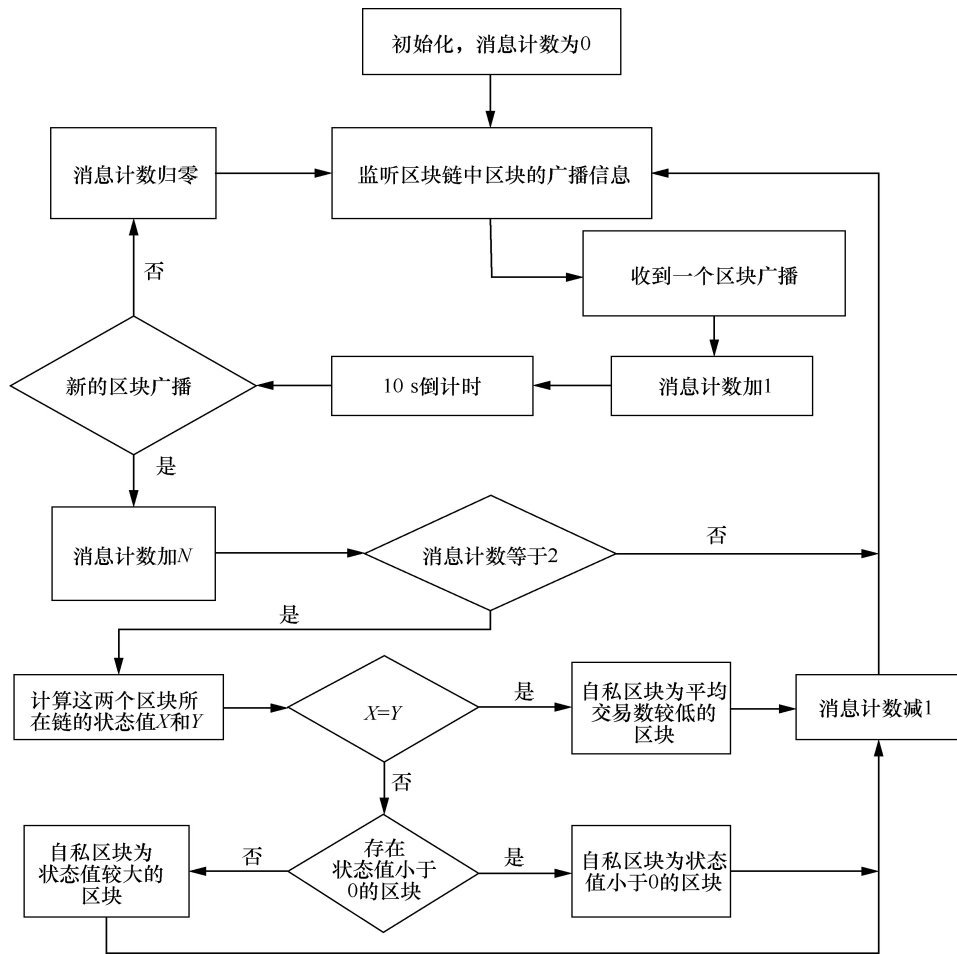


图4 检测流程

Figure 4 The detection process

内没有新的区块广播消息后,可以认为这个新的区块来自诚实矿工,因为这不符之前所描述的关于自私挖矿的行为特征。

2) 假如观测节点在接收到新消息的 10 s 内再次收到了一个新的区块广播消息,则需要通过算法进行判断。如果同时收到了多个广播,消息计数会加上相应的数量,当结果大于 2 时,说明同时接收到了多个独立的新区块,这种结果在正常情况下不会发生,因此不做判断,理论上观测节点的消息计数会变为 2,可以根据算法对这两个区块的来源进行判断,并拒绝接受被判断为来自自私矿工的区块,但没有足够的证据认为另一个区块来自诚实矿工。

3) 完成判断后,暂时不能将观测节点消息计数清零,因为尚不能判断在 10 s 内是否还会接收到其他新的区块广播信息。因此需要将消息计数

减 1, 然后重复上述过程。

4.3 数据获取

由于使用蒙特卡洛模型来获得交易的期望区块高度,需要合适的交易数据,为了保证训练数据的可靠性,本文实验从比特币网络中截取的部分已有的公开交易记录作为训练数据,本文认为在比特币网络中已有的交易数据是可信的。挑选交易数据时,这些交易所属区块高度为随机平均选择,数据量为 2 048 个区块,最后选择了在比特币区块高度范围为[539 715,541 762]中的区块。经过统计,这 2 048 个区块中共有 3 468 629 条交易记录,大小为 1.946 GB。平均每个区块的交易数量为 1 562,每个交易的平均大小为 788 byte,每个交易支付的平均费用为 18.51 聪^{注1}。其中,

注 1: “聪”是比特币中支付矿工费用时所用的单位,英文为 sat, 1sat=10⁻⁸ 比特币。

将高度位于[539 715, 541 507] 的数据作为 feesim 方法的训练数据, 将余下的 255 个区块中的交易数据作为要被打包的交易, 矿池中存在 10 000 条约 7 MB 未处理的交易记录。本文随机截取了一段对 feesim 方法的训练数

据, 其中交易支付的矿工费用与期望区块高度之间的关系如图 5 所示, 不同颜色的曲线表示交易可能出现的相对高度, 可以明显看出, 每字节所支付的交易费用越高, 期望区块高度越低。

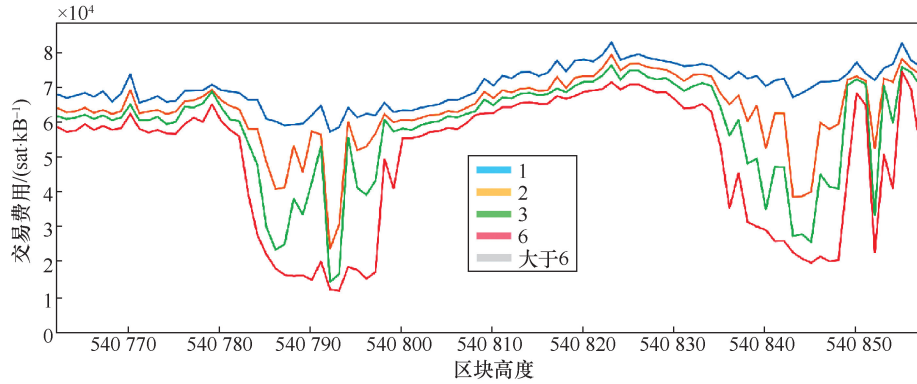


图5 feesim 方法交易支付的矿工费用与期望区块高度之间的关系
Figure 5 The correlation between fee rate and the block expectation height by feesim

5 实验结果

不同自私挖矿算力下的检测成功率如图 6 所示。本文从[0.25,0.5]范围中等差选取了 6 组算力比例进行测试, 可以看出, 在自私矿工算力占比较低时, 检测的成功率也比较低, 在算力占比为 0.25 时, 只有 86.02% 的检测成功率, 当自私矿工算力占比逐渐升高时, 实验的检测成功率也随之上升。

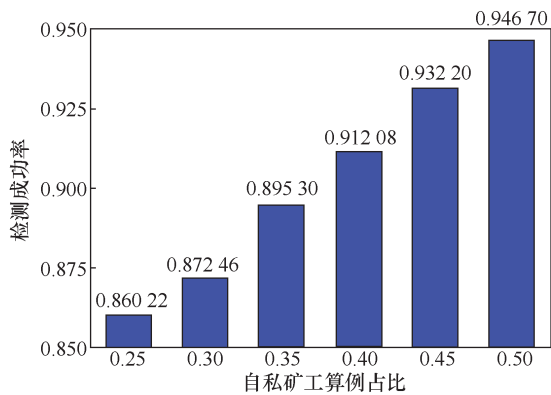


图6 不同自私挖矿算力下的检测成功率
Figure 6 The success rate under different hashrate

这种情况的出现, 符合之前在检测标准中的描述。自私矿工私链的长度越高, 领先的区块数量越多, 为其定义的状态值也越高。因此, 自私

矿工算力占比越高, 其领先于公链的区块数量就可能越多, 因此状态值越大, 这会使得判断更加容易, 从而在实验结果中表现出了检测效率随自私矿工算力的增加而升高的情况。为了改进提出的方案, 可以从交易的期望区块高度预测着手。图 7 展示了实验交易池中, 交易支付费用的分布。红色曲线表示交易支付费用所对应的期望区块高度, 这里用相对高度表示。可以看出, 当支付费用高于 80 sat/byte 时, 期望区块高度为 1, 即该交易可以立即得到打包。但是观察交易支付费用的累计分布可以看出, 有大量的交易所支付的费用低于 20 sat/byte, 这些交易对应的期望区块高度会很高。再观察累计支付的费用, 可以看到, 大多数再打包时会选择高于 75 sat/byte 的交易。但是在提出的方案中, 假设自私挖矿过程区块中的交易数量尽可能少, 同时选择高费用的交易进行打包, 由于自私矿工拥有较高比例的算力, 这就导致区块链中, 实际上可用的算力较低, 那么交易费的实际平均值可能升高, 导致 feesim 方法预测的期望区块高度不准确。由于假设了自私矿工打包的交易数量很少, 虽然这一假设可以适用于多数场景, 但不排除不适用的情况, 这需要对应用本文检测方案的区块链的实际情况具体分析。另外, 自私矿工可能通过打包较多的交易来避

免检测。本文提出的检测方案则拥有实时性和较高的准确率, 自私挖矿产生的区块可以在被全网接受前检测出来, 这保证了诚实矿工的权益, 从而可以成功地防御自私挖矿攻击, 因此可以被部署在实际的区块链网络中, 拥有较高的使用价值。

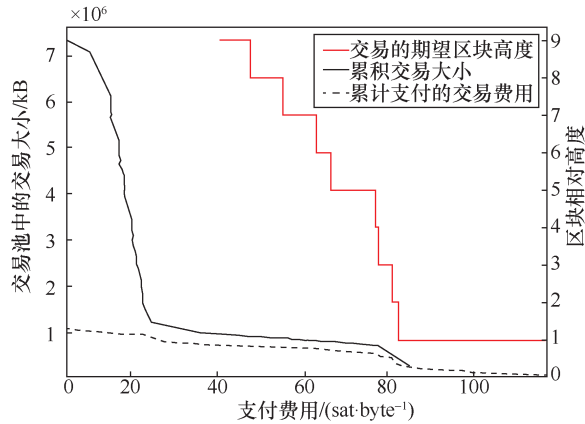


图7 交易支付费用的分布
Figure 7 The distribution of fee rate

6 结束语

本文结合区块中交易记录的特征提出了一种针对私自挖矿的检测方案。该方案为生成的新区块定义了一个状态值, 这个状态值与区块中包含交易数量、支付给矿工的交易费用等数据有关。根据这些特征值间的数学关系, 可以判断出该区块是否来自自私矿工。本文通过仿真实验测试了自私矿工在不同算力下, 该检测方案的检测情况。结果表明, 这一检测方案在判断区块是否来自自私挖矿时, 可达到 86.02% 以上的检测准确率, 可以有效地对私自挖矿产生的区块进行检测。本文提出的检测方案仍存在可以改进之处。首先是关于交易期望高度的改进; 然后由于本文提出的检测方案假设了自私挖矿者打包的交易数量很少, 导致方案在判断拥有正常交易数量的区块时准确度不足, 自私矿工可能利用这一缺陷来逃避检测方案, 为了弥补这一不足, 需要进一步研究自私挖矿与区块中交易记录的关系。

参考文献:

[1] EYAL I, SIRER E G. Majority is not enough: bitcoin mining is vulnerable[C]//Proceedings of International Conference on Financial Cryptography and Data Security. 2014: 436-454.
[2] HEILMAN E, KENDLER A, ZOHAR A, et al. Eclipse attacks on

{Bitcoin's} {peer-to-peer} network[C]//Proceedings of 24th USENIX Security Symposium (USENIX Security 15). 2015: 129-144.
[3] SAPIRSHTEIN A, SOMPOLINSKY Y, ZOHAR A. Optimal selfish mining strategies in bitcoin[C]//Proceedings of International Conference on Financial Cryptography and Data Security. 2016: 515-532.
[4] NAYAK K, KUMAR S, MILLER A, et al. Stubborn mining: generalizing selfish mining and combining with an eclipse attack[C]//Proceedings of 2016 IEEE European Symposium on Security and Privacy (EuroS&P). 2016: 305-320.
[5] DONG X, WU F, FAREE A, et al. Selfholding: a combined attack model using selfish mining with block withholding attack[J]. Computers & Security, 2019, 87:101584.
[6] ALBRECHER H, GOFFARD P O. On the profitability of selfish blockchain mining under consideration of ruin[J]. arXiv preprint arXiv:2010.12577, 2020.
[7] ZUR R B, EYAL I, TAMAR A. Efficient mdp analysis for selfish-mining in blockchains[C]//Proceedings of the 2nd ACM Conference on Advances in Financial Technologies. 2020: 113-131.
[8] MOTLAGH S G, MIŠIĆ J, MIŠIĆ V B. The impact of selfish mining on bitcoin network performance[J]. IEEE Transactions on Network Science and Engineering, 2021, 8(1): 724-735.
[9] LI Q, CHANG Y, WU X, et al. A new theoretical framework of pyramid markov processes for blockchain selfish mining[J]. Journal of Systems Science and Systems Engineering, 2021, 30(6): 667-711.
[10] HEILMAN E. One weird trick to stop selfish miners: Fresh bitcoins, a solution for the honest miner[C]//Proceedings of International Conference on Financial Cryptography and Data Security. 2014: 161-162.
[11] SOLAT S, POTOP-BUTUCARU M. Brief announcement: Zeroblock: Timestamp-free prevention of block-withholding attack in bitcoin[C]//Proceedings of International Symposium on Stabilization, Safety, and Security of Distributed Systems. 2017: 356-360.
[12] CHICARINO V, ALBUQUERQUE C, JESUS E, et al. On the detection of selfish mining and stalker attacks in blockchain networks[J]. Annals of Telecommunications, 2020, 75(3): 143-152.
[13] GERVAIS A, KARAME G O, WÜST K, et al. On the security and performance of proof of work blockchains[C]//Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. 2016: 3-16.
[14] WANG W, HOANG D T, HU P, et al. A survey on consensus mechanisms and mining strategy management in blockchain networks[J]. IEEE Access, 2019, 7: 22328-22370.
[15] GÖBEL J, KEELER H P, KRZESINSKI A E, et al. Bitcoin blockchain dynamics: The selfish-mine strategy in the presence of propagation delay[J]. Performance Evaluation, 2016, 104:23-41.
[16] SAAD M, NJILLA L, KAMHOUA C, et al. Countering selfish mining in blockchains[C]//Proceedings of 2019 International Conference on Computing, Networking and Communications (ICNC). 2019: 360-364.
[17] HOU C, ZHOU M, JI Y, et al. Squirrl: Automating attack analysis on blockchain incentive mechanisms with deep reinforcement learning[J]. arXiv preprint arXiv:1912.01798, 2019.
[18] DECKER C, WATTENHOFER R. Information propagation in the bitcoin network[C]//Proceedings of IEEE P2P 2013. 2013: 1-10.
[19] BITCOINFEEES. Bitcoinfees/feesim: model-based bitcoin fee estimation[EB].
[20] YCHARTS. Bitcoin average transactions per block[EB].
[21] BAG S, RUJ S, SAKURAI K. Bitcoin block withholding attack:

- analysis and mitigation[J]. IEEE Transactions on Information Forensics and Security, 2016, 12(8): 1967-1978.
- [22] GUTTERIDGE D. Japanese cryptocurrency monacoin hit by selfish mining attack[EB].
- [23] GÖBEL J, KRZESINSKI A E. Increased block size and bitcoin blockchain dynamics[C]//Proceedings of 2017 27th International Telecommunication Networks and Applications Conference (IT-NAC). 2017: 1-6.
- [24] BTC.COM. Poolstate-2021[EB].
- [25] HAYES A S. Cryptocurrency value formation: an empirical study leading to a cost of production model for valuing bitcoin[J]. Telecommunications and Informatics, 2017, 34(7):1308-1321.
- [26] COURTOIS N T, GRAJEK M, NAIK R. Optimizing SHA256 in bitcoin mining[C]//Proceedings of International Conference on Cryptography and Security Systems. 2014: 131-144.
- [27] CORON J S, DODIS Y, MALINAUD C, et al. Merkle-damgård revisited: how to construct a hash function[C]//Proceedings of Annual International Cryptology Conference. 2005: 430-448.

[作者简介]



王贺立（1994- ），男，湖北恩施人，深圳大学硕士生，主要研究方向为区块链安全。



闫巧（1972- ），女，重庆人，深圳大学教授、博士生导师，主要研究方向为网络安全、人工智能和物联网。