



(12) 发明专利申请

(10) 申请公布号 CN 117911028 A

(43) 申请公布日 2024. 04. 19

(21) 申请号 202311689049.8

(22) 申请日 2023.12.08

(71) 申请人 中国人民解放军国防科技大学

地址 410003 湖南省长沙市开福区德雅路
109号(72) 发明人 张经宇 侯星星 郭得科 罗来龙
任棒棒 张千桢

(74) 专利代理机构 北京风雅颂专利代理有限公司 11403

专利代理师 曾志鹏

(51) Int. Cl.

G06Q 20/40 (2012.01)

G06Q 20/38 (2012.01)

G06F 3/06 (2006.01)

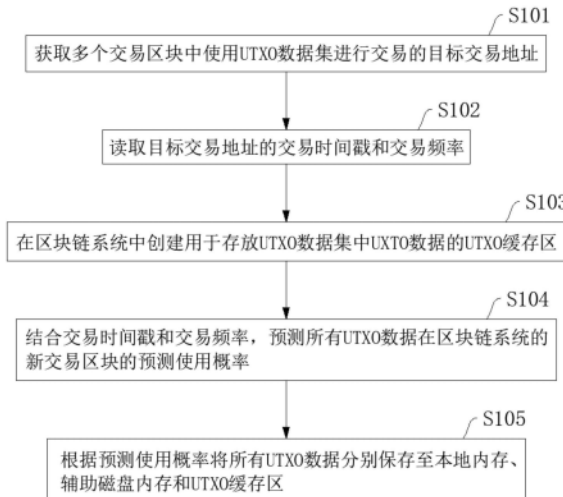
权利要求书4页 说明书14页 附图4页

(54) 发明名称

基于多分区的区块链数据管理优化方法及系统

(57) 摘要

本发明提供一种基于多分区的区块链数据管理优化方法及系统,该方法包括如下步骤:获取多个所述交易区块中使用所述UTXO数据集进行交易的目标交易地址;读取所述目标交易地址的交易时间戳和交易频率;在所述区块链系统中创建用于存放所述UTXO数据集中UTXO数据的UTXO缓存区;结合所述交易时间戳和所述交易频率,预测所有所述UTXO数据在所述区块链系统的新交易区块的预测使用概率;根据所述预测使用概率将所有所述UTXO数据分别保存至所述本地内存、所述辅助磁盘内存和所述UTXO缓存区。本发明具有降低交易数据集数据管理成本的同时可以提升区块验证效率的效果。



1. 一种基于多分区的区块链数据管理优化方法,其特征在于,应用于区块链节点设备,所述节点设备具有本地内存和辅助磁盘内存,所述节点设备用于搭载基于交易的区块链系统,所述区块链系统包含多个交易区块和未花费交易输出数据集,所述方法包括如下步骤:

获取多个所述交易区块中使用所述未花费交易输出数据集进行交易的目标交易地址;

读取所述目标交易地址的交易时间戳和交易频率;

在所述区块链系统中创建用于存放所述未花费交易输出数据集中未花费交易输出数据的未花费交易输出缓存区;

结合所述交易时间戳和所述交易频率,预测所有所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率;

根据所述预测使用概率将所有所述未花费交易输出数据分别保存至所述本地内存、所述辅助磁盘内存和所述未花费交易输出缓存区。

2. 根据权利要求1所述的基于多分区的区块链数据管理优化方法,其特征在于,所述结合所述交易时间戳和所述交易频率,预测所有所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率包括如下步骤:

基于所述交易时间戳计算每两个交易时间之间的交易时间间隔;

结合所述交易时间间隔和所述交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率;

根据所述交易时间间隔或所述交易频率将所有所述未花费交易输出数据进行排序;

结合所述交易时间间隔阈值和所述最优交易频率,按照排序顺序逐个预测所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率。

3. 根据权利要求2所述的基于多分区的区块链数据管理优化方法,其特征在于,所述结合所述交易时间间隔和所述交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率包括如下步骤:

按照所述交易时间间隔的大小将所有所述交易时间间隔进行排序;

利用多个预设的分位数对排序后的所有所述交易时间间隔进行分组,得到多种分组方式,每种所述分组方式中均将所述交易时间间隔分为两个数据组;

通过统计分析法计算各个所述分组方式中数据量最小的预选数据组所对应的未花费交易输出数据使用分布;

结合所述未花费交易输出数据使用分布和对应的所述交易频率,判断各个所述预选数据组中是否存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系;

若存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系,则将对对应的预选数据组作为目标数据组,并将所述目标数据组对应的所述分位数作为交易时间间隔阈值;

测试各个所述交易频率的数据命中率,并将所述数据命中率最高的所述交易频率作为最优交易频率。

4. 根据权利要求2所述的基于多分区的区块链数据管理优化方法,其特征在于,所述结合所述交易时间间隔阈值和所述最优交易频率,按照排序顺序逐个预测所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率包括如下步骤:

按照排序顺序逐个判断所述未花费交易输出数据的所述交易频率是否为所述最优交

易频率,且所述未花费交易输出数据的所述交易时间间隔是否小于所述交易时间间隔阈值;

若所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔小于所述交易时间间隔阈值,则判断所述未花费交易输出数据的数据缓存长度是否小于所述未花费交易输出缓存区的缓存空间;

若所述未花费交易输出数据的数据缓存长度小于所述未花费交易输出缓存区的缓存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第一预测使用概率;

若所述未花费交易输出数据的数据缓存长度大于等于所述未花费交易输出缓存区的缓存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第二预测使用概率,所述第一预测使用概率大于所述第二预测使用概率;

若所述交易频率大于预设的频率阈值,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值,则判断所述未花费交易输出数据的磁盘储存长度是否小于所述辅助磁盘内存的内存空间,所述频率阈值大于所述最优交易频率;

若所述未花费交易输出数据的磁盘储存长度小于所述辅助磁盘内存的内存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第三预测使用概率,所述第三预测使用概率小于所述第二预测使用概率;

若所述未花费交易输出数据的磁盘储存长度大于等于所述辅助磁盘内存的内存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为所述第二预测使用概率;

若所述交易频率小于所述频率阈值且不为所述最优交易频率,
或,

若所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为所述第二预测使用概率。

5. 根据权利要求4所述的基于多分区的区块链数据管理优化方法,其特征在于,所述根据所述预测使用概率将所有所述未花费交易输出数据分别保存至所述本地内存、所述辅助磁盘内存和所述未花费交易输出缓存区包括如下步骤:

判断所述未花费交易输出数据的所述预测使用概率为所述第一预测使用概率、所述第二预测使用概率或所述第三预测使用概率;

若所述未花费交易输出数据为所述第一预测使用概率,则将所述未花费交易输出数据保存至所述未花费交易输出缓存区;

若所述未花费交易输出数据为所述第二预测使用概率,则将所述未花费交易输出数据保存至所述本地内存;

若所述未花费交易输出数据为所述第三预测使用概率,则将所述未花费交易输出数据保存至所述辅助磁盘内存。

6. 一种基于多分区的区块链数据管理优化系统,其特征在于,搭载于区块链节点设备,所述节点设备具有本地内存和辅助磁盘内存,所述节点设备用于搭载基于交易的区块链系统,所述区块链系统包含多个交易区块和未花费交易输出数据集,所述系统包括:

地址获取子系统,用于获取多个所述交易区块中使用所述未花费交易输出数据集进行交易的目标交易地址;

地址读取子系统,用于读取所述目标交易地址的交易时间戳和交易频率;

缓存创建子系统,用于在所述区块链系统中创建用于存放所述未花费交易输出数据集中未花费交易输出数据的未花费交易输出缓存区;

概率预测子系统,用于结合所述交易时间戳和所述交易频率,预测所有所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率;

分区管理子系统,用于根据所述预测使用概率将所有所述未花费交易输出数据分别保存至所述本地内存、所述辅助磁盘内存和所述未花费交易输出缓存区。

7. 根据权利要求6所述的基于多分区的区块链数据管理优化系统,其特征在于,所述概率预测子系统包括:

间隔计算模块,用于基于所述交易时间戳计算每两个交易时间之间的交易时间间隔;

频率计算模块,用于结合所述交易时间间隔和所述交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率;

数据排序模块,用于根据所述交易时间间隔或所述交易频率将所有所述未花费交易输出数据进行排序;

概率预测模块,用于结合所述交易时间间隔阈值和所述最优交易频率,按照排序顺序逐个预测所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率。

8. 根据权利要求7所述的基于多分区的区块链数据管理优化系统,其特征在于,所述频率计算模块包括:

数据排序单元,用于按照所述交易时间间隔的大小将所有所述交易时间间隔进行排序;

数据分组单元,用于利用多个预设的分位数对排序后的所有所述交易时间间隔进行分组,得到多种分组方式,每种所述分组方式中均将所述交易时间间隔分为两个数据组;

数据计算单元,用于通过统计分析法计算各个所述分组方式中数据量最小的预选数据组所对应的未花费交易输出数据使用分布;

关联判断单元,用于结合所述未花费交易输出数据使用分布和对应的所述交易频率,判断各个所述预选数据组中是否存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系;

间隔阈值确认单元,当所述关联判断单元的判断结果为存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系时,用于将对应的预选数据组作为目标数据组,并将所述目标数据组对应的所述分位数作为交易时间间隔阈值;

最优频率确认单元,用于测试各个所述交易频率的数据命中率,并将所述数据命中率最高的所述交易频率作为最优交易频率。

9. 根据权利要求7所述的基于多分区的区块链数据管理优化系统,其特征在于,所述概率预测模块包括:

第一条件判断单元,用于按照排序顺序逐个判断所述未花费交易输出数据的所述交易频率是否为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔是否小于所述交易时间间隔阈值;

缓存判断单元,当所述第一条件判断单元的判断结果为所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔小于所述交易时间间隔阈值时,用于判断所述未花费交易输出数据的数据缓存长度是否小于所述未花费交易输出缓存区的缓存空间;

第一概率预测单元,当所述缓存判断单元的判断结果为所述未花费交易输出数据的数据缓存长度小于所述未花费交易输出缓存区的缓存空间时,用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第一预测使用概率;

第二概率预测单元,当所述缓存判断单元的判断结果为所述未花费交易输出数据的数据缓存长度大于等于所述未花费交易输出缓存区的缓存空间时,用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第二预测使用概率,所述第一预测使用概率大于所述第二预测使用概率;

第二条件判断单元,当所述第一条件判断单元的判断结果为所述交易频率大于预设的频率阈值,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值时,用于判断所述未花费交易输出数据的磁盘储存长度是否小于所述辅助磁盘内存的内存空间,所述频率阈值大于所述最优交易频率;

第三概率预测单元,当所述第二条件判断单元的判断结果为所述未花费交易输出数据的磁盘储存长度小于所述辅助磁盘内存的内存空间时,用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第三预测使用概率,所述第三预测使用概率小于所述第二预测使用概率;

当所述第二条件判断单元的判断结果为所述未花费交易输出数据的磁盘储存长度大于等于所述辅助磁盘内存的内存空间时,

或,

当所述第一条件判断单元的判断结果为所述交易频率小于所述频率阈值且不为所述最优交易频率时,

或,

当所述第一条件判断单元的判断结果为所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值时,所述第二概率预测单元用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为所述第二预测使用概率。

10. 根据权利要求9所述的基于多分区的区块链数据管理优化系统,其特征在于,所述分区管理子系统包括:

概率判断模块,用于判断所述未花费交易输出数据的所述预测使用概率为所述第一预测使用概率、所述第二预测使用概率或所述第三预测使用概率;

第一分配模块,当所述概率判断模块的判断结果为所述未花费交易输出数据为所述第一预测使用概率时,用于将所述未花费交易输出数据保存至所述未花费交易输出缓存区;

第二分配模块,当所述概率判断模块的判断结果为所述未花费交易输出数据为所述第二预测使用概率时,用于将所述未花费交易输出数据保存至所述本地内存;

第三分配模块,当所述概率判断模块的判断结果为所述未花费交易输出数据为所述第三预测使用概率时,用于将所述未花费交易输出数据保存至所述辅助磁盘内存。

基于多分区的区块链数据管理优化方法及系统

技术领域

[0001] 本发明属于区块链技术领域,具体是涉及到一种基于多分区的区块链数据管理优化方法及系统。

背景技术

[0002] 区块链技术的持续发展离不开它所特有的去中心化、可溯源、高隐私等特性,区块链技术能灵活运用于其他行业的各种应用场景。区块链的节点可以分为轻节点和全节点。轻节点负责处理和验证一部分交易和区块,不需要存储完整的区块链数据。全节点需要保存完整区块链数据以维护系统运行。特别的,一部分竞争记账权的全节点被称为记帐节点。区块链系统数据的分布式存储让各节点都能够在不通过第三方的情况下进行点对点沟通,但沟通便利实际上是以牺牲的全节点的存储空间作为交换的,全节点需要下载全部数据才能完成独立验证。相应的,随着区块链系统广受欢迎,其系统和全节点存储的数据注定会持续增长。区块链系统包含一个庞大的分布式账本,记帐节点不断记录账本上的数据。所有节点共同构建区块链,用户提出交易后上传到链上进行广播,接收到交易的交易者会对该交易进行交易验证,只有验证通过的交易才会被交易者存放在自己的内存池Mempool中。

[0003] 未花费交易输出(Unspent Transaction Output,UTXO)是交易数据集的一部分。未花费交易输出的数据管理能力、存储方式和占用空间将极大地影响区块链系统的运行效率和验证性能。特别是随着区块链技术的普及,相关的未花费交易输出数据集不断增长,所有存储的数据几乎不能再完整地存储在本地内存中。此时的未花费交易输出交易数据应该如何存储和管理,是区块链系统有待解决的问题。全节点为了保证验证性能,需要加载完整的未花费交易输出集到自己的本地内存中,这将占用了大量的内存空间。目前,随着未花费交易输出集的急剧增加,将整个未花费交易输出集完全存储在内存中的成本很高,所以有时会将部分未花费交易输出存储在低速硬盘中。但由于磁盘访问延迟比本地内存访问长几个数量级,所以需要花费100倍的时间来进行交易验证。

[0004] 区块验证包括版本验证、简单输入验证、签名验证、锁定时间验证等,其中未花费交易输出集访问,ECC(Elliptic Curves Cryptography)解密和SHA-2哈希是最耗时间的三个步骤。评估了这些操作在通用PC上的消耗时间得出,如果整个未花费交易输出集都在本地内存中,验证速度将以ECC解密时间为主,大约是105us,吞吐量可达到10000TPS。但是,如果从磁盘读取未花费交易输出数据库,则访问时间将增加到10毫秒,从而导致验证速度仅为100TPS。两者之间的较大差距影响着区块链节点的高效验证和数据管理。

发明内容

[0005] 本发明提供一种基于多分区的区块链数据管理优化方法及系统,以解决降低交易数据集数据管理成本的同时会降低区块验证效率的问题。

[0006] 第一方面,本发明提供一种基于多分区的区块链数据管理优化方法,应用于区块链节点设备,所述节点设备具有本地内存和辅助磁盘内存,所述节点设备用于搭载基于交

易的区块链系统,所述区块链系统包含多个交易区块和未花费交易输出数据集,所述方法包括如下步骤:

[0007] 获取多个所述交易区块中使用所述未花费交易输出数据集进行交易的目标交易地址;

[0008] 读取所述目标交易地址的交易时间戳和交易频率;

[0009] 在所述区块链系统中创建用于存放所述未花费交易输出数据集中未花费交易输出数据的未花费交易输出缓存区;

[0010] 结合所述交易时间戳和所述交易频率,预测所有所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率;

[0011] 根据所述预测使用概率将所有所述未花费交易输出数据分别保存至所述本地内存、所述辅助磁盘内存和所述未花费交易输出缓存区。

[0012] 可选的,所述结合所述交易时间戳和所述交易频率,预测所有所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率包括如下步骤:

[0013] 基于所述交易时间戳计算每两个交易时间之间的交易时间间隔;

[0014] 结合所述交易时间间隔和所述交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率;

[0015] 根据所述交易时间间隔或所述交易频率将所有所述未花费交易输出数据进行排序;

[0016] 结合所述交易时间间隔阈值和所述最优交易频率,按照排序顺序逐个预测所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率。

[0017] 可选的,所述结合所述交易时间间隔和所述交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率包括如下步骤:

[0018] 按照所述交易时间间隔的大小将所有所述交易时间间隔进行排序;

[0019] 利用多个预设的分位数对排序后的所有所述交易时间间隔进行分组,得到多种分组方式,每种所述分组方式中均将所述交易时间间隔分为两个数据组;

[0020] 通过统计分析法计算各个所述分组方式中数据量最小的预选数据组所对应的未花费交易输出数据使用分布;

[0021] 结合所述未花费交易输出数据使用分布和对应的所述交易频率,判断各个所述预选数据组中是否存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系;

[0022] 若存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系,则将对应的预选数据组作为目标数据组,并将所述目标数据组对应的所述分位数作为交易时间间隔阈值;

[0023] 测试各个所述交易频率的数据命中率,并将所述数据命中率最高的所述交易频率作为最优交易频率。

[0024] 可选的,所述结合所述交易时间间隔阈值和所述最优交易频率,按照排序顺序逐个预测所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率包括如下步骤:

[0025] 按照排序顺序逐个判断所述未花费交易输出数据的所述交易频率是否为所述最

优交易频率,且所述未花费交易输出数据的所述交易时间间隔是否小于所述交易时间间隔阈值;

[0026] 若所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔小于所述交易时间间隔阈值,则判断所述未花费交易输出数据的数据缓存长度是否小于所述未花费交易输出缓存区的缓存空间;

[0027] 若所述未花费交易输出数据的数据缓存长度小于所述未花费交易输出缓存区的缓存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第一预测使用概率;

[0028] 若所述未花费交易输出数据的数据缓存长度大于等于所述未花费交易输出缓存区的缓存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第二预测使用概率,所述第一预测使用概率大于所述第二预测使用概率;

[0029] 若所述交易频率大于预设的频率阈值,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值,则判断所述未花费交易输出数据的磁盘储存长度是否小于所述辅助磁盘内存的内存空间,所述频率阈值大于所述最优交易频率;

[0030] 若所述未花费交易输出数据的磁盘储存长度小于所述辅助磁盘内存的内存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第三预测使用概率,所述第三预测使用概率小于所述第二预测使用概率;

[0031] 若所述未花费交易输出数据的磁盘储存长度大于等于所述辅助磁盘内存的内存空间,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为所述第二预测使用概率;

[0032] 若所述交易频率小于所述频率阈值且不为所述最优交易频率,

[0033] 或,

[0034] 若所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值,则预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为所述第二预测使用概率。

[0035] 可选的,所述根据所述预测使用概率将所有所述未花费交易输出数据分别保存至所述本地内存、所述辅助磁盘内存和所述未花费交易输出缓存区包括如下步骤:

[0036] 判断所述未花费交易输出数据的所述预测使用概率为所述第一预测使用概率、所述第二预测使用概率或所述第三预测使用概率;

[0037] 若所述未花费交易输出数据为所述第一预测使用概率,则将所述未花费交易输出数据保存至所述未花费交易输出缓存区;

[0038] 若所述未花费交易输出数据为所述第二预测使用概率,则将所述未花费交易输出数据保存至所述本地内存;

[0039] 若所述未花费交易输出数据为所述第三预测使用概率,则将所述未花费交易输出数据保存至所述辅助磁盘内存。

[0040] 第二方面,本发明提供一种基于多分区的区块链数据管理优化系统,搭载于区块链节点设备,所述节点设备具有本地内存和辅助磁盘内存,所述节点设备用于搭载基于交易的区块链系统,所述区块链系统包含多个交易区块和未花费交易输出数据集,所述系统包括:

- [0041] 地址获取子系统,用于获取多个所述交易区块中使用所述未花费交易输出数据集进行交易的目标交易地址;
- [0042] 地址读取子系统,用于读取所述目标交易地址的交易时间戳和交易频率;
- [0043] 缓存创建子系统,用于在所述区块链系统中创建用于存放所述未花费交易输出数据集中未花费交易输出数据的未花费交易输出缓存区;
- [0044] 概率预测子系统,用于结合所述交易时间戳和所述交易频率,预测所有所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率;
- [0045] 分区管理子系统,用于根据所述预测使用概率将所有所述未花费交易输出数据分别保存至所述本地内存、所述辅助磁盘内存和所述未花费交易输出缓存区。
- [0046] 可选的,所述概率预测子系统包括:
- [0047] 间隔计算模块,用于基于所述交易时间戳计算每两个交易时间之间的交易时间间隔;
- [0048] 频率计算模块,用于结合所述交易时间间隔和所述交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率;
- [0049] 数据排序模块,用于根据所述交易时间间隔或所述交易频率将所有所述未花费交易输出数据进行排序;
- [0050] 概率预测模块,用于结合所述交易时间间隔阈值和所述最优交易频率,按照排序顺序逐个预测所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率。
- [0051] 可选的,所述频率计算模块包括:
- [0052] 数据排序单元,用于按照所述交易时间间隔的大小将所有所述交易时间间隔进行排序;
- [0053] 数据分组单元,用于利用多个预设的分位数对排序后的所有所述交易时间间隔进行分组,得到多种分组方式,每种所述分组方式中均将所述交易时间间隔分为两个数据组;
- [0054] 数据计算单元,用于通过统计分析法计算各个所述分组方式中数据量最小的预选数据组所对应的未花费交易输出数据使用分布;
- [0055] 关联判断单元,用于结合所述未花费交易输出数据使用分布和对应的所述交易频率,判断各个所述预选数据组中是否存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系;
- [0056] 间隔阈值确认单元,当所述关联判断单元的判断结果为存在所述未花费交易输出数据的使用频率与所述交易频率之间的关联关系时,用于将对应的预选数据组作为目标数据组,并将所述目标数据组对应的所述分位数作为交易时间间隔阈值;
- [0057] 最优频率确认单元,用于测试各个所述交易频率的数据命中率,并将所述数据命中率最高的所述交易频率作为最优交易频率。
- [0058] 可选的,所述概率预测模块包括:
- [0059] 第一条件判断单元,用于按照排序顺序逐个判断所述未花费交易输出数据的所述交易频率是否为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔是否小于所述交易时间间隔阈值;
- [0060] 缓存判断单元,当所述第一条件判断单元的判断结果为所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔小于所述交易时间间隔阈值

时,用于判断所述未花费交易输出数据的数据缓存长度是否小于所述未花费交易输出缓存区的缓存空间;

[0061] 第一概率预测单元,当所述缓存判断单元的判断结果为所述未花费交易输出数据的数据缓存长度小于所述未花费交易输出缓存区的缓存空间时,用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第一预测使用概率;

[0062] 第二概率预测单元,当所述缓存判断单元的判断结果为所述未花费交易输出数据的数据缓存长度大于等于所述未花费交易输出缓存区的缓存空间时,用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第二预测使用概率,所述第一预测使用概率大于所述第二预测使用概率;

[0063] 第二条件判断单元,当所述第一条件判断单元的判断结果为所述交易频率大于预设的频率阈值,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值时,用于判断所述未花费交易输出数据的磁盘储存长度是否小于所述辅助磁盘内存的内存空间,所述频率阈值大于所述最优交易频率;

[0064] 第三概率预测单元,当所述第二条件判断单元的判断结果为所述未花费交易输出数据的磁盘储存长度小于所述辅助磁盘内存的内存空间时,用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为第三预测使用概率,所述第三预测使用概率小于所述第二预测使用概率;

[0065] 当所述第二条件判断单元的判断结果为所述未花费交易输出数据的磁盘储存长度大于等于所述辅助磁盘内存的内存空间时,

[0066] 或,

[0067] 当所述第一条件判断单元的判断结果为所述交易频率小于所述频率阈值且不为所述最优交易频率时,

[0068] 或,

[0069] 当所述第一条件判断单元的判断结果为所述交易频率为所述最优交易频率,且所述未花费交易输出数据的所述交易时间间隔大于等于所述交易时间间隔阈值时,所述第二概率预测单元用于预测对应的所述未花费交易输出数据在所述区块链系统的新交易区块的预测使用概率为所述第二预测使用概率。

[0070] 可选的,所述分区管理子系统包括:

[0071] 概率判断模块,用于判断所述未花费交易输出数据的所述预测使用概率为所述第一预测使用概率、所述第二预测使用概率或所述第三预测使用概率;

[0072] 第一分配模块,当所述概率判断模块的判断结果为所述未花费交易输出数据为所述第一预测使用概率时,用于将所述未花费交易输出数据保存至所述未花费交易输出缓存区;

[0073] 第二分配模块,当所述概率判断模块的判断结果为所述未花费交易输出数据为所述第二预测使用概率时,用于将所述未花费交易输出数据保存至所述本地内存;

[0074] 第三分配模块,当所述概率判断模块的判断结果为所述未花费交易输出数据为所述第三预测使用概率时,用于将所述未花费交易输出数据保存至所述辅助磁盘内存。

[0075] 本发明的有益效果是:

[0076] 本发明利用同一地址下的历史交易的交易间隔时间和交易次数来预测地址相关

未花费交易输出数据会作为交易输出出现在下一个交易区块中的概率,并将概率大的未花费交易输出数据保存在缓存中,用于加速验证这些交易;将概率小的未花费交易输出数据存放在磁盘,用于减少内存占用量。通过这种方式,全节点就不需要频繁访问磁盘中的未花费交易输出数据,能够在减少内存占用量。同时利用一个小的未花费交易输出缓存,以达到加速区块验证时间并实现数据高效管理。

附图说明

- [0077] 图1为本发明中区块链系统的应用场景示意图。
- [0078] 图2为本发明中基于多分区的区块链数据管理优化方法的流程示意图。
- [0079] 图3为本发明中采用统计分析方法分析使用频率与交易频率之间关联关系的分析示意图。
- [0080] 图4为本发明中数据命中率的测试结果示意图。
- [0081] 图5为本发明中实验所采用的未花费交易输出数据年龄分布图。
- [0082] 图6为本发明中输入未花费交易输出相关地址的交易频率分布图。
- [0083] 图7为本发明中输出未花费交易输出相关地址的交易频率分布图。

具体实施方式

[0084] 本发明公开一种基于多分区的区块链数据管理优化方法。该方法应用于区块链节点设备,节点设备具有本地内存和辅助磁盘内存,节点设备用于搭载基于交易的区块链系统,区块链系统包含多个交易区块和未花费交易输出(Unspent Transaction Output, UTXO)数据集。未花费交易输出数据集的数据管理能力、存储方式和占用空间将极大地影响区块链系统的运行效率和验证性能。

[0085] 参照图1,图1为区块链系统的应用场景图,区块链的节点可以分为轻节点和全节点。轻节点负责处理和验证一部分交易和区块,不需要存储完整的区块链数据。全节点需要保存完整区块链数据以维护系统运行。特别的,一部分竞争记账权的全节点被称为记帐节点。区块链系统数据的分布式存储让各节点都能够在不通过第三方的情况下进行点对点沟通,但沟通便利实际上是以牺牲的全节点的存储空间作为交换的,全节点需要下载全部数据才能完成独立验证。

[0086] 参照图2,本发明所公开的基于多分区的区块链数据管理优化方法具体包括如下步骤:

[0087] S101. 获取多个交易区块中使用未花费交易输出数据集进行交易的目标交易地址。

[0088] 其中,未花费交易输出数据广泛应用于基于交易的区块链系统,它和基于账户的交易模型不同。在基于交易的交易模型中,它的一笔交易可以是一个或者多个输入,也可以是一个或者多个输出。每当发起一笔新的交易时,交易提出者需要选择一些之前的未花费交易输出数据作为输入,并使用它们为收款人创建出新的未花费交易输出数据。每个未花费交易输出数据可能属于不同的所有者,未花费交易输出数据可以在后续交易中作为输入继续使用。在传统的基于账户的系统中,系统直接记录每个账户的余额。但是,基于交易的模型获取特定账户余额的方法与此不同。它需要聚合所有与该账户关联的未花费交易输出

数据并将它们累加(使用数字钱包等)以此获取总余额。为了后续方便进行数据集分析,可以下载区块链中最新生成的交易区块,并收集该交易区块所包含的地址的最近五次交易记录。

[0089] S102.读取目标交易地址的交易时间戳和交易频率。

[0090] 其中,在收集到交易区块所包含的地址的最近五次交易记录后,可以读取这些目标交易地址的交易时间戳及该地址的交易频率。进而可以分析未花费交易输出相关地址的交易频率、时间间隔之间的关系。

[0091] S103.在区块链系统中创建用于存放未花费交易输出数据集中未花费交易输出数据的未花费交易输出缓存区。

[0092] 其中,区块链系统包含一个庞大的分布式账本,记帐节点不断记录账本上的数据。所有节点共同构建区块链。用户提出交易后上传到链上进行广播,接收到交易的交易者会对该交易进行交易验证,只有验证通过的交易才会被交易者存放在自己的内存池Mempool(Mempool是未打包交易的临时存储空间)中。同时交易者也在争夺下一个区块的记账权,获胜的交易者将获得创建新区块并打包交易的权利。交易者为了节省时间,通常会直接在Mempool中选择需要打包的交易,毕竟里面的交易已经完成过交易验证。同时,交易者可以从新建的区块中获得交易费和奖励。之后,获胜的交易者将向区块链网络广播新区块。当其他节点收到关于新区块的区块头时,它们将验证并存储这个区块。之后,交易者将在验证区块的基础上继续进行新一轮的交易。

[0093] 其中,区块验证包括版本验证、简单输入验证、签名验证、锁定时间验证等[20],其中未花费交易输出集访问,ECC(Elliptic Curves Cryptography)解密和SHA-2哈希是最耗时间的三个步骤。评估了这些操作在通用PC上的消耗时间得出,如果整个未花费交易输出集都在本地内存中,验证速度将以ECC解密时间为主,大约是105us,吞吐量可达到10000TPS。但是,如果从辅助磁盘读取未花费交易输出数据库,则访问时间将增加到10毫秒,从而导致验证速度仅为100TPS。因此,可以通过创建未花费交易输出缓存区,将少部分极有可能会被使用的未花费交易输出数据预存至未花费交易输出缓存区,从而加速区块验证的速度。

[0094] S104.结合交易时间戳和交易频率,预测所有未花费交易输出数据在区块链系统的新交易区块的预测使用概率。

[0095] 其中,由于未花费交易输出缓存区容量较小,因此只能存放少量的未花费交易输出数据,所以需要结合交易时间戳和交易频率,准确预测所有未花费交易输出数据在区块链系统的新交易区块的预测使用概率,将预测使用概率最高的一部分未花费交易输出数据存放在未花费交易输出缓存区。

[0096] S105.根据预测使用概率将所有未花费交易输出数据分别保存至本地内存、辅助磁盘内存和未花费交易输出缓存区。

[0097] 其中,可以将预测使用概率最高的一部分未花费交易输出数据存放在未花费交易输出缓存区,将预测使用概率次高的一部分未花费交易输出数据存放在本地内存,而剩下的数据则存放在辅助磁盘内存中,通过以上数据分区管理的方法可以加速区块验证时间并实现数据高效管理。

[0098] 本实施方式的实施原理为:

[0099] 本发明利用同一地址下的历史交易的交易间隔时间和交易次数来预测地址相关未花费交易输出数据会作为交易输出出现在下一个交易区块中的概率,并将概率大的未花费交易输出数据保存在缓存中,用于加速验证这些交易;将概率小的未花费交易输出数据存放在磁盘,用于减少内存占用量。通过这种方式,全节点就不需要频繁访问磁盘中的未花费交易输出数据,能够在减少内存占用量。同时利用一个小的未花费交易输出缓存,以达到加速区块验证时间并实现数据高效管理。

[0100] 在其中一种实施方式中,步骤S104具体包括如下步骤:

[0101] 基于交易时间戳计算每两个交易时间之间的交易时间间隔;

[0102] 结合交易时间间隔和交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率;

[0103] 根据交易时间间隔或交易频率将所有未花费交易输出数据进行排序;

[0104] 结合交易时间间隔阈值和最优交易频率,按照排序顺序逐个预测未花费交易输出数据在区块链系统的新交易区块的预测使用概率。

[0105] 在本实施方式中,结合交易时间间隔和交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率这一步骤具体包括如下步骤:

[0106] 按照交易时间间隔的大小将所有交易时间间隔进行排序;

[0107] 利用多个预设的分位数对排序后的所有交易时间间隔进行分组,得到多种分组方式,每种分组方式中均将交易时间间隔分为两个数据组;

[0108] 通过统计分析法计算各个分组方式中数据量最小的预选数据组所对应的未花费交易输出数据使用分布;

[0109] 结合未花费交易输出数据使用分布和对应的交易频率,判断各个预选数据组中是否存在未花费交易输出数据的使用频率与交易频率之间的关联关系;

[0110] 若存在未花费交易输出数据的使用频率与交易频率之间的关联关系,则将对应的预选数据组作为目标数据组,并将目标数据组对应的分位数作为交易时间间隔阈值;

[0111] 测试各个交易频率的数据命中率,并将数据命中率最高的交易频率作为最优交易频率。

[0112] 在本实施方式中,假设最新生成的区块高度为801213的交易区块,收集该区块所包含的地址的目标交易地址最近五次交易的时间戳及该地址的交易频率。按照交易时间间隔的大小将所有交易时间间隔进行排序,利用多个预设的分位数对排序后的所有交易时间间隔进行分组,得到多种分组方式,每种分组方式中均将交易时间间隔分为两个数据组。通过统计分析法计算各个分组方式中数据量最小的预选数据组所对应的未花费交易输出数据使用分布。结合未花费交易输出数据使用分布和对应的交易频率,判断各个预选数据组中是否存在未花费交易输出数据的使用频率与交易频率之间的关联关系。

[0113] 参照图3,图3所示为其中一种分组方式,左侧饼图所展示的时间间隔分布即为两个不同的预选数据组。可以发现最近两笔交易之间时间间隔小于72小时的输入地址占大多数,时间间隔大于72小时的输入地址只占一小部分。从图3的右侧所示可以发现,当时间间隔大于72小时,未花费交易输出数据的使用频率与交易频率之间的关联关系。如果地址交易频率较低就意味着产生新交易的可能性较高,反之亦然。因此,可以将对应的预选数据组作为目标数据组,并将目标数据组对应的分位数作为交易时间间隔阈值。

[0114] 参照图4,考虑到未花费交易输出缓存区的容量大小相对较小,应该提高数据命中率。假设目前有6个区块的未花费交易输出数据,若这些未花费交易输出都遵循一定的地址交易频率分布。在本实施方式中,为测试缓存命中率,使用4种不同交易频率进行4个简单测试。图4说明若将交易频率为1的地址相关未花费交易输出从内存复制到缓存会更有助于提高缓存命中率。

[0115] 在其中一种实施方式中,结合交易时间间隔阈值和最优交易频率,按照排序顺序逐个预测未花费交易输出数据在区块链系统的新交易区块的预测使用概率这一步骤具体包括如下步骤:

[0116] 按照排序顺序逐个判断未花费交易输出数据的交易频率是否为最优交易频率,且未花费交易输出数据的交易时间间隔是否小于交易时间间隔阈值;

[0117] 若交易频率为最优交易频率,且未花费交易输出数据的交易时间间隔小于交易时间间隔阈值,则判断未花费交易输出数据的数据缓存长度是否小于未花费交易输出缓存区的缓存空间;

[0118] 若未花费交易输出数据的数据缓存长度小于未花费交易输出缓存区的缓存空间,则预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第一预测使用概率;

[0119] 若未花费交易输出数据的数据缓存长度大于等于未花费交易输出缓存区的缓存空间,则预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第二预测使用概率,第一预测使用概率大于第二预测使用概率;

[0120] 若交易频率大于预设的频率阈值,且未花费交易输出数据的交易时间间隔大于等于交易时间间隔阈值,则判断未花费交易输出数据的磁盘储存长度是否小于辅助磁盘内存的内存空间,频率阈值大于最优交易频率;

[0121] 若未花费交易输出数据的磁盘储存长度小于辅助磁盘内存的内存空间,则预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第三预测使用概率,第三预测使用概率小于第二预测使用概率;

[0122] 若未花费交易输出数据的磁盘储存长度大于等于辅助磁盘内存的内存空间,则预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第二预测使用概率;

[0123] 若交易频率小于频率阈值且不为最优交易频率,

[0124] 或,

[0125] 若交易频率为最优交易频率,且未花费交易输出数据的交易时间间隔大于等于交易时间间隔阈值,则预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第二预测使用概率。

[0126] 在本实施方式中,步骤S105具体包括如下步骤:

[0127] 判断未花费交易输出数据的预测使用概率为第一预测使用概率、第二预测使用概率或第三预测使用概率;

[0128] 若未花费交易输出数据为第一预测使用概率,则将未花费交易输出数据保存至未花费交易输出缓存区;

[0129] 若未花费交易输出数据为第二预测使用概率,则将未花费交易输出数据保存至本

地内存;

[0130] 若未花费交易输出数据为第三预测使用概率,则将未花费交易输出数据保存至辅助磁盘内存。

[0131] 在本实施方式中,假设最优交易频率为1,未花费交易输出缓存区为cache,本地内存为memory,辅助磁盘内存为disk,交易频率为count,交易时间间隔为time,交易时间间隔阈值为 τ_{time} ,预设的频率阈值为 τ_{count} ,结合上述“结合交易时间间隔阈值和最优交易频率,按照排序顺序逐个预测未花费交易输出数据在区块链系统的新交易区块的预测使用概率”这一步骤的具体实施方式,以及上述步骤S105的具体实施步骤,可以得到以下数据管理优化算法:

	输入: 区块链用户的UTXO数据集: UTXO_list
	输出: 返回UTXO子集热区: cache, memory, disk
	1: 初始化
[0132]	<pre> 2: for utxo in utxo_list 3: if count=1 & time<τ_{time} 4: if cache 长度<cache 空间 5: 存入 cache 6: else 7: 存入 memory 8: elif time>τ_{time} & count > τ_{count} 9: if disk 的长度<disk 空间 10: 存入 disk 11: else 12: 存入 memory 14: else: </pre>
[0133]	15: 存入 memory
	16: end for
	17: 结束

[0134] 节点执行区块验证时的主要成本是交易验证。对于交易验证,它的主要时间花费在脚本验证和未花费交易输出访问。假设在一个新的区块中有N个交易输入,区块验证的时间是 T_{BV} ,那么有以下公式:

$$[0135] \quad T_{BV} = T_{SV} + T_A + T_{others}$$

[0136] 其中 T_{SV} 为脚本验证的总时间, T_A 为未花费交易输出数据访问的总时间。另一个时间 T_{others} ,包含验证工作证明、Merkle树等的时间,可以忽略不计。设新区块中有 P_M 的概率分布在本地图存,则分布在本地图存中的未花费交易输出个数为 N_M' ,且 $N_M' = N * P_M'$, P_M' 为优

化前的 P_M 。同样的,有 $P_D' = 1 - P_M'$ 的概率分布在辅助磁盘内存,那么分布在辅助磁盘内存的未花费交易输出个数为 $N_D' = N * P_D' = N * (1 - P_M')$ 。此时,新区块访问未花费交易输出所需时间为:

$$[0137] \quad T_A' = \sum_{i=1}^{N_M'} t_{AMi} + \sum_{i=1}^{N_D'} t_{ADi}$$

[0138] 其中, t_{AMi} 为第 i 个未花费交易输出访问本地内存所花费的时间, t_{ADi} 为第 i 个未花费交易输出访问辅助磁盘内存所花费的时间。且有 $N = N_M' + N_D'$ 。那我们能够得出含 N 个交易输入的新区块所花费的区块验证时间 T_{SV}' 为:

$$[0139] \quad T_{SV}' = \sum_{i=1}^N t_{svi} + T_A' + T_{others}$$

[0140] 其中 t_{svi} 是新块中第 i 个未花费交易输出的脚本验证时间。

[0141] 基于以上一轮,设新区块中在缓存的未花费交易输出的概率为 P_{cache} ,则其个数 $N_{cache} = N * P_{cache}$,设分布在本地图存中的未花费交易输出个数为 N_M ,设有 P_M 的概率分布在本地图存,则分布在本地图存中的未花费交易输出个数为: $N_M = N * P_M$,有 $P_D = 1 - P_M - P_{cache}$ 的概率分布在辅助磁盘内存,那么分布在辅助磁盘内存的未花费交易输出个数为 $N_D = N * P_D = N * (1 - P_M - P_{cache})$ 。故而针对使用分区优化机制之后的 T_A ,对于任意未花费交易输出,访问时间有以下公式:

$$[0142] \quad T_A = \sum_{i=1}^{N_{cache}} t_{Acachei} + \sum_{i=1}^{N_M} t_{AMi} + \sum_{i=1}^{N_D} t_{ADi}$$

[0143] 其中 $t_{Acachei}$ 为访问未花费交易输出缓存中第 i 个未花费交易输出的所花费的时间。且 $\Delta N_M = N_M' - N_M$, $\Delta N_D = N_D' - N_D$,然后可以得到优化时间差:

$$[0144] \quad \Delta T_A = \sum_{i=1}^{\Delta N_M} t_{AMi} + \sum_{i=1}^{\Delta N_D} t_{ADi} - \sum_{i=1}^{N_{cache}} t_{Acachei}$$

[0145] 可以看出,如果优化机制能够进行合理优化的话,是能够快速缩短区块验证中访问未花费交易输出所花费的时间的。

[0146] 在其中一种实施方式中,通过设计实验证明本发明中所采用的方法的效果。在设计的实验中,初始化了100000个未花费交易输出数据构成原始未花费交易输出数据集。使用不同的数据管理方法作用于相同的未花费交易输出数据集:随机实验是将未花费交易输出随机分配到内存和磁盘中,不配置缓存;LRU/MRU实验是通过LRU/MRU将未花费交易输出存储到磁盘/缓存中;LFU/MFU实验是通过LFU/MFU将未花费交易输出存储到磁盘/缓存中;最后一个Opt实验是通过本文所提出的机制将未花费交易输出存储到磁盘和缓存中。从原始数据集中选择1000个未花费交易输出来生成新的区块。此外,我们设置缓存中存储 N 个未花费交易输出的数据,同时占用 $x\%$ 的内存空间, t_{AMi} 的时间为 $16\mu s$, t_{ADi} 为 $10120\mu s$ 。同时,在下节实验中的访问时间也会以此设置,并将访问缓存的时间设置为 $1\mu s$ 。

[0147] 参照图5,从浏览器中下载相关未花费交易输出年龄分布数据,绝大部分未花费交易输出年龄已超过1周,甚至超过整体未花费交易输出数据集的99%。为更好控制整体数据集大小,在模拟实验生成未花费交易输出数据集时,将生成90%的年龄超过72小时的未花

费交易输出和10%的年龄未超过72小时的未花费交易输出。

[0148] 对于整个区块而言,交易输入和交易输出传达的意义并不一样,交易输入的打包规则更能契合生成新区块的规律,而交易输出的生成规则就更能清楚反馈未花费交易输出数据集内数据的生成情况。参照图6,以区块高度为801213的区块为例,交易输入未花费交易输出的相关地址交易次数为2的占据整个区块近乎一半的数据。在生成新区块时,根据输入未花费交易输出相关地址的交易次数分类打包交易,将未花费交易输出的地址交易次数为1、2-10、11-100、101以上分别打包未花费交易输出集的55%、15%、15%、15%。参照图7,在该区块,输出未花费交易输出相关地址的交易次数在101及以上的占据多半,而相关地址的交易次数为1的输出未花费交易输出占据不超1/5。可以从图6和图7看出新区块生成未花费交易输出的规律。

[0149] 对新生成的区块中进行了缓存命中率测试和未花费交易输出访问时间评估实验。对于每个数据管理方法,实验都有四种内存设置:实验使用70%、60%、50%和40%的内存空间,满内存空间可存储100000个未花费交易输出。为了研究交易数据管理,测试的缓存大小从300到600不等。在缓存命中率测试中,缓存命中率越高,磁盘访问次数越低,表示交易数据集管理越好。实验结果显示,尽管LRU/MRU和LFU/MFU均能取得较好的性能,但本发明所提出的方法在数据集管理方面仍优于上述方法。

[0150] 实验结果还显示,无论在不同的缓存大小和内存使用设置下,本发明所提出的方法在未花费交易输出访问时间方面都能获得更好的性能。结果表明,与LRU&MRU或LRU&MRU相比,本发明所提出的方法可以更好地让未花费交易输出数据集存储在不同的分区。当内存使用率为70%~40%时,本发明提出的方法的未花费交易输出访问时间约为0.12~0.58秒,而使用LRU&MRU和LFU&MRU的未花费交易输出访问时间分别为0.69~1.44秒和0.67~1.40秒。基于以上实验评估与分析,本发明提出的方法实现了高效的数据管理,减少了区块验证所需的未花费交易输出访问时间。

[0151] 本发明还公开一种基于多分区的区块链数据管理优化系统,搭载于区块链节点设备,节点设备具有本地内存和辅助磁盘内存,节点设备用于搭载基于交易的区块链系统,区块链系统包含多个交易区块和未花费交易输出(Unspent Transaction Output,UTXO)数据集,系统包括:

[0152] 地址获取子系统,用于获取多个交易区块中使用未花费交易输出数据集进行交易的目标交易地址;

[0153] 地址读取子系统,用于读取目标交易地址的交易时间戳和交易频率;

[0154] 缓存创建子系统,用于在区块链系统中创建用于存放未花费交易输出数据集中未花费交易输出数据的未花费交易输出缓存区;

[0155] 概率预测子系统,用于结合交易时间戳和交易频率,预测所有未花费交易输出数据在区块链系统的新交易区块的预测使用概率;

[0156] 分区管理子系统,用于根据预测使用概率将所有未花费交易输出数据分别保存至本地内存、辅助磁盘内存和未花费交易输出缓存区。

[0157] 本实施方式的实施原理为:

[0158] 通过程序的调取,上述子系统共同执行前述实施方式中所公开的基于多分区的区块链数据管理优化方法,以实现以下效果:利用同一地址下的历史交易的交易间隔时间和

交易次数来预测地址相关未花费交易输出数据会作为交易输出出现在下一个交易区块中的概率,并将概率大的未花费交易输出数据保存在缓存中,用于加速验证这些交易;将概率小的未花费交易输出数据存放在磁盘,用于减少内存占用量。通过这种方式,全节点就不需要频繁访问磁盘中的未花费交易输出数据,能够在减少内存占用量。同时利用一个小的未花费交易输出缓存,以达到加速区块验证时间并实现数据高效管理。

[0159] 在其中一种实施方式中,概率预测子系统包括:

[0160] 间隔计算模块,用于基于交易时间戳计算每两个交易时间之间的交易时间间隔;

[0161] 频率计算模块,用于结合交易时间间隔和交易频率,并通过统计分析法计算得到交易时间间隔阈值和最优交易频率;

[0162] 数据排序模块,用于根据交易时间间隔或交易频率将所有未花费交易输出数据进行排序;

[0163] 概率预测模块,用于结合交易时间间隔阈值和最优交易频率,按照排序顺序逐个预测未花费交易输出数据在区块链系统的新交易区块的预测使用概率。

[0164] 在其中一种实施方式中,频率计算模块包括:

[0165] 数据排序单元,用于按照交易时间间隔的大小将所有交易时间间隔进行排序;

[0166] 数据分组单元,用于利用多个预设的分位数对排序后的所有交易时间间隔进行分组,得到多种分组方式,每种分组方式中均将交易时间间隔分为两个数据组;

[0167] 数据计算单元,用于通过统计分析法计算各个分组方式中数据量最小的预选数据组所对应的未花费交易输出数据使用分布;

[0168] 关联判断单元,用于结合未花费交易输出数据使用分布和对应的交易频率,判断各个预选数据组中是否存在未花费交易输出数据的使用频率与交易频率之间的关联关系;

[0169] 间隔阈值确认单元,当关联判断单元的判断结果为存在未花费交易输出数据的使用频率与交易频率之间的关联关系时,用于将对应的预选数据组作为目标数据组,并将目标数据组对应的分位数作为交易时间间隔阈值;

[0170] 最优频率确认单元,用于测试各个交易频率的数据命中率,并将数据命中率最高的交易频率作为最优交易频率。

[0171] 在其中一种实施方式中,概率预测模块包括:

[0172] 第一条件判断单元,用于按照排序顺序逐个判断未花费交易输出数据的交易频率是否为最优交易频率,且未花费交易输出数据的交易时间间隔是否小于交易时间间隔阈值;

[0173] 缓存判断单元,当第一条件判断单元的判断结果为交易频率为最优交易频率,且未花费交易输出数据的交易时间间隔小于交易时间间隔阈值时,用于判断未花费交易输出数据的数据缓存长度是否小于未花费交易输出缓存区的缓存空间;

[0174] 第一概率预测单元,当缓存判断单元的判断结果为未花费交易输出数据的数据缓存长度小于未花费交易输出缓存区的缓存空间时,用于预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第一预测使用概率;

[0175] 第二概率预测单元,当缓存判断单元的判断结果为未花费交易输出数据的数据缓存长度大于等于未花费交易输出缓存区的缓存空间时,用于预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第二预测使用概率,第一预测使用概率大

于第二预测使用概率；

[0176] 第二条件判断单元,当第一条件判断单元的判断结果为交易频率大于预设的频率阈值,且未花费交易输出数据的交易时间间隔大于等于交易时间间隔阈值时,用于判断未花费交易输出数据的磁盘储存长度是否小于辅助磁盘内存的内存空间,频率阈值大于最优交易频率；

[0177] 第三概率预测单元,当第二条件判断单元的判断结果为未花费交易输出数据的磁盘储存长度小于辅助磁盘内存的内存空间时,用于预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第三预测使用概率,第三预测使用概率小于第二预测使用概率；

[0178] 当第二条件判断单元的判断结果为未花费交易输出数据的磁盘储存长度大于等于辅助磁盘内存的内存空间时，

[0179] 或，

[0180] 当第一条件判断单元的判断结果为交易频率小于频率阈值且不为最优交易频率时，

[0181] 或，

[0182] 当第一条件判断单元的判断结果为交易频率为最优交易频率,且未花费交易输出数据的交易时间间隔大于等于交易时间间隔阈值时,第二概率预测单元用于预测对应的未花费交易输出数据在区块链系统的新交易区块的预测使用概率为第二预测使用概率。

[0183] 在其中一种实施方式中,分区管理子系统包括：

[0184] 概率判断模块,用于判断未花费交易输出数据的预测使用概率为第一预测使用概率、第二预测使用概率或第三预测使用概率；

[0185] 第一分配模块,当概率判断模块的判断结果为未花费交易输出数据为第一预测使用概率时,用于将未花费交易输出数据保存至未花费交易输出缓存区；

[0186] 第二分配模块,当概率判断模块的判断结果为未花费交易输出数据为第二预测使用概率时,用于将未花费交易输出数据保存至本地内存；

[0187] 第三分配模块,当概率判断模块的判断结果为未花费交易输出数据为第三预测使用概率时,用于将未花费交易输出数据保存至辅助磁盘内存。

[0188] 所属领域的普通技术人员应当理解:以上任何实施例的讨论仅为示例性的,并非旨在暗示本申请的保护范围限于这些例子;在本申请的思路下,以上实施例或者不同实施例中的技术特征之间也可以进行组合,步骤可以以任意顺序实现,并存在如上的本申请中一个或多个实施例的不同方面的许多其它变化,为了简明它们没有在细节中提供。

[0189] 本申请中一个或多个实施例旨在涵盖落入本申请的宽泛范围之内的所有这样的替换、修改和变型。因此,凡在本申请中一个或多个实施例的精神和原则之内,所做的任何省略、修改、等同替换、改进等,均应包含在本申请的保护范围之内。

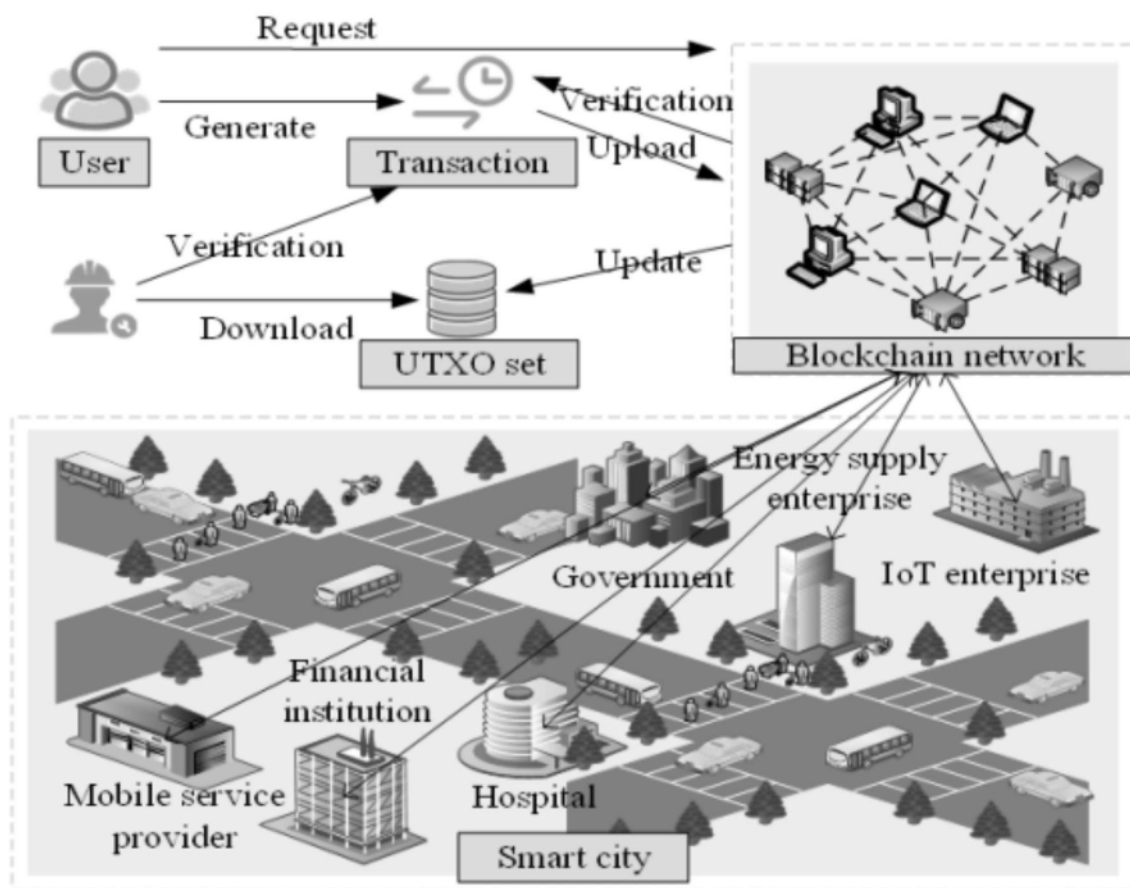


图1

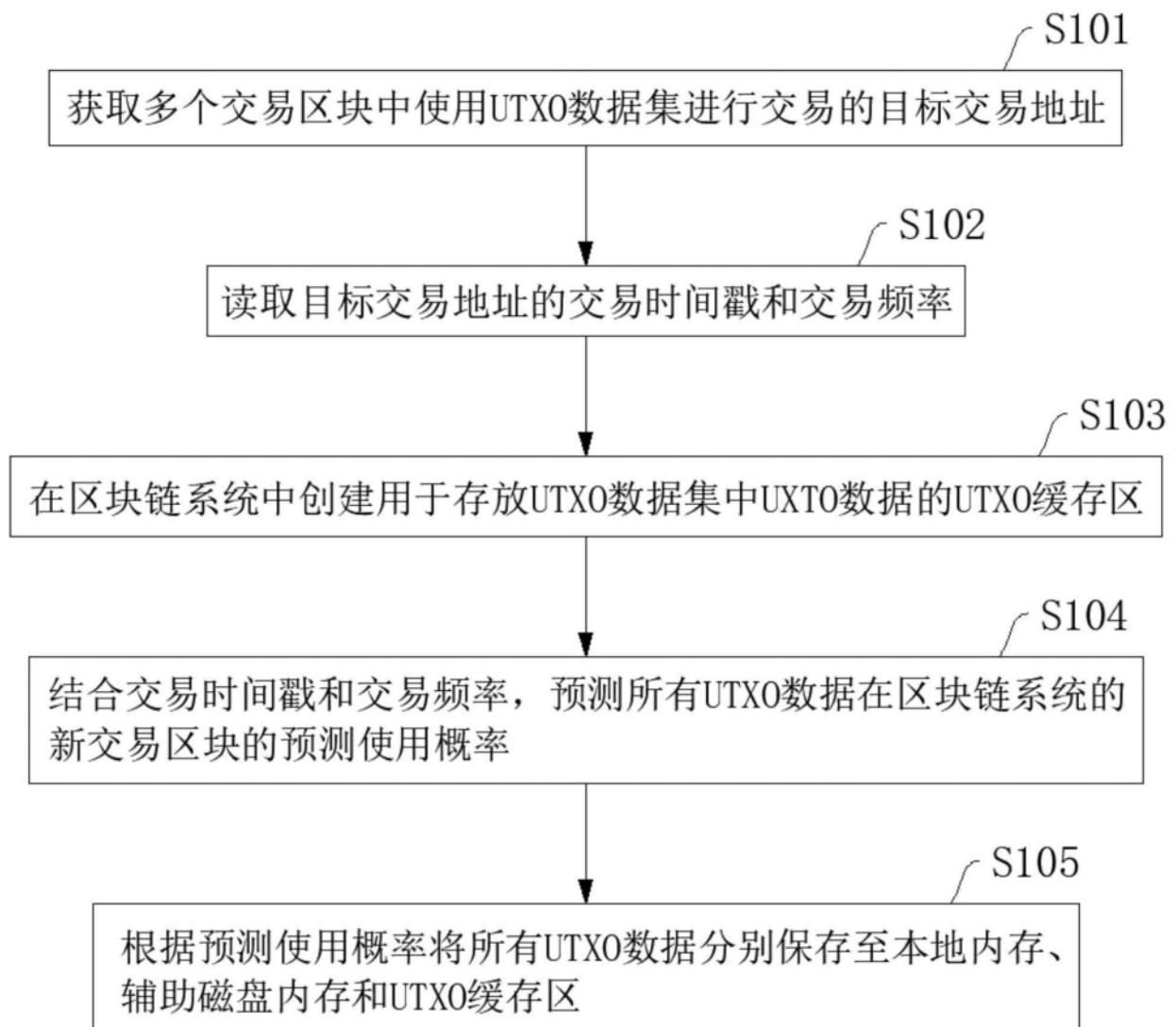


图2

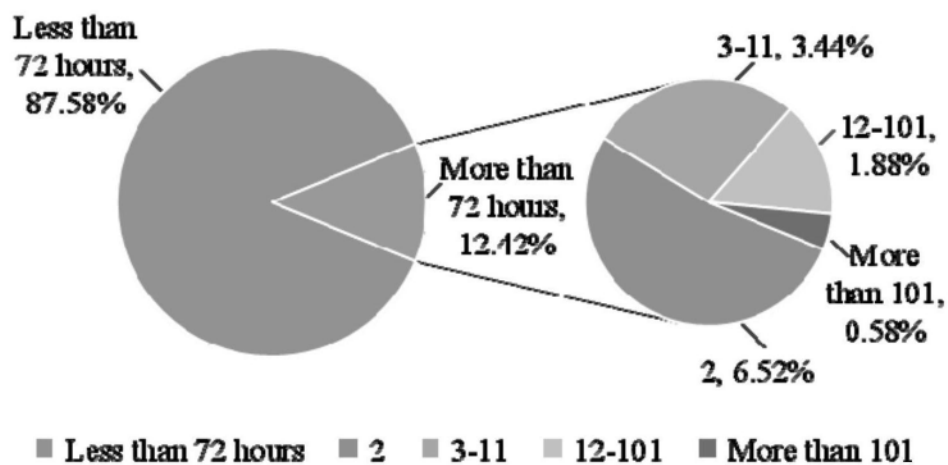


图3

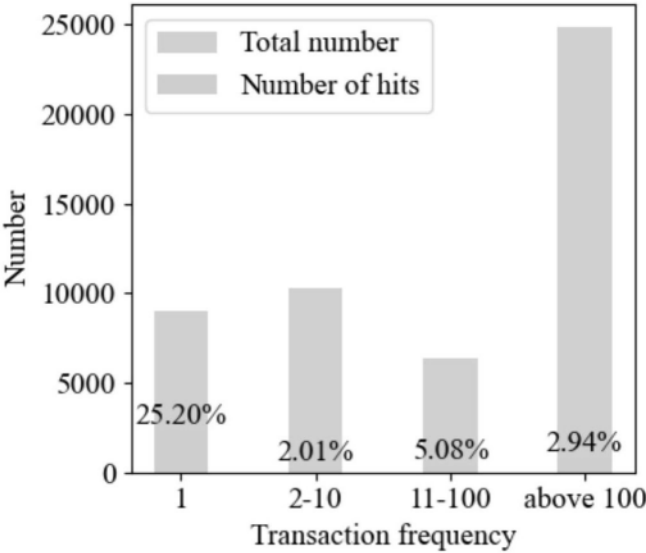


图4

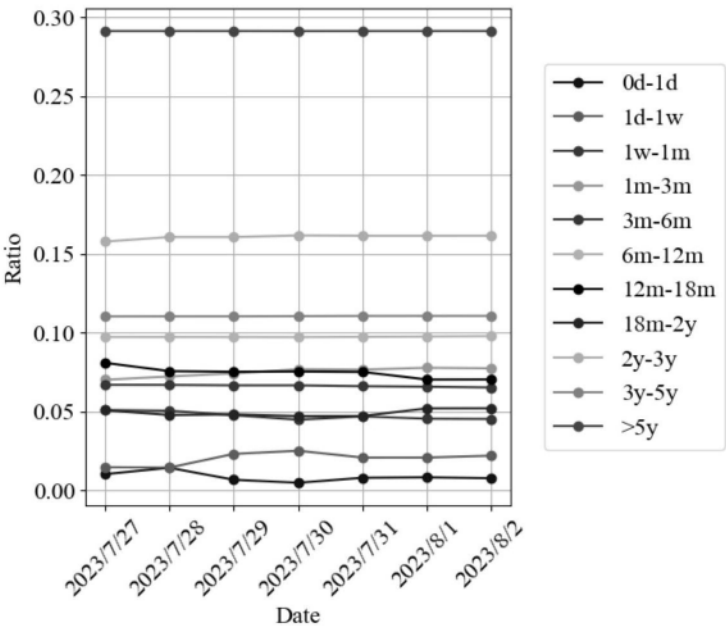


图5

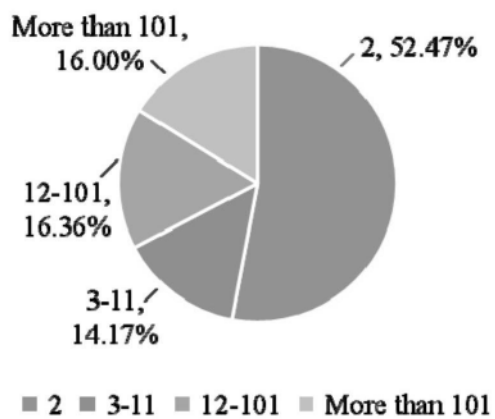


图6

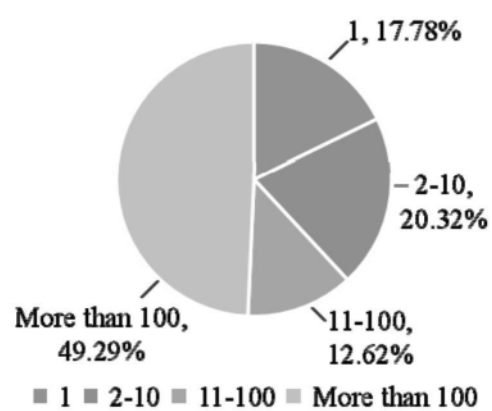


图7