基于信誉节点分类的 PBFT 共识算法优化研究

周岩龙[1] 宋淑彩[1]
ZHOU Yanlong SONG Shucai

摘 要 针对实用拜占庭容错算法（PBFT）随意选取主节点的方式会导致系统安全受攻击，且当节点较多时通信开销大的问题。文章提出了一种基于信誉节点分类的 PBFT 共识算法优化研究。首先，根据信誉模型为每一个节点进行信誉评估分类，高信誉值节点充当主节点和共识节点的概率更大，增加系统的安全；其次减少一致性协议参与共识的节点数量，提高系统的共识效率。仿真实验表明，SR-PBFT 共识算法能有效减少共识过程中的恶意节点，提高系统安全性，同时在共识时延和吞吐量方面比原 PBFT 共识算法具有更好的性能。

关键词 实用拜占庭容错共识算法（PBFT）；节点分类；信誉评估；共识算法；通信开销；吞吐量；共识时延

doi：10.3969/j.issn.1672-9528.2025.07.039

0 引言

实用拜占庭容错共识算法（practical byzantine fault tolerance, PBFT）在恶意节点不大于 (*n*-1)/3（*n* 为总节点数）的情况下，仍可以达成共识，使拜占庭将军问题的解决变得可能，应用在联盟链较为广泛，但仍存在需要优化的地方。首先 PBFT 主节点选取随意，增加了恶意节点成为主节点从而作恶的概率，不利于系统的安全性；其次 PBFT 三阶段的广播协议随着节点数的增多极大增加了通信开销。

针对 PBFT 提出的缺陷，学者们研究出不少研究成果。文献 [1] 提出将 PBFT 算法与 DPOS 算法相结合，根据股份投票的方式选出共识节点，减少共识过程中节点间的通信量，但此算法与 POX 类算法都存在相同的问题，即需要数字货币才能完成共识。文献 [2] 设置一个系统，用来同步拜占庭的数量，但存在错误可能将诚信节点同步为拜占庭节点。文献 [3] 提出基于随机预言模型验证的异步拜占庭一致算法，通信复杂度每一轮都会得到有效地降低，其缺点是随机选取可能会造成对系统的安全威胁。文献 [4-5] 则提出在网络中不存在拜占庭节点的情况下，简化 PBFT 算法的一致性协议，时间复杂度从 $O(n^2)$ 降到了 $O(n)$，但恶意节点的不可避免行使该算法难以应用。文献 [6-7] 通过随机哈希的方式对主节点进行选取，可一定概率阻止恶意节点的恶意行为。文献 [8-9] 通过对节点进行随机分组以及双节点的方式，均避免了单一节点的恶意行为，但成群作恶则无法避免。上述文献都对 PBFT 算法进行了不同层面的改进优化，但均没有涉及主节点的选取信誉问题，也就不能说明主节点选取的可靠性和安全性问题。且如果发现存在拜占庭节点，文献中的方法没有进行相应的处理操作，也就导致恶意节点在系统中可以持续进行破坏。

1 SR-PBFT 算法设计方案

1.1 信誉值机制设计

将区块链中每个节点的初始信誉分定为 0.6。通过引入正向更新、负向更新和信誉值下限对节点的信誉分进行增加、扣减和删除。不同程度的错误可设定不同的信誉值扣减幅度。信誉值下限即设定一个信誉值的下限，低于此值的节点可能被网络临时或永久性地隔离，以防止恶意行为对系统的影响。信誉机制流程图如图 1 所示。

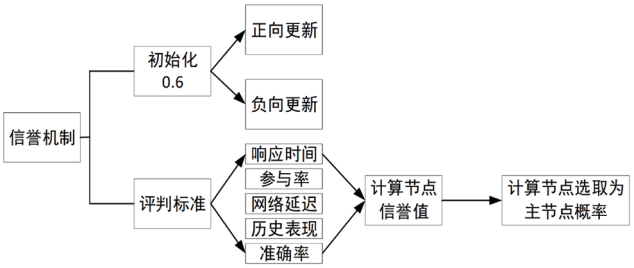

图 1 信誉机制流程图

通过引入评判标准来综合评估每个节点的信誉，响应时间（response time, RT）：节点从接收到请求到回应的时间。响应准确率（response accuracy, RA）：节点的响应被验证为正确的频率。参与率（participation rate, PR）：节点参与共识的次数与总共识次数的比例。网络延迟（network latency, NL）：节点在网络中的延迟时间。各项标准的计算公式为：

1. 河北建筑工程学院信息工程学院 河北张家口 75000
[基金项目] 基于信誉机制的区块链 PBFT 改进研究（XY2025021）

响应时间得分 S_{RT}：定义一个理想的响应时间 RT_{ideal} 和最大可接受时间 RT_{max} 分按公式计算为：

$$S_{RT=}\begin{cases}1-\dfrac{RT_i-RT_{ideal}}{RT_{max}-RT_{ideal}} & RT_{ideal}\le RT_i\le RT_{max}\\ 0 & RT_i>RT_{max}\end{cases} \quad (1)$$

响应准确率得分 S_{RA}：

$$S_{RA}=\frac{正确响应次数}{总响应次数} \quad (2)$$

参与率得分 S_{PR}：

$$S_{PR}=\frac{参与共识次数}{总共识次数} \quad (3)$$

网络延迟得分 S_{NL}：定义一个理想的网络延迟 NL_{ideal} 和最大可接受延迟 NL_{max} 分按公式计算为：

$$S_{NL}=\begin{cases}1-\dfrac{NL_i-NL_{ideal}}{NL_{max}-NL_{ideal}} & NL_{ideal}\le NL_i\le NL_{max}\\ 0 & NL_i>NL_{max}\end{cases} \quad (4)$$

式中：NL 为节点的延迟时间大概在 5~50 ms 之间。

历史表现得分 S_{HP}：根据节点的历史信誉值的移动平均值计算：

$$S_{HP}=\frac{\sum_{k=1}^{T}R_{i(k)}}{T} \quad (5)$$

式中：T 为时间窗口；$R_i(k)$ 为第 k 次共识后的信誉值。

使用随机前沿分析（stochastic frontier analysis, SFA）模型。

（1）输入与输出定义

输入：每个节点的输入变量矩阵 $\boldsymbol{X}$，包括响应时间、响应准确率等。节点数量为 n，每个节点有 m 个输入变量。

输出：节点的输出变量 Y，即该节点的信誉值或效率。

（2）SFA 模型的基本公式为：

$$Y_i=f(x_i;\beta)+v_i-u_i \quad (6)$$

式中：Y_i 是第 i 个节点的输出（如信誉值）；x_i 是输入变量（如响应时间、响应准确率等）；β 是回归系数（需要通过模型拟合获得）；v_i 是随机噪声，通常服从正态分布$v_i\sim N(0,\sigma_v^2)$；u_i 是无效率项，通常服从截断正态分布$u_i\sim N^{+}(0,\sigma_u^2)$，该项代表每个节点的效率损失。效率 E_i 的计算公式为：

$$E_i=\frac{y_i}{f(x_i;\beta)+v_i} \quad (7)$$

即效率是实际输出和预测输出之比。

具体计算完整方案步骤如下：

（1）数据准备：收集每个节点的输入和输出数据。输入数据包含多个输入变量，如响应时间、响应准确率等，输出数据是信誉值。输入变量将各项评判标准视为 SFA 模型的输入。此方案有以下几个标准：

X_1 为响应时间；X_2 为响应准确率；X_3 为参与率；X_4 为网络延迟；X_5 为历史表现。

输出变量：将节点的信誉值 R 作为 SFA 模型的输出。

SFA 模型建立：使用最小二乘法估计回归参数 β，同时使用极大似然估计（MLE）来估计随机噪声项 v_i 和无效率项 u_i。

效率计算：基于估计的回归系数和噪声项，计算每个节点的效率。

优化和结果输出：评估每个节点的效率，优化参数并输出结果。如节点作恶，假设节点的作恶次数为 L 次，则节点的最终信誉值 $S_i=E_i^{L+1}$。

1.2 节点分类标准

在本方案中，共分为 3 个角色：领导者角色、共识者角色、备份者角色。备份者角色的节点不参与共识但是会存储整个共识过程中所产生的区块。选择信誉值大于阈值共识组的临界值 S=0.6 的节点参与共识。设置一个次数阈值 L_{max}，如节点作恶的次数大于 L_{max} 则将节点剔除网络。会对所有备份组的节点进行观察组阶段的表现进行临时信誉值排序，只对观察阶段的表现进行排序。由于每个节点存储了所有的区块数据，因此只需对节点的延迟时间进行比较即可，延迟时间越短，会越快被加入正常节点中。如每轮的共识节点的数目不足总节点数目的 2/3，则将凑足备份组节点进行的表现进行排序，凑够总节点数目的 2/3 并且将其信誉值设置成 S，将其除作恶以外的其他数据清空，重新计算。

1.3 领导者节点选取

信誉值可以影响领导者（primary）选举过程。信誉值高的节点更有可能被选为领导者，从而增强整个系统的稳定性。

（1）计算领导者资格分数（LES）：每个节点根据其信誉值 R_i 计算其领导者资格分数（leader eligibility score, LES）。具体来说，节点的 LES 直接等于其信誉值：$LES_i=S_i$。

（2）确定选举概率：节点根据其 LES 计算成为领导者的概率。选举概率 $Prob_i$ 通过节点的 LES 相对于所有节点的 LES 总和计算得出：

$$Prob_i=\frac{LES_i}{\sum_{k=1}^{n}LES_k} \quad (8)$$

（3）领导者选举过程：在领导者选举过程中，所有节点参与选举轮次，并广播各自的 $Prob_i$。选举通过加权随机选择的方式进行，具体步骤如下：

为每个节点计算累积概率。例如，假如有 A、B、C 三个节点，ProbA=0.3、ProbB=0.5、ProbC=0.2，通过加权的方式计算出各自的概率。

AccumulatedProbA=0.3

AccumulatedProbB=0.3+0.5=0.8

AccumulatedProbC=0.8+0.2=1.0

生成一个在 [0.5~1] 之间均匀分布的随机数 rand，确保其通常大于 0.5。假设 rand=0.6。比较随机数 rand 和累积概率，找到使累积概率首次超过随机数的节点。例如，在上述例子中，累积概率达到 0.8 时首次超过随机数 0.6，因此节点 B 被选为领导者。若该节点的选举概率与随机数的差值过大，重新生成随机数并进行选择，直到符合条件。这种机制确保高信誉值的节点更有可能被选为领导者，同时也引入了随机性以防止单一节点反复当选。

2 实验分析

本实验采用 GO 编程语言仿真一个多节点的区块链系统，在实验过程中对 PBFT 及本文的 SR-PBFT 共识算法进行对比分析，集中在时延、开销、吞吐量和安全 4 个方面进行分析。实验设施配置信息如表 1 所示。

表 1 实验设施配置信息

对象	配置
软件环境	Go、Intellij IDEA
操作系统	Windows 10
处理器	lnte164 Family6 Model 183 Stepping 1 GenuinelIntel

2.1 共识时延分析

共识时延：客户端从提交请求开始计时到完成确认结束期间记录的时间。实验中节点数量从 10 递增到 55，步长为 5。为避免偶然性的发生，在不同节点数下重复进行 10 次实验，最终取 10 次平均值，此方案 SR-PBFT 与原 PBFT 共识算法共识时延对比结果如图 2 所示。

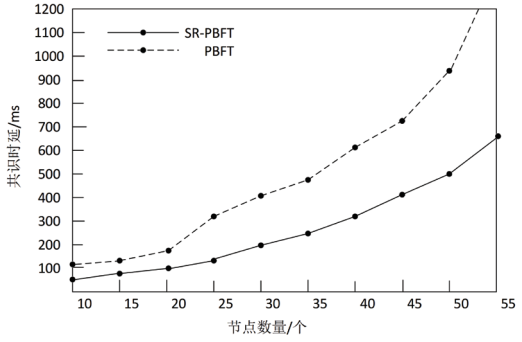

图 2 共识时延对比结果

由实验结果可知，在相同节点数的情况下，SR-PBFT 共识算法的共识时延均小于 PBFT 共识算法，当节点数增加到 20 个节点 PBFT 共识算法的共识时延增长率剧增，而 SR-PBFT 共识算法的共识时延增长率较稳定。

2.2 交易吞吐量分析

交易吞吐量：单位时间内处理请求的数量。 实验中设置节点数量由 10 增加到 55，步长为 5，同样为避免偶然性的发生，在每个设定节点下进行 10 次实验测试，最终取 10 次重复测试的平均值。PBFT 与 SR-PBFT 共识算法的吞吐量对比结果如图 3 所示。

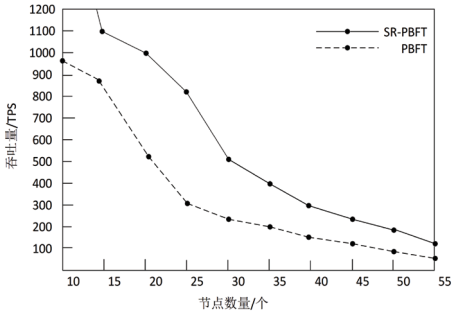

图 3 吞吐量对比结果

由实验结果可知，随着节点数量从 10 持续增加到 55，同一节点下 SR-PBFT 的吞吐量始终大于 PBFT，且可以看到，当节点数小于 30 时，两种共识算法的吞吐量差距较大，更能说明 SR-PBFT 在吞吐量方面比 PBFT 具有更大的优势，故 SR-PBFT 在吞吐量方面比 PBFT 做出了很大的改进。

3 总结展望

本设计通过信誉模型改进的方式将节点进行信誉评估，通过信誉分的高低对节点进行初步信誉判断，大大降低了恶意节点作恶的概率，且通过信誉分对节点进行分类，通过筛选部分共识节点参与共识过程大大减少节点之间的通信量，提高共识效率。通过实验分析可知，在共识时延和交易吞吐量两方面性能均优于原始 PBFT，此方案对多节点场景下的区块链应用具有重要的理论价值和实际意义。但仍存在不足，PBFT 繁琐的共识流程依然是解决问题的重中之重，三阶段的广播协议复杂度高，通信量呈指数增长，通过降低节点数的方式间接提高共识效率，但终究没有解决根本问题的所在，对于 PBFT 共识流程的改进是区块链进步的一大关键。

参考文献：

[1] 许艳艳 . 基于 DPoS 与评分制的 PBFT 算法研究 [D]. 淮南 : 安徽理工大学 , 2022.

[2] MALKHI D, REITER M.Byzantine quorum systems[C]// STOC '97: proceedings of the twenty-ninth annual ACM symposium on theory of computing.NewYork:ACM,1997: 569-578.

[3] ABRAHAM I, MALKHI D, SPIEGELMAN A, et al. Validated asynchronous byzantine agreement with optimal resilience and asymptotically optimal time and word communication[EB/OL].(2018-11-04)[2024-11-10]. https://doi.org/10.48550/arXiv.1811.01332.

[4] 唐朕 , 戴欢 , 王陆平 , 等 . 面向动态网络的改进 PBFT 共识算法 [J]. 微电子学与计算机 ,2025,42(6):97-104.

[5] VOJISLAV B M, JELENA M, CHANG X L. Performance of PBFT vonsensus under voting by groups[J]. Blockchains,

基于模糊理论的局部空间同位模式挖掘方法

段韶鹏[1] 娄文璐[2] 张金鹏[3] 张 占[4]
DUAN Shaopeng LOU Wenlu ZHANG Jinpeng ZHANG Zhan

摘 要 地理空间具有显著的异质性，不同区域的地理属性分布不均衡，导致全局分析方法难以精准定位具有相似地理属性的局部区域，挖掘结果的准确性不足。为此，文章提出了一种基于模糊理论的局部空间同位模式挖掘方法。首先，采用多层次融合整合地理属性和特征，形成全面的地理属性空间表达。通过参与指数评估空间同位模式普遍性，筛选候选局部同位模式集合。然后，针对该集合，运用模糊理论处理地理空间异质性和不确定性，构建模糊上下近似描述地理属性，提取相似地理属性的局部感兴趣区域。利用图结构模型挖掘局部空间同位模式，将区域视为节点，通过边和权重体现相似性或空间关联，构建邻接矩阵。最后，采用谱聚类技术分析图结构模型，观察聚类簇特性，深入了解局部空间同位模式分布规律和特征，从而实现挖掘。实验结果表明，所提方法准确率为 0.945，决定系数为 0.932，相较于对照组方法显著提升，证明了其可行性和可靠性。

关键词 模糊理论；同位模式；地理属性；图结构模型；谱聚类技术；模式挖掘

doi：10.3969/j.issn.1672-9528.2025.07.040

0 引言

在空间数据分析领域，局部空间同位模式的挖掘对于理解地理现象的空间分布和相互关联具有重要意义。随着空间数据的快速增长，如何高效地挖掘这些模式成为研究热点。然而，现有的同位模式挖掘方法在处理复杂空间数据时，往往存在效率低下的问题。刘宇情等人[1]提出了基于网格空间团的多级同位模式挖掘方法，通过划分网格并构建空间团来挖掘同位模式。在地理空间异质性显著的情况下，固定的网格划分难以适应不同区域的地理属性分布不均衡，导致密集区域出现过多噪声数据，影响挖掘的准确性和全面性。王靖涵等人[2]则采用了基于图结构的空间同位模式挖掘方法，利用图模型表示空间关系，并通过图算法进行挖掘。在定位具有相似地理属性的局部区域时，该方法依赖于节点或子图的相似性度量。由于相似性度量的不精确，导致局部区域的相似性被错误评估。吴静等人[3]从全局视角筛选不频繁同位模式为候选，通过粗糙集处理其实例位置特点，对局部自然分

1. 河南开放大学 河南郑州 450008
2. 云南财经大学商学院 云南昆明 650221
3. 云南财经大学信息学院 云南昆明 650221
4. 长沙民政职业技术学院软件学院 湖南长沙 410004

[基金项目] 2025 年河南省科技攻关项目（252102210149）；云南省基础研究计划面上项目（202501AT070455）；云南省教育厅科学研究基金项目青年人才基础研究专项（2024J0643）；湖南省自然科学基金资助项目（2024JJ8025）

2024, 2 (2):134-149.
[6] 翟社平，霍媛媛，杨锐，等 . 基于一致性哈希和随机选取的 PBFT 算法改进 [J]. 计算机工程与应用 ,2024,60(12):294-302.
[7] CASTRO M, LISKOV B. Practical byzantine fault tolerance [C]//OSDI '99: Proceedings of the Third Symposium on Operating Systems Design and Implementatio.NewYork: ACM, 1999: 173-186.
[8] 宋宇哲，郑广海，张鑫 . RG-BFT: 基于随机分组的拜占庭容错算法 [J]. 计算机工程与设计 , 2024, 45 (6): 1661-1667.
[9] 陈珩，黄世成，郑明辉 . 一种双主节点的 PBFT 共识算法 [J]. 湖北民族大学学报 (自然科学版), 2024, 42 (2): 185-190.

【作者简介】

周岩龙（2001—），男，河北廊坊人，硕士研究生，研究方向：区块链共识算法、大数据计算，email: 1164460643@qq.com。

宋淑彩（1970—），女，河北张家口人，博士，研究生导师、教授，研究方向：大数据技术、计算机视觉，email: ssc1330@hebiace.edu.cn。

（收稿日期：2025-02-04 修回日期：2025-07-11）