



西安交通大学

硕士研究生学位论文选题报告

学 号： 3122358153

姓 名： 王璟

导 师： 齐赛宇

论文题目： 基于模糊测试路径探索的区块链共识算法研究

学科专业： 软件工程

学 院： 软件学院

填写时间： 2024 年 3 月 1 日

西安交通大学研究生院制

硕士研究生学位论文选题报告填写说明及管理规定

硕士学位论文选题报告是做好学位论文的基础，为了完善硕士研究生过程质量监控体系，提高硕士研究生培养质量，要求在校硕士生应在第三学期结束前（两年毕业试点学院的硕士生应在第二学期结束前）完成学位论文选题报告。

一、硕士生在查阅一定的国内外文献资料基础上，填写完成《硕士研究生学位论文选题报告》。

二、《硕士研究生学位论文选题报告》完成以后，应组织公开的学位论文选题报告会。

三、选题报告会由学院或系、所负责组织，选题报告会的评审专家组一般由 3-5 名副高以上（含副高）人员组成。评审专家在选题报告会后负责就选题的意义、文献综述、研究内容、可能遇到的问题、是否通过选题等写出结论性的审查意见，并将结果和相关材料留学院备案。

四、《硕士研究生学位论文选题报告》必须采用 A4 纸双面打印，左侧装订成册，各栏空格不够时，请自行加页。本表可在研究生院主页 <http://gs.xjtu.edu.cn/> 下载。

五、《硕士研究生学位论文选题报告》由学院归档。

论文题目：基于模糊测试路径探索的区块链共识算法研究

论文类型：（1）基础研究；（2）应用基础研究；（3）应用研究；（4）其它

课题来源：（1）纵向课题；（2）横向课题；（3）自选课题；（4）其它

一、选题的科学依据（1、选题背景；2、理论意义和应用价值；3、国内外研究现状及发展趋势。附主要参考文献）

1、选题背景

2008 年中本聪发表了一篇点对点的数字货币系统的论文^[1]，该论文提出了比特币的概念并将区块链技术作为比特币的底层技术，截至今日，比特币是最常用的加密货币，区块链也逐渐走进大众视野并获得了越来越多的关注度。近些年随着各界对其研究与应用的不断深入，区块链技术已经开始独立于加密数字货币，发展成为一门新的研究领域。

区块链作为一种底层技术和基础架构，本质上是一个存放在非安全环境中的分布式的、去中心化的、篡改难度极高的数据库系统，它采用密码学方法保证数据不被篡改，采用共识算法对新增数据达成共识。区块链技术建立了新的信任机制，允许在没有权威节点的去中心化情况下，各网络节点之间达成可信共识，是一项从思想到技术的重大飞跃。

区块链相关技术有很多方面，大致可分为 4 类：共识算法、隐私保护技术和相应密码学技术、智能合约相关技术、面向应用相关技术。其中，共识算法是区块链技术最核心，也是整个技术发展和学术界最热衷的领域。目前常见的共识算法有工作量证明（PoW，Proof-of-Work）、权益证明（PoS，Proof-of-Stake）、委托权益证明（DPoS，Delegated-Proof-of-Stake）等。

PoW 是比特币使用的共识算法，属于最早问世的共识算法，时至今日仍处于主流之列。PoW 的原理如图 1 所示。网络上的节点（即“矿工”）需要执行大量计算工作来解决谜题，通常是寻找一个随机数（nonce），使得整个区块的哈希值小于给定的阈值，寻找随机数的过程称为“挖矿”。在节点成功找到满足条件的哈希值后，就会立即向全网广播，网络上的节点收到后则会立即验证，验证通过表示成功解谜。挖矿成功的矿工们获得奖励和下一个区块的记账权。然而，传统 PoW 需要较高的处理器能力和能耗，环境成本高，且这些算力只是用于挖矿，并没有其他用途。

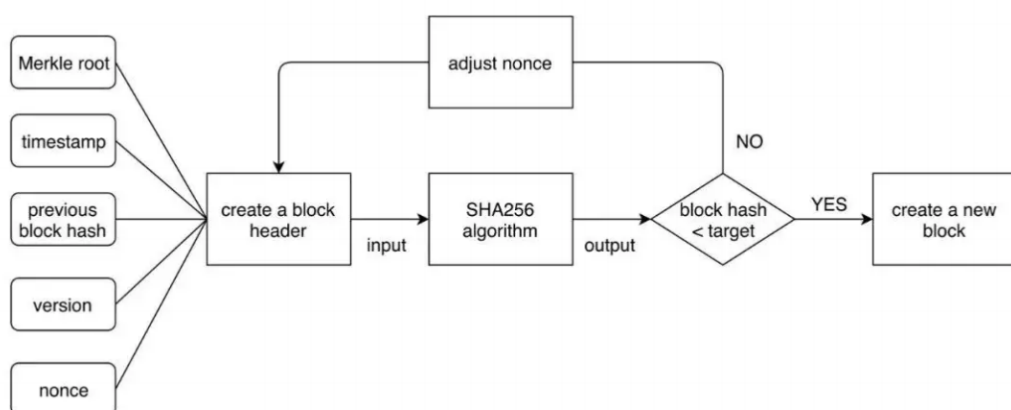


图 1 工作量证明原理图

随着信息技术的发展，计算机软件已逐步渗透到社会生活的方方面面，软件正在爆发式地增长，随之而来的软件质量问题也大大影响着我们的生活。软件缺陷是影响软件质量的关键因素，同时它也是软件开发过程中不可避免的。

1990 年，模糊测试（Fuzzing）首次出现，并被用于 UNIX 系统的软件测试，以测试各种 UNIX 应用程序的鲁棒性^[2]。模糊测试是一种软件测试技术，它会根据一定的规则自动或半自动地生成随机数据，然后将这些产生的随机数据输入到动态运行的被测程序入口，同时监控被测程序是否有异常情况出现，如系统崩溃、断言失败等来发现软件的缺陷。随着模糊技术的发展，研究人员引入了一些最先进的技术来进行漏洞挖掘，模糊技术的应用场景也在不断扩大。模糊测试简单而有效，已成为迄今挖掘最多软件漏洞的一种测试方法。许多软件制造商在软件发布之前进行模糊测试，以发现潜在的漏洞^[3]。

覆盖引导（Coverage-guided）的模糊测试是 Fuzzing 最有用的技术之一，其关键元素是跟踪代码覆盖率信息。现有的覆盖引导模糊器通常使用探索到的基本块或边的数量来度量代码覆盖率。AFL（American Fuzzy Lop）是由 Google 安全工程师 Michal Zalewski 开发的一款基于覆盖引导的开源模糊测试工具，它通过记录输入样本的代码覆盖率，从而调整输入样本以提高覆盖率，增加发现漏洞的概率。基于覆盖率反馈的模糊测试方法是目前模糊测试的主要研究方向^[3]。AFL 的基本流程如图 2 所示。AFL 采用哈希函数和位图来记录执行路径以提高执行效率。路径信息可以提供比基本块和边覆盖更准确的执行跟踪信息，然而，路径的数量随着程序规模的增加呈指数增长，跟踪实际程序的所有路径几乎是不可能的^[4]。但是我们可以利用路径信息来辅助 Fuzzer，使 Fuzzer 能够更有效地识别和利用新路径。因此，路径探索在 Fuzzing 中十分重要，是提升 Fuzzing 性能的一大法宝。

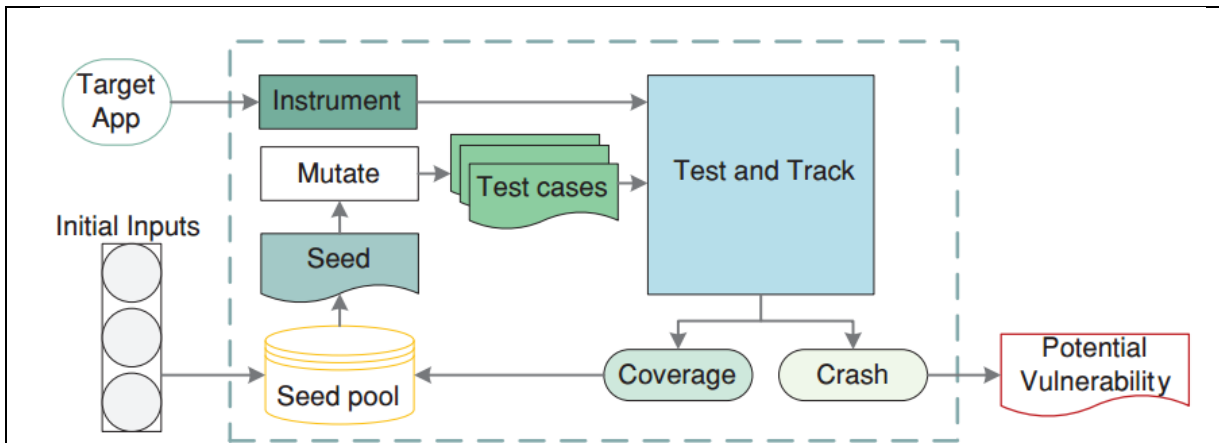


图 2 AFL 基本流程图

2、理论意义和应用价值

随着区块链应用场景的不断拓展，区块链从局限的数字货币职能过渡到其他非金融领域，一步步扩大自身影响，向其他行业不断渗透，朝着“区块链+”的方向不断延伸^[5]。

在共识机制方面，比特币首先提出了中本共识，该共识将 PoW 应用于区块链结构，以建立一个无需信任的大型分类账。中本共识解决了公共网络的分布式一致性问题。一般来说，只有最快解决谜题的矿工才能赢得比赛并获取奖励，其他人则不得不放弃当前的难题来计算下一个难题，这浪费了太多的计算能力。PoW 会消耗大量的自然资源：据估计，从 2017 年 12 月开始，比特币网络每年消耗超过 30 太瓦时，超过了丹麦的能源消耗。此外，目前大多数挖矿都是由专门的 ASIC 完成的，除了比特币挖矿之外没有其他用途。这种计算是昂贵的，对环境不友好的。鉴于这些问题，人们开始寻找 PoW 的替代方案，目前很多研究为了将 PoW 中的资源利用起来，将 PoW 和许多领域结合，把传统 PoW 中的寻找符合要求的哈希值这一难题替换为其他有意义的难题，如：通过对给定任务的机器学习系统进行排名来达到分布式共识、将人工智能模型的训练作为难题，等等。

考虑到现有的 Fuzzing 方法在寻找漏洞时需要进行大量的路径探索，是一个消耗大量算力的过程，而区块链中传统 PoW 存在大量算力浪费的问题，因此可以将传统 PoW 中浪费的算力应用到 Fuzzing 中，两者相结合，达到将浪费的算力利用起来的目的。为了让更多有用的算力投入到漏洞检测上，本文把利用 Fuzzing 寻找新的程序漏洞作为要解决的难题来替换 PoW 中基于哈希的难题，以此和区块链结合。这样做不仅将原本浪费的算力利用起来，而且能够提供一个有激励机制的 Fuzzing 平台，以便测试人员们更好地进行漏洞检测。

3、国内外研究现状及发展趋势

在区块链方面，目前已经有研究人员提出了很多改进的共识方案，这些研究工作都将区块链传统共识机制中浪费的算力利用到其他有意义的事情中。WekaCoin^[6]，是

一种基于被称为 Proof-of-Learning 的新型分布式共识协议的点对点加密货币，其工作原理与比特币类似，但使用 Proof-of-Learning 而不是基于哈希的谜题作为工作量证明。Proof-of-Learning 通过对给定任务的机器学习系统进行排名来实现分布式共识。除此之外，还有一些方案通过利用矿工的存储资源来解决浪费算力的问题。Proof-of-Continuous-Work(POCW)^[7]是一种具有存储相关激励机制的变体共识算法，矿工可以通过不断提交存储证明来积累挖矿优势。SpaceMint^[8]是一种基于空间证明的加密货币。SpaceMint 中的矿工比拼磁盘空间而不是算力。Coin.AI^[9]是一种基于训练深度学习模型的工作量证明方案，它的挖矿过程相当于具有许多潜在应用的人工智能模型的训练。

在过去的十年中，Fuzzing 已经被证明是一种非常有效的发现软件漏洞的方法。在 AFL 推广了轻量级覆盖反馈的突破性概念之后，Fuzzing 领域出现了大量的科学工作，提出了新技术，改进了现有策略的方法学方面，或者将现有方法移植到新的领域。在 AFL 中，探索到新的执行路径的测试用例在下次变异(mutation)中被选中的几率更大。因此，路径探索在 Fuzzing 中是非常重要的，一些研究将重点放在如何让种子触发更多的独特路径。SmartSeed + AFL^[10]比现有策略多发现 30.7% 的独特路径，这表明与当时最先进的种子选择策略相比，SmartSeed 产生了更好的性能。PathAFL^[4]在 AFL 的基础上提出了一种新的路径覆盖辅助 Fuzzing 解决方案，在跟踪路径粒度和 Fuzzing 性能之间取得了平衡。EcoFuzz^[11]利用一种变体的 Adversarial Multi-Armed Bandit (VAMAB)模型来模拟基于覆盖率的灰盒模糊测试。VAMAB 模型和自适应功率调度的集成放大了 EcoFuzz 在有效探索各种执行路径方面的潜力，从而提高了模糊过程的整体有效性。Wang 等^[12]利用 LSTM 模型学习脆弱程序路径的隐藏模式，识别脆弱路径，指导种子选择。这些选择的种子可以探索更多更容易受到攻击的路径。

然而，现有的变体共识算法中还未出现将传统 PoW 和 Fuzzing 结合的工作，因此本课题聚焦于将两者结合，以实现新的变体共识算法。

参考文献

- [1] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[J]. Decentralized business review, 2008.
- [2] Miller B P, Fredriksen L, So B. An empirical study of the reliability of UNIX utilities[J]. Communications of the ACM, 1990, 33(12): 32-44.
- [3] Yu Z, Liu Z, Cong X, et al. Fuzzing: Progress, Challenges, and Perspectives[J]. Computers, Materials & Continua, 2024, 78(1).
- [4] Yan S, Wu C, Li H, et al. Pathafl: Path-coverage assisted fuzzing[C]//Proceedings of the 15th ACM Asia Conference on Computer and Communications Security. 2020: 598-609.
- [5] 曾诗钦,霍如,黄韬等.区块链技术研究综述: 原理、进展与应用[J].通信学报,2020,41(01):134-151.
- [6] Bravo-Marquez F, Reeves S, Ugarte M. Proof-of-learning: a blockchain consensus mechanism based on machine learning competitions[C]//2019 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON). IEEE, 2019: 119-124.
- [7] Yin H, Zhang Z, He J, et al. Proof of Continuous Work for Reliable Data Storage Over Permissionless

- Blockchain[J]. IEEE Internet of Things Journal, 2021, 9(10): 7866-7875.
- [8] Park S, Kwon A, Fuchsbaue G, et al. Spacemint: A cryptocurrency based on proofs of space[C]//Financial Cryptography and Data Security: 22nd International Conference, FC 2018, Nieuwpoort, Curaçao, February 26–March 2, 2018, Revised Selected Papers 22. Springer Berlin Heidelberg, 2018: 480-499.
- [9] Baldominos A, Saez Y. Coin.AI: A proof-of-useful-work scheme for blockchain-based distributed deep learning[J]. Entropy, 2019, 21(8): 723.
- [10] C. Lv, S. Ji, Y. Li, J. Zhou, J. Chen et al., “SmartSeed: Smart seed generation for efficient fuzzing,” arXiv:1807.02606, 2018.
- [11] T. Yue, P. Wang, Y. Tang, E. Wang, B. Yu et al., “EcoFuzz: Adaptive energy-saving greybox fuzzing as a variant of the adversarial multi-armed bandit,” in Proc. of the USENIX Security, Boston, MA, USA, pp.2307–2324, 2020.
- [12] Y. Wang, Z. Wu, Q. Wei and Q. Wang, “NeuFuzz: Efficient fuzzing with deep neural network,” IEEE Access, vol. 7, no. 1, pp. 36340–36352, 2019.

二、主要研究内容和方案

由于本课题将区块链的共识机制和 Fuzzing 相结合，因此研究内容主要分为这两个大部分，在这之下又可以细分为一些小步骤，如下所示。

1、搭建一个区块链框架

理解区块链原理，用 Java 实现一个初始区块链，共识机制部分待修改。

2、构建 Fuzzing 平台

AFL 是目前比较常用的 Fuzzing 工具，它首次采用源码编译插桩和 QEMU 模式来实现代码覆盖引导 Fuzzing 的方式，这绝对是 Fuzzing 技术发展历程中最重要的一次里程碑，也是技术分水岭，它开启了 Fuzzing 技术的新篇章。因此本课题使用 AFL 作为 Fuzzing 的工具。

2.1 修改 AFL 源码

该步骤的目的是获取每次执行程序时的输入和执行路径，为了实现该目标，需要阅读并理解 AFL 源码并合理修改源码。AFL 源码中的核心代码有 afl-gcc.c、afl-as.c、afl-as.h、afl-fuzz.c 等，其中 afl-gcc.c、afl-as.c、afl-as.h 用于普通模式插桩，针对源码进行插桩，适用于 gcc 和 clang，afl-fuzz.c 是 Fuzzing 的核心实现代码，是 AFL 的主体。

afl-as.c 和 afl-as.h 会记录程序执行路径之类的关键信息，对程序的运行情况进行反馈。afl-as 在函数入口、条件跳转指令后、指令的标签之后进行插桩，插桩时会赋予每个桩一个随机数作为标识符，于是获取执行路径这个问题就转换为获取每次程序执行时被命中的桩的标识符，并顺序输出。由于随机数是在插桩时生成，因此需要重点修改的是插桩部分的代码，也就是 afl-as.h 中的代码，以获取桩被命中时的随机数，再将每次程序运行时代表被命中桩的随机数顺序输出。

afl-fuzz 主要作用是通过不断变异测试用例来影响程序的执行路径，期间涉及到输入处理、覆盖率记录等等，因此可以通过增加一些代码来获取每次 Fuzzing 的输入。

2.2 实现执行窗口

考虑到挖矿的公平性，需要保证每个矿工每次挖矿的物理时间是一致的，这样系统就能比较同样的时间内哪位矿工的贡献最大，可以通过实现一个执行窗口来达到上述目的。该执行窗口以时间为界限，在被测程序执行固定时间后，对本次挖矿所运行的所有输入、对应的执行路径（可能此时并没有完全执行结束）添加随机数后进行哈希运算获得该轮执行窗口中每次执行的哈希值，以便后续处理。即使窗口时间到了但程序并未执行结束，也必须挂起程序，输出挂起前运行的一部分路径，待下一轮窗口接着执行。

3、实现 Proof-of-Fuzzing

共识机制要明确获取奖励和记账权的条件，本课题中获取的条件有两个且优先级不同。条件一：如果某个矿工最先发现了新路径，那么该矿工将获得奖励和记账权；条件二：如果在某个执行窗口中，没有矿工发现新路径，那么则选择拥有最先满足哈希条件（即小于某个哈希值）的矿工赋予记账权；条件一的优先级要高于条件二。如果在该轮执行窗口中，没有任何矿工符合这两个条件中的任意一个，则该轮置空，没有人获得奖励和记账权，系统需要动态调整条件二中满足哈希范围的难度使得下一轮不被置空。

上述的共识机制将替换传统 PoW，以达到将区块链和 Fuzzing 结合的目的，同时还可以激励矿工们发现更多的漏洞。

4、评测工作

在实现了 Proof-of-Fuzzing 后，需要对该工作进行评估和测试，使用现有的不同程序作为 Fuzzing 对象，模拟多个矿工挖矿过程，统计挖掘出的漏洞，以对该共识机制的效果进行评价。

三、研究计划及预期进展

时间	研究内容	预期效果
2023.9-2023.10	搭建一个区块链框架	用 java 实现一个简单的区块链框架，包括工作量证明、P2P 网络
2023.10-2023.12	阅读并理解 AFL 源码，对其进行修改	修改 AFL 源码，获得每次运行的输入和路径
2024.1-2024.3	实现执行窗口	正确获得每轮执行窗口的所有输入、执行路径以及处理后的

2024.4-2024.7	实现 Proof-of-Fuzzing	哈希值 成功将传统 PoW 替换为基于 Fuzzing 的共识机制
2024.8-2024.9	评测工作	生成评估结果

四、指导教师意见

该生通过与课题组成员和老师充分讨论，参考了许多文献，确定了具有一定理论价值和现实意义的课题。本课题研究计划合理，思路基本明确，将区块链的共识算法和模糊测试结合，难度合适，学生能够在预定时间内完成该课题的设计。

同意开题。

签名： _____

日期： 2024 年 3 月 4 日

五、选题报告会记录（着重记录评审专家的质疑问题与研究生的回答要点，以及专家对选题的具体修改意见）

六、论文选题评价结果（请评审专家在相应等级后的“（ ）”内打“√”，并给出评语）

评价结果： 通过（ ）； 修改后通过（ ）； 不通过（ ）

评语：

评审专家小组签名：组长_____

成员_____

时间： 年 月 日