**软院做工程系统的要求：**  
开发一个系统，要有7、8个模块功能

包括3~4个主要功能模块和其他的辅助功能模块

主要功能中要有算法，可以是调用现有算法库，也可以是自己设计的，但一定要有自己设计的能够表明工作的新意，自己设计的算法越深越好，有2、3个算法是自己设计的，不是直接调用库，工作量就很饱满

辅助功能模块比如报告生成、管理模块之类的

有专利或者论文，创新性这块答辩老师就不太会深究，没有的话就会问的比较细

要检查代码，现场跑

Motivation：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。

* 关于论文的背景和意义：

1、问题和提出的解决方案耦合度不高：fuzzing和区块链共识机制是两个不相关的任务，需要在这两者之间建立共生关系，但是目前只觉得是区块链对fuzzing的单向需要，fuzzing并不一定需要区块链。

2、软院其他做系统的论文首先都有一个非常具体的现实问题背景，其次能应用到生活的各种方面，如电商、公安局、公交车、智能客服、工业产线机器人、英汉机器翻译、智能安防小区

感觉目前的这个背景有点薄弱，因为它只是提供给矿工fuzzing的功能，每个节点并不是实际场景中的角色，系统设计的时候就比较单薄，无法给身份做权限管理之类的，我的理解是它只是一个区块链的改版，fuzzing和区块链的联系就是它有一个激励机制而不是因为它有什么亟待解决的问题

3、现在的方法对模糊测试也没有任何改进

4、需要有个针对fuzzing的实际意义，比如机器学习和区块链结合是提供一个开放的**存储机器学习模型和数据集的存储库**，那这个也得提供一个什么，否则的话为什么要选择fuzzing呢

* 题目是否应改为“基于区块链的xxx分布式系统设计与实现”
* 改比特币源码还是自己搭一个呢，如果自己搭的话，功能就要先确定一下
* **数据共识模块功能**：根据要求，目前的基于AFL模糊测试的共识算法只能算一个功能模块的一个算法，还差至少1个

对Proof-of-Fuzzing的设计需要复杂一些

* 是否可以改进一下AFL算法并和AFL做对比实验（代码覆盖率、新路径、漏洞发现），我希望这块可以做得深一点，可以调研一下最近的模糊测试算法论文，不知道这样可行吗？
* 是否可以多加几个fuzzing算法进去，让矿工选择用哪个
* **安全保障，检测恶意节点、保证数据一致性的功能**，这里可以添加一个算法
* **数据同步功能**：区块链基本功能，应该没啥算法，就代码实现
* **难题发布功能：**我觉得需要有难题发布者这个角色，如果没有发布者这个角色，而是由系统发被测代码，那被测代码如何获得，如果随便选择，那这和传统PoW并没有区别，因为这样一来fuzzing的结果就没有意义，找到漏洞然后就没有别的贡献了。
* **生成报告功能：**对每个被测程序生成漏洞报告功能（输入、执行路径、漏洞原因）、矿工查看自己漏洞挖掘报告
* **用户角色：**

**发布者：**发布任务的人，供应商/雇主，如公司需要对某个代码做漏洞检测或者某个学生要检测自己的代码等等，但是企业又涉及到隐私安全问题

**执行者（矿工）**

**存储者（？）：**是否可以添加漏洞存储数据库功能：把发现的漏洞都存下来，提高一下这个系统的价值，但是不知道存下来能干啥，或许可以用来统计高频漏洞？

* 这个系统和只用afl挖漏洞比有什么优势呢？我目前的理解是它可以提供一个大平台让大家一起fuzzing，找漏洞的效率更高，并且可以记录漏洞生成漏洞报告。矿工费半天劲fuzzing是为了拿钱，但是发布者为什么要选择这个系统发布任务，自己用AFL又不是不行。
* 怎么把前一个区块的hash值和挖矿结合起来
* 是否满足共识机制中“很难解决，很容易验证”的特点呢？

Fuzzing到底算不算一个很难解决的数学问题呢？前人的机器学习训练、矩阵运算、搜索Cunningham链和素数的双孪生链、正交向量、3SUM或全对最短路径问题等等可以算作难解决的问题，但是fuzzing在刚测的时候找到新路径还是非常容易的，不确定算不算复杂工作。

* 挖矿难度如何调整呢
* 获胜条件：

如果多个矿工，每个人都发现了新路径，但是由于每次只有一个人有记账权，那其他人也可能fuzzing出不一样的新路径，比如一个人发现了5条新路径，另一个人发现了4条新路径，他们发现的路径中有3条重合，4条的那个人先发现新路径，那把记账权给他，那发现5条的那个人怎么处理呢？同理，把记账权给发现5条的人的话，那发现4条的人先发现的，而且他还发现了1条不一样的，这怎么算呢

两个人都发现了漏洞，只是时间不一样，另一个晚一些的矿工发现的漏洞也是有价值的，并不是传统PoW一样算力消耗了就扔掉了，他发现的漏洞也是要记录下来的，那这个奖励怎么分配？同理，没有获得记账权却发现了不一样的新路径的矿工，怎么分配奖励

（发现漏洞、新路径的数量大小、新路径的时间先后、执行路径总条数）

* 所有人每次都fuzzing同一个代码吗？

发布者们会有很多需要fuzzing的代码，后面发布任务的人要排队，但是每次执行fuzzing时间都不短，后来的人自己都不知道要等多久？这对发布者来说非常不合理，那发布者为什么要来这个平台发布任务呢，对他也没有好处。

但是如果让一部分人执行这个代码，一部分人执行那个代码，这又有违挖矿的公平性，无法给记账权

--共识协议，面向公链，以区块链为背景，把算力用在fuzzing，既可以挖掘漏洞，又可以记账，奖金问题（google、微软大型软件有做fuzzing的需求，是现金）

* Fuzzing终止条件如何确定

因为afl-fuzz永远不会停止，所以何时停止测试很多时候就是依靠afl-fuzz提供的状态来决定的

* 测试用例的输入文件怎么给？不能大家都用一样的输入，如果要挖出漏洞的话，种子就很重要，那种子文件怎么生成呢？如果让矿工自己生成，是不是得先让矿工知道这个代码的输入应该是什么？但是这样会不会涉及到隐私问题呢（如果发布者是企业、研究所等）？
* 如何获得连续的任务供应？
* 评测工作如何开展呢

**软院毕业论文（来源：学校图书馆数据库）：**

**基于区块链的物联网数据安全同步系统**

随着物联网产业应用的逐步深化，物联网技术在各行各业扮演者重要角色，然而物联网数据来源众多、地域分散、传输信道复杂而导致的数据共享困难和真实性难以保证等问题，并且数据极易受网络波动的影响，在处于网络信号极差的环境下，会造成数据丢失、数据残缺问题。因此且企业为了保证提供给用户服务的连续性，往往需要对服务器数据进行备份，保障数据的安全，从而确保当网络或服务器出现故障时，能够在短时间内恢复业务的提供。数据备份是容灾的基础，但建立一个标准的数据备份中心价格高昂，而且用于数据备份的服务器使用率低，因此使用其他企业数据中心闲置的存储空间成为解决该问题的一种方法。但在企业间对数据进行协同存储备份过程中，存在企业之间相互不信任、数据隐私难以保障的问题，为了保证在不同数据中心间数据备份的安全性，引入区块链技术为解决该问题提供了一种可行的技术手段。 本文使用区块链技术将不同企业的数据中心相互连接，共同构成区块链网络进行数据的协同备份。由于多个节点共同参与到区块链的管理中，链上数据无法被恶意节点篡改，真实可靠，当某节点出现故障时，也能从其他节点重新获取完整数据。

**1、设备与节点身份认证模块**：通过数字签名、路由表同步与安全证书的机制对系统内的各个物联网设备与节点进行身份认证

**2、数据安全通信与校验模块**：提出基于校验码的对等节点安全通信方法，保障数据在传输过程中的安全性

**3、节点恶意行为检测模块**：提出基于信誉值模型的节点行为检测方法，通过对节点信誉值的计算检测恶意节点

**4、区块生成与同步模块**：以安全通信的方式将生成的区块向系统广播，对全局验证完毕的区块进行并发上链与事务数据持久化

**5、数据一致性保障模块**：提出基于粗粒度默克尔树的链状态生成方法，通过对链状态的比较，快速检测区块链的完整性，向其他节点请求和替换异常区块，保障数据的一致性

**6、用户交互模块**：使用Django框架实现后端系统，为用户提供设备与节点认证功能、节点状态展示功能与数据可视化功能

**7、对系统进行功能性测试与非功能性测试**。测试结果表明，系统能够达到平均3400条/秒的数据处理速度，具有针对物联网数据的高效处理能力，能够保障各个节点之间的数据安全同步。

**基于区块链的农产品溯源系统的设计与实现**

1. fabric联盟链作为区块链底层
2. PBFT算法作为共识机制，针对传统PBFT算法在农产品溯源的场景下存在的性能上的不足，设计了一种根据节点硬件性能与历史共识表现计算节点可信度的方法，并基于该方法分别对算法的一致性协议与视图切换协议的流程进行了优化，设计了改进后的基于节点可信度的IPBFT算法
3. 溯源企业注册审核功能
4. 溯源信息上链功能
5. 溯源信息查询功能
6. 用户角色管理功能
7. 日志管理功能
8. 功能的测试以及性能的测试

**基于区块链的物联网数据协同存储备份系统的设计与实现**

1. 为了提高区块链系统性能，本文采用了基于星际文件系统的数据日志区块存储方法，结合Gossip网络，实现了更快更高效的数据同步管理
2. 为了保证节点在恶劣网络环境中系统的正常使用，本文提出了基于链式哈希回溯的漏块补全方法，使得节点在网络波动或掉线重连后能够补全丢失区块，达成各节点间数据的一致性。
3. 数据同步功能：系统的单个节点与传统数据备份中心一般，在对数据进行增删改查等操作后，其余节点也会进行相同的操作，能够完成数据同步。
4. 身份认证：统为用户提供一体化的密钥分配和身份认证控制，严格管理用户的数据访问等权限
5. 日志：每次对数据的操作记录都会作日志上传到区块链中，由其余节点同步，恶意用户无法对备份的数据进行伪造以及删除。

6、适用于数据协同备份的多子链并行区块链解决方案，实现了多节点间的协同数据备份，数据的安全保护，数据的溯源查询以及节点用户的身份控制，在充分利用了闲置资源的同时，进一步提高了数据的安全性。

7、系统架构设计和编码实现，在去中心化的环境中对系统进行了测试，测试结果表明该系统能够完成不同节点之间的数据备份与数据保真，同时相较于传统的区块链架构有更高的上链性能，具有较高的实用价值。

**基于区块链的园区级虚拟电厂分布式交易平台的设计与实现**

用户准入管理模块、信息总览模块、设备管理模块、电力交易模块和设备数据采集与控制模块

1. 针对交易信息难以追溯问题，结合Hyperledger Fabric特点设计本文区块链网络、基于Fabric CA的用户身份认证方法和基于IPFS的大文件存储上链方案。
2. 针对交易机制不明确问题，将平台交易类型划分为预交易、自由交易、撮合交易和利益分配交易；对虚拟电厂内部电费惩罚机制和利益分配机制进行建模，并设计为智能合约。
3. 搭建区块链网络和IPFS私有化集群，将智能合约部署到区块链网络上，使用ASP.NET Core MVC框架实现本文平台并对其进行完整全面的测试。
4. 通过中国西部科技创新港园区实际数据测试结果表明，本文提出的电费惩罚机制和利益分配机制，能很好地调节并激励用户参与虚拟电厂的电力交易。同时，基于区块链的园区级虚拟电厂分布式交易平台在满足基础功能的前提下，做到身份认证、交易信息可追溯。

**基于分层模型的联盟链共识算法研究与应用**

作为近十年工业界和学术界的明星技术，从比特币和以太坊为代表的公有链平台，到现在落地最为广泛的联盟链系统，区块链始终紧随时代的步伐。相对于公有链平台，联盟链侧重于吞吐量更大的共识机制，如底层的PBFT算法。但随着服务流量的不断提升，当前的拜占庭容错算法已然不能满足日益增长的高服务质量的需求。 本文设计了一种主要应用于联盟链的基于分层模型的拜占庭容错算法，并以超级账本Hyperledger下的 Fabric平台作为框架在联盟链中做了迁移应用。**新的共识算法**延续了Raft算法中节点三种角色状态的设定，吸收了PBFT算法的数字签名技术进行拜占庭验证，并参考Gossip传播协议的特点将集群划成分层的模型，在每层内根据日志的完备性选出代理节点，利用代理节点优先参与主节点选举，缩短选举的时间消耗，同时在层与层传播过程中进行拜占庭验证；在Fabric平台的迁移应用中，验证区块环节增添了缓存交易验证，提高了系统的安全性。本文的主要工作内容有：首先研究了大量的共识算法的设计特性，参考部分算法以及协议的特性设计了一种**基于分层模型的共识算法**，在本地利用单机模拟多节点通信完成新共识算法的性能测试，同其他主流共识算法做了**相关性能指标的对比**；然后研究了联盟链开源平台Fabric的设计模式，结合Fabric可插拔的设计理念，根据平台要求和新的共识算法的设计机制实现了Fabric共识模块提供的系列接口，完成了**新的共识算法在联盟链的迁移**，将**新的分层共识机制与Fabric内置的共识机制做了性能的对比**；最后利用Fabric新的共识机制，结合实际的应用场景，设计了包含两条链的相关的区块链网络，实现一个**应用于二手房交易的原型**并完成了相关的测试，进一步验证了之前的工作。 本文的实验环节验证了新共识算法的特性，实验得出，在应用了新的共识算法的满足节点数目要求的集群中，主节点选举环节得到了优化，共识效率较高。同时迁移到区块链网络中的共识机制也有着较好的性能表现。

粗略统计22年毕业的

共400篇论文，做系统的大概138人，12人延期答辩

做算法的大概262人，7人延期答辩