北京理工大学

本科生毕业设计（论文）开题报告

学　　院： 徐特立学院

专　　业： 徐特立英才班

班　　级： 30081702

姓　　名： 邢智博

指导教师： 张子剑

二〇二一年一月九日

毕业设计（论文）开题报告评审表

<table>
<tr><td colspan="2">姓名</td><td></td><td>学号</td><td></td><td>班级</td><td></td><td>专业</td><td></td></tr>
<tr><td colspan="2">导师</td><td></td><td colspan="2">校外导师（职称）</td><td colspan="2"></td><td>校外导师单位</td><td></td></tr>
<tr><td rowspan="3">论文选题</td><td>题目名称</td><td colspan="7"></td></tr>
<tr><td>题目性质</td><td colspan="7">软件开发（　　）　　理论研究（　　）
工程设计（　　）　　技术科学研究与工程技术研究（　　）</td></tr>
<tr><td>题目来源</td><td colspan="7">结合科研　　（　　）　　结合生产实际（　　）
结合实验室建设（　　）　　自拟题目　　（　　）</td></tr>
<tr><td rowspan="4">评审组成员</td><td>姓　名</td><td colspan="2">职　　称</td><td colspan="3">工作单位及职务</td><td colspan="2">签　　字</td></tr>
<tr><td></td><td colspan="2"></td><td colspan="3"></td><td colspan="2"></td></tr>
<tr><td></td><td colspan="2"></td><td colspan="3"></td><td colspan="2"></td></tr>
<tr><td></td><td colspan="2"></td><td colspan="3"></td><td colspan="2"></td></tr>
<tr><td>评审意见</td><td colspan="8">（含：选题意义；选题是否满足毕业要求；技术方案是否可行；进度安排是否合理等）</td></tr>
<tr><td>成　绩</td><td colspan="8"></td></tr>
<tr><td colspan="9">评审组长签字：　　年　　月　　日</td></tr>
</table>

注：成绩以“合格”“不合格”记；评审组长为高级职称人员。

1. 毕业设计（论文）选题的内容

我们正在步入数字化时代，根据 IWS(Internet World Stats) 提供的数据，截至 2020 年 6 月，全球互联网用户数量达到了 46.48 亿人，占世界人口的 59.6%，意味着全球有超过一半的人口在使用互联网并产生数据，同时全球用户产生的数据量是巨大的，根据 IDC(International Data Corporation) 的预测，到 2025 年，全球的数据量将高达 175ZB，即 10^{21} 字节。另一方面，企业和公司采用各种不同的面向数据驱动的算法来提升自己的服务质量，拓宽自己的服务领域，这使得数据本身成为了一类特殊的商品，具有极大的价值，同时初创公司往往可以通过购买数据来快速构建起有效可靠的服务，但是数据的生产者和使用者往往属于不同的组织，使用者需要数据，生产者提供数据，这就产生了数据的供求关系。以上种种原因使得数据交易逐渐成为市场的必然需求。

然而，通过对现有的数据交易模式进行调研，本文发现其中存在着一些问题和安全隐患。目前的交易模式大多是买方和卖方分别将钱款和数据提交给第三方进行委托保管，若双方确认进行交易，则交易平台完成钱款与数据的交换并结束交易，双方都信任该第三方会严格执行交易流程。可以看出上述交易模式强烈依赖一个可信第三方，但是可信第三方假设在方案设计中是不合理的。一方面，在实际场景中难以对第三方的行为进行约束或验证，无法确定其是否有勾结或窃取数据的行为，另一方面，可信第三方的引入也增加了方案实现的难度。而在抛弃可信第三方这个假设之后，会产生如下问题：其一，如何保证数据不被交易平台窃取？数据作为一种特殊的商品，具有“看过即拥有”的属性，要保证交易平台无法在交易过程中窃取拥有者的数据，需要对数据进行加密，但卖方也需要向买方共享数据的解密密钥，如若把解密密钥当作一种数据商品，则又引入了数据安全共享的问题；其二，如何保证交易能够公平可靠地进行？交易的公平性是指在交易过程中不会出现某一方在提供钱或数据后未收到应得的报酬，现有的交易模式中交易的公平性完全依赖于交易平台，而交易平台可能与某一方参与者勾结，导致另一方参与者利益受损。其三，如何实现高效便捷的交易？现有的撮合交易模式下需要数据的卖方在本地存放数据，并且在每次交易时发送数据，给卖方带来了较大的存储以及通信成本。因此，如何在不引入可信第三方假设的情况下构建一个数据公平交易、安全共享的外包数据交易系统是一个亟待解决的问题。

本课题研究基于智能合约与数据安全共享的外包数据公平交易系统，利用智能合约自动控制交易流程，利用安全共享方案保证数据在交易流程中不外泄，研究一种不依赖于可信第三方的数据公平交易系统，最终，通过仿真实验对该系统的效果进行验证和评价分析，验证方案的有效性。

2. 研究方案

2.1 本选题的主要任务

1. 学习智能合约的基本概念和工作原理，了解其与一般程序语言的区别和局限性；
2. 调研现有的数据安全共享方案，结合外包数据场景进行比较；
3. 设计基于智能合约和数据安全共享方案的外包数据公平交易系统；
4. 实现外包数据公平交易系统，验证数据安全共享方案的可行性；
5. 完成毕业设计论文并提交相关程序和文档。

2.2 技术方案的分析、选择

本部分首先对该系统的功能性需求进行分析和总结，之后分别对每个需求涉及到的技术进行详细的分析与选择。

2.2.1 系统功能

通过对任务书进行分析并调研了一些现有的数据交易系统后，本文认为一个外包数据公平交易系统应包含的功能点如下：

1. 安全数据外包，即数据的卖方可以将待出售的数据委托给外包存储方进行售卖，这需要卖方事先对数据进行加密，以避免外包存储方非法窃取数据，同时为使得卖方能够对数据进行一次加密并上传后多次售卖，可以使用代理重加密技术构造加密方案，使得卖方不必为不同的买方对数据进行反复加密，实现数据外包；
2. 数据搜索，即数据的买方可以根据关键词在外包存储方存储的数据中进行搜索，最终选择其中符合条件的数据进行购买，这需要加密方案具有一定程度的可搜索加密性质；
3. 公平交易，即保证交易的公平性，避免出现某一方在提供了金钱或数据后没有得到应得的数据或金钱的情况发生，这需要设计一个能够保证公平性的交易流程，同时引入智能合约对交易流程进行自动控制，确保交易流程的正确执行。

2.2.2 代理重加密

代理重加密适用于包含外包存储方的三方数据交易场景，该技术能够安全地进行数据内容的传递，而不向第三方泄露数据的额外信息。代理重加密的概念是由Blaze、Bleumer 和 Strauss[1] 三位密码学家在 Eurocrypt98 会议上首先提出的。在代理重加密中，重加密的角色由一个半可信代理者扮演。重加密方案要求代理者必须拥有一个由委托者授权的、针对被委托者的密文转换密钥，即代理重加密密钥。委托者首先将第一次加密后的密文上传给代理者，称为二级密文。当需要将密文共享给被委托者时，委托者授权代理者进行重加密，代理者使用密文转换密钥将密文重加密后发送给被委托者，转换后的密文称为一级密文。被委托者使用自身密钥可以解密重加密后的密文。在这个过程中，代理者无法获知任何关于密文中对应明文的任何信息。

该技术与本文的应用场景十分契合，我们认为外包存储方是一个理性参与者，卖方对密文加密后将该二级密文上传至外包存储方，在有买方提出购买需求后，卖方生成一个针对该买方的代理重加密密钥，并授权外包存储方通过该密钥进行重加密，而后外包存储方将该一级密文发送给买方，买方可以使用自身密钥解密。在该过程中，既满足了卖方没有对数据进行反复加密，实现了数据外包，也满足了外包存储方无法得到该密文对应的明文的任何信息，实现了数据安全。

2.2.3 可搜索加密

可搜索加密是在数据加密的情况下对数据进行搜索。2000 年，Song 等[2] 首次提出了可搜索加密的概念。作为一种新型的密码原语，可搜索加密技术使用户具有在密文域上进行关键词搜索的能力。数据以密文方式存储在云服务器上时，利用云服务器的强大计算能力进行关键词的检索，而不会向服务器泄露任何用户的隐私。可搜索加密的工作流程如下: 首先数据拥有者把加密的文件数据以及相关的关键词密文上传到云服务器, 然后用户利用私钥生成搜索陷门, 并把该陷门信息发送给云服务器, 云服务器通过使用该陷门信息搜索到用户感兴趣的数据, 并把数据发回给用户。该技术实现了用户在不可信赖云服务器环境下进行快速有效的密文关键词检索, 同时不泄漏任何关于数据的信息。

该技术与本文的应用场景十分契合，我们认为外包存储方是一个理性参与者，卖方根据私钥和买方的购买请求生成搜索陷门，外包存储方利用该陷门搜索相关关键词的密文，从而确定买方需要的数据。在该过程中，满足了买方可以根据关键词对数据进行筛选，实现了数据交易的基本要求。

2.2.4 智能合约

为了保证交易的公平性以及交易流程的精确执行，本文引入了智能合约这一技术。智能合约最早由 Szabo[3] 在 1994 年提出，他认为智能合约是一套以数字形式定义的承诺，包括合约参与方可以在其上执行这些承诺的协议。其本质是通过自动执行若干个交易方之间确定的合约，从而消除对合约条款的争议，自动实现资产的转移。智能合约具有去中心化，自动执行等特点，在区块链平台上能够较好实现。目前与智能合约技术结合的区块链平台主要有以太坊和 Hyperledger Fabric。

通过对以上两个区块链平台进行初步调研，本文拟选择以太坊上的智能合约对交易流程进行控制，原因有三：其一，以太坊平台相对于超级账本而言较易部署与使用，使用以太坊可以更快地进入智能合约设计与开发的工作；其二，以太坊是公有链，超级账本是联盟链，以太坊的公有链环境更贴合本文的假设，即买方与卖方没有交易前的了解，没有建立信任关系；其三，以太坊智能合约提供了一些预编译合约，这些预编译合约中包含一些密码学相关的函数，如 sha256，ripemd160，以及双线性配对运算操作，使得我们可以以较低的实现代价和调用代价去使用以上函数来实现对于交易流程的控制，其中双线性配对是代理重加密中常用的工具之一。

智能合约可以实现的操作包含：暂存买方支付的金钱，暂存卖方提供的数据，在双方都进行了相应的提供后向双方发送数据与金钱。以上操作保证了交易的原子性，避免了某一方未获得相应的钱或物的情况的发生，同时智能合约对交易流程进行了自动控制，在合约设计合理的情况下交易流程不会受到外部非法操作的干扰，进而保证了交易的公平性。

2.2.5 系统架构

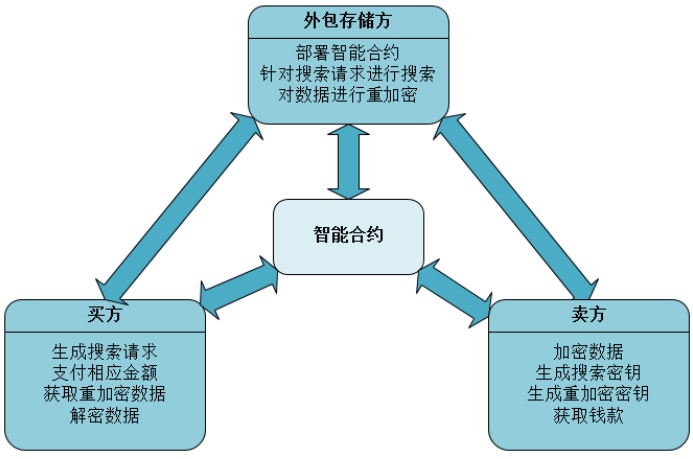

图 2-1 系统整体结构

表 2-1 硬件、软件环境

	指标	版本参数
硬件环境	处理器	Intel i7-7700HQ
	内存	8 GB
软件环境	操作系统	Windows 10
	solidity	0.4.26
	Python	3.7.2

通过以上分析，本文初步确定了系统架构，并给出如图 2-1 的结构图。本系统中共有三方，分别是出售数据的卖方，购买数据的买方，以及提供外包存储以及其他相应服务的外包存储方。卖方的能力有加密数据，生成搜索密钥，生成重加密密钥，获取应得钱款；买方的能力有生成搜索请求，支付相应金额，获取重加密数据，解密数据；外包存储方的能力有部署智能合约，针对搜索请求进搜检索，对数据进行重加密。这三方通过以上能力互相交互，同时也与智能合约进行交互，通过智能合约控制交易流程。

2.3 实施技术方案所需的条件

所需的软硬件环境在表 2-1 中列出，包括处理器，内存，操作系统，以及使用的语言及其版本。

2.4 存在的主要问题和技术关键

目前存在的主要问题有以下几点：

1. 如何设计基于智能合约的公平交易流程。一个公平的交易流程应能保证在理性参与者的假设下，不会有参与者的利益受到损害。由于使用了智能合约对交易流程进行自动控制，所以需要对可能出现的攻击进行尽可能全面的列举，考虑以下场景：1. 某一方参与者攻击智能合约以改变交易流程，损害其他参与方的利益；2，外包存储方侵吞买方应得钱款；3. 外包存储方拒绝给出或给出错误的重加密数据。需要考虑到的攻击手段以及方式包括但不限于以上场景。

2. 如何构造有效的支持代理重加密与可搜索加密的加密方案。为了同时支持密文搜索和解密授权这两个功能，Shao 等人[4] 提出了带关键字搜索的代理重加密方

案 (PRES)，构造了一个在随机预言模型下可证明安全的双向 PRES 方案。他们将具有关键字搜索性质的公钥加密和代理重加密方案结合，由于该方案的重加密与加密算法使用的是同一套加密方法，因此在安全性上有所缺陷。同时该方案的计算效率也十分低下。在该系统中，我们不仅要求支持代理重加密与可搜索加密的加密方案是安全且有效的，同时也需要该方案能够与数据公平交易流程有机结合，确保流程的公平性。

2.5 预期能够达到的研究目标

实现一个基于智能合约和数据安全共享方案的的外包数据公平交易系统，并完成相应的毕业论文。

3. 课题计划进度表

(2021.1-2021.2)：学习并掌握基本的密码学原语和安全知识，以及智能合约相关知识；

(2021.2-2021.3)：阅读相关文献，对现有的数据安全共享方案进行研究，结合外包数据场景进行分析，设计基于智能合约和数据安全共享方案的外包数据公平交易系统；

(2021.3-2021.4)：实现基于智能合约和数据安全共享方案的外包数据公平交易系统，完成对系统的测试与实验；

(2021.4-2021.5)：完成本科生毕业设计（论文）文献翻译，完成毕业论文并提交代码及相关文档。

4. 参考文献

[1] BLAZE M, BLEUMER G, STRAUSS M. Divertible Protocols and Atomic Proxy Cryptography[C]//Springer. International Conference on the Theory and Applications of Cryptographic Techniques, May 31 - June 4, 1998, Espoo, Finland. Heidelberg: Springer-Berlin, 1998: 127-144.

[2] SONG D X, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[C]// IEEE Computer Society. 2000: 44.

[3] SZABO N．Formalizing and securing relationships on public networks [EB/OL]. https://journals.uic.edu/ojs/index.php/fm/article/view/548/469.

[4] SHAO J, CAO Z, LIANG X, et al. Proxy re-encryption with keyword search[J]. Information Sciences, 2010, 180(13): 2576-2587.