# 区块链技术调研报告

当前正在筹备、规划有关区块链研究实验室（研究中心）的工作，为了更好地评估有关设想，利用假期一个多月时间，对区块链相关技术做了一个较为深入全面的调研，主要是关于密码学最新发展以及隐私计算领域。在这个过程中又特别留意和关注到有关区块链教学问题。本文对调研内容和相关结论做一个总结报告。

由于需要面向科研和教学双重任务，一些优秀的在线公开课就非常有意义，调研中针对隐私计算这个领域问题认真听了麻省理工、斯坦福、哈佛的一些优秀的公开课，主要是密码学、密码学相关的数学课程以及一个反映密码学最新成果的国际性在线论坛课程[1]。学术论文部分的阅读主要针对全同态加密算法，从Gentry开创性的博士论文到最近阿里研究者的改进做了比较全面的了解。开源是了解算法技术落地最重要途径，有关全同态加密的几大开源都有涉及。

## 有关隐私计算的调研

### 1.1 选择隐私计算方向做调研的基本考虑

大概基于以下几点：

* 隐私计算目前处在区块链相关技术领域的前沿，并且同时成为了国家级的技术战略重点之一，可参见中国信通院出的《隐私保护计算与合规应用研究报告》[2]。开展区块链研究，隐私计算是一个需要密切关注的领域。
* 隐私计算的可能的应用前景和影响非常巨大和深远，不仅仅是区块链应用的关键技术之一，更主要的是涉及未来大数据、深度学习、云计算的方式。

关于这两点在调研之前已经比较清楚，通过调研对这个领域的发展历史、当前不同技术选择、相关技术和算法的成熟度、学术研究的状况和业界的行动等方面做了比较全面的了解，以便于从我们自身条件出发，从哪个角度、何种方式开展工作。

### 1.2有关全同态加密的调研

全同态加密算法和技术被称之为隐私计算技术的圣杯，所以对此投入相当时间对算法和技术做了深入学习和了解。

#### 1.2.1关于算法

全同态加密技术的密码学基础部分主要是lattice cryptography，格加密理论，主要是通过几个公开课《Lattice Theory》《The second bar Ilan winter school on cryptography: lattice-based cryptography application》[3]了解这个领域的一些进展状况。

全同态加密的算法突破是2009年斯坦福博士生gentry的论文《A FULLY HOMOMORPHIC ENCRYPTION SCHEME》，初步了解了一下算法的主要机制。

* 它是一种基于理想格（ideal lattice）的全同态加密算法
* 算法的核心是用一种bootstrap机制解决了噪声放大问题而使得circuit的深度可以不受限，达成了全同态加密。

在对全同态加密的密码学基础和主要算法做了比较深入的了解之后，又对算法的近期进步做了一些了解。Gentry的ideal lattice算法诞生已有十多年，陆续有比较重大的改进，这个改进被认为经过三代。

#### 1.2.2业界的行动

目前在全同态加密领域采取了比较大的行动的大公司主要是：IBM、谷歌和微软，其中Gentry在领导IBM的研发团队[4]，谷歌在今年刚刚开源的它的C语言全同态编译迁移器[5]，这个编译工具是基于IBM全同态引擎HELib。微软也有自己的开源产品[6]；国内公司是阿里巴巴在开展相关研究[7]，并有自己的开源[8]。

#### 1.2.3关于全同态加密调研的结论与考虑

大概有如下结论性的认识：

1. 应用前景巨大，影响深远，因为全同态可以被认为是隐私计算终极意义上的算法技术（关于这一点与TEE的比较部分会提到）。所以毫无疑问是学术热点。
2. 在算法效率上虽经十多年提升，目前距离实用仍然遥远。这一点需要特别值得注意。

所以对于全同态加密这个方向，立足于学术研究合适，从应用应用、实用的角度，这是一个需要观望和等待的领域

### 1.3可信计算（TEE）的调研及结论

目前隐私计算的解决方案大致有三大类：

* 全同态加密FHE
* 多方安全计算MPC
* 可信计算TEE

其中MPC方法有特定场景，通用性不足；全同态加密距离实用还是一个未知的问题。重点对TEE技术目前的发展，特别是在业界的使用做了比较全面的了解。我的结论性认知有以下几点：

1. 云平台服务选择了基于芯片的tee技术来做私密计算，谷歌用的AMD有虚拟机[9]，阿里用的Intel sgx没有虚拟机[10]。
2. tee跟fhe的本质区别在于，fhe是真正trustless的，但tee不是。tee是把信任转移到芯片厂商。芯片厂商比云平台服务提供者作弊的动机要小很多，小到足够让用户放心
3. 其中的逻辑是，tee虽然做不到去中心化，但是云平台服务这种模式本来就是高度中心化的，在这个语境下fhe的去中心化也就没有独立的价值

## 有关区块链数学基础教学的调研和思考

之前的区块链课程规划中规划了《区块链的数学基础》这门课，在了解全同态加密的过程中需要了解密码学的一些新进展，比较全面完整地听了几门相关课程，主要有

《Number Theory》[11]

《[Introduction to Cryptography](https://www.youtube.com/watch?v=2aHkqB2-46k&list=PL6N5qY2nvvJE8X75VkXglSrVhLv1tVcfy)》[12]

《[Abstract Algebra](https://www.youtube.com/watch?v=VdLhQs_y_E8&list=PLelIK3uylPMGzHBuR3hLMHrYfMqWWsmx5)》[13]

[《Theory of Computation](https://www.youtube.com/playlist?list=PL7HjUNIdk93ThXvz2Oa_g30Jt3Owwm4HZ)》[14]

区块链涉及到的数学知识面非常广泛，以上课程都有涉及到，但是针对职业本科甚至专科教育，独立开设以上数学课程显然是不现实的，这是考虑规划《区块链数学基础》这门课程的原因，它无疑是区块链教学最核心的课程。而且经过调研，目前尚未见到有类似教材出现。通过学习、分析以上课程的内容，对《区块链数学基础》这门可得结构有了一个大体把握：

1. 数论部分，基本模运算；欧拉函数、欧拉定理的内容
2. 基础的群论，特别是循环群的内容
3. 计算理论中的复杂度理论的基础部分
4. 基础密码学公钥体系中的RSA、离散对数、椭圆曲线加密的内容，更高级内容，格密码学不适合包括在内。
5. 共识算法部分

以上内容基本能够形成《区块链的数学基础》的内容结构。后续可以考虑通过讲座的方式试水，评估可行性。

## 零知识证明

零知识证明部分在本次调研之前已经做过比较充分的工作，并且出了一篇论文目前发在中科院预印平台上和arxiv预印平台上[15, 16]，后续考虑选择合适学术期刊投稿。这篇论文对区块链技术与传统应用结合的方式、核心问题、业务模型影响以等问题做了全面深入的分析，提出了比较完整的区块链产业应用的方法论，可作为学术研究的核心主张和路线。其具体内容在此不赘述。

## 小结

一个非常重要的认知是，当前谈到区块链技术，其实已经不限于构建区块链本身的技术，更重要的是，由于区块链的出现和巨大影响，有一些被长期研究但在应用领域并未产生像样影响的技术板块被快速激活，其中包括零知识证明和同态加密计算领域。对于零知识证明，我们已经基本确定作为实验室主体方向之一[15]；隐私计算部分相当前沿，一方面对研究团队的数学、密码学算法能力有相当高的要求，另一方面，虽然在基础算法上突破已有12年之久并经历次效率改进，目前仍然无法断言何时达到实用化程度。这一点是需要本着大胆设想、小心求证的态度加以认真考量的。本次调研对这个方向的基本情况提供一个较为全面的了解和分析，以供进一步的探讨和决策之用。简单地说，有利于提升科研水平与层次，但充满挑战。

有关区块链数学基础教学这个问题同样存在一个需要充分认识的问题，而且这个问题与刚刚谈到的区块链技术激活问题相关：在IT信息化迄今为止的几十年时间里，业界的主要活动集中在各个行业业务过程信息化、自动化这个初级任务上，而区块链和大数据的出现及其广泛影响，意味着业界已经逐渐超越既有的、表层业务信息化，进入了一个用IT技术颠覆既有业态的阶段，这个趋势极为明显。这件事情对计算机教育意味着什么？密码学（以及其数学基础）以前其实是局限在一个IT行业的特殊工种：安全人员的知识结构要求中，一般开发人员对密码学的接触限于签名与HTTPS，但是以上谈到的IT任务的转换变迁，使得密码学和相关数学基础成为从业者知识结构中必备的部分。这一点对计算机教育的影响是深远的，国内高校并未充分意识到，清华本科计算机教育还在把数论当作选修类课程。我们是否有必要在这方面有所作为，也是一个需要认真考量的问题。

## 参考资料

1. *The BIU Research Center on Applied Cryptography and Cyber Security*. Available from: <https://www.youtube.com/channel/UCQlhkGqNiyEwJO8HN48SLTw>.

2. 中国信通院. *隐私保护计算与合规应用研究报告*. 2021.3; Available from: <http://www.caict.ac.cn/kxyj/qwfb/ztbg/202104/P020210401408574284747.pdf>.

3. Security, T.B.R.C.o.A.C.a.C. *The 2nd Bar Ilan Winter School on Cryptography Lattice- Based Crypography and Applications*. 2016.7; Available from: <https://www.youtube.com/playlist?list=PL8Vt-7cSFnw2OmpCmPLLwSx0-Yqb2ptqO>.

4. IBM. *HELib*. Available from: <https://github.com/homenc/HElib>.

5. Google. *Google FHE Transpiler*. Available from: <https://github.com/google/fully-homomorphic-encryption>.

6. Microsoft. *Seal*. Available from: <https://github.com/microsoft/SEAL>.

7. Lu, W.-j., et al., *PEGASUS: Bridging Polynomial and Non-polynomial Evaluations in Homomorphic Encryption.* IACR Cryptol. ePrint Arch., 2020. **2020**: p. 1606.

8. Alibaba. *OpenPEGASUS*. Available from: <https://github.com/Alibaba-Gemini-Lab/OpenPEGASUS>.

9. cloud, g. *机密虚拟机和 Compute Engine* Available from: <https://cloud.google.com/compute/confidential-vm/docs/about-cvm>.

10. 阿里云. *ACK-TEE机密计算介绍*. 2021.4; Available from: <https://help.aliyun.com/document_detail/164536.html>.

11. Penn, M., *Number Theory.*

12. Paar, C. *Introduction to Cryptography*. 2014.6; Available from: <https://www.youtube.com/playlist?list=PL6N5qY2nvvJE8X75VkXglSrVhLv1tVcfy>.

13. Opencourse, H. *Abstract Algebra*. 2014.6; Available from: <https://www.youtube.com/playlist?list=PLelIK3uylPMGzHBuR3hLMHrYfMqWWsmx5>.

14. Simonson, S. *Theory of Computation*. 2015.4; Available from: <https://www.youtube.com/playlist?list=PL7HjUNIdk93ThXvz2Oa_g30Jt3Owwm4HZ>.

15. Bai, Y., *Zero-knowledge Based Proof-chain: A methodology for blockchain-partial system.* ChinaXiv, 2021.

16. Bai, Y. and L. Luo, *Zero-knowledge Based Proof-chain--A methodology for blockchain-partial system.* arXiv preprint arXiv:2107.14142, 2021.