基于可加密搜索的云存储方案

阅读文献：

[1]: Song, D. X., Wagner, D., & Perrig, A. (2000). Practical techniques for searches on encrypted data. In *Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on* (pp. 44-55). IEEE.

[2]: Kamara, S., & Lauter, K. (2010, January). Cryptographic cloud storage. In *International Conference on Financial Cryptography and Data Security* (pp. 136-149). Springer, Berlin, Heidelberg.

[3]: Li, J., Wang, Q., Wang, C., Cao, N., Ren, K., & Lou, W. (2010, March). Fuzzy keyword search over encrypted data in cloud computing. In *Infocom, 2010 proceedings ieee* (pp. 1-5). IEEE.

[4]: Awad, A., Matthews, A., Qiao, Y., & Lee, B. (2018). Chaotic searchable encryption for mobile cloud storage. *IEEE Transactions on Cloud Computing*, *6*(2), 440-452.

介绍：

随着云计算的迅速发展，和智能手机的普及，越来越多的用户选择将大规模数据迁移到云服务器中，以此来节省本地数据管理开销和系统维护开支。由于数据脱离了用户的物理控制，云端服务器管理员和非法用户（如黑客）可以在未经用户授权的情况下访问数据，这将造成数据信息和用户隐私的泄露，而现实中已经发生过这类事件。为了数据的安全性，应该将数据加密后上传云端。但当用户需要寻找并下载特定数据时，将会遇到在加密数据中进行搜索的问题。

为了很好地解决这个问题，基于可加密搜索(searchable encryption)的云存储是较好的选择。在可搜索加密机制下用户需要搜索某关键词时将该关键字的搜索凭证发给云服务器，云端将接收到的搜索凭证对每个文件进行试探匹配并返回结果。用户接收到搜索结果后，对结果进行解密即可。在这种机制下，云端服务器将无法获得除访问模式以外的任何明文数据，可以保证安全性。但目前商业化低，没有很多部署案例。

在2000年，[Song D等人[1]](#a)就提出了一种基于对称算法的可搜索加密的实现方法，开创了实用性可搜索加密的先例。2010年，微软公司的[Kamara, S等人[2]](#b)实现了基于对称密码的密文搜索功能，并针对云计算环境做了一些具有现实指导意义的工作。[Li, J.等人[3]](#c)也研究了支持模糊查询的可搜索加密，最近2018年，[Awad, A等人[4]](#d)提出了针对移动数据的混沌可搜索加密算法，以达到支持模糊查询的目的。

我们想要基于阿里云平台提供的云服务器ECS服务实现一个支持可搜索加密的云存储解决方案，在使用尽量小的网络开销以及本地存储开销的情况下，支持对敏感数据的多用户访问、关键字查询，并在算法上创新以提高效率，并能够合理、安全地管理搜索凭证。

我们计划在第一阶段(12月1-8日)进行前期准备，包括开通并熟悉阿里云的相关服务，论文收集并进行初步阅读。第二阶段(12月8-15日)进行文献阅读，设计算法和架构方案并对相关算法进行测试。第三阶段(12月15日-1月3日)进行方案的实现与开发。第四阶段(1月3-10日)进行测试并进行论文的撰写。

我们计划从搜索时间、搜索命中率、网络开销、数据安全性这几个指标进行测试。我们将随机生成上千万个用户及其对应的身份信息(身份证号、居住地址等)，然后进行超过十万次搜索，统计命中率、平均搜索时间、使用最大带宽和网络流量使用量。然后尝试使用当下常见的破密方法攻击我们的解决方案以测试安全性。