标题：VeriDedup: A Verifiable Cloud Data Deduplication Scheme With Integrity and Duplication Proof

主要问题：

先前的解决方案采用消息锁和可恢复性证明的方式来检查去重加密数据的完整性。存在两个问题：

1. 忽略了在数据上传阶段提供重复校验的正确性；
2. 相同的文件派生到相同的验证标签中，会遭受到暴力攻击；限制了用户灵活生成个体验证标签。

本篇论文的主要工作：

提出了一种新的重复数据删除方案Veri Dedup：包含TDICP协议和UDDCP协议。

1. 通过引入一个新的验证标签——note set（符合函数f的随机化比特序列，将note set插入到文件中），提出一个基于私有信息检索(PIR)的灵活标签支持去重的完整性检测协议(TDICP)，该协议允许持有相同文件的多个用户生成各自的验证标签，并且仍然支持CSP的重复数据删除；（解决问题2）
2. 通过引入基于私有集交叉（PSI）提出一个新型用户确定重复检查协议（UDDCP）来保证数据重复检查的正确性（让用户代替CSP先判断文件是否重复，从而使CSP在文件上传过程中无法根据重复检查的结果欺骗用户），该协议可以抵制CSP向用户提供虚假的重复检查结果。（解决问题1）

相关工作：

发现现有的工作要么不能对验证标签进行去重，要么不允许用户在去重过程中灵活地创建自己的验证标签。特别地，现有的方案都没有考虑到重复校验的正确性保证的必要性，这使得CSP可以欺骗用户以获得利润 。

相关知识：

- Proxy Re-Encryption (PRE)L：用于将文件密钥分配给授权的数据持有者
- RSA-PSI：使得数据持有者首先判断一个文件是否重复，而不是CSP
- PIR：使数据持有者能够在不将集合的位置暴露给CSP的情况下检索笔记集

威胁模型：

数据拥有者是诚实的；CSP是半可信的。

CSP危害：1)窥探数据持有者的私人数据;2)欺骗数据持有人，提供错误的重复检查结果，以索取更高的存储费用;3)由于数据维护疏忽造成数据丢失。

假设问题1已经解决，AA和CSOP没有串通。

AA是半可信的，它对存储在云上的数据很好奇，假设数据持有者、CSP和AA通过一些安全协议(例如，开放安全套接字层(SSL))通过安全通道相互通信。

在系统设置或初始化阶段，所有系统参数都可以安全地与所有相关方共享。

设计目标：

1. 重复数据去重时的独立完整性检查:VeriDedup允许数据持有者检查存储在CSP上的文件的完整性，而无需下载整个文件并与相应的数据所有者进行交互
   （对于去重数据存储需要保证的基本要求）
2. 灵活的标签生成:VeriDedup允许每个数据持有者创建自己的单独的验证标签，同时仍然可以对这些标签执行数据去重。
   （解决问题2）
3. 重复检查的正确性保证:VeriDedup可以保证重复检查的正确性。因此，一个半可信的CSP永远不能欺骗数据持有者上传任何已经被CSP存储的数据。
   （解决问题1）

TDICP协议

系统设置：输入安全参数 输出一个应用于标记生产的隐藏函数
标记生成和插入：输入隐藏函数和数据拥有者密钥，输出一个随机标记集s和位置集p，将标记集插入到加密区块相应的位置上。
检查初始化：输入区块检查索引，数据拥有者输出一个系数集e，计算一个挑战集v
相应计算：输入挑战集v，CSP输出Resp。
完整性检查：输入Resp，数据拥有者通过计算得Res输出检查结果，挑选出笔记集并验证这些笔记是否符合隐藏函数f。如果验证通过，数据持有者确认存储文件的完整性。

UDDCP协议

系统设置：输入安全参数，数据拥有者输出RSA密钥对，AA初始化一个空的杜鹃过滤器
过滤器生成：输入CSP维护的标记集x，输出到过滤器
检查初始化：输入安全参数，数据拥有者输出三个系数集来维护标记集y，计算挑战A
响应计算：输入挑战集A，CSP计算C，响应给数据拥有者。
数据去重检查：输入响应集C，数据拥有者输出重复标记
过滤器更新：输入更新的标记集y'，AA更新杜鹃过滤器。