计算机研究与发展
Journal of Computer Research and Development
DOI:10.7544/issn1000-1239.2014.20140115
51(7):1397-1409, 2014

公共云存储服务数据安全及隐私保护技术综述

李 晖[1] 孙文海[1] 李凤华[2] 王博洋[1]
[1](综合业务网络理论与关键技术国家重点实验室(西安电子科技大学) 西安 710071)
[2](中国科学院信息工程研究所 北京 100093)
(lihui@mail.xidian.edu.cn)

Secure and Privacy-Preserving Data Storage Service in Public Cloud

Li Hui[1], Sun Wenhai[1], Li Fenghua[2], and Wang Boyang[1]
[1](*State Key Laboratory of Integrated Services Networks* (*Xidian University*), *Xi'an* 710071)
[2](*Institute of Information Engineering*, *Chinese Academy of Sciences*, *Beijing* 100093)

Abstract Cloud computing has been gradually considered the most significant turning point in the development of information technology during past few years. People reap the benefits from cloud, such as ubiquitous and flexible access, considerable capital expenditure savings, pay-as-you-go computing resources configuration, etc. Many companies, organizations, and individual users have adopted the public cloud storage service to facilitate their business operations, research, or everyday needs. However, in the outsourcing cloud computing model, users' physical control of the underlying infrastructure including the system hardware and lower levels of software stack, is shifted to third-party public cloud service providers, such as Dropbox, Google Drive, Microsoft SkyDrive and so on. In addition, the sensitive data of users are also outsourced to and stored in the cloud, e.g., they may upload emails, photos, financial reports, and health records to the cloud. Thus, the potential private information leakage and integrity of the outsourced data is one of the primary concerns for the cloud users. To build users' confidence in such cloud storage service paradigm, tons of attentions have been drawn and a number of related problems have been studied extensively in the literature, such as fine-grained cloud data access control mechanism, secure search over encrypted cloud data, outsourced data integrity auditing, secure deletion for cloud data, etc., which ensure that cloud users enjoy the convenience the cloud offers in a privacy-preserving way. Otherwise, the cloud will become merely a remote storage which provides limited values to all parties. This paper focuses on the enabling and critical cloud computing security protection techniques and surveys on the recent researches in these areas. In addition, we further point out some unsolved but important challenging issues and hopefully provides insight into their possible solutions.

Key words cloud storage; attribute-based encryption; encrypted data search; data integrity auditing; secure data deletion

摘 要 随着云计算技术的日益普及,公共云存储服务已经得到普遍的应用,DropBox、Google Drive、金山快盘等公共云存储应用的用户飞速增长.然而,用户对他们保存在云端数据的隐私性、完整性以及可

收稿日期:2014-02-17;**修回日期**:2014-04-04
基金项目:国家自然科学基金项目(61272457, 61170251);国家"八六三"高技术研究发展计划基金项目(2012AA013102);中央高校基本科研业务费专项基金项目(K50511010001)

控分享的关切也日益增长.针对这一问题,近年来学术界对公共云存储服务中数据的细粒度访问控制、密文搜索、数据完整性审计以及云存储数据的安全销毁等问题展开了大量研究.对云存储数据安全和隐私保护技术方向的研究进展进行综述,并指出待解决的关键问题.

关键词 云存储;属性基加密;密文搜索;数据完整性审计;安全数据销毁

中图法分类号 TP309

随着网络带宽的增加以及移动互联网的普及,云存储已经成为云计算最广泛的应用之一.由于用户可能同时拥有 PC、笔记本电脑、平板电脑和智能手机等多种终端,并且在不同的地方和不同的终端上访问数据,云存储则为这些设备间共享数据提供了一种最适合的解决方案.Dropbox、GoogleDrive、金山快盘等云存储应用用户数飞速增长,目前,Dropbox 的用户数已经达到 1.75 亿.云存储在为广大用户带来方便性的同时,也造成了数据所有权和管理权分离的问题.云服务提供商(cloud service provider, CSP)可以获取、搜索用户存储在云端的数据;CSP 也可能因为系统故障使用户的数据丢失;攻击者也有可能通过攻击 CSP 的服务器获取用户的数据,这些都为用户带来了信息泄露或数据丢失的担忧.

为保护云存储应用的用户数据隐私,数据所有者可以在将其数据上传到 CSP 之前对数据进行加密.然而加密云存储在应用过程中还存在许多需要解决的问题.

1) 要选择适当的数据加密机制,以适应只包含数据所有者的单用户场景和需要数据共享的多用户场景.

2) 当大量的数据以密文方式存储在云端之后,数据所有者或者被授权的数据使用者如何有效地查找或定位这些数据.因为对密文数据的搜索要比对明文的搜索要困难得多.

3) 存储在 CSP 的数据完整性和可用性是否能够高效地被第三方公开审计,在审计过程中是否可以保护数据的隐私以及数据所有者的身份隐私?

4) 当需要删除保存在 CSP 的数据时,数据所有者如何确信 CSP 真正删除了这些数据.

目前已经提出了一些兼容现有公共云存储服务的安全解决方案.Yun 等人[1]提出了一个密码文件系统,使用基于通用 Hash 的消息认证码树对外包数据提供加密和完整性保证,并实现了一个原型系统,通过修改的文件系统可以和不可信的云存储服务器互操作.JungleDisk[2] 和 Cumulus[3] 都保护外包数据的机密性,他们的原型系统都以 Amazon S3 作为后端存储.这 2 个系统都只是实现了简单的加密操作.

本文将以公共云存储服务数据安全和隐私保护为背景,对前述 4 方面的研究进展进行综述.

1 公共云存储服务的数据加密机制

针对公共云存储服务,使用混合加密机制来保证公共云存储数据安全是一个合适的选择.混合加密机制包括密钥封装机制(key encapsulation mechanism, KEM)和数据封装机制(data encapsulation mechanism, DEM).其中 DEM 采用对称密码算法加密数据量较大的数据文件,以保证加解密运算具有高速度和低复杂性.而 KEM 则以公钥密码算法封装了用户加密数据文件的对称密钥 K,如图 1 所示.文献[4-5]给出了抗适应性选择密文攻击(adaptive chosen ciphertext attack, CCA2)的混合加密方案.

数据所有者使用混合加密方案将加密的文件存储到 CSP,而数据使用者通过解封装 KEM,得到文件的对称加密密钥 K,并解密数据文件进行访问.混合加密机制的优点是每个数据文件使用随机的密钥加密,随机密钥 K 只以加密方式保存在云端,本地只保存 KEM 的密钥对,降低了客户端密钥管理的复杂性.通过对 KEM 方式的控制,可以实现多种灵活的加密文件访问控制方式.

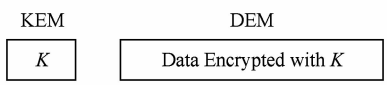


Fig. 1 Overview of KEM and DEM.

图 1 KEM 和 DEM 混合加密示意图

1.1 传统非对称加密封装

在单用户环境下,数据所有者自己使用云存储服务,可以使用基于传统非对称密码体制实现 KEM.即自己产生一个加解密的密钥对,用公钥实现 KEM,自己保存私钥.当从 CSP 取回加密文件后,可以用保存的私钥解开 KEM,从而能够解密数据.

1.2　代理重加密机制

代理重加密是指允许第三方(代理 Trent)改变由 Alice 加密的密文,使得 Bob 可以解密,而 Trent 并不知道原来的明文.在公共云存储应用中,如果数据所有者 Alice 想将其加密的文件分享给一个特定的数据使用者 Bob 时,可以使用代理重加密机制,委托 CSP 将由 Alice 公钥封装的 KEM 转换为 Bob 公钥封装的 KEM,而 CSP 并不能解开 KEM.重加密的操作由 CSP 完成,以节省 Alice 的运算开销.

代理重加密除与普通的对称和非对称加密相同的步骤之外,还包括 2 个步骤:

1) 指派.允许 Bob 基于其私密钥和 Alice 的公钥产生重加密的密钥.该重加密密钥由 CSP 用于重加密函数的输入,该函数由 CSP 运行将密文转换为 Bob 的密钥加密的密文.非对称代理重加密方案分为双向和单向两类.双向方案中重加密是可逆的,而单向重加密则是不可逆的.

2) 传递.可传递的代理重加密方案允许密文重加密的次数无限制.不可传递的代理重加密方案只允许重加密一次.目前,还没有已知的单向的、可传递的代理重加密方案.这样的构造是否可能仍然是一个公开问题.

1998 年 Blaze,Bleumer,Strauss(BSS)[6] 提出了一个 atomic proxy re-encryption 方案,该方案基于 Elgamal 体制,允许一个半可信的代理者将 Alice 的密文转换为 Bob 可以解密的密文,同时不泄漏 Alice 的明文消息.该方案是一个双向方案,并且将单向方案的构造遗留为一个公开问题.2005 年 Ateniese 等人[7] 提出了一个更为有效的代理重加密方案,这是一个单向方案,但是代价是该方案是一个不可传递的重加密方案.这 2 个方案都是选择明文安全的(CPA).Green 和 Ateniese[8] 提出了基于身份的代理重加密方案,Hohenberger 等人[9] 提出了安全混淆重加密方案.2007 年 Canetti 和 Hohenberger[10] 提出了选择密文安全(CCA)的双向代理重加密方案.Weng 等人[11] 提出了条件代理重加密方案,引入了访问控制机制,只有当满足条件时才能完成委托.Libert 等人[12] 提出了一个在标准模型双线性映射 Diffie-Hellman 假设下 CCA 安全的单向代理重加密方案.Matsuda 等人[13] 提出了一个标准模型不使用双线性对 DDH 假设下 CCA 安全的双向代理重加密方案.Chow 等人[14] 提出随机预言机模型下基于 CDH 假设的单向代理重加密方案.Wang 等人[15] 则提出了随机预言机模型下基于 DBDH 假设的 IND-CCA2 安全单向代理重加密方案.Xagawa 等人[16] 则提出了一个基于 LWE(learning with error)问题的代理重加密方案.

1.3　广播加密机制

当数据所有者希望向多个数据使用者分享数据时,可以使用广播加密机制进行 KEM.在广播加密过程中,广播者为动态选择的接收用户子集加密消息,子集中的任意用户都可以用自己的私钥解密消息,而子集外的用户即使合谋也无法解密消息.接收用户应是无状态的,即每个用户持有的密钥在系统的整个生存期内都不需要更新.

第 1 个广播加密方案由 Fiat 和 Naor[17] 提出,文献[18]提出了 2 个适用于无状态接收者的子集覆盖方案.在这种方案中,每个用户属于一个子集的集合,并且存储有所有这些子集对应的密钥.发送者选择适当的子集覆盖所有合法用户而不包括被撤销的用户,用这些子集对应的密钥加密消息.文献[19-20]提出了“分层子集差”方法以减小传输开销.

也有一些方案不采用子集覆盖,如文献[21-24]在广播中指出所有撤销的用户,而文献[25-29]则包含广播中所有接收者的集合,并且使用被撤销用户或者目的接收用户的信息来推导会话密钥.

文献[30-32]还考虑了用户拥有不同权利的情况,文献[30]中用户的权利是静态的,而文献[31-32]考虑了线型结构的动态权利广播加密方案.

广播加密的关键指标包括 KEM 的密文尺寸、公钥尺寸和每个用户存储的私钥数量.目前最好的广播加密方案公钥尺寸为 $O(n)$ 量级,而密文和私钥尺寸均为常数,n 为接收者的数目;或者公钥和密文尺寸均为 $O(\sqrt{n})$ 量级,私钥尺寸为常数.近期的研究结果大多关注于标准模型以及更为常用的安全假设下的方案构造,对于如何提高成员动态变化时的处理效率仍然需要加强研究,文献[22,27]在一定程度上能适应用户动态增加和注销的情况,但仍不理想.

1.4　基于属性的加密机制

基于属性的加密机制(attribute-based encryption,ABE)是另一种控制接收者对加密数据的解密能力的密码机制,当用户所拥有的属性满足一定的接入策略时就可以解密信息.接入策略可以用逻辑表达式表示,也可转化为树型结构表示.数据所有者可以使用 ABE 实现 KEM,以控制加密数据文件在满足接入策略的用户集合中共享.

第 1 个 ABE 方案由 Sahai 等人[33] 提出,发送者

加密一个数据,并且指定一个属性集合和数值 d,只用拥有至少 d 个给定属性的接收者才能恢复对 KEM 解密的密钥并恢复 KEM 中封装的数据. 基于这个方案,Goyal 等人[34]提出了一个结合细粒度访问控制的 ABE 方案,可以支持任意单调的包含与/或限门的接入公式. 这个方案又被称为密钥策略的 ABE(KP-ABE),因为接入策略是嵌入在私钥当中,当用户属性满足接入公式时可以恢复密钥,从而解密消息,其本质是一个多级的秘密共享机制. 随后 Ostrocsky 等人[35]提出了非单调接入结构的 ABE 方案. KP-ABE 方案需要根据接入策略的不同,向具有权限的用户分配解密私钥份额,尤其是当接入策略改变时,需要重新分配解密私钥份额,这样对用户带来了较大的密钥管理开销.

Bethencourt 等人[36]提出了密文策略的 ABE(CP-ABE),接入策略嵌入在 KEM 的密文当中,而解密私钥只与属性集合相关,因此当密文的接入策略改变时只需要重新加密密文,不需要重新分配属性对应的解密私钥,因此 CP-ABE 方案更加类似于基于角色的访问控制,也与广播加密的接收者无状态的要求相一致. 为了抗合谋攻击,每个用户拥有属性的对应私钥对应于不同的随机数,拥有这个随机数信息才能解密信息. 因此将不同用户的属性私钥组合起来无法解密密文.

CP-ABE 方案需要对属性以及属性对应的解密私钥进行管理,通常需要一个属性授权中心. 数据所有者也可以自己管理属性和分配属性管理密钥. Chase[37]提出了一个多授权中心 ABE 系统,每个授权中心管理不同的属性域. 随后 Chase 等人[38]又提出了一个实用的多中心 ABE 系统,该系统去除了可信的授权中心同时能保护用户隐私. Muller 等人[39]提出了一个有效的支持析取范式策略的分布式 ABE 方案. Yu 等人[40]则提出了一个安全、可扩展的细粒度访问控制的 ABE 方案. Li 等人[41]提出了一个支持用户审计的细粒度访问控制的 ABE 方案,以防止云当中非法的密钥共享. 上述 CP-ABE 方案密文尺寸与密文接入策略涉及到的属性数目呈线性关系. 文献[42-49]分别提出了具有常数量级密文尺寸的 CP-ABE 方案,文献[43]提出了一个常数密文尺寸的 KP-ABE 方案,Liu 和 Cao 等人[45]提出了一个多中心的基于标准模型的 CP-ABE 方案,这些方案的密文尺寸大小与属性数目无关.

ABE 加密和解密的计算比较复杂,通常涉及多个双线性对运算,特别是在接入策略的表述比较复杂的情况下,将部分计算外包到 CSP 有助于减少数据使用者终端的开销. 文献[47]给出了一个 ABE 外包解密方案,将复杂的解密操作由 CSP 转化为一个普通的 ElGamal 解密问题,终端只需要一次模指数运算,可有效降低终端的解密计算量. Lai 等人[50]在文献[47]的基础上对 CSP 解密转换正确性增加了可验证属性.

CP-ABE 的另一个问题是当属性撤销或者属性中一个用户撤销时的密钥更新问题. 文献[46]讨论了这一问题并分析了现有撤销方案的不足,但是其提出的方案引入了对称密钥的群组密钥管理方案,也不是十分理想. CP-ABE 的高效动态权利管理仍是一个需要解决的问题.

一般基于属性的加密只提供文件内容的保密性,而很多情形下,除了提供文件的保密性之外,还需要保护密文中的属性和相关策略的隐私性,Kapadia 等人[51]提出了匿名 ABE 以解决这个问题. 现有的匿名 ABE 方案大部分仅支持直接的解密方式,即用户只有在执行了包含大量计算的解密过程之后,才能判断接入策略匹配是否成功. 这种方法在用户端带来巨大的解密开销,存在严重的效率缺陷. 为了解决以上问题,我们提出了一个新技术"match-then-decrypt"[52]. 该技术在匿名 ABE 的解密阶段之前加入了一个属性匹配阶段,在密文中引入了一些专用分量,用于实现解密之前的属性匹配检测,以判定属性私钥是否与密文中潜在的访问策略相匹配. 其代价远远小于一次解密的代价,从而在根本上避免了传统的直接解密方式带来的巨大开销.

总体说来,ABE 更加适应于具有较多用户的群组加密数据共享场合. 根据应用场景的需要,用户可以选择合适的 KEM 机制.

2 隐私保护的密文搜索

对加密数据的搜索是安全云存储应用中的一个基本功能,典型的具备隐私保护功能的密文关键词搜索系统如图 2 所示,其包括数据所有者、数据使用者和 CSP 三个参与者. 数据所有者将加密数据和安全索引上传存储在 CSP,数据一般使用诸如 AES 之类的密码算法加密,而安全索引则使用特定的可搜索加密机制生成. 可搜索加密机制可以分为非对称密钥可搜索加密和对称密钥可搜索加密. 当用户想搜索云服务器中的加密数据时,数据使用者会产生陷门(采用非对称可搜索加密时)或者向数据所有者

请求陷门(采用对称可搜索加密时),然后将陷门发送给 CSP. CSP 执行搜索算法,向数据使用者返回搜索的结果. 搜索可以基于一定的排序准则,将最匹配的前 n 个相关的结果返回.

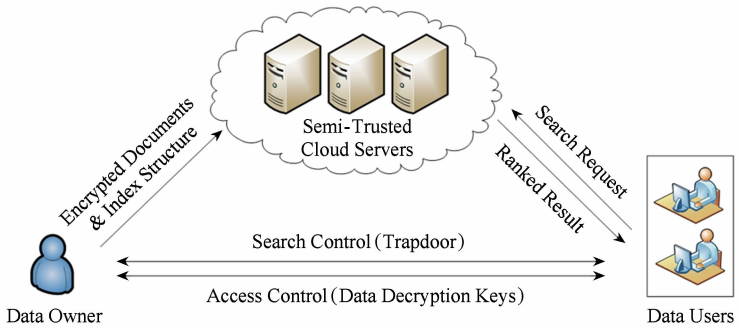


Fig. 2　Framework of privacy-preserving encrypted data search in the cloud.

图 2　云中隐私保护的密文搜索示意图

2.1　威胁模型

现有大多数搜索方案的典型威胁模型是考虑云服务器是半可信的,即它诚实地按照预定的协议和步骤完成操作,但是会试图分析用户的数据、索引以及搜索协议过程中交互的消息以获得额外的信息.

2.2　搜索隐私

搜索过程中数据所有者和数据使用者所关心的隐私大致包括 3 个方面:

1) 关键词隐私

如何在陷门中对 CSP 保护数据使用者感兴趣关键词的隐私是密文关键词搜索隐私保护的一个重要课题. 尽管关键词的陷门都是被加密的,但是 CSP 可以通过其他的边信息攻击方式获取用户感兴趣的关键词,比如词频分析攻击[53-55]. 给定一个特定数据集中特定关键词文件频率(包含特定关键词的文件数)或者关键词频率(一个文件中含有多少个关键词)分布,攻击者可以通过逆向工程恢复出陷门中的关键词.

2) 陷门不可关联性

陷门应该以随机方式产生,否则给定任意两个陷门,攻击者可以容易地确定他们之间的关系,比如限门里是否包含有相同的关键词集合. 如果不满足这一要求,也会导致关键词隐私的泄露,因为 CSP 可以累加不同搜索请求中不同关键词的频率.

3) 接入模式

接入模式是指返回文件的次序. 可以使用秘密信息获取(private information retrieval)技术[56-57]来保护接入模式的隐私,但是这一方式代价很大,其要求搜索算法遍历所有存储到云上的数据,这对大规模的云存储系统是不现实的. 所以目前大多数密文搜索方案都不将保护接入模式作为安全目标.

2.3　基于对称密钥的可搜索加密方案

其基本模式为数据的所有者和对这些数据进行搜索的用户共享相同的密钥信息. Song 等人[58]首次提出基于对称密钥的可搜索加密方案,用一种特殊双层结构的加密方式对文档中的每一个单词分别进行加密,对密文进行搜索时需要遍历这个文档,来检查其是否存在某一特定关键词,这样导致搜索效率低下. 其次,所使用的加密方案同实际应用中数据加密所用的标准加密技术截然不同,存在兼容性问题,不能处理压缩和二进制数据,不能抵抗对密文的频率攻击. Goh[59]提出了一种使用伪随机函数和布隆过滤器(bloom filter)对每个文件构建一个安全索引的方案,搜索时间同文件数量成正比,由于使用布隆过滤器时不可避免地引入的误差,使得搜索结果不完全正确. Chang 等人[60]和 Curtmola 等人[61]几乎同时提出了利用伪随机技术构建关键词索引和查询请求,使得搜索效率大为提高,但对动态的数据更新支持不够,更新时所需的计算量极大,或需要重构整个索引. Wang 等人[62]使用反向索引(inverted index),利用顺序保持加密(order-preserving encryption)技术对文件中的关键词频率进行加密,同时,能够使得用户对返回的搜索结果进行验证. Kamara 等人[63]利用同文献[61]相同的安全索引构造方法,提出了一种支持文件更新的加密数据检索技术.

应当指出以上技术方案仅适用于单一关键词搜索,无法满足明文搜索中已广泛支持的多关键词搜索. 当然可以利用多次单一关键词搜索后进行交集操作来实现多关键词搜索,但需要较大的运算和通信开销,效率极低. Cao 等人[64]通过对每一个文档建立索引向量,利用矩阵加密,运用查询搜索后内积

值的大小排列来对密文进行多关键词搜索,并返回包含关键词最多的前 n 个文件,但搜索效率不高,需要遍历所有文件,且搜索结果的精确性较低. Vimercati 等人[65]提出一种将访问控制和可搜索加密相融合的技术方案,通过访问控制列表对用户进行划分,用加盐操作使得相同关键词在不同用户组中的索引不同,来保护关键词的隐私安全,但这一方案的使用场景特殊,不具有普遍意义. Sun 等人[66]提出了一种安全多关键词密文搜索方案,通过使用多维 B 树(multi-dimensional B-tree)技术大大提高了密文搜索效率,同时利用余弦相似性测量(cosine similarity measure)方法提高了返回结果的精确度,并且能够对搜索结果进行验证[67].

应当指出,现有的基于对称密钥的可搜索加密方案,均无法支持系统中存在多个数据拥有者,或实现代价极高,且效率和可用性低.

2.4 基于非对称密钥的可搜索加密方案

Boneh 等人[68]提出了第 1 个非对称可搜索加密方案,在这一方案中,任何公钥所有者,都可以向服务器中写入数据,但只有授权的拥有私钥的用户才能够对密文进行搜索. 在文献[69]中,Abdalla 等人对基于非对称密钥的可搜索加密方案进行了改进. 随后,学术界提出了支持布尔关键词操作[70]、子集操作[71]、范围查询[71-72]等一系列密文搜索技术方案. Kerschbaum 等人[73]提出了基于身份加密的公钥可搜索加密方案. Lin 等人[74]提出一种对单一关键词搜索的公钥加密方案,减少了对双线性对的使用,效率大为提高,并提出了适用于网络鉴证的应用场景. Hwang 等人[75]在企业多用户场景下提出了一种支持固定关键词并集搜索的方案,同时使服务器利用用户列表对用户的搜索权限进行控制,但此方案的扩展性不好,当系统中存在大量用户时,方案的效率会大幅降低,同时用户不能对多关键词进行自由搜索. Li 等人[76]利用谓词加密(predicate encryption)技术设计了一种支持多用户、并可实现授权搜索的公钥可搜索加密方案,同样此方案存在着扩展性不好的问题,同时由于用户的搜索请求需要由可信的第三方产生,方案的效率较低,且不支持对文本、二进制等其他文件类型的搜索,限制了方案的应用场景. Sun 等人[77]第 1 次在基于属性加密(attribute-based encryption)技术的基础上,提出了一种灵活、扩展性强、支持授权搜索的安全密文搜索技术,大大扩展了密文搜索的应用场景. 数据拥有者利用访问控制策略构建安全索引,当且仅当数据使用者的陷门中包含的用户属性满足某一索引中的访问控制策略,且这一索引包含用户查询的关键词时,相关的文件才能被搜索到. 此方案不仅支持单一关键词搜索,也支持固定多关键词的并集搜索.

需要指出的是,虽然非对称可搜索加密技术更适用于多用户的互联网模型,但因现有方案不可避免地存在对双线性对的操作,搜索效率不高,同时现有方案并不支持多关键词的自由搜索,这也影响了这一技术在实际场景中的使用,另外现有工作缺少对密文数据更新和对搜索结果进行认证的支持. 基于对称密钥的可搜索加密方案搜索效率高,且能够支持对多关键词的自由搜索,但当系统存在大量用户,且需要对用户进行访问控制时,同非对称密文搜索技术相比,方案的灵活性和可扩展性较低,应用场景的局限性较强.

另外需要指出的是,从对查询关键词的精确程度上划分,以上 2 种类型的可搜索加密方案均为针对精确关键词的搜索,除此以外,还有一类支持模糊关键词搜索的加密方案. 当用户查询的关键词存在拼写错误时,系统仍能够辨识出其搜索目标,返回所需的搜索结果. Li 等人[78]提出利用编辑距离来衡量关键词的相似程度,设计出基于通配符的模糊关键词搜索的加密方案,其方案只针对单一关键词搜索. Chuah 等人[79]利用布隆过滤器构建关键词索引,提出一种支持多关键词模糊查询的可搜索加密方案,但应当指出这不是传统意义下的多关键词概念,而是利用概率分布将相关单词置于同一索引中. Wang 等人[80]利用位置敏感的散列函数(locality-sensitive Hash)技术提出了一种支持多关键词模糊查询的密文搜索方案,此方案真正实现了多关键词的自由搜索,方案的灵活性较强. 但这类方案大多基于对称密钥实现,并不支持复杂的多用户场景.

3 隐私保护的云存储数据完整性审计

当用户将其数据存储到云端之后,他们最关心的问题之一就是存储的文件是否完好. 最直接的办法是将文件全部取回检查,但是这样将耗费大量的网络带宽,特别当存储在云端的数据量很大时是很不现实的.

Juels 等人[81]提出了“可回收证明技术”(PoR)模型,通过抽样和纠错码技术来保证服务器正确保存了数据,使用户不用下载全部数据就可以验证保存的数据是完整的、可取回的. 其主要思想是将文件

分成多个数据块,对每一个数据块纠错编码后计算由密钥产生的消息认证码,密钥由用户保存,数据和消息认证码保存到 CSP. 校验时,用户随机产生要检查的数据块位置作为挑战,CSP 返回相应位置的数据块和消息认证码,用户可以验证数据的完整性. 由于该方案采用对称密码机制,不支持第三方的公开审计;此外,审计过程需要用户和云服务器之间的多轮交互,对用户还是会造成一定的负担.

由于在一些情况下,用户需要频繁的验证数据的完整性(例如需要频繁使用的数据,在每次使用之前都需要保证数据的完整性),而对应所产生的验证开销会给用户带来过多的负担. 因此,人们自然地会想到将审计委托给一个可信任的第三方(TPA)来完成,用户只需向 TPA 发出审计请求,TPA 向 CSP 进行多轮的审计,将审计结果通知用户. 第三方审计必然要求用户对外包数据块产生的消息认证码是可公开验证的,这就需要采用非对称密码机制. 图 3 是公开完整性审计示意图:

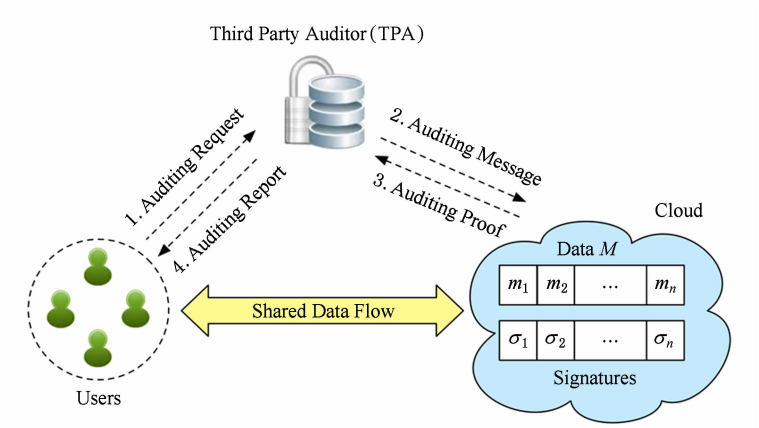


Fig. 3　Framework of data integrity auditing in the cloud.

图 3　云中完整性审计系统示意图

用户将文件分组为(m_1,m_2,…,m_n),并对每个分组产生认证标签(σ_1,σ_2,…,σ_n),将文件和对应标签上传到一个半可信的 CSP 中. TPA 是一个第三方审计机构,它通过与 CSP 之间的挑战和响应交互来审计数据的完整性. Ateniese 等人[82]在他们定义的"可证明的数据所有权"(PDP)模型中第 1 次考虑到了可公开审计性. 该方案利用基于 RSA 的同态标签技术来对外包数据完整性进行审计. 方案对待审计的数据文件进行随机采样,为了降低审计过程中的通信开销,方案中 CSP 将采样数据块的随机线性组合和认证标签提交给 TPA. 但是,在此方案中 TPA 可以通过多次审计请求的结果构造一个方程组解得用户的数据,因此用户数据的隐私无法得到保证. Shacham 等人[83]提出了一种改进的 PoR 方案,利用了可证明安全的 BLS 短签名构造的可公开审计的同态标签技术. 同文献[81]的方案一样,这一方法并不能够实现支持隐私保护的验证,TPA 仍然可以获得用户的数据内容.

为解决这一隐私问题,Shah 等人[84-85]建议通过先加密数据,然后将许多对加密数据预计算的、基于对称密钥的 Hash 值发送给审计方的方法来使 TPA 在保证验证数据文件的完整性的同时,无法获得用户的隐私数据.

以上方案考虑的都是静态数据的完整性审计,如果考虑数据频繁进行插入、修改或者部分删除等编辑操作,将导致整个标签或认证码的重新计算,这样的更新方式并不有效. 为有效地解决对动态变化数据的完整性审计支持,Ateniese 等人[86]提出了一个以之前 PDP 方案为基础的部分动态版本,此方案仅支持动态的删除和修改,但不能有效地支持插入. 此外,该方案使用对称加密技术,并不能支持公开的数据审计. Wang 等人[87]类似地提出了一个在云计算场景中支持部分动态数据存储的方案,且此方案具有数据错误定位的特性,即在数据的多个备份存储在多个服务器的情况下,可以定位持有错误数据的服务器. 在接下来的工作中,Wang 等人[88]于 2009 年提出了一个安全的数据存储审计方案,然而这一方案并不提供对用户数据的隐私保护,用户的数据信息可能会泄露给提供审计服务的第三方. 同时,在审计过程的实施中,所需的通信量和计算量较大,影响了协议的运行效率. 与此同时,Erway 等人[89]发展出了一个基于 skip lists 的方案,使得可证明数据所有权的拥有全动态支持. 然而,由于需要采样数据块的线性组合,这一方案同样不能保证在公开数据

验证过程中的用户的数据隐私安全.

Wang 等人[90]于 2010 年提出了一个支持隐私保护的可公开审计的方案.在此方案中,通过利用基于公钥的同态标签技术,TPA 无需从 CSP 上得到用户的全部数据来进行审计,这样就大大减少了协议运行的通信量和计算量.该方案在每次 CSP 向 TPA 返回验证数据时增加了 1 个一次性随机数,这样就使得 TPA 无法解出对应的方程组,因此不会向 TPA 泄露用户的数据内容,保证了用户数据的隐私安全.另外,此方案支持对不同用户的数据进行批量审计,即在有多个审计任务的情况下,同时对这多个任务进行审计.相比逐一审计的方式,批处理可以大大提高审计的总体效率.

以上讨论的云端存储数据完整性审计都是针对数据属于单一所有者的情况.当考虑群组中多个用户可以对数据进行编辑、修改的情况下,群组共享数据的完整性公开审计面临新的问题.其中,一个新的问题是 TPA 可以根据多次验证过程中获得的验证信息,来分析验证信息的更新,从而分析出哪个用户修改了哪部分数据,并分析哪个用户在群组中具有更重要的作用,最终泄露群组用户的身份隐私;另一个问题是当群组成员撤销时,如何对被撤销用户产生的认证标签进行有效地处理以保证数据继续能够被 TPA 进行完整性审计,同时能够尽量降低终端用户的处理开销.

针对第 1 个问题,我们首先提出了一个解决方案 Oruta[91-92].在该方案中,通过构造基于环签名的同态可认证标签,对 TPA 保护了用户的身份隐私.这一方案的不足之处是认证标签的数据量以及审计过程的额外通信开销随群组中用户数目线性增加,不适于大规模的群组共享数据的应用场景,且新用户的增加必需重新计算认证标签;另一个特点是由于环签名的特性,任何人都没有办法区分对某一数据块的认证标签是哪个用户产生的.随后我们又基于群签名构造的同态可认证标签,设计了群组共享数据完整性审计方案 Knox[93],该方案标签长度与群组用户数无关,当必要时,群控制者(一般是数据的原始创建者)可以揭示对某个数据块的签名者身份.

在上述方案的基础上,我们又基于代理重签名的思想,设计了一个可以有效地支持群组用户撤销的云端群组数据的完整性审计方案[94].该方案在群组用户的撤销过程中,因维护数据完整性所产生的开销主要由云端而不是用户来承担,从而极大地减轻了群组在用户撤销过程中的计算和通信开销.在此基础上,我们还对最初的方案进行了扩展,使其能够支持批处理的多个任务审计(即同时审计多个任务),从而能提高了在多个任务情况下的审计效率.

4 云存储数据的确定性删除技术

数据安全销毁(secure data deletion)是近年来云数据安全中的新的热点问题.由于用户在使用云数据服务的过程中,不再真正意义(物理)上拥有数据,如何保证存储在云端、不再需要的隐私数据能够安全销毁成为新的难点问题.传统的保护存储在云端隐私数据的方法是在将数据外包之前进行加密.那么(云端)数据的安全销毁实际上就转化为(用户端)对应密钥的安全销毁.一旦用户可以安全销毁密钥,那么即使不可信的云服务器仍然保留用户本该销毁的密文数据,也不能破坏用户数据的隐私.

现有大量的系统是通过覆盖来删除所存储的数据[95-96].但是使用覆盖的方法严重依赖于基本的物理存储介质的性质.对现在广泛使用的云计算以及虚拟化模型来说,数据所有者失去了对数据存储位置的物理控制.因此,基于存储介质的物理性质的安全数据删除方法并不能满足现在的需求.张逢喆等人[97]提出了一种基于可信计算数据销毁机制.Boneh 和 Lipton[98]提出了利用加密数据来销毁信息的机制.Di Crescenzo 等人[99]针对文件集合中的任意一个文件的有效安全删除引进了树状结构.在树的根节点上的主密钥保持可擦除,树中的每一个密钥都加密下面的多个密钥,而叶子节点的密钥加密文件本身.Mitra 和 Winslett[100]提出了通过反向索引来组织数据,利用加密并假设加密密钥能够被破坏,进而允许可选择地删除数据记录以及相关的索引中的关键字.Perlman[101]首次提出了基于时间的数据可信删除方法.在该方法中,数据可被安全删除,并在预定的时间后数据将永久不可被访问.FADE 系统[102]使用公钥密码并引入简单的利用布尔操作调整删除的策略.但是,FADE 的策略仅支持一层或两层的布尔表达,并且需要使用复杂的公钥密码系统.Peterson 等人[103]在数据块层使用全有或全无的转换技术(AONT)来实施安全删除.该方法通过 AONT 存储每一个数据块,然后覆盖其中的一部分,这使得整个数据块不可用.Geambasu 等人给出了一个基于时间的数据可信删除方法的原型 Vanish[104].

Vanish将数据密钥划分为多个部分,然后被存储在P2P网络的不同节点中,数据在每个节点的缓存中保存8 h,之后节点将删除所分享的密钥部分.当文件在8 h后需要被访问时,文件拥有者需要更新节点缓存中的密钥部分.王丽娜等人[105]提出了一种适用于云的数据安全删除方案,利用秘钥生成树,门限秘密共享(threshold secret sharing)来组织和管理秘钥,同时利用分布式散列表(distributed Hash table)周期性地删除相关秘钥来实现云数据的安全删除.Cachin等人[106]首次严格的定义了基于加密的数据安全销毁的安全模型和安全定义.该方案将图论(graph),对称加密算法和门限秘密共享技术相结合来实现基于策略的安全数据销毁(policy-based secure deletion).除了针对云数据以外,安全数据销毁在通过物理媒介传播数据的过程中也被广泛的关注[96,107].这些方案通过图论、数据结构(如B-tree)等方法来实现数据的安全销毁.这些方法同样也适用于云数据的安全销毁.

上述确定性删除技术是在假设数据使用者不保存数据加密密钥这样一个强的安全假设下设计的,无法满足数据的后向安全性.若数据使用者成功访问过一次数据并保存数据加密密钥,即使密钥管理者回收控制策略、删除与其相关联的控制密钥,数据访问者依旧可以恢复明文数据,这样就不能达到数据确定性删除的效果.一种解决办法是数据所有者可以周期性地更新数据加密密钥,但这需要消耗大量的计算能力和通信带宽.

5　结　　论

本文介绍了公共云存储服务安全和隐私保护的主要研究方向和当前的进展.值得指出的是这些研究很多是独立开展的,而安全和隐私保护的公共云存储服务需要综合采用加密机制、密文搜索、完整性审计机制以及安全数据删除等技术.要组合应用这些机制还需要进一步加以研究,因为本文所述不同安全机制的安全目标、威胁模型尚未统一加以考虑,密码机制也存在不相容的情况.此外,针对群组应用的公共云存储服务在用户撤销、属性变更与撤销的情况下还存在很多问题,特别是如何降低密钥重新分发、重加密的运算开销等都需要研究实用的方法.支持高效加密、快速方便搜索、完整性审计、群组共享应用、安全数据销毁的公共云存储服务系统投入实用还有待进一步努力.

参　考　文　献

[1] Yun A, Shi C, Kim Y. On protecting integrity and confidentiality of cryptographic file system for outsourced storage [C] //Proc of ACM Workshop Cloud Computing Security (CCSW 2009). New York: ACM, 2009: 67-76

[2] Rackspace. JungleDisk [EB/OL]. 2010 [2011-01-09]. http://www.jungledisk.com/

[3] Vrable M, Savage S, Voelker G M. Cumulus: Filesystem Backup to the cloud [J]. ACM Trans on Storage, 2009, 5(4): 1-28

[4] Cramer R, Shoup V. Design and analysis of practical public-key encryption schemes secure against adaptive chosen ciphertext attack [J]. SIAM Journal on Computing, 2004, 33(1): 167-226

[5] Hofheinz D, Eike K. Secure hybrid encryption from weakened key encapsulation [G] //LNCS 4622: Proc of CRYPTO 2007. Berlin: Springer, 2007: 553-571

[6] Blaze M, Bleumer G, Strauss M. Divertible protocols and atomic proxy cryptography [G] //LNCS 1403: Proc of EUROCRYPT 1998. Berlin: Springer, 1998: 127-144

[7] Ateniese G, Fu K, Green M, et al. Improved proxy re-encryption schemes with applications to secure distributed storage [J]. ACM Trans on Information and System Security, 2006, 9(1): 1-30

[8] Green M, Ateniese G. Identity-based proxy re-encryption [G] //LNCS 4521: Proc of the 5th Applied Cryptography and Network Security Conf. Berlin: Springer, 2007: 288-306

[9] Hohenberger S, Rothblum G, Shelat A, et al. Securely obfuscating re-encryption [C] //Proc of the 4th Theory of Cryptography Conf. Berlin: Springer, 2007: 233-252

[10] Canetti R, Hohenberger S. Chosen-ciphertext secure proxy re-encryption [C] //Proc of ACM CCS 2007. New York: ACM, 2007: 185-194

[11] Weng J, Deng R, Ding X, et al. Conditional proxy re-encryption secure against chosen-ciphertext attack [C] //Proc of ASIACCS 2009. New York: ACM, 2009: 322-332

[12] Libert B, Vergnaud D. Unidirectional chosen-ciphertext secure proxy re-encryption [G] //LNCS 4939: Proc of PKC 2008. Berlin: Springer, 2008: 360-379

[13] Matsuda T, Nishimaki R, Tanaka K. CCA proxy re-encryption without bilinear maps in the standard model [C] //Proc of PKC 2010. Berlin: Springer, 2010: 261-278

[14] Chow S, Weng J, Yang Y, et al. Efficient unidirectional proxy re-encryption [G] //LNCS 6055: Proc of Progress in Cryptology-AFRICACRYPT 2010. Berlin: Springer, 2010: 316-332

[15] Wang Hongbing, Cao Zhenfu, Wang Licheng. Multi-use and unidirectional identity-based proxy re-encryption [J]. Information Sciences, 2010, 180(20): 4042-4059

[16] Xagawa K, Tanaka K. Proxy re-encryptions based on learning with errors [EB/OL]. 2009[2010-12-03]. http://www.nishizeki.ecei.tohoku.ac.jp/LA2009/proceedings_winter/07.pdf

[17] Fiat A, Naor M. Broadcast encryption [G] //LNCS 773: Proc of CRYPTO 1993. Berlin: Springer, 1993: 480-491

[18] Naor D, Naor M, Lotspiech J. Revocation and tracing schemes for stateless receivers [G] //LNCS 2139: Proc of CRYPTO 2001. Berlin: Springer, 2001: 41-62

[19] Halevy D, Shamir A. The LSD broadcast encryption scheme [G] //LNCS 2442: Proc of CRYPTO 2002. Berlin: Springer, 2002: 47-60

[20] Goodrich M T, Sun J Z, Tamassia R. Efficient tree-based revocation in groups of low-state devices [G] //LNCS 3152: Proc of CRYPTO 2004. Berlin: Springer, 2004: 511-527

[21] Naor M, Pinkas B. Efficient trace and revoke schemes [C] //Proc of Financial Cryptography (FC 2000). Berlin: Springer, 2000: 1-20

[22] Delerabl′ee C, Paillier P, Pointcheval D. Fully collusion secure dynamic broadcast encryption with constant-size ciphertexts or decryption keys [G] //LNCS 4575: Proc of Pairing-Based Cryptography (Pairing 2007). Berlin: Springer, 2007: 39-59

[23] Kusakawa M, Hiwatari H, Asano T, et al. Efficient dynamic broadcast encryption and its extension to authenticated dynamic broadcast encryption [G] //LNCS 5339: Proc of CANS 2008. Berlin: Springer, 2008: 31-48

[24] Lewko A B, Sahai A, Waters B. Revocation systems with very small private keys [C] //Proc of the 31st IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2010: 273-285

[25] Boneh D, Gentry C, Waters B. Collusion resistant broadcast encryption with short ciphertexts and private keys [G] //LNCS 3621: Proc of CRYPTO 2005. Berlin: Springer, 2005: 258-275

[26] Sakai R, Furukawa J. Identity-based broadcast encryption [OL]. IACR Cryptology ePrint Archive, 2007[2009-11-08]. http://eprint.iacr.org/2007/217

[27] Delerabl′ee C. Identity-based broadcast encryption with constant size ciphertexts and private keys [G] //LNCS 4833: Proc of ASIACRYPT 2007. Berlin: Springer, 2007: 200-215

[28] Gentry C, Waters B. Adaptive security in broadcast encryption systems (with short ciphertexts) [G] //LNCS 5479: Proc of EUROCRYPT 2009. Berlin: Springer, 2009: 171-188

[29] Ren Y, Gu D. Fully cca2 secure identity based broadcast encryption without random oracles [J]. Information Processing Letters, 2009, 109 (11): 527-533

[30] Adelsbach A, Huber U, Sadeghi A R. Property-based broadcast encryption for multi-level security policies [G] //LNCS 3935: Proc of ICISC 2005. Berlin: Springer, 2005: 15-31

[31] Zhao Xingwen, Zhang Fangguo. Traitor tracing for dynamic privileges against pirate rebroadcast [J]. Computer & Security, 2012, 31(1): 59-69

[32] Jin H, Lotspiech J. Broadcast encryption for differently privileged [G] //Emerging Challenges for Security, Privacy and Trust 297: IFIP Advances in Information and Communication Technology. Berlin: Springer, 2009: 283-293

[33] Sahai A, Waters B. Fuzzy identity-based encryption [G] //LNCS 3494: Proc of EUROCRYPT 2005. Berlin: Springer, 2005: 457-473

[34] Goyal V, Pandey O, Sahai A et al. Attribute-based encryption for fine-grained access control of encrypted data [C] //Proc of ACM CCS 2006. New York: ACM, 2006: 89-98

[35] Ostrovsky R, Sahai A, Waters B. Attribute-based encryption with nonmonotonic access structures [C] //Proc of ACM CCS 2007. New York: ACM, 2007: 195-203

[36] Bethencourt J, Sahai A, Waters B. Ciphertext-policy attribute based encryption [C] //Proc of IEEE ISSP 2007. Piscataway, NJ: IEEE, 2007: 321-334

[37] Chase M. Multi-authority attribute based encryption [G] //LNCS 4392: Proc of TCC 2007. Berlin: Springer, 2007: 515-534

[38] Chase M, Chow S. Improving privacy and security in multi-authority attribute-based encryption [C] //Proc of ACM CCS 2009. New York: ACM, 2009: 121-130

[39] Muller S, Katzenbeisser S, Eckert C. Distributed attribute-based encryption [G] //LNCS 5461: Proc of ICISC 2008. Berlin: Springer, 2008: 20-36

[40] Yu S, Wang C, Ren K, et al. Achieving secure, scalable, and fine-grained data access control in cloud computing [C] //Proc of IEEE INFOCOM 2010. Piscataway, NJ: IEEE, 2010: 15-19

[41] Li Jin, Zhao Ganshen, Chen Xiaofeng, et al. Fine-grained data access control systems with user accountability in cloud computing [C] //Proc of IEEE CloudCom 2010. Piscataway, NJ: IEEE, 2010: 89-96

[42] Herranz J, Laguillaumie F, Ràfols C. Constant size ciphertexts in threshold attribute-based encryption [G] //LNCS 6056: Proc of Public Key Cryptography (PKC 2010). Berlin: Springer, 2010: 19-34

[43] Attrapadung N, Libert B, Panafieu E. Expressive key-policy attribute-based encryption with constant-size ciphertexts [G] //LNCS 6571: Proc of Public Key Cryptography (PKC 2011). Berlin: Springer, 2011: 90-108

[44] Attrapadung N, Herranz J, Laguillaumie F, et al. Attribute-based encryption schemes with constant-size ciphertexts [J]. Theoretical Computer Science, 2012, 422 (9): 15-38

[45] Liu Zhen, Cao Zhenfu, Huang Qiong, et al. Fully secure multi-authority ciphertext-policy attribute-based encryption without random oracles [G] //LNCS 6879: Proc of ESORICS 2011. Berlin: Springer, 2011: 278-297

[46] Hur J, Noh D K. Attribute-based access control with efficient revocation in data outsourcing systems [J]. IEEE Trans on Parallel and Distributed Systems, 2011, 22(7): 1214-1221

[47] Green M, Hohenberger S, Waters B. Outsourcing the decryption of ABE ciphertexts [C] //Proc of the 9th USENIX Security Symp. Berkeley: USENIX Association, 2011: 3-18

[48] Zhou Zhibin, Huang Dijiang. On efficient ciphertext-policy attribute based encryption and broadcast encryption [C] // Proc of ACM CCS 2010. New York: ACM, 2010: 753-755

[49] Chen Cheng, Zhang Zhenfeng, Feng Dengguo. Efficient ciphertext policy attribute-based encryption with constant-size ciphertext and constant computation-cost [G] //LNCS 6980: Proc of Provable Security (ProvSec 2011). Berlin: Springer, 2011: 84-101

[50] Lai Junzuo, Deng R, Guan Chaowen, et al. Attribute-based encryption with verifiable outsourced decryption [J]. IEEE Trans on Information Forensics and Security, 2013, 8 (8): 1343-1354

[51] Kapadia A, Tsang P P, Smith S W. Attribute-based publishing with hidden credentials and hidden policies [C] // Proc of NDSS 2007. Reston, VA: ISOC, 2007: 179-192

[52] Zhang Yinghui, Chen Xiaofeng, Li Jin, et al. Anonymous attribute-based encryption supporting efficient decryption test [C] //Proc of ASIACCS 2013. New York: ACM, 2013: 511-516

[53] Swaminathan A, Mao Y, Su, G M, et al. Confidentiality-preserving rank-ordered search [C] //Proc of the 3rd ACM Workshop on Storage Security and Survivability. New York: ACM, 2007: 7-12

[54] Wang C, Cao N, Ren K, et al. Enabling secure and efficient ranked keyword search over outsourced cloud data [J]. IEEE Trans on Parallel and Distributed Systems, 2012, 23(8): 1467-1479

[55] Zerr S, Olmedilla D, Nejdl W, et al. Top-k retrieval from a confidential index [C] //Proc of the 12th Int Conf on Extending Database Technology: Advances in Database Technology. New York: ACM, 2009: 439-449

[56] Chor B, Kushilevitz E, Goldreich O, et al. Private information retrieval [J]. Journal of the ACM, 1998, 45 (6): 965-981

[57] Ishai Y, Kushilevitz E, Ostrovsky R, et al. Cryptography from anonymity [C] //Proc of the 47th Annual IEEE Symp on Foundations of Computer Science. Piscataway, NJ: IEEE, 2006: 239-248

[58] Song D, Wagner D, Perrig A. Practical techniques for searches on encrypted data [C] //Proc of the 21st IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2000: 44-55

[59] Goh E J. Secure indexes [OL]. IACR Cryptology ePrint Archive. 2003 [2004-02-15]. http://eprint.iacr.org/2003/216

[60] Chang Y, Mitzenmacher M. Privacy preserving keyword searches on remote encrypted data [G] //LNCS 3531: Proc of ACNS 2005. Berlin: Springer, 2005: 442-455

[61] Curtmola R, Garay J A, Kamara S, et al. Searchable symmetric encryption: Improved definitions and efficient constructions [C] //Proc of ACM CCS 2006. New York: ACM, 2006: 79-88

[62] Wang C, Cao N, Li J, et al. Secure ranked keyword search over encrypted cloud data [C] //Proc of ICDCS 2010. Piscataway, NJ: IEEE, 2010: 253-262

[63] Kamara S, Papamanthou C, Roeder T. Dynamic searchable symmetric encryption [C] //Proc of ACM CCS 2012. New York: ACM, 2012: 965-976

[64] Cao N, Wang C, Li M, et al. Privacy-preserving multi-keyword ranked search over encrypted cloud data [C] //Proc of IEEE INFOCOM 2011. Piscataway, NJ: IEEE, 2011: 829-837

[65] Vimercati S, Foresti S, Jajodia S, et al. Private data indexes for selective access to outsourced data [C] //Proc of the 10th Workshop on Privacy in the Electronic Society (WPES 2011). New York: ACM, 2011: 69-79

[66] Sun W, Wang B, Cao N, et al. Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking [C] //Proc of ACM ASIACCS 2013. New York: ACM, 2013: 71-82

[67] Sun W, Wang B, Cao N, et al. Verifiable privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking [J/OL]. IEEE Trans on Parallel and Distributed Systems, 2013 [2013-12-12]. http://ieeexplore.ieee.org

[68] Boneh D, Crescenzo G D, Ostrovsky R, et al. Public key encryption with keyword search [G] //LNCS 3027: Proc of EUROCRYP 2004. Berlin: Springer, 2004: 506-522

[69] Abdalla M, Bellare M, Catalano D. et al. Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions [G] //LNCS 3621: Proc of CRYPTO 2005. Berlin: Springer, 2005: 205-222

[70] Golle P, Staddon J, Waters B. Secure conjunctive keyword search over encrypted data [G] //LNCS 3089: Proc of ACNS'04. Berlin: Springer, 2004: 31-45

[71] Boneh D, Waters B. Conjunctive, subset, and range queries on encrypted data [G] //LNCS 4392: Proc of TCC 2007. Berlin: Springer, 2007: 535-554

[72] Shi E, Bethencourt J, Chan H, et al. Multi-dimensional range query over encrypted data [C] //Proc of the 28th IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2007: 350-364

[73] Kerschbaum F, Sorniotti A. Searchable encryption for outsourced data analytics [C] //Proc of the 7th European Conf on Public Key Infrastructures, Services and Applications. Berlin: Springer, 2010: 61-76

[74] Lin X, Lu R, Foxton K. An efficient searchable encryption scheme and its application in network forensics [G] // LNICST 56: Proc of e-Forensics 2010. Berlin: Springer, 2011: 66-78

[75] Hwang Y, Lee P. Public key encryption with conjunctive keyword search and its extension to a multi-user system [C] //Proc of Pairing-Based Cryptography (Pairing 2007). Berlin: Springer, 2007: 2-22

[76] Li M, Yu S, Cao N, et al. Authorized private keyword search over encrypted data in cloud computing [C] //Proc of IEEE ICDCS 2011. Piscataway, NJ: IEEE, 2011: 383-392

[77] Sun W, Yu S, Lou W, et al. Protecting your right: Attribute-based keyword search with fine-grained owner enforced search authorization in the cloud [C] //Proc of IEEE INFOCOM 2014. Piscataway, NJ: IEEE, 2014 [2014-03-12]. http://www.cnsr.ictas.vt.edu/publication/ABKS.pdf

[78] Li J, Wang Q, Wang C, et al. Fuzzy keyword search over encrypted data in cloud computing [C] //Proc of IEEE INFOCOM Mini Conf 2010. Piscataway, NJ: IEEE, 2010: 1-5

[79] Chuah M, Hu W. Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data [C] //Proc of the 31st Int Conf on Distributed Computing Systems Workshops. Piscataway, NJ: IEEE, 2011: 273-281

[80] Wang B, Yu S, Lou W, et al. Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud [C] // Proc of IEEE INFOCOM 2014. Piscataway, NJ: IEEE, 2014 [2014-03-12]. http://www.cnsr.ictas.vt.edu/publication/ver3_Bing.pdf

[81] Juels A, Burton J, Kaliski S. Pors: Proofs of retrievability for large files [C] //Proc of ACM CCS 2007. New York: ACM, 2007: 584-597

[82] Ateniese G, Burns R, Curtmola R, et al. Provable data possession at untrusted stores [C] //Proc of ACM CCS 2007. New York: ACM, 2007: 598-609

[83] Shacham H, Waters B. Compact proofs of retrievability [G] //LNCS 5350: Proc of ASIACRYPT 2008. Berlin: Springer, 2008: 90-107

[84] Shah M, Swaminathan R, Baker M. Privacy-preserving audit and extraction of digital contents [OL]. 2008 [2009-01-05]. http://www.hpl.hp.com/techreports/2008/HPL-2008-32R1.pdf

[85] Shah M, Baker M, Mogul J, et al. Auditing to keep online storage services honest [C] //Proc of HotOS'07. Berkeley: USENIX Association, 2007: 1-6

[86] Ateniese G, Pietro R, Mancin L V, et al. Scalable and efficient provable data possession [C] //Proc of SecureComm 2008. New York: ACM, 2008: 1-10

[87] Wang C, Wang Q, Ren K, et al. Ensuring data storage security in cloud computing [C] //Proc of IWQoS 2009. Piscataway, NJ: IEEE, 2009: 1-9

[88] Wang Q, Wang C, Li J, et al. Enabling public verifiability and data dynamics for storage security in cloud computing [G] //LNCS 5789: Proc of ESORICS'09. Berlin: Springer, 2009: 355-370

[89] Erway C, Kupcu A, Papamanthou C. et al. Dynamic provable data possession [C] //Proc of ACM CCS 2009. New York: ACM, 2009: 213-222

[90] Wang C, Wang Q, Ren K. et al. Privacy-preserving public auditing for data storage security in cloud computing [C] // Proc of IEEE INFOCOM 2010. Piscataway, NJ: IEEE 2010: 1-9

[91] Wang Boyang, Li Baochun, Li Hui. Oruta: Privacy-preserving public auditing for shared data in the cloud [C] // Proc of IEEE Cloud 2012. Piscataway, NJ: IEEE, 2012: 295-302

[92] Wang Boyang, Li Baochun, Li Hui. Oruta: Privacy-preserving public auditing for shared data in the cloud [J/OL]. IEEE Trans on Cloud Computing, 2014 [2014-02-22]. http://ieeexplore.ieee.org

[93] Wang Boyang, Li Baochun, Li Hui. Knox: Privacy-preserving auditing for shared data with large groups in the cloud [G] //LNCS 7341: Proc of ACNS 2012. Berlin: Springer, 2012: 507-525

[94] Wang Boyang, Li Baochun, Li Hui. Panda: Public auditing for shared data with efficient user revocation in the cloud [J/OL]. IEEE Trans on Service Computing, 2014[2013-12-22]. http://ieeexplore.ieee.org

[95] Joukov N, Papaxenopoulos H, Zadok E. Secure deletion myths, issues, and solutions [C] //Proc of StorageSS'06. New York: ACM, 2006: 61-66

[96] Reardon J, Capkun S, Basin D. SoK: Secure data deletion [C] //Proc of the 34th IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2013: 301-315

[97] Zhang Fengzhe, Chen Jin, Chen Haibo, et al. Lifetime privacy and self-destruction of data in the cloud [J]. Journal of Computer Research and Development, 2011, 48(7): 1155-1167 (in Chinese)
(张逢喆, 陈进, 陈海波, 等. 云计算中的数据隐私性保护与自我销毁[J]. 计算机研究与发展, 2011, 48(7): 1155-1167)

[98] Boneh D, Lipton R. A revocable backup system [C] //Proc of the 4th USENIX Security Symp. Berkeley: USENIX Association, 1996: 91-96

[99] Di Crescenzo G, Ferguson N, Impagliazzo R, et al. How to forget a secret [C] //Proc of the 16th Symp on Theoretical Aspects of Computer Science (STACS 1999). Berlin: Springer, 1999: 500-509

[100] Mitra S, Winslett M. Secure deletion from inverted indexes on compliance storage [C] //Proc of StorageSS'06. New York: ACM, 2006: 67-72

[101] Perlman R. File system design with assured delete [C] // Proc of SISW'05. Piscataway, NJ: IEEE, 2005: 83-88

[102] Tang Yang, Lee P, Lui J, et al. FADE: Secure overlay cloud storage with file assured deletion [C] //Proc of Security and Privacy in Communication Networks (SecureComm 2010). Berlin: Springer, 2010: 380-397

[103] Peterson Z N, Burns R C, Herring J, et al. Secure deletion for a versioning file system [C] //Proc of the 4th USENIX Conf on File and Storage Technologies (FAST 2005). Berkeley: USENIX Association, 2005: 4-11

[104] Geambasu R, Kohno T, Levy A, et al. Vanish: Increasing data privacy with self-destructing data [C] //Proc of the 17th USENIX Security Symp. Berkeley: USENIX Association, 2009: 299-316

[105] Wang Lina, Ren Zhengwei, Yu Rongwei, et al. A data assured deletion approach adapted for cloud storage [J]. Acta Electronica Sinica, 2012, 40(2): 266-272 (in Chinese)
(王丽娜, 任正伟, 余荣威, 等. 一种适用于云存储的数据确定性删除方法[J]. 电子学报, 2012, 40(2): 266-272)

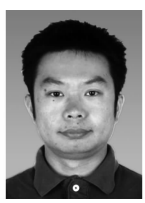
[106] Cachin C, Haralambiev K, Hsiao H C, et al. Policy-based Secure Deletion [C] //Proc of ACM CCS 2013. New York: ACM, 2013: 259-270

[107] Reardon J, Ritzdorf H, Basin D, et al. Secure data deletion from persistent media [C] //Proc of ACM CCS 2013. New York: ACM, 2013: 271-284

Li Hui, born in 1968. Received his BSc degree from Fudan University in 1990, MSc and PhD degrees from Xidian University in 1993 and 1998. Since 2005, he has been professor in the School of Telecommunications Engineering, Xidian University, China. His research interests include cryptography, wireless network security, information theory and network coding.

Sun Wenhai, born in 1985. Received his BSc degree in information security from Xidian University, Xi'an, China, in 2007. Since 2009, he has been PhD student in a combined MS/PhD program in the School of Telecommunications Engineering at Xidian University. His research interests include cloud computing security, wireless network security and applied cryptography.

Li Fenghua, born in 1966. Received his BSc, MSc and PhD degrees from Xidian University. Currently professor of the Institute of Information Engineering, Chinese Academy of Sciences. His research interests include cryptography, network security and computer system security.



Wang Boyang, born in 1985. Received his BSc in information security from Xidian University in 2007. Currently PhD candidate from the School of Telecommunications Engineering, Xidian University, Xi'an, China. His current research interests include security and privacy issues in cloud computing, big data, and applied cryptography.