面向 DaaS 的隐私保护机制研究综述*

杨 进[1,2]，王亮明[2]，杨英仪[2]
（1. 广东药学院 医药信息工程学院，广州 510006；2. 华南理工大学 计算机学院，广州 510006）

摘 要：DaaS 是基于云基础设施对外提供数据库服务的云服务模式，能有效地解决个人和企业处理海量数据所带来的存储、管理压力，但隐私泄露极大地阻碍了 DaaS 的发展，如何增强现有 DaaS 模式的隐私保护成为亟需解决的问题。首先通过文献分析的方法剖析了 DaaS 的服务框架及其隐私泄露模型，接着对 DaaS 中实现委托数据的机密性、对机密数据查询过程中的隐私保护及查询结果的验证、委托数据完整性验证过程中隐私保护三个方面的发展现状进行了综合分析。通过分析得出，现有的隐私保护方法对 DaaS 中数据更新和查询效率方面的支持及其实用性都存在不同程度的缺陷，如何设计高效的机密性算法和保护隐私的数据查询及查询结果验证仍是未来研究的重点。最后展望了未来的研究方向。

关键词：云计算；数据库即服务；隐私保护；数据机密性；隐私信息检索
中图分类号：TP393 **文献标志码**：A **文章编号**：1001-3695(2013)09-2565-05
doi:10.3969/j.issn.1001-3695.2013.09.002

Research on privacy preservation mechanism for DaaS

YANG Jin[1,2], WANG Liang-ming[2], YANG Ying-yi[2]
(1. *College of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou* 510006, *China*; 2. *School of Computer Science & Engineering, South China University of Technology, Guangzhou* 510006, *China*)

Abstract: DaaS is a kind of cloud service which can provide the database services for others based on cloud infrastructure. It can effectively solve the difficulty of storage and management for massive data of individuals and enterprises. However, privacy leaks greatly hinder the development of DaaS. How to enhance the privacy protection of the existing DaaS becomes an urgent problem needed to be solved. Adopting the literature analysis method, this paper firstly summarized service architecture of DaaS and its privacy leaks mode. Then comprehensively analyzed the present researches on data confidentiality, query privacy preserving and query result validation, privacy preserving during data integrity validation of DaaS. From the analysis, it can be inferred that existing privacy preserving methods exist varying degrees of impairment in the aspects of efficiency of data update and query of DaaS and its practical applications. How to design more efficient algorithms for data confidential and the data queries with privacy preserving and query results validation are still the future research focus. Finally, it gave the he future research directions of in the privacy preserving in DaaS.

Key words: cloud computing; database as a service(DaaS); privacy preservation; data confidentiality; privacy information retrieve

0 引言

信息技术的不断推进使得企业积累了各种形式的海量数据，海量数据管理和维护所需的软、硬件成本压力促使企业寻找可靠第三方进行托管。

DaaS 基于云计算基础设施对外提供数据管理服务以减少企业维护、管理海量数据成本[1]。Software as a service(软件即服务，SaaS)、infrastructure as a service(基础设施即服务，IaaS)、platform as a service (平台服务，PaaS)等云计算服务的成功应用为 DaaS 成功实施提供了必要的硬件基础和后台支持。

DaaS 的出现可有效地满足上述需求，企业或个人用户可用两种方式租用 DaaS：a)将其已有数据托管给 DaaS；b)直接租用 SaaS，而 SaaS 将自身数据库置于 DaaS。不管哪种方式都需用户将其数据置于云端而失去对数据的物理控制，这将引起用户担忧自己数据的安全和隐私保护。近期云计算领跑者 Google、Salesforce 等发生的用户隐私泄露事件[2,3]加剧了人们对这一服务模式的担忧。

因此增强 DaaS 中数据安全和用户隐私保护成为其走向实用所亟需解决的问题。隐私保护需从法律制度、管理和技术等多方面有机结合，本文主要从技术角度对目前 DaaS 中隐私保护机研究现状和相关技术进行分析与总结。

1 DaaS 服务框架

1.1 DaaS 体系结构

DaaS 一般处于云计算的 PaaS 层，向上对 SaaS 提供数据库服务，向下建立在 IaaS 所提供的存储池、CPU 等硬件设施上。本文在田秀霞[4]提出的 DaaS 架构的基础上，结合云存储中数

收稿日期：2013-01-20；**修回日期**：2013-03-11 **基金项目**：广东省医学基金资助项目(A2012295)；广东省战略性新兴产业核心攻关项目(2012A010701005)；广东省计算机网络重点实验室开放基金资助项目(CCNL201105)

作者简介：杨进(1977-)，男，湖南邵阳人，博士，主要研究方向为云安全、隐私保护(goodskyfly@163.com)；王亮明(1976-)，男，江西金溪人，讲师，博士，主要研究方向为云安全；杨英仪(1982-)，男，广东揭阳人，博士，主要研究方向为云存储.

据完整性保护最新研究成果[5]引入可信第三方,提出了增强隐私的 DaaS 服务架构,如图 1 所示。架构中主要包括四个角色即 DaaS 提供商(DaaS provider,DSP)、数据拥有者(data owner,DO)、数据请求者(data requestor,DReq)及可信第三方(trust third part,TTP)。

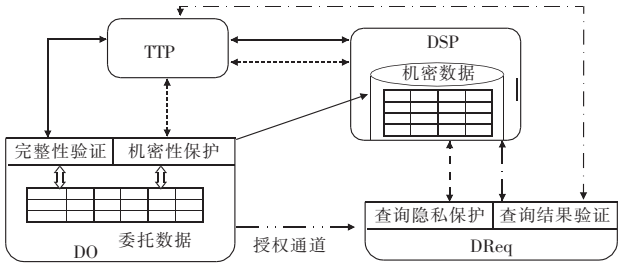

图 1 DaaS 服务框架

1.1.1 DO 向 DSP 委托数据

DO 利用机密性保护模块将需委托的数据转换成不可识别状态然后上传至 DSP。DO 在进行机密性保护时需要 TTP 保存或生成关键参数(如密钥等)[6~9]。此外 DO 还需提供一些辅助手段如建立保护隐私的索引,以加快对转换数据的查询与更新效率。DSP 可接受多个 DO 的委托数据。

1.1.2 委托数据的完整性验证

为了确保 DaaS 中数据来源的真实性(确实是取自 DO 的数据)且不被未授权地修改,需定期或不定期进行完整性验证。其验证流程如下:DO 或者 DReq 通过完整性验证模块向 TTP 提交验证请求,TTP 接受请求后生成验证消息并将之发送给 DSP;DSP 接受验证信息后根据实际情况生成一个完整性证明消息并返回给 TTP;TTP 对证明信息进行校验并将校验结果返回给 DO 或 DReq,如果证明信息有效则表示 DSP 中数据是完整的,否则表示其被破坏。TTP 可以是独立于 DSP 的第三方[10~14];也可以基于可信技术置于 DSP 中作为隐式第三方[5];甚至在验证过程中不需要第三方[15]。

1.1.3 委托数据的查询及查询结果验证

DReq 获得 DO 访问授权后(即获得机密数据的解析参数),调用查询隐私保护模块将其查询转换成 DaaS 可识别的查询,将 DaaS 返回的保护隐私查询结果进行解析,调用查询结果验证模块验证其完整性(查询结果确实来自于 DO 且未经删减)和完备性(所有满足要求的结果都已返回)。

1.2 DaaS 隐私泄露模型

在上述模型中,从隐私保护的角度假定一个角色认为其他角色都是非可信的(其中 TTP 只能保证相关机密参数的生成不会被干扰,不能确保数据审计过程中不会发生数据泄露)。因此 DaaS 中隐私泄露主要在如下方面:

a)非可信 DSP 泄露用户隐私。由于 DSP 对委托数据具备完全控制权,且目前 SLA(service level agreement, 服务等级协议)缺少隐私保护相关的条款,存在 DSP 利用特权无意或者恶意泄露用户隐私的隐患。即使用户委托数据经过隐私处理,DSP 也可能与获得授权(拥有隐私数据解析参数)的 DReq 合谋获得用户敏感数据并恶意泄露。

b)完整性验证过程中的隐私泄露。在完整性验证过程中,由于 TTP 需验证原始数据的完整性,不可避免地要获得诸如每个数据块的签名信息等内容,甚至是用户的明文敏感数据,TTP 成为隐私泄露的极大隐患;此外授权 DReq 也可以进行完整性验证,DReq 可能利用其授权解析参数恶意获取敏感数据。

c)查询用户的查询隐私泄露。DSP 或者 DO 能根据用户的查询关键词、方式、频率来推测其具体查询内容隐私,即使在采取隐私增强的情况下 DSP 与 DO 也可合谋来解析用户的查询隐私。

d)网络传输的隐私泄露。在 DaaS 中网络传输不可避免,恶意攻击者可以通过非法截取数据包、中间人攻击、重放攻击等手段来获取用户敏感数据,尤其在 DO 将机密数据解析参数传给 DReq 过程中,参数隐私泄露的后果极为严重。

2 DaaS 隐私保护关键技术及研究进展

综上所述,由于在网络中传输的数据都是经过隐私保护转换后的非明文数据,且现有的 SSL、IPSec、VPN 等众多网络安全技术能最大限度地保护用户敏感数据的安全及隐私,因此数据机密性保护、查询隐私保护和查询结果验证、完整性验证过程中隐私保护是 DaaS 隐私保护的三大关键技术。

2.1 委托数据机密性保护方法及研究现状

DaaS 中的数据机密性保护要从不被未授权的 DReq 访问和不被不可信的 DSP 访问两个层面上同时进行。目前主要采用基于数据加密和基于数据分布两类方法。

2.1.1 基于数据加密的方法

对数据库数据加密的研究始于 1981 年[16],加密的对象主要为元组和元组属性值。一般用对称算法来加密数据,而用非对称算法来加密对称加密密钥。加密是一种有效地保护隐私的方法,但数据加密后往往丢失了可操作性,因此提高密文检索速度和密文处理效率是现阶段数据隐私保护研究的热点。

Song 等人[17]于 2000 年提出的带单关键字可检索的加密,实现了在密文上的查询,将数据库加密技术转为带检索操作的加密。

Liu 等人[18]于 2009 年根据云计算的特征提出了一种基于对称加密实现隐私保护的密文关键词检索方法,它支持 DSP 可参与部分解密工作以减少 DReq 计算与网络通信负担,同时支持密文关键词检索以保护 DO 和 DReq 的隐私;此外,Wang 等人[19]提出了基于非对称加密的密文检索方法;Huang 等人[20]提出了基于 Bloom Filter 优化的密文检索方法。但这些方法只支持精确的字符串匹配,即两字符串是否相等。然而在许多实际的情况下,错别字和格式不一致不可避免。因此,Li 等人[21]设计了一个支持加密字符串模糊检索的方案,它使用编辑距离来量化字符串的相似度,并为每个字符串附加一个基于通配符的模糊字符串组,用多个精确匹配来实现模糊检索。该方法的不足是它仅仅针对“all or nothing”的查询方式,不能对查询结果进行排序。

Wang 等人[22]考虑关键词词频信息,提出基于对称密钥保序加密技术 OPSE 的单关键词分级密文排序查询方法,能够根据某一指标对检索的关键字分级,并按用户的要求返回最符合要求的前 N 个结果。然而单关键词查询不足以表达和满足用户的个性化查询需求,且涉及过多查询结果而导致巨大的网络通信负载。为了进一步满足用户个性化查询需求,Cao 等人[23]第一次提出多关键词密文排序查询问题,并基于安全 KNN 查询技术[7]中索引向量与查询向量间内积相似度来实现

排序查询并进行隐私保护增强。但该方法未考虑数据中索引关键词及查询关键词权重等因素,使得查询结果排序不太准确。程芳权等人[24]提出了一种支持隐私保护的高效密文排序查询方法(RQED),通过设计无证书认证的 PKES(支持关键词检索的公钥加密),并构建 RQED 框架来实现强隐私保护的密文查询。基于该框架,设计了更合理的多属性多关键词密文查询排序函数,并提出了基于层次动态布隆过滤器的 RQED 索引机制,提高了密文查询时空效率。

对部分常用检索加密方法的总结如表 1 所示。

表 1　常见检索加密方法比较

算法	加密方式	单关键词	多关键词	精确检索	模糊检索	结果排序
文献[14]	对称	√	√	√	√	√
文献[17]	对称	√		√		×
文献[18]	对称	√		√		×
文献[19]	非对称	√		√		×
文献[20]	非对称	√		√		×
文献[21]	非对称	√	√	√	√	√
文献[23]	非对称	√	√	√	√	√
文献[24]	非对称	√	√	√	√	√

从表 1 中可以得知,最近几年对数据库的加密一般都支持多关键字检索且对查询结果进行排序。

为了确保对密文数据处理的效率,上述可检索的加密方法需建立针对密文的额外的索引信息用来对密文数据进行检索,数据处理效率较低。为进一步提高密文数据段处理效率,一些可直接密文上进行操作的方法被相继提出。

Agrawal 等人[25]提出一个基于桶划分和分布概率映射思想的保序对称加密算法 OPES,支持对加密数值数据的各种比较操作,SQL 语句中诸如 MAX、MIN、COUNT、GROUP BY 和 ORDER BY 等操作可以在加密的数据上进行重写和处理。但是 OPES 不支持在加密数据上进行 SUM 和 AVG 操作,这两种操作必须在数据解密之后进行处理。Boldyreva 等人[26]提出一个基于折半查找和超几何概率分布的保序对称加密算法 OPSE,支持对加密数据的各种比较操作,但计算超几何概率需进行多次组合运算使得计算负载较大。上述保序加密算法为确定性的加密方案,不具语义安全性。Wong 等人[7]设计了一个基于向量标量积的对称加密方案,该方案支持对加密数据库进行 KNN(K-nearest neighbor) 计算。IBM 研究员 Gentry[27]、Dijk 等人[28,29]分别利用理想格和整数算术设计了全同态加密方案,该方案能同时支持密文数据上加法和乘法同态。在理论上取得了一定突破,但现有全同态方案都因太复杂且计算量太大而不适用于 DaaS。

黄汝维等人[29]设计了一个基于向量和矩阵运算的可计算加密方案 CESVMC,通过运用向量和矩阵的各种运算,实现了对数据的加密,并支持对加密字符串的模糊检索和对加密数值数据的加、减、乘、除四种算术运算。

2.1.2　基于数据切分的方法

针对加密隐私保护技术的不足,研究者提出通过保护数据间关系而不是数据值的方式防止隐私泄露。

毛剑等人[9]基于求有限域中多项式的解的理论提出了适用于云存储的二次混淆数据分割方案及基于可信第三方的存储架构,并设计了隐私保护的读、写、删除流程。该方案将用户的个人信息(存于可信第三方)与所需委托存储的数据分开,利用可信技术保护个人信息免遭泄露而用分割的方法保护数据的隐私。

张坤[31]提出一种根据隐私约束条件来切分 SaaS 数据属性组合的隐私保护方法。该方法支持租户自定义 SaaS 数据中哪些组合可能泄露用户隐私。然后将这些属性切分到不同的数据分块中进行物理存储,借助可信第三方来实现数据切片间关联关系的混淆和重构,并设计伪造数据算法来保证同一数据分块内部数据切片分布的均衡化,从而达到 SaaS 数据组合隐私保护和实用性的目的。

陈钊[32]基于数据拆分思想设计了一种在三维空间中进行数据置乱和切分的算法——ESSA 来实现数据机密性保护。ESSA 不对字节流的数据格式进行修改,而在二维空间中进行字节级的数据置乱和拆分,在三维空间中进行字节级的数据置乱,将原始字节流中相邻字节分散到空间的不同平面而使得最终每一个数据块都不会包含有用的原始信息。攻击者仅凭部分拆分块也无法获取到任何原始数据当中的信息。该方法在机密性、计算效率方面比加密、纠删码的方式有明显的优势。

田秀霞[4]将$\{k,n\}$门限方案对秘密密钥的保护特性引入 DaaS 的隐私保护机制中,提出了基于秘密共享的数据存储模式来保护 DO 的数据机密性:DO 首先将源数据库利用秘密共享函数分成 n 分,并分别存储在不同的 $DSP_i(1\leqslant i\leqslant n)$上,这样单个 DSP 便无法获取用户的敏感数据。

2.2　查询隐私保护及结果验证

2.2.1　查询隐私保护

查询隐私即 DReq 从 DSP 处检索数据时,不想让 DSP、DO 或者其他 DReq 知道其查询意图和查询内容。现有的机密性保护方法都具备一定的查询隐私保护能力。

带检索加密方法由于数据传输本身为加密数据,已经具备一定的隐私保护,但仍存在不足:基于对称密钥加密的查询隐私保护[7,17,23]具有较高的执行效率,但因需同一个密钥进行所有的查询授权控制,无法抵御 DReq 与 CSP 的合谋攻击。若实施选择性查询授权控制,即 DO 对不同 DReq 授权不同的查询密钥,则密钥安全分发和管理又极其复杂。文献[23]试图通过扰乱排序实现查询隐私保护,但掌握密钥的 DReq 与 DSP 合谋可推理出其他 DReq 的查询请求隐私。针对上述缺陷,文献[32]首次基于 DH 密钥交换协议提出关键词可检索的公钥加密方法(PKES),用公钥加密和私钥查询虽然很好地避免了上述方法中复杂密钥管理问题,但其效率有待进一步提升。Boneh 等人[33]更进一步提出了支持隐私信息的公钥加密方法,但需要 DReq 与 DSP 频繁交互。与基于对称密钥的方法相比,PKES 较具优势。

在基于数据切分的机密性保护方法方面,田秀霞[4]基于秘密共享数据存储模式构建了保护隐私的 B + 树索引,提出了保护隐私的查询处理方法,该方法支持精确查询和范围查询。

2.2.2　查询结果验证

DReq 需要保证其查询结果是完整和完备的。到目前为止,实现数据的完整性或完备性的机制主要有三种,即基于数字签名(digital signature)的方法、基于挑战—响应(challenge-respond)的方法和基于概率(probability)的方法。

基于数字签名的方法引入验证树,主要为 Merkle hash tree[34]及其改进,例如对于包括 SUM、MAX、MIN 等聚集查询结果的验证,Li 等人[35]提出了 AAB tree 和 AAR tree。Wen 等人[36]改进文献[35]中的 AAB tree 中只能对一维数据进行 SUM、MAX、MIN 等聚集查询的结果进行验证的不足,提出能处

理二维数据的 M^2T tree。

Merkle hash tree 的构建过程如下：a）首先对数据排序，设排序后的数据为 $d_1, d_2, \cdots, d_n$；b）计算 MH tree 中叶子节点的值，每个叶子节点都与一个 d_i 对应，其值为对应数字的 hash 函数值，即 $h_i = H(d_i)$，H 为抗碰撞攻击的单向 hash 函数，如 MD5、SHA 等；c）计算 MH tree 中非叶子节点的值，每个非叶子节点都有两个子节点，其值为两个子节点值连接的 hash 函数值，即 $h_N = H(h_{N.L} h_{N.R})$，其中 $h_{N.L}$ 和 $h_{N.R}$ 分别为节点 N 的左子节点和右子节点的值，符号“”为连接符，将两个子节点的值连接在一起；d）最后得到根节点的值 h_{root}，数据所有者使用公开密钥签名方法对 h_{root} 签名，得到 $sig(h_{root})$。

基于 MH tree 的 DaaS 整个工作流程如下：a）DO 在发布数据前，首先根据其数据表构建 MH tree，然后将数据和 MH tree 根节点的签名 $sig(h_{root})$ 一起发给 DSP；b）DSP 收到数据后，用同样的算法构建 MH tree，但不签名根节点；c）DSP 收到 DReq 的查询请求后，先搜索符合条件的数据，然后在 MH tree 中搜索相关节点（即为验证对象），最后将符合条件的数据和验证对象以及从 DO 处收到的 MH tree 根节点签名 $sig(h_{root})$ 一起返回给 DReq；d）DReq 收到所有数据后，先用符合条件的数据和验证对象构建 MH tree，得到一个根节点 h'_{root}，如果 h'_{root} 与 $sig(h_{root})$ 相符，说明 DSP 发回的数据为正确的，否则表示被窜改。

为了确保非可信 DSP 能够完整地执行 DReq 的查询，文献[37]提出了一种建立在挑战—响应协议上的运行时查询证明机制。但是该方法并不能保证返回的查询结果的正确性以及完全性，DSP 完全可以通过修改返回的查询结果攻击该验证模式。另外该方法需要修改 DBMS 核心以返回查询证明，且还需要用户在本地维护一份镜像数据信息，因此并不适合在实际场合使用。

Xie 等人[38]提出了一种基于概率的查询结果完整性验证方法。DO 在委托的数据库中插入一组特别的监测元组，DReq 查询该数据库时，这些混在原始数据中的监测元组就会以一定概率包含在查询结果中返回给 DReq，因而可以通过有效地分析返回结果中的监测元组实现数据的完备性验证。这种方法对于完整性的认证包括数据认证（验证元组是否被篡改）和元组认证（检验元组是伪造元组还是真实元组）。如果一个满足查询条件的监控元组没有被返回，那么用户就可以肯定完整性被破坏；反之，如果所有满足查询条件的监控元组都完整地返回，则以一定概率断定完整性没有受到攻击。

三种方法的比较如表 2 所示。其中 D_S 表示基于数字签名的方法；C_R 表示基于挑战—响应的方法；Pb 表示基于概率的方法。

表 2 三种查询结果验证机制的比较

方法	完整性	完备性	数据库刷新	实现难度	备注
D_S	√	√	×	中	
C_R	√	×	×	高	不能保证结果正确
Pb	√	√	√	低	不能确保未被攻击

从表 2 中可以看出，Pb（基于概率的方法）相对较好，但在 DaaS 中，如何在保证存储效率和完整性验证性能的前提下生成监测元组是研究难点。

2.3 数据完整性验证过程中隐私保护

数据的完整性验证是数据安全 C. I. A. 三大特性之一，DaaS 中 DSP 非可信或半可信特征要求必须提供数据完整性验证机制以确保用户能知晓其置于 DSP 中数据未被修改或者删除。完整性验证有两种：a）一般完整性验证，即将数据从 DSP 下载到本地后进行完整性验证；b）远程完整性验证，即不需要取回全部数据，通过类似知识证明的协议，判断存储在 DSP 处数据是否完整。在 DaaS 中出于性能考虑一般采用后一种方式。

DaaS 中对数据完整性验证要求如下：a）如果 DSP 中数据未被修改则通过验证，否则不能通过验证；b）用户可以对同一数据执行无数次远程验证；c）支持用户对数据的插入、修改、删除动态更新操作；d）不仅 DO 可以对数据完整性验证，任何第三方（如授权的 DReq）皆可进行验证；e）完整性验证过程要确保 DO 敏感数据不被 TTP 或第三方泄露；f）支持对 DSP 中存储副本的验证。本文主要关注在数据完整性验证过程中隐私保护的研究现状和进展。

自文献[38]基于 Diffie-Hellman 密钥交换验证协议首次提出远程非可信服务器上数据完整性验证协议后，研究人员围绕这一方向进行了大量相关研究。例如 Ateniese 等人[39]提出的 PDP 验证协议支持数据的公开可验证和动态的数据追加操作；Erway 等人[40]基于认证跳跃表和 RSA 树提出了两种支持完全动态数据更新的远程数据完整性验证协议；Wang 等人[12]提出了讨论完整性验证过程中隐私保护泄露问题并基于双映射理论给出了一种解决方案；Wang 等人[41]提出一种支持大量用户的、可追踪的、保护隐私保护远程数据完整性验证，前述方案均需 TTP 的支持；郝卓[15]基于 RSA 的同态验证标志提出了一种不需要 TTP 参与的保护隐私的远程数据完整性验证方案，安宝宇[5]用防篡改可信硬件作为隐式可信第三并集成在 DSP 服务器中，代替用户进行完整性验证，利用可信技术实现验证过程中的隐私保护。

3 存在的问题及研究方向

从上述综述可以看出，虽然在数据库的机密性保护、用户查询隐私保护及查询结果验证、数据数据完整性验证过程中的隐私保护等方面都有一定的解决方案和研究成果，但这些研究成果一般集中于云储存领域，而 DaaS 模式中虽然也基于云存储，但也有其自身的特点：以二维关系表存储数据，既包含字符数据又包含数字数据；需频繁地更新操作；需要较为丰富的查询功能且性能要求较高。目前的研究成果在一定程度上解决了 DaaS 中隐私保护的一些问题，但如何将这些方案有机结合，并根据 DaaS 的特点对其改进，还需在以下方面进行加强：

a）现有的委托数据机密性保护方法中，大多是针对文件型的数据，在少数适用关系数据库机密性保护方法中，其对数据的操作方面考虑得最多的是数据查询，而较少考虑数据库的增、删、改等更新操作。在 DaaS 环境中，更新操作是必需的，且对更新操作性能方面有一定的要求，因此研究一种适用于动态更新的数据库机密性方法是 DaaS 中急需解决的难题之一。

b）在 DaaS 模式中，需要满足用户的多样化的查询需求，如精确查询、模糊查询、范围查询、聚集查询、连接查询及查询结果排序等。现有保护隐私的查询方法都只能满足其中一部分查询要求，如现有对加密数据的检索算法或者只适用于字符数据，或者只适用于数字数据，少数适用两者的算法在性能方面又有待加强，且不同的查询基于不同的数据机密性保护方法。另外，现有查询方法大多忽略了查询刷新问题，即使查询

结果可以反映数据库的最新变化,保护隐私的多样化查询技术是 DaaS 主要特色之一。

c)可信技术是保护数据安全的一种强有力的方法,如何将可信技术应用到隐私信息检索、查询结果验证及数据完整性验证以提高其效率是下一步的研究方向。

d)人工智能技术解决问题的能力越来越被认可,可将其自动学习的能力引入 DaaS,如根据请求频率自动调整数据机密性算法来优化存储和更新效率。相信人工智能技术的引入会使得 DaaS 中隐私保护更加智能化。

4　结束语

隐私泄露已成为阻碍 DaaS 发展的重要因素,委托数据的机密性、查询隐私保护和查询结果验证、数据完整验证中的隐私保护是 DaaS 中隐私保护的三大关键技术。目前,不少学者对云存储中隐私保护技术进行了深入的研究,本文对相关研究现状进行了调研与总结分析,指出了目前实用云存储的隐私保护方法还需要在性能与实用方面进行进一步的改进才能适用于 DaaS 模式,同时也指出可信技术和人工智能技术与 DaaS 中隐私保护相结合可以为 DaaS 隐私保护研究提供一个宽广的前景。

参考文献:

[1] HACIGÜMÜS H, MEHROTRA S, IYER B. Providing database as a service[C]//Proc of the 18th International Conference on Data Engineering. Washington DC: IEEE Computer Society Press, 2002: 29-38.

[2] JESSICA E V. Google discloses privacy glitch[EB/OL]. [2009]. http://blogs. wsj. com/digits/2009/03/08/1214/.

[3] GREENBERG A. Cloud computing's stormy side[EB/OL]. [2008]. http://www. forbes. com/2008/02/17/web-application-cloud-tech-intel-cx_ag_0219cloud. html.

[4] 田秀霞. 数据库服务中保护隐私的访问控制与查询处理[D]. 上海:复旦大学,2012.

[5] 安宝宇. 云存储中数据完整性保护关键技术研究[D]. 北京:北京交通大学,2012.

[6] AGRAWAL R, KIERNAN J, SRIKANT R, *et al*. Order-preserving encryption for numeric data[C]//Proc of ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2004: 563-574.

[7] WONG W K, CHEUNG D W, KAO Ben, *et al*. Secure KNN computation on encrypted databases[C]//Proc of the 35th SIGMOD International Conference on Management of Data. Rhode Island: ACM Press, 2009: 139-152.

[8] GENTRY C. Fully homomorphic encryption using ideal lattices [C]// Proc of the 41st Annual ACM Symposium on Theory of Computing. New York: ACM Press, 2009: 169-178.

[9] 毛剑, 李坤, 徐先栋. 云计算环境下隐私保护方案[J]. 清华大学学报, 2011, 51(10): 1357-1362.

[10] ERWAY C, KUPCU A, PAPAMANTHOU C, *et al*. Dynamic provable data possession [C]//Proc of the 16th ACM Conference on Computer and Communications Security. New York: ACM Press, 2009: 213-222.

[11] WANG Qian, WANG Cong, LI Jin, *et al*. Enabling public verifiability and data dynamics for storage security in cloud computing [C]//Proc of the 14th European Symposium on Research in Computer Security. Berlin: Springer, 2009: 355-370.

[12] WANG Cong, WANG Qian, REN Kui, *et al*. Privacy-preserving public auditing for data storage security in cloud computing[C]//Proc of IEEE INFOCOM. Piscataway: IEEE Press, 2010: 1-9.

[13] WANG Cong, CHOW S S M, WANG Qian, *et al*. Privacy-preserving public auditing for secure cloud storage[J]. IEEE Trans on Computers, 2013, 62(2): 362-375.

[14] ZHU Yan, HU Hong-xin, AHN G J, *et al*. Cooperative provable data possession for integrity verification in multicloud storage[J]. IEEE Trans on Parallel and Distributed Systems, 2012, 23(12): 2231-2244.

[15] 郝卓. 远程数据完整性和认证技术研究[D]. 合肥:中国科学技术大学, 2011.

[16] DAVIDA G I, WELLS D L, KAM J B. A database encryption system with subkeys[J]. ACM Trans on Database Systems, 1981, 6(2): 312-328.

[17] SONG D X, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[C]//Proc of IEEE Symposium on Research in Security and Privacy. Washington DC: IEEE Computer Society Press, 2000: 44-55.

[18] LIU Qin, WANG Guo-jun, WU Jie. An efficient privacy preserving keyword search scheme in cloud computing[C]//Proc of International Conference on Computational Science and Engineering. Washington DC: IEEE Computer Society, 2009: 715-720.

[19] WANG Wei-chao, LI ZW, OWENS R, *et al*. Secure and efficient access to outsourced data[C]//Proc of ACM Workshop on Cloud Computing Security. New York: ACM Press, 2009: 55-66.

[20] HUANG Ru-wei, GUI Xiao-lin, YU Si, *et al*. Study of privacy preserving framework for cloud storage[J]. Computer Science and Information Systems, 2011, 8(3): 801-819.

[21] LI Jin, WANG Qian, WANG Cong, *et al*. Fuzzy keyword search over encrypted data in could computing[C]//Proc of the 29th Conference on Computer Communications. New York: ACM Press, 2010: 1-5.

[22] WANG Cong, CAO Ning, LI Jin, *et al*. Securerenked keyword search over encrypted could data[C]//Proc of the 30th International Conference on Distributed Computing Systems. Washington DC: IEEE Computer Society Press, 2010: 253-262.

[23] CAO Ning, WANG Cong, LI Ming, *et al*. Privacy preserving multi keyword ranked search over encrypted cloud data[C]//Proc of the 30th IEEE International Conference on Computer Communications. Washington DC: IEEE Computer Society Press, 2011: 829-837.

[24] 程芳权, 彭智勇. 云环境下一种隐私保护的高效密文排序查询方法[J]. 计算机学报, 2012, 35(11): 2216-2227.

[25] AGRAWAL R, KIERNAN J, SRIKANT R, *et al*. Order-preserving encryption for numeric data[C]//Proc of ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2004: 563-574.

[26] BOLDYREVA A, CHENETTE N, LEE Y, *et al*. Order-preserving symmetric encryption[C]//Proc of the 28th Annual International Conference on Advances in Cryptolog. Berlin: Springer, 2009: 224-241.

[27] GENTRY C. Fully homomorphic encryption using ideal lattices[C]// Proc of the 41st ACM Symposium on Theory of Computing. New York: ACM Press, 2009: 168-178.

[28] DIJK M, GENTRY C, HALEVI S, *et al*. Fully homomorphic encryption over the integers[C]//Proc of the 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques Advances in Cryptology. Berlin: Springer, 2010: 24-43.

(下转第 2582 页)

[68] FREEMAN W T, TAPPEN M F, ADELSON E H. Recovering intrinsic images from a single image[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(9): 1459-1472.

[69] YEDIDIA J S, FREEMAN W T, WEISS Y. Understanding belief propagation and its generalizations[M]//LAKEMEYER G, NEBEL B. Exploring Artificial Intelligence in the New Millennium. San Francisco, CA: Morgan Kaufmann Publishers, 2003: 239-269.

[70] FELZENSZWALB P, HUTTENLOCHER D. Efficient belief propagation for early vision [J]. International Journal of Computer Vision, 2006, 70(1): 41-54.

[71] 卢阿丽, 唐振民, 杨静宇. 基于信任度传播的体视算法[J]. 模式识别与人工智能, 2010, 23(1): 84-90.

[72] YANG Qing-xiong, WANG Liang. Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31(3): 1-12.

[73] 赵亮, 李昌华, 徐胜军, 等. 基于多尺度信念传播的混凝土 CT 图像分割[J]. 计算机工程, 2012, 38(8): 195-197.

[74] KOMODAKIS N, TZIRITAS G. Image completion using efficient belief propagation via priority scheduling and dynamic pruning[J]. IEEE Trans on Image Processing, 2007, 16(11): 2649- 2661.

[75] CHAN Jing-chu, YEN N, CHANG C, *et al*. Local belief propagation aggregation for MRF-based color image segmentation[C]//Proc of IPPR Conference on Computer Vision, Graphics and Image Processing. 2008.

[76] SCOTT G G, KAMBHAMETTU C. Hierarchical belief propagation to reduce search space using CUDA for stereo and motion estimation [C]//Proc of IEEE Workshop on Applications of Computer Vision. Washington DC: IEEE Computer Society, 2009: 1-8.

[77] LAN Xiang-yang, ROTH S, HUTTENLOCHER D, *et al*. Efficient belief propagation with learned higher-order Markov random fields [C]//Proc of the 9th European Conference on Computer Vision. Berlin: Springer-Verlag, 2006: 269-282.

[78] 徐胜军, 刘欣, 赵亮. 基于快速收敛 LBP 算法的图像分割[J]. 计算机应用, 2011, 31(8): 2229-2235.

[79] XU Sheng-jun, LIU Guang-hui, LIU Xin. Image segmentation via ant colony algorithm and loopy belief propagation algorithm[C]//Proc of International Joint Conference on Neural Networks. Piscataway: IEEE Press, 2012: 1-7.

[80] FUKUNAGA K, HOSTETLER L D. The estimation of the gradient of a density function, with applications in pattern recognition[J]. IEEE Trans on Information theory, 1975, 21(1): 32-40.

[81] CHENG Yi-zong. Mean shift, mode seeking, and clustering[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1995, 17(8): 790-799.

[82] COMANICIU D, MEER P. Mean shift: a robust approach toward feature space analysis [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.

[83] COMANICIU D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using mean shift [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2000: 142-149.

[84] COLLINS R T. Mean-shift blob tracking through scale space [C]// Proc of Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2003: 18-20.

[85] CORSO J J, TU Zhuo-wen, YUILLE A, *et al*. Segmentation of Sub-cortical structures by the graph-shifts algorithm[C]//Proc of Information Processing in Medical Imaging. Berlin: Springer-Verlag, 2007: 183-197.

[86] CORSO J J, YUILLE A, TU Zhuo-wen. Graph-shifts: natural image labeling by dynamic hierarchical computing[C]//Proc of IEEE Conference Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2008: 320-327.

[87] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59(2): 167-181.

(上接第 2569 页)

[29] Van DIJK M, GENTRY C, HALEVI S, *et al*. Fully homomorphic encryption over the integers[C]//Proc of the 29th Annual International Conference on Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2010: 24-43.

[30] 黄汝维, 桂小林, 余思. 云环境中支持隐私保护的可计算加密[J]. 计算机学报, 2011, 34(12): 2391-2402.

[31] 张坤. 面向多租户应用的云数据集隐私保护机制研究[D]. 济南: 山东大学, 2012.

[32] 陈钊. 基于云备灾的数据安全存储关键技术研究[D]. 北京: 北京邮电大学, 2012.

[33] BONEH D, CRESCENZO G, OSTROVSKY R. Public key encryption with keyword search[C]//Lecture Notes in Computer Science. Berlin: Springer, 2004: 506-522.

[34] BONEH D, KUSHILEVITZ E, OSTROVSKY R, *et al*. Public key encryption at allows PIR queries[C]//Proc of the 27th Annual International Cryptology Conference. Berlin: Springer, 2007: 50-59.

[35] MERKLE R. A certified digital signature[C]//Proc of Advance in Cryptology. Berlin: Springer, 1990: 218-238.

[36] LI Fei-fei, HADJIELEFTBERIOU M, KOLLIONS G, *et al*. Dynamic authenticated index structures for aggregation queries [J]. ACM Trans on Information and System Security, 2010, 13(4): 1-30.

[37] WEN Tao, SHENG Gang, GUO Quan, *et al*. Query results authentication of outsourced append-only database[J]. Journal of Computer Research and Development, 2012, 49(10): 2077-2085.

[38] XIE Min, WANG Hai-xun, YIN Jian, *et al*. Integrity audit of outsourced data[C]//Proc of the 33rd International Conference on Very Large Data Bases. New York: ACM Press, 2007: 782-793.

[39] ATENIESE G, BURNS R, CURTMOLA R, *et al*. Provable data possession at untrusted stores [C]//Proc of the 14th ACM Conference on Computer and Communications Security. New York: ACM Press, 2007: 598-609.

[40] ERWAY C, KUPCU A, PAPAMANTHOU C, *et al*. Dynamic provable data possession [C]//Proc of the 16th ACM Conference on Computer and Communications Security. New York: ACM Press, 2009: 213-222.

[41] WANG Bo-yang, LI Bao-chun, LI Hui. Knox: privacy-preserving auditing for shared data with large groups in the cloud[C]//Proc of the 10th International Conference on Applied Cryptography and Network Security. Berlin: Springer, 2012: 507-525.

[42] SION R. Query execution assurance for outsourced database[C]// Proc of the 31st International Conference on Very Large Data Bases. New York: ACM Press, 2005: 601-612.

[43] DESWARTE Y, QUISQUATER J J. Remote integrity checking[C]// Proc of the 6th Working Conference on Integrity and Internal Control in Information Systems. [S. l.]: Kluwer Academic Publishers, 2004: 1-11.