(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115459967 A
(43)申请公布日 2022.12.09

(21)申请号 202211027665.2

(22)申请日 2022.08.25

(66)本国优先权数据
202211007665.6 2022.08.22 CN

(71)申请人 翼方健数(北京)信息科技有限公司
地址 100037 北京市海淀区阜成路73号A座五层507,508,509,510,511,512号
申请人 翼健(上海)信息科技有限公司

(72)发明人 张李军 潘光明 张浩

(74)专利代理机构 北京华清迪源知识产权代理有限公司 11577
专利代理师 陈晨

(51)Int.Cl.
H04L 9/40(2022.01)
H04L 9/32(2006.01)
H04L 9/08(2006.01)
H04L 9/06(2006.01)
G06F 16/2458(2019.01)
G06F 16/242(2019.01)
G06F 16/22(2019.01)

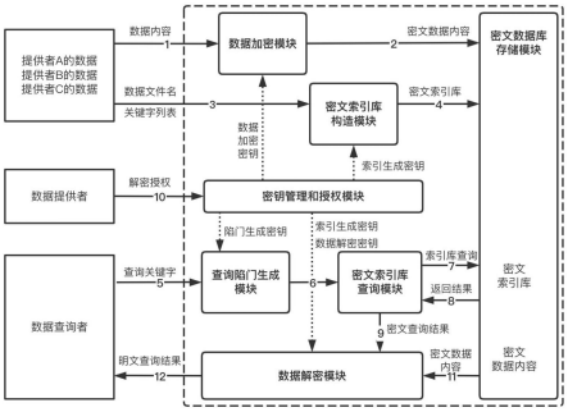
权利要求书2页 说明书8页 附图3页

(54)发明名称

一种基于可搜索加密的密文数据库查询方法及系统

(57)摘要

本申请公开了一种基于可搜索加密的密文数据库查询方法及系统。首先获取明文源数据，明文源数据中包括数据文件名和数据内容;对数据内容进行关键字提取得到关键字列表,并根据数据文件名基于索引生成密钥得到密文索引库,并将密文数据内容和密文索引库对应存储进密文数据库;然后获取目标查询关键字,通过陷门生成密钥对目标查询关键字进行处理生成查询密文;基于查询密文在密文索引库中进行查询得到密文查询结果并获取其对应的密文数据内容;最后基于数据解密密钥对密文数据内容进行数据解密得到明文数据内容。本申请中所有算法采用对称密码原语进行设计,保证了算法执行的高效性,同时以关键字作为键值采用了倒排索引技术,提高了搜索效率。



CN 115459967 A

1.一种基于可搜索加密的密文数据库查询方法,其特征在于,所述方法包括:

获取明文源数据,所述明文源数据中包括数据文件名和数据内容;基于数据加密密钥对所述数据内容进行数据加密得到密文数据内容;

对所述数据内容进行关键字提取得到关键字列表,并根据所述数据文件名基于索引生成密钥得到密文索引库,并将所述密文数据内容和所述密文索引库对应存储进密文数据库;

获取目标查询关键字,通过陷门生成密钥对所述目标查询关键字进行处理生成查询密文;

基于所述查询密文在密文索引库中进行查询得到密文查询结果并获取其对应的密文数据内容;

基于数据解密密钥对所述密文数据内容进行数据解密得到明文数据内容。

2.根据权利要求1所述的方法,其特征在于,在基于数据解密密钥对所述密文数据内容进行数据解密得到明文数据内容之前,所述方法还包括:

响应于明文源数据提供方的解密授权请求,生成该明文源数据的数据解密密钥。

3.根据权利要求1所述的方法,其特征在于,基于数据加密密钥对所述数据内容进行数据加密得到密文数据内容,包括:

通过分组加密算法对数据内容进行加密,其中,所述分组加密算法至少包括AES算法和SM4算法。

4.根据权利要求1所述的方法,其特征在于,所述数据加密密钥和所述索引生成密钥,包括:

生成一个随机值作为密钥生成的种子,然后利用这个种子和一个密钥导出函数为每个文件生成对应的加密密钥。

5.根据权利要求1所述的方法,其特征在于,将所述密文数据内容和所述密文索引库对应存储进密文数据库,还包括:

在密文数据库中建立数据删除表,所述数据删除表用于记录被删除的数据,当查询结果包含被删除的数据,则在解密时通过访问这个表跳过被删除数据的解密操作,从而返回的查询结果中不会包含已删除的数据。

6.根据权利要求1所述的方法,其特征在于,在通过陷门生成密钥对所述目标查询关键字进行处理生成查询密文之前,包括:

基于token/nonce的陷门生成算法得到陷门生成密钥。

7.一种基于可搜索加密的密文数据库查询系统,其特征在于,所述系统包括:

数据加密模块,用于获取明文源数据,所述明文源数据中包括数据文件名和数据内容;基于数据加密密钥对所述数据内容进行数据加密得到密文数据内容;

密文索引库构造模块,用于对所述数据内容进行关键字提取得到关键字列表,并根据所述数据文件名基于索引生成密钥得到密文索引库,并将所述密文数据内容和所述密文索引库对应存储进密文数据库;

密文数据库存储模块,用于对密文数据内容和密文索引库进行存储;

查询陷门生成模块,用于获取目标查询关键字,通过陷门生成密钥对所述目标查询关键字进行处理生成查询密文;

密文索引库查询模块，基于所述查询密文在密文索引库中进行查询得到密文查询结果并获取其对应的密文数据内容；

数据解密模块，用于基于数据解密密钥对所述密文数据内容进行数据解密得到明文数据内容。

8.根据权利要求7所述的系统，其特征在于，所述系统还包括：

密钥管理和授权模块，响应于明文源数据提供方的解密授权请求，生成该明文源数据的数据解密密钥。

9.根据权利要求7所述的系统，其特征在于，所述数据加密模块包括：

通过分组加密算法对数据内容进行加密，其中，所述分组加密算法至少包括AES算法和SM4算法。

10.根据权利要求7所述的系统，其特征在于，所述密文数据库存储模块，还包括：

建立数据删除表，所述数据删除表用于记录被删除的数据，当查询结果包含被删除的数据，则在解密时通过访问这个表跳过被删除数据的解密操作，从而返回的查询结果中不会包含已删除的数据。

一种基于可搜索加密的密文数据库查询方法及系统

技术领域

[0001] 本发明涉及数据处理领域，特别涉及一种基于可搜索加密的密文数据库查询方法及系统。

背景技术

[0002] 随着互联网的深入应用和大数据时代的来临，企业甚至个人都有大量的数据需要存储。目前许多云服务厂商为数据存储提供了性价比较高的外包存储服务，因此企业或个人通常都会购买这些存储服务来存储自己的数据。为了保护这些数据的安全性或隐私性，人们倾向于将这些数据加密后再存储在云服务器中，然而密文形式的数据给数据的使用带来了很大的阻碍。比如，常见的数据查询场景中，数据拥有者想通过某个关键字查询云服务器上存储的相关数据(数据可以是文档、音频、视频等)，但通常的加密算法并不支持在数据密文上进行查询。若是把待查询的数据下载到本地进行解密后查询，则通信代价和解密时间往往是不可接受的。为了解决密文数据上的查询问题，可搜索加密技术应运而生。该技术通过设计特殊的加密算法，将明文数据加密成具有索引结构的密文数据。将这些密文数据形成一个密文数据库，其中的索引结构具有可搜索的能力。查询时通过一个特定的密钥对查询的关键字生成查询陷门T，然后利用T来对密文索引进行搜索找到匹配关键字的数据，最后输出查询结果。

[0003] 根据生成密文索引的密钥和查询陷门的密钥是否相同，可搜索加密技术具体分为两类：对称可搜索加密SSE(Symmetric Searchable Encryption)和公钥可搜索加密PEKS(Public key Encryption with Keyword Search)。

[0004] 然而，现有可搜索加密算法搜索时需要扫描数据库全文，搜索时间与数据库大小成线性关系，算法效率较低，在数据分享场景下没有兼顾安全性和高效性，并且无法对密文查询请求进行有效的控制，难以干预密文数据的解密和感知高价值数据资源，影响数据提供者的数据分享收益。

发明内容

[0005] 基于此，本申请实施例提供了一种基于可搜索加密的密文数据库查询方法及系统，解决了现有可搜索加密算法在密文数据库查询中存在的问题。

[0006] 第一方面，提供了一种基于可搜索加密的密文数据库查询方法，该方法包括：

[0007] 获取明文源数据，所述明文源数据中包括数据文件名和数据内容；基于数据加密密钥对所述数据内容进行数据加密得到密文数据内容；

[0008] 对所述数据内容进行关键字提取得到关键字列表，并根据所述数据文件名基于索引生成密钥得到密文索引库，并将所述密文数据内容和所述密文索引库对应存储进密文数据库；

[0009] 获取目标查询关键字，通过陷门生成密钥对所述目标查询关键字进行处理生成查询密文；

[0010] 基于所述查询密文在密文索引库中进行查询得到密文查询结果并获取其对应的密文数据内容；

[0011] 基于数据解密密钥对所述密文数据内容进行数据解密得到明文数据内容。

[0012] 在基于数据解密密钥对所述密文数据内容进行数据解密得到明文数据内容之前，所述方法还包括：响应于明文源数据提供方的解密授权请求，生成该明文源数据的数据解密密钥。

[0013] 可选地，基于数据加密密钥对所述数据内容进行数据加密得到密文数据内容，包括：

[0014] 通过分组加密算法对数据内容进行加密，其中，所述分组加密算法至少包括AES算法和SM4算法。

[0015] 可选地，数据加密密钥和所述索引生成密钥，包括：

[0016] 生成一个随机值作为密钥生成的种子，然后利用这个种子和一个密钥导出函数为每个文件生成对应的加密密钥。

[0017] 可选地，将所述密文数据内容和所述密文索引库对应存储进密文数据库，还包括：

[0018] 在密文数据库中建立数据删除表，所述数据删除表用于记录被删除的数据，当查询结果包含被删除的数据，则在解密时通过访问这个表跳过被删除数据的解密操作，从而返回的查询结果中不会包含已删除的数据。

[0019] 可选地，在通过陷门生成密钥对所述目标查询关键字进行处理生成查询密文之前，包括：

[0020] 基于token/nonce的陷门生成算法得到陷门生成密钥。

[0021] 第二方面，提供了一种基于可搜索加密的密文数据库查询系统，该系统包括：

[0022] 数据加密模块，用于获取明文源数据，所述明文源数据中包括数据文件名和数据内容；基于数据加密密钥对所述数据内容进行数据加密得到密文数据内容；

[0023] 密文索引库构造模块，用于对所述数据内容进行关键字提取得到关键字列表，并根据所述数据文件名基于索引生成密钥得到密文索引库，并将所述密文数据内容和所述密文索引库对应存储进密文数据库；

[0024] 密文数据库存储模块，用于对密文数据内容和密文索引库进行存储；

[0025] 查询陷门生成模块，用于获取目标查询关键字，通过陷门生成密钥对所述目标查询关键字进行处理生成查询密文；

[0026] 密文索引库查询模块，基于所述查询密文在密文索引库中进行查询得到密文查询结果并获取其对应的密文数据内容；

[0027] 数据解密模块，用于基于数据解密密钥对所述密文数据内容进行数据解密得到明文数据内容。

[0028] 可选地，系统还包括：

[0029] 密钥管理和授权模块，响应于明文源数据提供方的解密授权请求，生成该明文源数据的数据解密密钥。

[0030] 可选地，数据加密模块包括：

[0031] 通过分组加密算法对数据内容进行加密，其中，所述分组加密算法至少包括AES算法和SM4算法。

[0032] 可选地，密文数据库存储模块，还包括：

[0033] 建立数据删除表，所述数据删除表用于记录被删除的数据，当查询结果包含被删除的数据，则在解密时通过访问这个表跳过被删除数据的解密操作，从而返回的查询结果中不会包含已删除的数据。

[0034] 本申请实施例提供的技术方案带来的有益效果至少包括：

[0035] (1)采用一数据一密钥的高强度加密方式、基于对称密码原语以及倒排索引搜索等技术设计了一种高效安全的密文查询方法和系统，尤其适用于数据的安全共享和数据价值的流通场景。

[0036] (2)支持数据新增或删除等数据动态变化下的密文查询，适配数据的可更新性，满足实际应用场景中数据变化的需求。

[0037] (3)利用token机制保证了数据查询请求的即时性和合法性，避免敌手发起非法查询或重放攻击，进一步提高了数据查询的业务安全性。

[0038] (4)数据提供者可以根据授权信息统计出被频繁查询的数据，从而能够感知数据价值并进行数据标价的灵活调整。在实际场景中，这有利于数据提供者在数据共享中获得更加合理的收益。

附图说明

[0039] 为了更清楚地说明本发明的实施方式或现有技术中的技术方案，下面将对实施方式或现有技术描述中所需要使用的附图作简单地介绍。显而易见地，下面描述中的附图仅仅是示例性的，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据提供的附图引申获得其它的实施附图。

[0040] 图1为本申请实施例提供的一种基于可搜索加密的密文数据库查询方法流程图；

[0041] 图2为本申请实施例提供的明文源数据的加密示意图；

[0042] 图3为本申请实施例提供的密文索引数据库EDB示意图；

[0043] 图4为本申请实施例提供的明文数据形式的查询结果示意图；

[0044] 图5为本申请实施例提供的一种基于可搜索加密的密文数据库查询系统框图。

具体实施方式

[0045] 为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

[0046] 在本发明的描述中，术语“包括”、“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包括了一系列步骤或单元的过程、方法、系统、产品或设备不必限于已明确列出的那些步骤或单元，而是还可包含虽然并未明确列出的但对于这些过程、方法、产品或设备固有的其它步骤或单元，或者基于本发明构思进一步的优化方案所增加的步骤或单元。

[0047] 本申请针对有多个数据提供者的数据共享场景设计了一种可搜索加密的密文数据库查询方法和系统架构，解决了以下技术问题：

[0048] 1、查询的高效性：设计的密文数据库查询方法从架构设计上规避了公钥加密技

术,全部采用对称密码原语进行设计,利用系统级的密钥共享完成明文数据的索引结构和密文数据生成,搜索时采用倒排索引技术,保证了数据加解密的效率和密文数据库搜索效率。

[0049] 2、查询的安全性:算法采用不同的密钥生成索引结构和数据密文,密文数据库中存储的索引和数据都是密文形式。此外,查询时关键字也会首先加密转换成查询陷门,密文数据库无法从查询陷门中恢复出关键字。这样保证了原始数据和查询关键字的安全性。

[0050] 3、查询的可控制性:算法中通过查询令牌token或是随机数nonce来对查询请求进行有效控制,即使针对同一个关键字的查询请求,生成的查询陷门也是不同的。重放的查询请求也可以准确地被检测出来。

[0051] 4、解密需要授权,数据价值可以进行排序:数据查询者获得密文形式的查询结果后,发起解密请求,系统向数据提供者请求解密授权,授权通过后系统将解密后的明文数据发送给数据查询方。数据提供者可以根据数据解密的授权记录进行统计,发现查询频率较高的高价值数据资源,从而可以调整数据的定价来提高数据共享的收益。

[0052] 具体地,请参考图1,其示出了本申请实施例提供的一种基于可搜索加密的密文数据库查询方法的流程图,该方法可以包括以下步骤:

[0053] S1,获取明文源数据,基于数据加密密钥对数据内容进行数据加密得到密文数据内容。

[0054] 其中,明文源数据中包括数据文件名和数据内容。本步骤对应于图中箭头1和2。步骤S1中涉及到了明文源数据的加密算法PlainEnc。具体地:

[0055] 明文源数据的加密算法对数据提供者提供的明文源数据进行加密,产生对应的密文数据。

[0056] 算法输入:明文源数据M,加密密钥K;

[0057] 算法输出:密文数据C=Enc(M,K),其中Enc为一个分组加密算法,如AES算法、SM4算法等。可选地,本申请对称加解密算法不局限于AES、SM4等分组密码算法,实际上可以使用其他任何分组密码算法或者流密码算法进行替换。

[0058] 算法详细描述:

[0059] 以明文源数据为文件为例,假设一共有三个数据提供者A,B,C,他们提供的文件列表分别为{A-01,A-02,A-03},{B-01,B-02},{C-01}.对这些文件分别生成明文数据的加密密钥{K_{A01},K_{A02},K_{A03}},{K_{B01},K_{B02}},{K_{C01}}.采用分组密码算法Enc对这些文件内容分别进行加密,如对明文A-01的数据内容加密得到密文文件Cipher_A01=Enc(A01,K_{A01}).类似地,加密所有的明文源数据后得到对应的密文文件{Cipher_A01,Cipher_A02,Cipher_A03},{Cipher_B01,Cipher_B02}以及{Cipher_C01},如图2所示。

[0060] 具体实现时,为了减少明文数据加密密钥的存储量,可以只为每个数据提供者生成一个随机值作为密钥生成的种子,然后利用这个种子和一个密钥导出函数KDF为每个数据提供者的每个文件生成对应的加密密钥,比如,为A用户设置密钥种子seed_A,利用KDF生成A用户的3个明文数据加密密钥为:

[0061] K_{A0i}=KDF(seed_A,A-0i),其中i=1,2,3。

[0062] S2,对数据内容进行关键字提取得到关键字列表,并根据数据文件名基于索引生成密钥得到密文索引库,并将密文数据内容和密文索引库对应存储进密文数据库。

[0063] 其中，本步骤对应于图中箭头3和4。步骤S2中涉及到密文索引数据库构造算法BuildIndex，具体地：

[0064] 密文索引数据库构造算法是对所有数据提供者的明文数据包含的关键字进行处理，生成密文索引，这些密文索引共同形成密文索引数据库，如图3所示。

[0065] 算法输入：所有的明文源数据{M_1,M_2,…,M_n}，每个源数据M_i提取的关键字列表{W_i1,W_i2,…,W_im}，即假设总共有n个源数据，第i个源数据包含有im个关键字。

[0066] 算法输出：密文索引生成密钥K，密文索引数据库EDB。

[0067] 算法详细描述：

[0068] (a)设所有数据的关键字集合为W，安全参数为λ，随机选择λ比特长度的关键字密文索引生成密钥K^*；

[0069] (b)对关键字集合中的每个关键字$w\in W$，用DB(w)表示包含关键字w的文件名的集合，利用密钥K^*和一个伪随机函数$F:\{0,1\}^{\lambda}\times\{0,1\}^{*}\rightarrow\{0,1\}^{\lambda}$，计算出两个密钥K1和K2：

[0070] $K1=F(K^*,w||key1)$，$K2=F(K^*,w||key2)$，

[0071] 其中符号||表示w与字符串“key1”或“key2”的级联。

[0072] 首先初始化密文索引数据库EDB为空，执行下面的循环：

[0073] 1)初始化计数器c＝0；

[0074] 2)计算标签l＝F(K1,c)，选择$id\in DB(w)$，计算其密文d＝Enc(K2,id)；

[0075] 3)计数器递增c++；

[0076] 4)将标签和密文对(l,d)添加到密文索引数据库中(l,d)→EDB.

[0077] (c)输出密文索引生成密钥K^*和密文索引数据库EDB。

[0078] S3，获取目标查询关键字，通过陷门生成密钥对目标查询关键字进行处理生成查询密文。

[0079] 其中，本步骤对应于图中箭头5和6，在本步骤中包括了查询陷门生成算法TrapGen，具体地：

[0080] 查询陷门生成算法是利用陷门生成密钥(与密文索引生成密钥K^*相同)对数据查询者的查询关键字w生成密文的查询条件(即查询陷门)，并提交到密文索引数据库EDB进行查询。

[0081] 本申请设计了基于token/nonce的陷门生成算法，这样数据查询者每次得到的查询陷门是不同的，即使对于同一个关键字w的两次查询得到的陷门也不相同。为了记号方便，本申请统一使用token来表示查询令牌token或随机数nonce.

[0082] 算法输入：查询关键字和token对(w,token)；

[0083] 算法输出：查询陷门T.

[0084] 算法描述：

[0085] (a)利用索引生成密钥K^*计算$K1=F(K^*,w||key1)$，$H=F(K^*,token)$；

[0086] (b)计算$Kt=K1\oplus H$，符号$\oplus$表示异或运算；

[0087] (c)输出查询陷门T＝(Kt,H)。

[0088] S4，基于查询密文在密文索引库中进行查询得到密文查询结果并获取其对应的密文数据内容。

[0089] 由于密文索引只是用来寻找到对应的密文数据，最后对密文数据解密得到明文数

据即可，只有最终的明文数据才是数据查询者需要的，所以在本申请中无需返回给数据查询者明文索引，仅返回明文数据内容。

[0090] 在本实施例中，本步骤对应于图中箭头7、8和9。本步骤涉及密文索引数据库搜索算法Search，具体地：

[0091] 密文索引数据库搜索算法是利用查询陷门搜索密文索引数据库，得到对应标签的密文文件名集合。

[0092] 算法输入：查询陷门T，密文索引数据库EDB；

[0093] 算法输出：密文文件名的集合Cipher_ID；

[0094] 算法描述：

[0095] (a)初始化计算器c＝0，Cipher_ID＝empty；

[0096] (b)利用查询陷门T计算出密钥$K1=Kt\oplus H$，执行下面的循环：

[0097] 1)计算标签l＝F(K1,c)，在EDB里查询对应标签l的密文索引，查询结果d＝find(EDB,l)；

[0098] 2)如果上一步EDB的查询结果d不为空，则将d添加到Cipher_ID中，并且递增计数器c++；若查询结果d为空，则退出循环。

[0099] (c)输出密文文件名的集合Cipher_ID。

[0100] 在本申请一个可选的实施例中，在步骤S4还包括

[0101] 响应于明文源数据提供方的解密授权请求，生成该明文源数据的数据解密密钥。在本实施例中，本步骤对应于图中箭头10。

[0102] S5，基于数据解密密钥对密文数据内容进行数据解密得到明文数据内容。

[0103] 在本实施例中，本步骤对应于图中箭头11和12。本步骤涉及密文索引解密算法DecryptIndex以及密文数据解密算法CipherDec，具体地：

[0104] 密文索引解密算法是对密文文件名的集合Cipher_ID进行解密，得到对应的明文文件名集合Plain_ID。

[0105] 算法输入：本次查询信息(w,token,H,Cipher_ID)；

[0106] 算法输出：明文文件名集合Plain_ID.

[0107] 算法描述：

[0108] (a)利用查询信息中的token和H验证本次查询是否有效，计算$F(K^*,token)$，若和本次查询信息中的H相等，且token与之前查询历史查询的token值都不同，则查询有效，否则输出无效查询，算法退出；

[0109] (b)初始化Plain_ID为空，恢复出解密密钥$K2＝F(K^*,w||key2)$；

[0110] (c)对每个d∈Cipher_ID，计算：

[0111] 1)解密出明文文件名id＝Dec(K2,d)，其中Dec表示Enc对应的解密算法。

[0112] 2)将id添加到明文文件名集合Plain_ID。

[0113] (d)输出明文文件名集合Plain_ID。

[0114] 密文数据解密算法是对密文数据内容进行解密，首先根据Plain_ID中的文件名确认明文数据来源，然后通过数据提供者授权获取到对应的解密密钥，解密出明文的文件内容，过程如图4所示。

[0115] 算法输入：明文文件名集合Plain_ID；

[0116] 算法输出:明文数据集合PlainData.

[0117] 算法描述:

[0118] (a)初始化PlainData为空,对每个id∈Plain_ID,执行:

[0119] 1)获取id对应的解密密钥DK,例如id为A-02,则数据来源于提供者A,密文数据内容的解密密钥为DK=KDF(seed_A,A-02).

[0120] 2)解密得到明文数据内容Plain=Dec(Cipher,DK),其中Cipher表示对应于本次id的密文数据内容。

[0121] 3)将明文数据Plain添加到集合PlainData中。

[0122] (b)输出明文数据集合PlainData。

[0123] 如图5,利用上面给出的6个算法,本申请设计了基于可搜索加密的密文数据库查询系统,系统由密钥管理和授权模块、数据加密模块、密文索引库构造模块、查询陷门生成模块、密文索引库查询模块、数据解密模块以及密文数据库存储模块等7个模块构成。

[0124] (1)密钥管理和授权模块:负责生成和管理数据加解密密钥和密文索引生成密钥,在数据提供者授权的前提下,提供密文索引生成密钥和数据解密密钥。

[0125] (2)数据加密模块:负责从密钥管理和授权模块获取数据加密密钥,运行PlainEnc算法对明文数据内容进行加密,产生密文数据内容。

[0126] (3)密文索引库构造模块:负责从密钥管理和授权模块获取密文索引生成密钥,运行BuildIndex算法对所有的数据名称和关键字列表生成密文索引库,并存储到密文数据库存储模块。此外,可以看出给出的密文索引数据库构造算法BuildIndex能够应对数据动态变化的情况(新增或删除),新增的数据只需要把数据文件名添加到其包含的关键字w的文件名列表DB(w)中,然后计算一个新的索引对(标签,文件名密文)即可,注意这时算法中计数器c是已有的最大值加1.数据删除时最简单的处理方式就是添加一个数据删除表,记录被删除的数据,若是查询结果包含被删除的数据,则在解密时通过访问这个表就可以跳过被删除数据的解密操作,从而返回的查询结果中不会包含已删除的数据。

[0127] (4)查询陷门生成模块:负责接收数据查询者的查询关键字,从密钥管理和授权模块获取陷门生成密钥,运行TrapGen算法生成对应的查询陷门。

[0128] (5)密文索引库查询模块:负责从查询陷门生成模块接收查询陷门,运行Search算法从密文存储模块中密文索引库进行查询,获得密文形式的查询结果。

[0129] (6)数据解密模块:负责从密文索引库查询模块接收密文查询结果,从密钥管理和授权模块获取索引生成密钥,运行DecryptIndex算法解密得到明文的文件名列表。解析文件名列表,发起内容解密授权请求,请求通过后运行CipherDec算法从密钥管理和授权模块获取到数据内容解密密钥,解密出明文数据内容,返回给数据查询者。

[0130] (7)密文数据库存储模块:负责统一存储密文索引库和密文数据内容,供密文索引库查询模块进行查询以及为数据解密模块提供密文数据内容。

[0131] 如图1,是实现上述系统的运行流程(用流程箭头上数字表示),从数据提供者提供的数据及关键字列表开始处理,首先生成密文数据内容和密文索引库,存储在密文存储模块中。然后数据查询者利用查询关键字发起查询,查询陷门生成模块产生出查询陷门,密文索引库查询模块利用查询陷门进行密文查询获得密文结果。数据解密模块解密密文查询结果,向相应的数据提供者发起授权并解密出明文结果返回给查询者。

[0132] (1)所有算法采用对称密码原语进行设计,避免采用双线性对或格密码等复杂和耗时的公钥密码,保证了算法执行的高效性。

[0133] (2)密文索引库搜索算法采用了倒排索引技术,即以关键字作为键值,其对应的数据值为包含该关键字的文件名id,这样能够提高搜索效率。因为正排的索引(文件名,关键字)在搜索时只能按照文件名来搜索,搜索时间为文件总个数的线性复杂度,而使用倒排索引技术,则在搜索时的循环次数为包含有该关键字的文件个数,搜索时间为亚线性时间复杂度,提高了搜索效率。

[0134] 同时,对不同的数据提供者的不同数据文件采用了不同的加密密钥(一数据一密钥),高强度地保护了数据内容的安全性。

[0135] 密文存储模块中存储的都是密文形式的索引库和数据内容,即使密文存储模块被攻破,敌手也无法解密。

[0136] 数据查询者提供的查询关键字会被陷门生成模块加密,密文索引库查询模块和密文存储模块都不能知道查询的关键字,保证了关键字的安全性。

[0137] 本申请实施例提供的基于可搜索加密的密文数据库查询系统用于实现上述基于可搜索加密的密文数据库查询方法,关于基于可搜索加密的密文数据库查询系统的具体限定可以参见上文中对于基于可搜索加密的密文数据库查询方法的限定,在此不再赘述。上述基于可搜索加密的密文数据库查询系统中的各个部分可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于设备中的处理器中,也可以以软件形式存储于设备中的存储器中,以便于处理器调用执行以上各个模块对应的操作。

[0138] 以上所述实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0139] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但并不能因此而理解为对申请专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

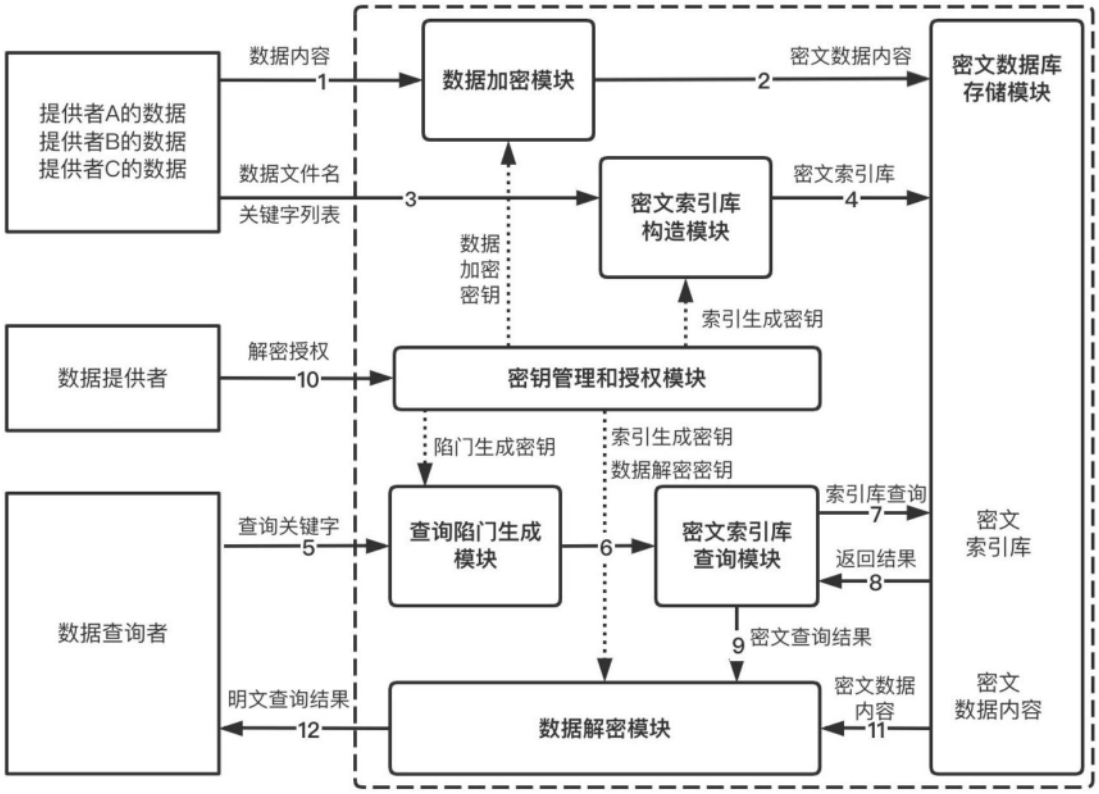

提供者A的数据
提供者B的数据
提供者C的数据
数据内容
1
数据加密模块
2
密文数据内容
密文数据库
存储模块
数据文件名
关键字列表
3
密文索引库
构造模块
4
密文索引库
数据
加密
密钥
索引生成密钥
数据提供者
解密授权
10
密钥管理和授权模块
陷门生成密钥
索引生成密钥
数据解密密钥
索引库查询
7
查询关键字
5
查询陷门生成
模块
6
密文索引库
查询模块
返回结果
8
密文
索引库
数据查询者
9
密文查询结果
明文查询结果
12
数据解密模块
密文数据
内容
11
密文
数据内容

图1

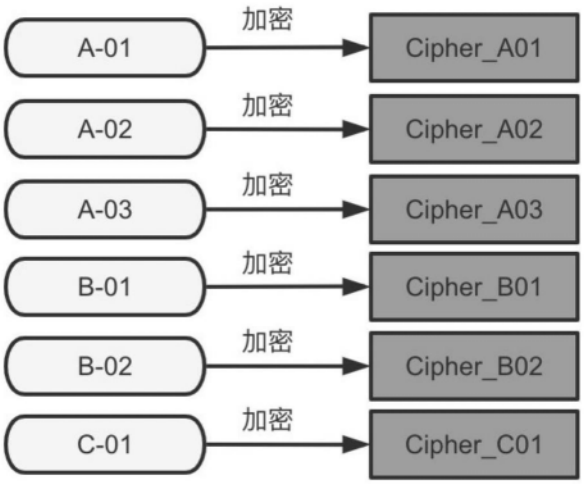

A-01
加密
Cipher_A01
A-02
加密
Cipher_A02
A-03
加密
Cipher_A03
B-01
加密
Cipher_B01
B-02
加密
Cipher_B02
C-01
加密
Cipher_C01

图2

索引编号	标签	id 密文
1	l_1	d_1
2	l_2	d_2
3	l_3	d_3
4	l_4	d_4
5	l_5	d_5

图3

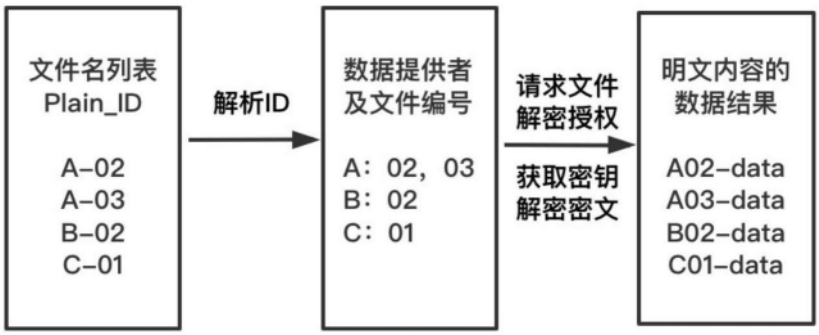


图4


密文数据库查询系统
密钥管理和授权模块
数据加密模块
密文索引库构造模块
查询陷门生成模块
密文索引库查询模块
数据解密模块
密文数据库存储模块

图5