(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 108108633 B
(45)授权公告日 2021.07.13

(21)申请号 201711386927.3
(22)申请日 2017.12.20
(65)同一申请的已公布的文献号
申请公布号 CN 108108633 A
(43)申请公布日 2018.06.01
(73)专利权人 中国科学院深圳先进技术研究院
地址 518000 广东省深圳市南山区深圳大学城学苑大道1068号
(72)发明人 胡希平 韩问寒 张佳 王飞 程俊
(74)专利代理机构 深圳中一专利商标事务所 44237
代理人 陈宇
(51)Int.Cl.
G06F 21/62(2013.01)

(56)对比文件
CN 102833346 A,2012.12.19
CN 102833346 A,2012.12.19
CN 103268455 A,2013.08.28
CN 106469281 A,2017.03.01
CN 102841902 A,2012.12.26
CN 1517906 A,2004.08.04

审查员 王慧敏

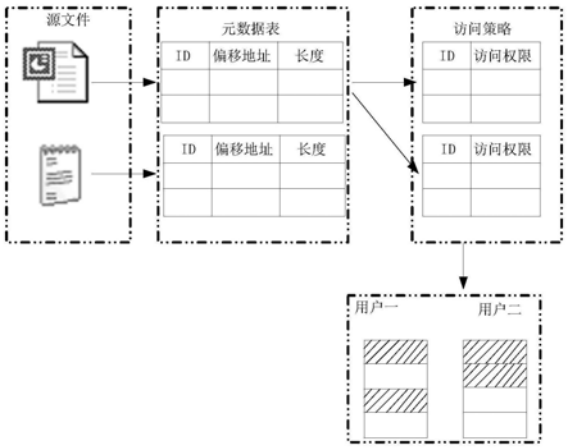
权利要求书2页 说明书8页 附图3页

(54)发明名称

一种数据文件及其访问方法、装置及设备

(57)摘要

一种数据文件包括源文件和元数据，其中：在所述源文件中包括有一个或者多个敏感数据；所述元数据包括用于记录所述源文件中的敏感数据的存储位置的元数据表，以及根据所述元数据表与敏感数据的存储位置的对应关系，结合用户对所述敏感数据的访问权限生成的访问策略。当不同的用户访问所述数据文件中的敏感数据时，只需要根据数据文件中的访问策略即可控制用户对敏感数据的访问，不需要对所述数据文件进行删除等修改，保证了分享过程中的可控制性和数据的一致性，可以根据所述访问策略有效的对数据文件进行控制。



CN 108108633 B

1.一种数据文件,其特征在于,所述数据文件包括源文件和元数据,其中:

在所述源文件中包括有一个或者多个敏感数据;所述源文件与一个或多个元数据关联;

所述元数据包括用于记录所述源文件中的敏感数据的存储位置的元数据表,以及根据所述元数据表与敏感数据的存储位置的对应关系,结合用户对所述敏感数据的访问权限生成的访问策略;其中,所述存储位置为字节位置,所述元数据表中包括一个或者多个元组,所述元组记录有敏感数据的偏移量、敏感数据的长度。

2.根据权利要求1所述的数据文件,其特征在于,所述访问策略根据用户的权限选择所述元数据中的一个或者多个元组。

3.根据权利要求1所述的数据文件,其特征在于,所述数据文件通过加密的方式存储在文件服务器,通过文件服务器中的访问表记录用户访问数据文件的访问记录。

4.一种数据文件的访问方法,其特征在于,所述数据文件包括源文件和元数据,在所述源文件中包括有一个或者多个敏感数据;所述源文件与一个或多个元数据关联;所述数据文件的访问方法包括:

接收文件数据的读取指令,获取所述文件数据对应的元数据,所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略,所述元数据包括元数据表,所述存储位置为字节位置,所述元数据表中包括一个或者多个元组,所述元组记录有敏感数据的偏移量、敏感数据的长度;

根据所获取的元数据,查找用户所对应的访问策略,所述访问策略根据用户对所述敏感数据的访问权限生成;

根据所述访问策略,控制所述用户对所述数据文件的访问。

5.根据权利要求4所述的数据文件的访问方法,其特征在于,所述根据所述访问策略,控制所述用户对所述数据文件的访问的步骤包括:

当所述访问策略允许用户访问所述数据文件中的敏感数据时,则开始读取所述数据文件中的敏感数据;

当所述访问策略不允许用户所述数据文件中的敏感数据时,则跳过该敏感数据的访问。

6.根据权利要求4所述的数据文件的访问方法,其特征在于,所述方法还包括:

当在所述数据文件中写入数据时,将写入数据后的数据文件与写入数据前的数据文件比较,获取新写入的数据信息;

将新写入的数据信息记录在所述数据文件对应的元数据中。

7.根据权利要求4所述的数据文件的访问方法,其特征在于,所述数据文件存储在中心文件服务器,所述数据文件的访问方法还包括:

接收到终端的发送的加密的访问请求,记录所述访问请求的用户,以及所访问的文件信息;

根据所述访问请求查找对应的数据文件,将所述数据文件发送给终端。

8.一种数据文件的访问装置,其特征在于,所述数据文件包括源文件和元数据,在所述源文件中包括有一个或者多个敏感数据;所述源文件与一个或多个元数据关联;所述数据文件的访问装置包括:

元数据获取单元，用于接收文件数据的读取指令，获取所述文件数据对应的元数据，所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略，所述元数据包括元数据表，所述存储位置为字节位置，所述元数据表中包括一个或者多个元组，所述元组记录有敏感数据的偏移量、敏感数据的长度；

访问策略查找单元，用于根据所获取的元数据，查找用户所对应的访问策略，所述访问策略根据用户对所述敏感数据的访问权限生成；

访问控制单元，用于根据所述访问策略，控制所述用户对所述数据文件的访问。

9.一种数据文件的访问设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现如权利要求4至7任一项所述数据文件的访问方法的步骤。

10.一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求4至7任一项所述数据文件的访问方法的步骤。

一种数据文件及其访问方法、装置及设备

技术领域

[0001] 本发明属于数据安全领域，尤其涉及一种数据文件及其访问方法、装置及设备。

背景技术

[0002] 随着移动互联网、智能硬件、传感器的发展，数据量越来越多，各个公司搜集的个人数据安全无法得到保障，数据泄露、买卖的事件层出不穷，已经对普通人的正常生活造成了巨大的影响。在个人逐渐被数据化的时代，国家开始立法管制，16年底颁布的《中华人民共和国网络安全法》对数据采集、使用都有明确的规定。立法一定程度上对个人数据的使用进行了限制，但是在数据平台上对敏感数据的保护的技术还有待提高，因为在进行数据共享时，也极有可能因为技术本身的缺陷造成敏感数据的泄露或者数据共享不便。

[0003] 在现在的云服务环境，尤其是PaaS(英文全称为Platform-as-a-Service，中文全称为平台即服务)或SaaS(英文全称为Software-as-a-service，中文全称为软件即服务)中，敏感数据是非常常见的。例如消费者的姓名、账户信息以及各种各样的数据形式例如文档、服务工单描述、报告等等。在服务提供的过程中，即使可能并不是故意为之，也很容易造成敏感数据的泄露。

[0004] 在隐私数据保护领域，最新的方案仅仅将数据保护推进到了文件级别(对于非结构化数据)或数据表级别(对于结构化数据库)。在这样的粗粒度下，数据的一致性将很难维护，因为数据源平台为了保护敏感数据，可能直接将其删除后再把数据分享给其他平台，这就造成了数据不一致。而为了达到更细的粒度，现有的方法是强行在数据库中将数据拆分到多个不同访问权限的数据表中，但是这将丢失数据操作的弹性。除此之外，在数据传输到其他平台(例如通过邮箱发送)的过程中，对数据的控制将会失效。

发明内容

[0005] 有鉴于此，本发明实施例提供了敏感数据访问方法、装置及设备，以解决现有技术中敏感数据分享时，容易造成数据不一致，或者丢失数据操作弹性，或者对数据失去控制权，无法管理数据获取的权限的问题。

[0006] 本发明实施例的第一方面提供了一种数据文件，所述数据文件包括源文件和元数据，其中：

[0007] 在所述源文件中包括有一个或者多个敏感数据；

[0008] 所述元数据包括用于记录所述源文件中的敏感数据的存储位置的元数据表，以及根据所述元数据表与敏感数据的存储位置的对应关系，结合用户对所述敏感数据的访问权限生成的访问策略。

[0009] 结合第一方面，在第一方面的第一种可能实现方式中，所述元数据表中包括一个或者多个元组，所述元组记录有敏感数据的偏移量、敏感数据的长度，所述访问策略根据用户的权限选择所述元数据中的一个或者多个元组。

[0010] 结合第一方面，在第一方面的第二种可能实现方式中，所述数据文件通过加密的

方式存储在文件服务器，通过文件服务器中的访问表记录用户访问数据文件的访问记录。

[0011] 本发明实施例的第二方面提供了一种数据文件的访问方法，所述数据文件的访问方法包括：

[0012] 接收文件数据的读取指令，获取所述文件数据对应的元数据，所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略；

[0013] 根据所获取的元数据，查找用户所对应的访问策略，所述访问策略根据用户对所述敏感数据的访问权限生成；

[0014] 根据所述访问策略，控制所述用户对所述数据文件的访问。

[0015] 结合第二方面，在第二方面的第一种可能实现方式中，所述根据所述访问策略，控制所述用户对所述数据文件的访问的步骤包括：

[0016] 当所述访问策略允许用户访问所述数据文件中的敏感数据时，则开始读取所述数据文件中的敏感数据；

[0017] 当所述访问策略不允许用户所述数据文件中的敏感数据时，则跳过该敏感数据的访问。

[0018] 结合第二方面，在第二方面的第二种可能实现方式中，所述方法还包括：

[0019] 当在所述数据文件中写入数据时，将写入数据后的数据文件与写入数据前的数据文件比较，获取新写入的数据信息；

[0020] 将新写入的数据信息记录在所述数据文件对应的元数据中。

[0021] 结合第二方面，在第二方面的第三种可能实现方式中，所述数据文件存储在中心文件服务器，所述数据文件的访问方法还包括：

[0022] 接收到终端的发送的加密的访问请求，记录所述访问请求的用户，以及所访问的文件信息；

[0023] 根据所述访问请求查找对应的数据文件，将所述数据文件发送给终端。

[0024] 本发明实施例的第三方面提供了一种数据文件的访问装置，所述数据文件的访问装置包括：

[0025] 元数据获取单元，用于接收文件数据的读取指令，获取所述文件数据对应的元数据，所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略；

[0026] 访问策略查找单元，用于根据所获取的元数据，查找用户所对应的访问策略，所述访问策略根据用户对所述敏感数据的访问权限生成；

[0027] 访问控制单元，用于根据所述访问策略，控制所述用户对所述数据文件的访问。

[0028] 本发明实施例的第四方面提供了一种数据文件的访问设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现如第二方面任一项所述数据文件的访问方法的步骤。

[0029] 本发明实施例的第四方面提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时实现如第二方面任一项所述数据文件的访问方法的步骤。

[0030] 本发明实施例与现有技术相比存在的有益效果是：当所述数据文件中包括敏感数据时，将所述敏感数据的存储位置记录在所述数据文件中的元数据中，并且所述元数据中还包括根据用户对所述敏感数据的访问权限生成的访问策略，当不同的用户访问所述数据

文件中的敏感数据时，只需要根据数据文件中的访问策略即可控制用户对敏感数据的访问，不需要对所述数据文件进行删除等修改，保证了分享过程中的可控制性和数据的一致性，可以根据所述访问策略有效的对数据文件进行控制。

附图说明

[0031] 为了更清楚地说明本发明实施例中的技术方案，下面将对实施例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本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

[0032] 图1是本发明实施例提供的数据文件的结构示意图；

[0033] 图2是本发明实施例提供的数据文件的表现形式的示意图；

[0034] 图3是本发明实施例提供的数据文件的访问方法的实现流程示意图；

[0035] 图4是本发明实施例提供的数据文件的访问装置的示意图；

[0036] 图5为本发明实施例提供的数据加密机制结构示意图；

[0037] 图6是本发明实施例提供的数据文件的访问设备的示意图。

具体实施方式

[0038] 以下描述中，为了说明而不是为了限定，提出了诸如特定系统结构、技术之类的具体细节，以便透彻理解本发明实施例。然而，本领域的技术人员应当清楚，在没有这些具体细节的其它实施例中也可以实现本发明。在其它情况中，省略对众所周知的系统、装置、电路以及方法的详细说明，以免不必要的细节妨碍本发明的描述。

[0039] 为了说明本发明所述的技术方案，下面通过具体实施例来进行说明。

[0040] 如图1所示为本申请实施例所述数据文件的结构示意图，所述数据文件包括源文件，所述源文件可以为不同格式的文件。比如，所述源文件可以WORD文件、POWERPOINT文件、TXT文件等等。在所述源文件中可以包括一个或者多个敏感数据。所述敏感数据可以为所述数据文件中的用户隐私数据，所述敏感数据可以根据敏感数据的特征关键词，或者数据的类型，由系统自动的查找确定，也可以由用户指定。

[0041] 在数据文件中，还包括与源文件绑定的元数据，所述元数据用于记录所述源文件中的敏感数据的存储位置，如图1所示，所述元数据中包括元数据表，所述元数据表中可以记录有所述敏感数据的存储位置。比如，所述敏感数据的存储位置可以通过敏感数据的偏移量和敏感数据的长度来表示。通过所述敏感数据的偏移量，可以确定敏感数据的起始位置的存储地址，根据所述敏感数据的偏移量和所述敏感数据的长度，可以确定所述敏感数据的结束位置的存储地址。根据所述起始位置的存储地址和结束位置的存储地址，即可确定所述敏感数据的存储地址。

[0042] 每个数据文件可以和一个或多个元数据关联，元数据通过元数据表记录敏感数据的字节位置，从而可以达到在字节级别上对敏感数据进行保护。所述元数据表可以包括多个元组，每个元组可以标记一个敏感数据的存储地址，所述元组的个数可以与所述敏感数据的个数对应。如图1所示，所述元组可以包括ID号和标记的存储地址。在生成访问策略时，可以通过所述ID号的选择，确定所选择的敏感数据。如图1所示，用户1对应于所选择元数据

表中的元组1(ID1)和元组3(ID3)所生成的访问策略，用户2对应于所选择元数据表中的元组1(ID2)和元组3(ID3)所生成的访问策略。

[0043] 由多个用户的访问策略可以生成访问策略组，不同的访问策略组可以通过不同的元数据关联同一个源文件，在组中用户访问数据时，通过相应的元数据表进行权限的管理。因此在数据分享的过程中，不需要生成文件的副本，更不需要对源文件进行修改，保证了分享过程的可控制性和数据的一致性。字节级别下的细粒度标记使得权限访问更加灵活，实现对不同用户的动态管理，对嵌入在文件中的复杂敏感数据的保护也更加全面。

[0044] 通过元数据记录数据文件中的零星的敏感数据的偏移量、敏感数据的长度以及敏感数据的访问权限，使得应用程序读取文件时，会产生相应的转化操作。当应用程序读取文件时，系统会把读取数据文件中的源文件所绑定的元数据，根据元数据中的访问策略，查询用户是否具备对敏感数据的访问权限。其中，所述权限可以为系统账户的权限，或者应用程序的权限，或者应用程序的账号所对应的权限等。所述权限还可以在访问策略中约定接收到所述数据文件的任意用户所具有的通用的访问权限。

[0045] 当通过访问策略检测到用户具有对敏感数据的访问权限时，可以授权用户具有访问权限。当通过访问策略检测到用户不具有对敏感数据的访问权限时，则跳过该敏感数据，对该敏感数据之后的其它非敏感数据进行访问，或者访问其它具有访问权限的敏感数据。

[0046] 当应用程序对数据文件写入数据时，可以将写入数据文件中的数据中的记录标记在元数据中。具体可以通过将写入数据后的数据文件与写入数据前的数据文件(上一版本的数据文件)进行比较，记录写入数据的位置，或者还可以包括记录写入数据的长度，或者还可以包括写入数据的敏感数据的存储位置等。通过元数据记录写入数据的记录，可以使得数据分享过程中保持数据的一致性，省去对敏感信息检索和模糊化的资源消耗，对系统性能有一定的提升。

[0047] 另外，所述数据文件的敏感数据在系统中的表现形式分为应用层、文件系统层和硬件层的表现形式，如图2所示，各层分别描述如下：

[0048] 在应用层，所述敏感数据会通过图形化方式高亮显示，即在图形化的界面中，高亮显示所述敏感数据，方便用户或者管理人员了解数据的结构。

[0049] 在文件系统层，表现的是源文件和元数据的虚拟地址，在这一层数据文件和元数据形成关联，转化为文件缓冲地址，所述数据文件的源文件和元数据存储在一块连续的虚拟地址，共同读出写入。

[0050] 在硬件层，源文件和元数据都将作为字节数据在磁盘存储，不存在逻辑意义上的关联。

[0051] 如图3所示为本申请实施例提供的一种数据文件的访问方法的实现流程示意图，详述如下：

[0052] 在步骤S301中，接收文件数据的读取指令，获取所述文件数据对应的元数据，所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略；

[0053] 具体的，所述元数据为图1所述的元数据基本相同，所述元数据包括标记敏感数据的存储地址的元数据表，以及用于确定用户是否具有对元数据表所标记的敏感数据的访问权限的访问策略。

[0054] 可以根据需要为不同的用户动态的分配访问策略，通过访问策略控制用户对数据

文件的访问。当发送数据文件或者分享数据文件时，只需要动态的调整数据文件中的元数据，不需要生成文件的副本，也不需要对源文件进行修改，即可实现对不同用户的动态管理，并且对嵌入在文件中的复杂敏感数据的保护也更加全面。

[0055] 在步骤S302中，根据所获取的元数据，查找用户所对应的访问策略，所述访问策略根据用户对所述敏感数据的访问权限生成；

[0056] 所述访问策略与用户对应，根据用户的不同，从而可以动态的生成不同的访问策略，实现对数据文件的访问权限的灵活控制。在某一用户的访问策略中，可以记录敏感数据的ID和权限信息，并且在元数据表中相应的记录有敏感数据的ID以及敏感数据的存储地址的对应关系。比如，所述敏感数据的存储地址可以包括敏感数据的偏移地址和敏感数据的长度等。

[0057] 在所述元数据表中记录有敏感数据的ID，因此，可以根据所述访问策略中的ID，在所述元数据表中查找到敏感数据的存储地址，从而可以确定对该存储地址所存储的敏感数据是否具有访问权限。

[0058] 在步骤S303中，根据所述访问策略，控制所述用户对所述数据文件的访问。

[0059] 根据所述访问策略，对所述数据文件进行访问时，如果检测用户具有对敏感数据的访问权限，则可以得到对所述敏感数据访问的授权，如果检测用户不具有对所述敏感数据的访问权限，则跳过该敏感数据。

[0060] 另外，在对所述数据文件进行访问时，还可以通过应用程序对所述数据文件写入数据。在写入数据时，可以通过和上一版本的数据文件进行比较，通过所述元数据记录写入数据的位置。因此，通过元数据的记录，可以保证数据文件的分享过程中保持数据的致性。

[0061] 作为本申请优选的一种实施方式，所述数据文件中的源文件还可以块的形式存储和传送。如图4所示为所述数据文件存储在中心文件服务器时，与终端的文件系统交互的示意图，在图中，所述中心文件服务器与终端的文件系统之间的数据文件的输入和输出，均通过加密的方式实现，从而可以保护用户的访问操作的隐私。

[0062] 如图4所示，在所述中心文件服务器中可以存储两张表格，包括访问表和块表，其中，所述访问表中记录有访问数据文件的用户ID、访问的数据文件的ID以及所述数据文件中的数据块ID，在所述块表中记录有数据文件ID、数据文件中的数据块ID和具体的源文件。通过所述访问表可以记录用户的访问操作，通过所述块表，可以快速的查找用户所需要访问的源文件的位置。

[0063] 所述中心文件服务器与终端的文件系统以块为单位进行输入和输出的操作，有利于提高读写性能，减少磁盘碎片。

[0064] 应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不应对本发明实施例的实施过程构成任何限定。

[0065] 图5为本申请实施例提供的一种数据文件访问装置的结构示意图，所述数据文件访问装置包括：

[0066] 元数据获取单元501，用于接收文件数据的读取指令，获取所述文件数据对应的元数据，所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略；

[0067] 访问策略查找单元502，用于根据所获取的元数据，查找用户所对应的访问策略，

所述访问策略根据用户对所述敏感数据的访问权限生成；

[0068] 访问控制单元503，用于根据所述访问策略，控制所述用户对所述数据文件的访问。

[0069] 优选的，所述访问控制单元包括：

[0070] 读取子单元，用于当所述访问策略允许用户访问所述数据文件中的敏感数据时，则开始读取所述数据文件中的敏感数据；

[0071] 跳过子单元，用于当所述访问策略不允许用户所述数据文件中的敏感数据时，则跳过该敏感数据的访问。

[0072] 优选的，所述装置还包括：

[0073] 写入数据信息获取单元，用于当在所述数据文件中写入数据时，将写入数据后的数据文件与写入数据前的数据文件比较，获取新写入的数据信息；

[0074] 记录单元，用于将新写入的数据信息记录在所述数据文件对应的元数据中。

[0075] 优选的，所述数据文件存储在中心文件服务器，所述数据文件的访问装置还包括：

[0076] 请求接收单元，用于接收到终端的发送的加密的访问请求，记录所述访问请求的用户，以及所访问的文件信息；

[0077] 数据文件发送单元，用于根据所述访问请求查找对应的数据文件，将所述数据文件发送给终端。

[0078] 图5所述数据文件的访问装置，与图3所述的数据文件的访问方法对应。

[0079] 图6是本发明一实施例提供的数据文件的访问设备的示意图。如图6所示，该实施例的数据文件的访问设备6包括：处理器60、存储器61以及存储在所述存储器61中并可在所述处理器60上运行的计算机程序62，例如数据文件的访问程序。所述处理器60执行所述计算机程序62时实现上述各个数据文件的访问方法实施例中的步骤，例如图3所示的步骤301至303。或者，所述处理器60执行所述计算机程序62时实现上述各装置实施例中各模块/单元的功能，例如图5所示模块501至503的功能。

[0080] 示例性的，所述计算机程序62可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器61中，并由所述处理器60执行，以完成本发明。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段，该指令段用于描述所述计算机程序62在所述数据文件的访问设备6中的执行过程。例如，所述计算机程序62可以被分割成元数据获取单元、访问策略查找单元和访问控制单元，各单元具体功能如下：

[0081] 元数据获取单元，用于接收文件数据的读取指令，获取所述文件数据对应的元数据，所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略；

[0082] 访问策略查找单元，用于根据所获取的元数据，查找用户所对应的访问策略，所述访问策略根据用户对所述敏感数据的访问权限生成；

[0083] 访问控制单元，用于根据所述访问策略，控制所述用户对所述数据文件的访问。

[0084] 所述数据文件的访问设备6可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述数据文件的访问设备可包括，但不仅限于，处理器60、存储器61。本领域技术人员可以理解，图6仅仅是数据文件的访问设备6的示例，并不构成对数据文件的访问设备6的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件，例如

所述数据文件的访问设备还可以包括输入输出设备、网络接入设备、总线等。

[0085] 所称处理器60可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0086] 所述存储器61可以是所述数据文件的访问设备6的内部存储单元,例如数据文件的访问设备6的硬盘或内存。所述存储器61也可以是所述数据文件的访问设备6的外部存储设备,例如所述数据文件的访问设备6上配备的插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。进一步地,所述存储器61还可以既包括所述数据文件的访问设备6的内部存储单元也包括外部存储设备。所述存储器61用于存储所述计算机程序以及所述数据文件的访问设备所需的其他程序和数据。所述存储器61还可以用于暂时地存储已经输出或者将要输出的数据。

[0087] 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,仅以上述各功能单元、模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能单元、模块完成,即将所述装置的内部结构划分成不同的功能单元或模块,以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中,上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。另外,各功能单元、模块的具体名称也只是为了便于相互区分,并不用于限制本申请的保护范围。上述系统中单元、模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0088] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中没有详述或记载的部分,可以参见其它实施例的相关描述。

[0089] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0090] 在本发明所提供的实施例中,应该理解到,所揭露的装置/终端设备和方法,可以通过其它的方式实现。例如,以上所描述的装置/终端设备实施例仅仅是示意性的,例如,所述模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以是通过一些接口,装置或单元的间接耦合或通讯连接,可以是电性,机械或其它的形式。

[0091] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0092] 另外，在本发明各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

[0093] 所述集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明实现上述实施例方法中的全部或部分流程，也可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器执行时，可实现上述各个方法实施例的步骤。。其中，所述计算机程序包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM，Read-Only Memory)、随机存取存储器(RAM，Random Access Memory)、电载波信号、电信信号以及软件分发介质等。需要说明的是，所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减，例如在某些司法管辖区，根据立法和专利实践，计算机可读介质不包括是电载波信号和电信信号。

[0094] 以上所述实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围，均应包含在本发明的保护范围之内。

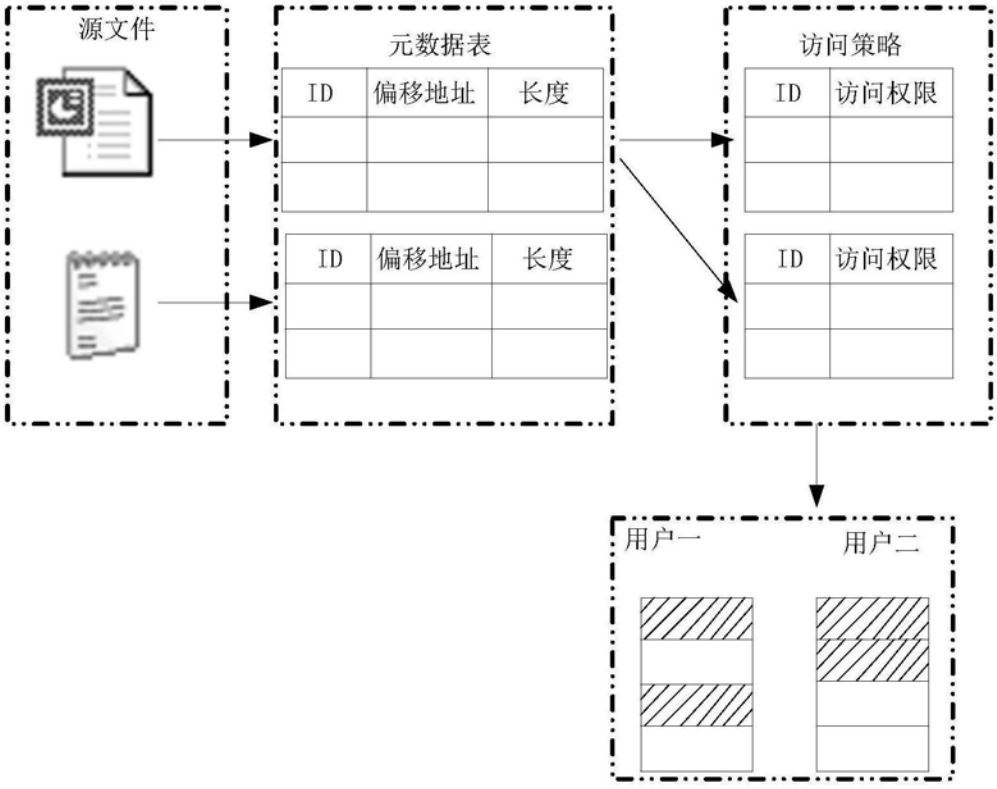

源文件
元数据表
ID
偏移地址
长度
ID
偏移地址
长度
访问策略
ID
访问权限
ID
访问权限
用户一
用户二

图1

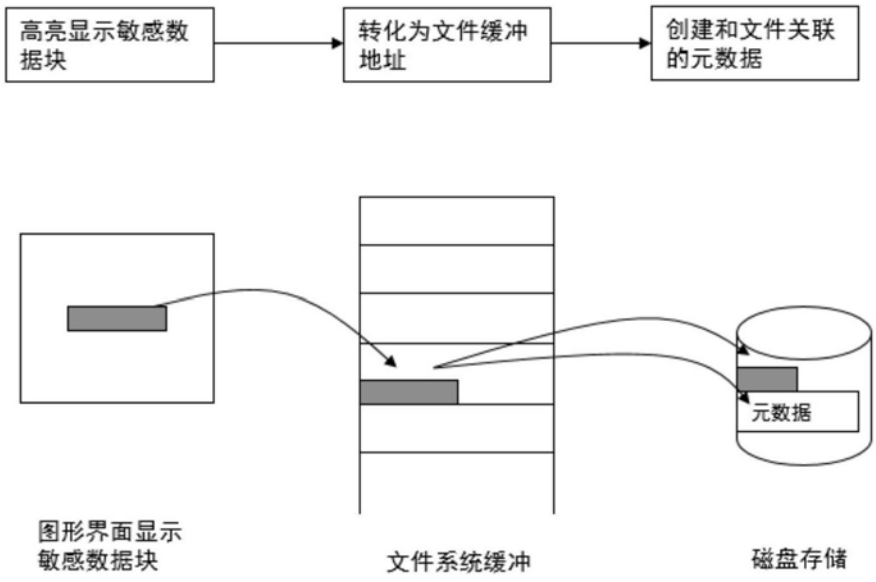

高亮显示敏感数据块
转化为文件缓冲地址
创建和文件关联的元数据
元数据
图形界面显示敏感数据块
文件系统缓冲
磁盘存储

图2

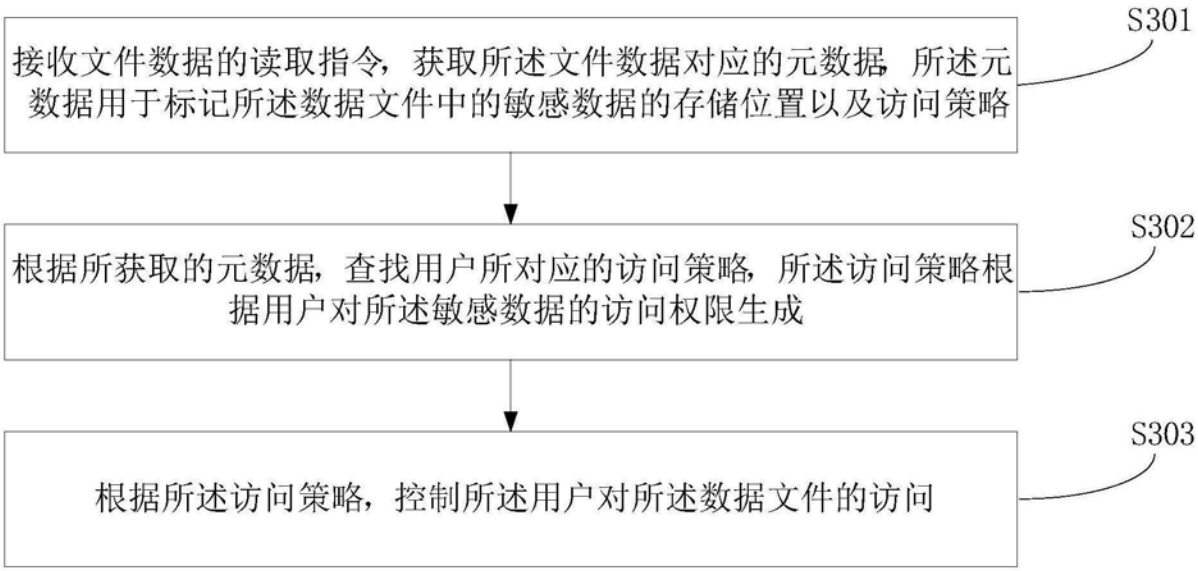

S301
接收文件数据的读取指令，获取所述文件数据对应的元数据，所述元数据用于标记所述数据文件中的敏感数据的存储位置以及访问策略
S302
根据所获取的元数据，查找用户所对应的访问策略，所述访问策略根据用户对所述敏感数据的访问权限生成
S303
根据所述访问策略，控制所述用户对所述数据文件的访问

图3

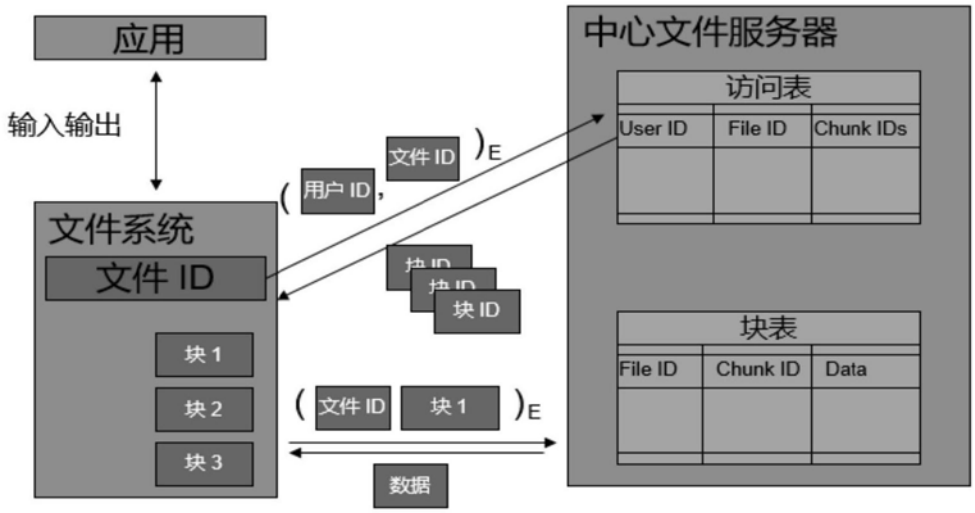

应用
输入输出
文件系统
文件 ID
块 1
块 2
块 3
(用户 ID , 文件 ID)E
块 ID
(文件 ID 块 1)E
数据
中心文件服务器
访问表
User ID
File ID
Chunk IDs
块表
File ID
Chunk ID
Data

图4

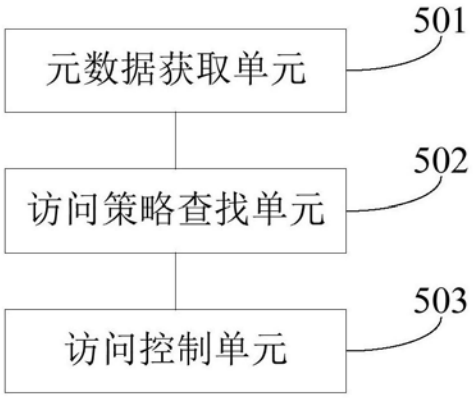

501
元数据获取单元
502
访问策略查找单元
503
访问控制单元

图5

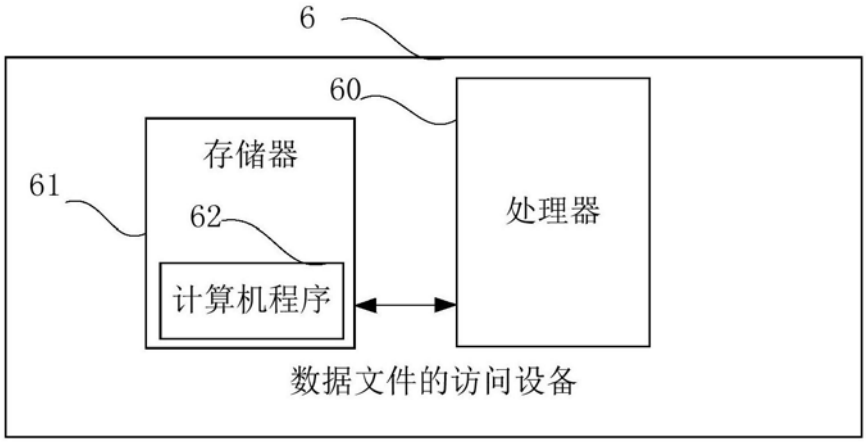

6
60
存储器
61
62
计算机程序
处理器
数据文件的访问设备

图6