



(12)发明专利申请

(10)申请公布号 CN 106844091 A

(43)申请公布日 2017.06.13

(21)申请号 201611125510.7

(22)申请日 2016.12.08

(71)申请人 深圳市金立通信设备有限公司

地址 518040 广东省深圳市福田区深南大道7028号时代科技大厦东座21楼

(72)发明人 张超

(74)专利代理机构 广州三环专利代理有限公司

44202

代理人 郝传鑫 熊永强

(51) Int. Cl.

G06F 11/14(2006.01)

G06F 17/30(2006.01)

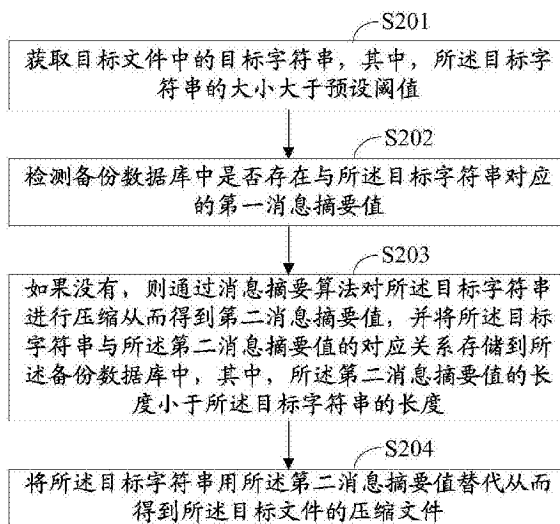
权利要求书2页 说明书10页 附图4页

(54)发明名称

一种压缩、还原方法及终端

(57)摘要

本发明实施例公开了一种压缩、还原方法及终端。一种压缩方法包括：获取目标文件中的目标字符串，其中，所述目标字符串的大小大于预设阈值；检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值；如果没有，则通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值，并将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中，其中，所述第二消息摘要值的长度小于所述目标字符串的长度；将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。所述方法能够对大字符串进行压缩存储，节省存储空间，将大字符串和消息摘要值的对应关系备份存储，方便还原所述压缩文件。



1. 一种压缩方法,其特征在于,包括:

获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设阈值;

检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值;

如果没有,则通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,并将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中,其中,所述第二消息摘要值的长度小于所述目标字符串的长度;

将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

2. 根据权利要求1所述的方法,其特征在于,在获取目标文件中的目标字符串之后,在通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值之前,还包括:

检测所述目标字符串在所述目标文件中的重复次数;

通过消息摘要算法对所述目标字符串进行压缩从而得到所述第二消息摘要值具体为:在所述目标字符串的重复次数大于数量阈值的情况下,则通过消息摘要算法对所述目标字符串进行压缩从而得到所述第二消息摘要值。

3. 根据权利要求1所述的方法,其特征在于,如果有所述第一消息摘要值,从所述备份数据库中获取与所述目标字符串对应的所述第一消息摘要值,并将所述目标字符串用所述对应的所述第一消息摘要值替代从而得到所述目标文件的压缩文件。

4. 根据权利要求1所述的方法,其特征在于,所述消息摘要算法包括:MD5、sha-1、RIPEMD、Haval中的任意一个。

5. 一种还原方法,其特征在于,包括:

获取压缩文件中的消息摘要值,其中,所述消息摘要值是通过消息摘要算法对目标字符串进行压缩得到的,所述消息摘要值的长度小于所述目标字符串的长度;

从备份数据库中获取与所述消息摘要值对应的目标字符串,其中,所述目标字符串与所述消息摘要值的对应关系存储在所述备份数据库中;

将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

6. 一种终端,其特征在于,包括:

第一获取单元,用于获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设阈值;

第一检测单元,用于检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值;

压缩单元,用于通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,其中,所述第二消息摘要值的长度小于所述目标字符串的长度;

存储单元,将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中;

替代单元,用于将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

7. 根据权利要求5所述的终端,其特征在于,还包括:

第二检测单元,用于检测所述目标字符串在所述目标文件中的重复次数;

所述压缩单元具体用于,在所述目标字符串的重复次数大于数量阈值的情况下,则通

过消息摘要算法对所述目标字符串进行压缩从而得到所述第二消息摘要值。

8. 根据权利要求5所述的终端,其特征在於,还包括:

第二获取单元,用于从所述备份数据库中获取与所述目标字符串对应的所述第一消息摘要值;

所述替代单元具体用于,将所述目标字符串用所述对应的所述第一消息摘要值替代从而得到所述目标文件的压缩文件。

9. 根据权利要求6所述的终端,其特征在於,所述消息摘要算法包括:MD5、sha-1、RIPEMD、Haval中的任意一个。

10. 一种终端,其特征在於,包括:

第一获取单元,用于获取压缩文件中的消息摘要值,其中,所述消息摘要值是通过消息摘要算法对目标字符串进行压缩得到的,所述消息摘要值的长度小于所述目标字符串的长度;

第二获取单元,用于从备份数据库中获取与所述消息摘要值对应的目标字符串,其中,所述目标字符串与所述消息摘要值的对应关系存储在所述备份数据库中;

替代单元,用于将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

一种压缩、还原方法及终端

技术领域

[0001] 本发明涉及电子技术领域,特别涉及一种压缩、还原方法及终端。

背景技术

[0002] 随着互联网的高速发展,每天都会产生数量庞大的数据,而其中又有很多是大字节的重复数据,在存储时浪费了存储空间,还会影响查询性能。为了节省存储空间,一般使用压缩算法对数据进行压缩存储。

[0003] 现有的压缩方法通过Quicklz,Snappy,zlib等算法来压缩数据,在需要用到数据时进行解压缩,以达到节省存储空间的目的,但现有的压缩算法都有一定的压缩百分比,受限于原始数据的大小,对原始数据的压缩效果较差,而且在解压缩的过程中还有可能损坏原始数据。

发明内容

[0004] 本发明实施例提供一种压缩、还原方法及终端,能够用摘要值替代大字符串进行存储,节省存储空间。

[0005] 第一方面,本发明实施例提供了一种压缩方法,包括:

[0006] 获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设阈值;

[0007] 检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值;

[0008] 如果没有,则通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,并将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中,其中,所述第二消息摘要值的长度小于所述目标字符串的长度;

[0009] 将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

[0010] 第二方面,本发明实施例提供了一种还原方法,包括:

[0011] 获取压缩文件中的消息摘要值,其中,所述消息摘要值是通过消息摘要算法对目标字符串进行压缩得到的,所述消息摘要值的长度小于所述目标字符串的长度;

[0012] 从备份数据库中获取与所述消息摘要值对应的目标字符串,其中,所述目标字符串与所述消息摘要值的对应关系存储在所述备份数据库中;

[0013] 将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

[0014] 第三方面,本发明实施例提供一种终端,包括:

[0015] 第一获取单元,用于获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设阈值;

[0016] 第一检测单元,用于检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值;

[0017] 压缩单元,用于通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息

摘要值,其中,所述第二消息摘要值的长度小于所述目标字符串的长度;

[0018] 存储单元,将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中;

[0019] 替代单元,用于将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

[0020] 第四方面,本发明实施例提供一种终端,包括:

[0021] 第一获取单元,用于获取压缩文件中的消息摘要值,其中,所述消息摘要值是通过消息摘要算法对目标字符串进行压缩得到的,所述消息摘要值的长度小于所述目标字符串的长度;

[0022] 第二获取单元,用于从备份数据库中获取与所述消息摘要值对应的目标字符串,其中,所述目标字符串与所述消息摘要值的对应关系存储在所述备份数据库中;

[0023] 替代单元,用于将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

[0024] 本发明实施例,通过获取目标文件中的目标字符串,检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值,如果没有,则通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,并将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中,将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件,能够节省存储空间,将大字符串备份存储,方便还原所述压缩文件。

附图说明

[0025] 为了更清楚地说明本发明实施例或现有技术中的技术方案,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的前提下,还可以根据这些附图获得其他的附图。

[0026] 图1是现有技术中日志记录的示意图;

[0027] 图2是本发明实施例提供的一种压缩方法的流程图;

[0028] 图3是本发明实施例提供的压缩后的日志记录的示意图;

[0029] 图4是本发明实施例提供的另一种压缩方法的流程图;

[0030] 图5是本发明实施例提供的一种还原方法的流程图;

[0031] 图6是本发明实施例提供的一种终端的结构示意图;

[0032] 图7是本发明实施例提供的另一种终端的结构示意图;

[0033] 图8是本发明实施例提供的另一种终端的结构示意图;

[0034] 图9是本发明实施例提供的另一种终端的结构示意图。

具体实施方式

[0035] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施

例,都属于本发明保护的范围内。

[0036] 需要说明的是,在本发明实施例中使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本发明。在本发明实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0037] 随着互联网的高速发展,每天都有大量的数据产生。在服务器中,日志数据库中保存有和终端进行交互的日志记录,在每一条日志记录中都包含有一个UserAgent字段,用于存储终端的基本信息(例如浏览器标识、操作系统标识、浏览器语言、版本信息等),所述UserAgent字段有200多个字节大小,而一条日志记录中其他的字段加起来不到80字节。针对同一个终端,所述日志数据库中存储有多条日志记录,每一条日志记录中的UserAgent字段都是相同的,相当于所述UserAgent字段重复存储了多次,浪费了存储空间。如图1所示,图1为一条日志记录中各字段字节大小的示意图。

[0038] 为了解决上述问题,本发明实施例公开了一种压缩方法。请参阅图2,图2是本发明实施例提供的一种压缩方法的流程图,所述方法包括以下步骤:

[0039] S201:获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设阈值。

[0040] 在本实施方式中,所述目标文件为需要进行压缩处理的文件,例如服务器中的日志数据库。所述目标字符串为所述目标文件中大小大于预设阈值的字符串,其中,所述预设阈值为用户根据所述目标文件的具体情况设置的或者默认的。

[0041] 在一具体实施例中,在服务器的日志数据库中,假设默认的预设阈值为200字节,一条日志记录中的UserAgent字段的大小超过了预设阈值200字节,那么所述UserAgent字段就为目标字符串。

[0042] 在本实施方式中,所述目标文件可以是已经存储好的文件,所述获取目标文件中的目标字符串具体为:定期地检测所述目标文件中是否有大小大于预设阈值的字符串,若有,则所述字符串为目标字符串,获取所述目标文件中的目标字符串;在其他实施方式中,所述目标文件可以是正在进行存储操作的文件,所述获取目标文件中的目标字符串具体为:在进行存储时,先检测所述目标文件中是否有大小大于预设阈值的字符串,若有,则所述字符串为目标字符串,获取所述目标文件中的目标字符串。

[0043] S202:检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值。

[0044] 具体的,所述备份数据库用于存储字符串和消息摘要值的对应关系,其中,所述消息摘要值为通过消息摘要算法对所述字符串进行压缩从而得到的。

[0045] 在本实施方式中,检测所述备份数据库中是否存在与所述目标字符串对应的第一消息摘要值,若没有,执行步骤S203-S204的步骤,保证所述备份数据库中针对同一个目标字符串的目标字符串和消息摘要值的对应关系只存储一次,不会重复存储,减少存储量,方便查找。

[0046] 在另一实施方式中,若所述备份数据库中不存在与所述目标字符串对应的第一消息摘要值,则执行步骤:从所述备份数据库中获取与所述目标字符串对应的所述第一消息摘要值,并将所述目标字符串用所述对应的所述第一消息摘要值替代从而得到所述目标文件

的压缩文件。具体的,若所述备份数据库中存在与所述目标字符串对应的第一消息摘要值,那么说明所述目标字符串在之前存储的文件中已经出现过,并且经过了压缩操作,存储在所述备份数据库中,为了保证所述备份数据库中针对同一个目标字符串的目标字符串和消息摘要值的对应关系只存储一次,所述目标字符串不需要再次压缩进行存储,只需从所述备份数据库中直接查找所述目标字符串对应的第一消息摘要值,使用所述第一消息摘要值替代所述目标文件中的目标字符串,得到所述目标文件的压缩文件。

[0047] S203:如果没有,则通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,并将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中,其中,所述第二消息摘要值的长度小于所述目标字符串的长度。

[0048] 在本实施方式中,所述消息摘要算法可以为MD5 (Message Digest Algorithm MD5,消息摘要算法第五版)、sha-1 (Secure Hash Algorithm,安全哈希算法)、RIPEMD (RACE Integrity Primitives Evaluation Message Digest,RACE原始完整性校验消息摘要)、Haval算法中的任意一个,还可以为其他的能够有效压缩大字符串的压缩算法。

[0049] 在本实施方式中,所述消息摘要算法将任意大小的目标字符串压缩成固定长度的第二消息摘要值,两个不同的目标字符串不会产生相同的消息摘要值,保证了目标字符串和消息摘要值之间唯一的对应关系。

[0050] S204:将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

[0051] 具体的,将所述目标文件中的所述目标字符串用所述第二消息摘要值替代,减小所述目标文件的大小,从而达到压缩所述目标文件的目的。

[0052] 在一具体实施例中,如图3所示,图3为日志记录中使用消息摘要值替代目标字符串的示意图,服务器的日志数据库中,对于一个终端有多条日志记录,每条日志记录中都包含相同的200多个字节的UserAgent字段,将所述UserAgent字段使用MD5算法压缩为固定的16字节的消息摘要值,所述消息摘要值在图3中以UAMd5表示,所述终端下的多条日志记录中的UserAgent字段都用所述16字节的消息摘要值替代进行存储。若所述日志数据库中对于所述终端有1000条日志记录,那么使用所述16字节的消息摘要值替代所述UserAgent字段进行存储后,总的存储量减少了约200000字节。

[0053] 上述可知,本发明实施例通过获取目标文件中的目标字符串,检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值,如果没有,则通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,并将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中,将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件,能够节省存储空间,将大字符串和消息摘要值的对应关系备份存储,方便还原所述压缩文件。

[0054] 请参阅图4,图4为本发明实施例提供的另一种压缩方法的流程图,所述方法包括以下步骤:

[0055] S401:获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设阈值。

[0056] 在本实施方式中,所述目标文件为已经存储好的文件,例如服务器中的日志数据库。所述目标字符串为所述目标文件中大小大于预设阈值的字符串,其中,所述预设阈值为

用户根据所述目标文件的具体情况设置的或者默认的。

[0057] 在一具体实施例中,在服务器的日志数据库中,默认的预设阈值为200字节,一条日志记录中的UserAgent字段的大小超过了预设阈值200字节,那么所述UserAgent字段就为目标字符串。

[0058] S402:检测所述目标字符串在所述目标文件中的重复次数。

[0059] S403:检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值。

[0060] 在本实施方式中,检测所述备份数据库中是否存在与所述目标字符串对应的第一消息摘要值,若没有,执行步骤S404-S406的步骤,保证所述备份数据库中针对同一个目标字符串的目标字符串和消息摘要值的对应关系只存储一次,不会重复存储,减少存储量,方便查找。

[0061] S404:在所述目标字符串的重复次数大于数量阈值,且所述数据备份库中不存在与所述目标字符串对应的第一消息摘要值的情况下,通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,其中,所述第二消息摘要值的长度小于所述目标字符串的长度。

[0062] 具体的,检测所述目标字符串在所述目标文件中的重复次数,即,检测所述目标字符串在所述目标文件中重复存储的次数,在所述目标字符串的重复次数大于数量阈值的情况下,则说明所述目标字符串重复存储的次数达到一定数量,重复存储而浪费的存储空间也达到一定的程度,在这种情况下对所述目标字符串进行压缩,能够节约更多的存储空间。其中,所述数量阈值为用户根据所述目标文件的具体情况设置的或者默认的。

[0063] 在本实施方式中,所述消息摘要算法可以为MD5(Message Digest Algorithm MD5,消息摘要算法第五版)、sha-1(Secure Hash Algorithm,安全哈希算法)、RIPEMD(RACE Integrity Primitives Evaluation Message Digest,RACE原始完整性校验消息摘要)、Haval算法中的任意一个,还可以为其他的能够有效压缩大字符串的压缩算法。在本实施方式中,所述消息摘要算法将任意大小的目标字符串压缩成固定长度的消息摘要值,两个不同的目标字符串不会产生相同的消息摘要值,保证了目标字符串和消息摘要值之间唯一的对应关系。

[0064] S405:将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中。

[0065] S406:将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

[0066] 具体的,将所述目标文件中的所述目标字符串用所述消息摘要值替代,减小所述目标文件的大小,从而达到压缩所述目标文件的目的。

[0067] 在一具体的实施例中,服务器的日志数据库中,对于一个终端有1000多条日志记录,每条日志记录中都包含相同的200多个字节的UserAgent字段,重复存储的UserAgent字段超过了数量阈值,将所述UserAgent字段使用MD5算法压缩为固定的16字节的消息摘要值,所述终端下的1000多条日志记录中的UserAgent字段都用所述16字节的消息摘要值替代进行存储。若所述日志数据库中对于所述终端有1000条日志记录,那么使用所述16字节的消息摘要值替代所述UserAgent字段进行存储后,总的存储量减少了约200000字节。

[0068] 上述可知,所述压缩方法通过获取目标文件中的目标字符串,在所述目标字符串

在所述目标文件中的重复此处大于数量阈值的情况下,通过消息摘要算法对所述目标字符串进行压缩从而得到消息摘要值,并将所述目标字符串与所述消息摘要值的对应关系存储到所述备份数据库中,将所述目标字符串用所述消息摘要值替代从而得到所述目标文件的压缩文件,能够节省存储空间,将大字符串和消息摘要值的对应关系备份存储,方便还原所述压缩文件。

[0069] 参阅图5,图5是本发明实施例提供的一种还原方法,所述方法针对上述图2或图4所示的压缩方法得到的压缩文件进行还原。所述方法包括以下步骤:

[0070] S501:获取压缩文件中的消息摘要值,其中,所述消息摘要值是通过消息摘要算法对目标字符串进行压缩得到的,所述消息摘要值的长度小于所述目标字符串的长度。

[0071] 在本实施方式中,在需要查看数据时,需要将压缩文件还原,首先获取所述压缩文件中的消息摘要值,所述消息摘要值是通过上述图2或图4所示方法实施例中通过消息摘要算法对目标字符串进行压缩得到的。

[0072] 在本实施方式中,所述消息摘要算法可以为MD5(Message Digest Algorithm MD5,消息摘要算法第五版)、sha-1(Secure Hash Algorithm,安全哈希算法)、RIPEMD(RACE Integrity Primitives Evaluation Message Digest,RACE原始完整性校验消息摘要)、Haval算法中的任意一个。在本实施方式中,所述消息摘要算法将任意大小的目标字符串压缩成固定长度的消息摘要值,两个不同的目标字符串不会产生相同的消息摘要值,保证了目标字符串和消息摘要值之间唯一的对应关系。

[0073] S502:从备份数据库中获取与所述消息摘要值对应的目标字符串,其中,所述目标字符串与所述消息摘要值的对应关系存储在所述备份数据库中。

[0074] S503:将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

[0075] 具体的,将所述压缩文件中的所述消息摘要值用所述对应的目标字符串替代,得到所述压缩文件的原始文件,从而达到还原所述压缩文件的目的。

[0076] 在一具体实施例中,服务器的日志数据库中,对于一个终端有多条日志记录,每条日志记录中都包含相同的200多个字节的UserAgent字段,所述多条日志记录在使用图2或图4所示的方法将UserAgent字段使用消息摘要值替代存储后,此时所述日志记录如图3所示,若需要对所述日志数据库中针对所述终端的日志记录进行分析或者其他的操作,就需要对所述压缩后的日志记录进行还原。首先,获取所述日志记录中的消息摘要值,在备份数据库中查找到所述消息摘要值对应的UserAgent字段,然后再用所述UserAgent字段代替日志记录中的消息摘要值,从而得到所述日志记录的原始文件,此时所述日志记录如图1所示。

[0077] 上述可知,所述还原方法通过获取压缩文件中的消息摘要值,从备份数据库中获取与所述消息摘要值对应的目标字符串,将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

[0078] 上述详细阐述了本发明实施例的方法,下面为了便于更好地实施本发明实施例的上述方案,相应地,下面还提供用于配合实施上述方案的终端。

[0079] 参阅图6,图6为本发明实施例提供的一种终端,所述终端包括以下功能单元:第一获取单元610,用于获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设

阈值；

[0080] 第一检测单元620,用于检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值；

[0081] 压缩单元630,用于通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,其中,所述第二消息摘要值的长度小于所述目标字符串的长度；

[0082] 存储单元640,将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中；

[0083] 替代单元650,用于将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

[0084] 可选地,还包括:第二检测单元660,用于检测所述目标字符串在所述目标文件中的重复次数;所述压缩单元具体用于,在所述目标字符串的重复次数大于数量阈值的情况下,则通过消息摘要算法对所述目标字符串进行压缩从而得到所述第二消息摘要值。

[0085] 可选地,还包括:第二获取单元670,用于从所述备份数据库中获取与所述目标字符串对应的所述第一消息摘要值;所述替代单元具体用于,将所述目标字符串用所述第一消息摘要值替代从而得到所述目标文件的压缩文件。

[0086] 可选地,所述消息摘要算法包括:MD5、sha-1、RIPEMD、Haval中的任意一个。

[0087] 可以理解的是,本发明实施例中的终端600中各功能单元的功能可根据上述方法实施例中的方法具体实现,其具体实现过程可以参照上述图2或图4方法实施例中的相关描述,此处不再赘述。

[0088] 参阅图7,图7是本发明实施例提供的一种终端,所述终端针对上述图6所示终端压缩处理后的压缩文件进行还原,所述终端700包括以下功能单元:

[0089] 第一获取单元710,用于获取压缩文件中的消息摘要值,其中,所述消息摘要值是通过消息摘要算法对目标字符串进行压缩得到的,所述消息摘要值的长度小于所述目标字符串的长度；

[0090] 第二获取单元720,用于从备份数据库中获取与所述消息摘要值对应的目标字符串,其中,所述目标字符串与所述消息摘要值的对应关系存储在所述备份数据库中；

[0091] 替代单元730,用于将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

[0092] 可以理解的是,本发明实施例中的终端700中各功能单元的功能可根据上述方法实施例中的方法具体实现,其具体实现过程可以参照上述图5方法实施例中的相关描述,此处不再赘述。

[0093] 请参见图8,图8为本发明实施例公开的一种终端的结构示意图。本实施例的终端包括:至少一个处理器801、通信接口802、用户接口803和存储器804,处理器801、通信接口802、用户接口803和存储器804可通过总线或者其它方式连接,本发明实施例以通过总线805连接为例。其中,

[0094] 处理器801可以是通用处理器,例如中央处理器(Central Processing Unit, CPU)。

[0095] 通信接口802可以为有线接口(例如以太网接口)或无线接口(例如蜂窝网络接口或使用无线局域网接口),用于与其他终端或网站进行通信。

[0096] 用户接口803具体可为触控面板,包括触摸屏和触控屏,用于检测触控面板上的操作指令,用户接口803也可以是物理按键或者鼠标。用户接口803还可以为显示屏,用于输出、显示图像或数据。

[0097] 存储器804可以包括易失性存储器(Volatile Memory),例如随机存取存储器(Random Access Memory, RAM);存储器也可以包括非易失性存储器(Non-Volatile Memory),例如只读存储器(Read-Only Memory, ROM)、快闪存储器(Flash Memory)、硬盘(Hard Disk Drive, HDD)或固态硬盘(Solid-State Drive, SSD);存储器804还可以包括上述种类的存储器的组合。存储器804用于存储一组程序代码,处理器801用于调用存储器804中存储的程序代码,执行如下操作:

[0098] 获取目标文件中的目标字符串,其中,所述目标字符串的大小大于预设阈值;

[0099] 检测备份数据库中是否存在与所述目标字符串对应的第一消息摘要值;

[0100] 如果没有,则通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值,并将所述目标字符串与所述第二消息摘要值的对应关系存储到所述备份数据库中,其中,所述第二消息摘要值的长度小于所述目标字符串的长度;

[0101] 将所述目标字符串用所述第二消息摘要值替代从而得到所述目标文件的压缩文件。

[0102] 可选地,在获取目标文件中的目标字符串之后,在通过消息摘要算法对所述目标字符串进行压缩从而得到第二消息摘要值之前,还包括:

[0103] 检测所述目标字符串在所述目标文件中的重复次数;

[0104] 通过消息摘要算法对所述目标字符串进行压缩从而得到所述第二消息摘要值具体为:在所述目标字符串的重复次数大于数量阈值的情况下,则通过消息摘要算法对所述目标字符串进行压缩从而得到所述第二消息摘要值。

[0105] 可选地,如果有所述第一消息摘要值,从所述备份数据库中获取与所述目标字符串对应的所述第一消息摘要值,并将所述目标字符串用所述对应的所述第一消息摘要值替代从而得到所述目标文件的压缩文件。

[0106] 可选地,所述消息摘要算法包括:MD5、sha-1、RIPEMD、Haval中的任意一个。

[0107] 请参阅图9,图9为本发明实施例提供的一种终端的结构示意图。本实施例的终端包括:至少一个处理器901、通信接口902、用户接口903和存储器904,处理器901、通信接口902、用户接口903和存储器904可通过总线或者其它方式连接,本发明实施例以通过总线905连接为例。其中,

[0108] 处理器901可以是通用处理器,例如中央处理器(Central Processing Unit, CPU)。

[0109] 通信接口902可以为有线接口(例如以太网接口)或无线接口(例如蜂窝网络接口或使用无线局域网接口),用于与其他终端或网站进行通信。

[0110] 用户接口903具体可为触控面板,包括触摸屏和触控屏,用于检测触控面板上的操作指令,用户接口903也可以是物理按键或者鼠标。用户接口903还可以为显示屏,用于输出、显示图像或数据。

[0111] 存储器904可以包括易失性存储器(Volatile Memory),例如随机存取存储器(Random Access Memory, RAM);存储器也可以包括非易失性存储器(Non-Volatile

Memory),例如只读存储器(Read-Only Memory,ROM)、快闪存储器(Flash Memory)、硬盘(Hard Disk Drive,HDD)或固态硬盘(Solid-State Drive,SSD);存储器904还可以包括上述种类的存储器的组合。存储器904用于存储一组程序代码,处理器901用于调用存储器904中存储的程序代码,执行如下操作:

[0112] 获取压缩文件中的消息摘要值,其中,所述消息摘要值是通过消息摘要算法对目标字符串进行压缩得到的,所述消息摘要值的长度小于所述目标字符串的长度;

[0113] 从备份数据库中获取与所述消息摘要值对应的目标字符串,其中,所述目标字符串与所述消息摘要值的对应关系存储在所述备份数据库中;

[0114] 将所述消息摘要值用所述对应的目标字符串替代,从而得到所述压缩文件的原始文件。

[0115] 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、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。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。

[0116] 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,上述描述的终端和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0117] 在本申请所提供的几个实施例中,应该理解到,所揭露的终端和方法,可以通过其它的方式实现。例如,以上所描述的终端实施例仅仅是示意性的,例如,所述单元的划分,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。另外,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口、装置或单元的间接耦合或通信连接,也可以是电的,机械的或其它的形式连接。

[0118] 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发明实施例方案的目的。

[0119] 本发明实施例方法中的步骤可以根据实际需要进行顺序调整、合并和删减。

[0120] 本发明实施例终端中的单元可以根据实际需要进行合并、划分和删减。

[0121] 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,也可以是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。

[0122] 所述集成的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,本发明的技术方案本质上或者说对现有技术做出贡献的部分,或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only

Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0123] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到各种等效的修改或替换,这些修改或替换都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以权利要求的保护范围为准。

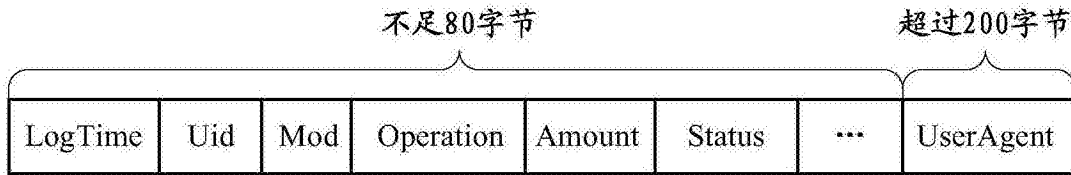


图1

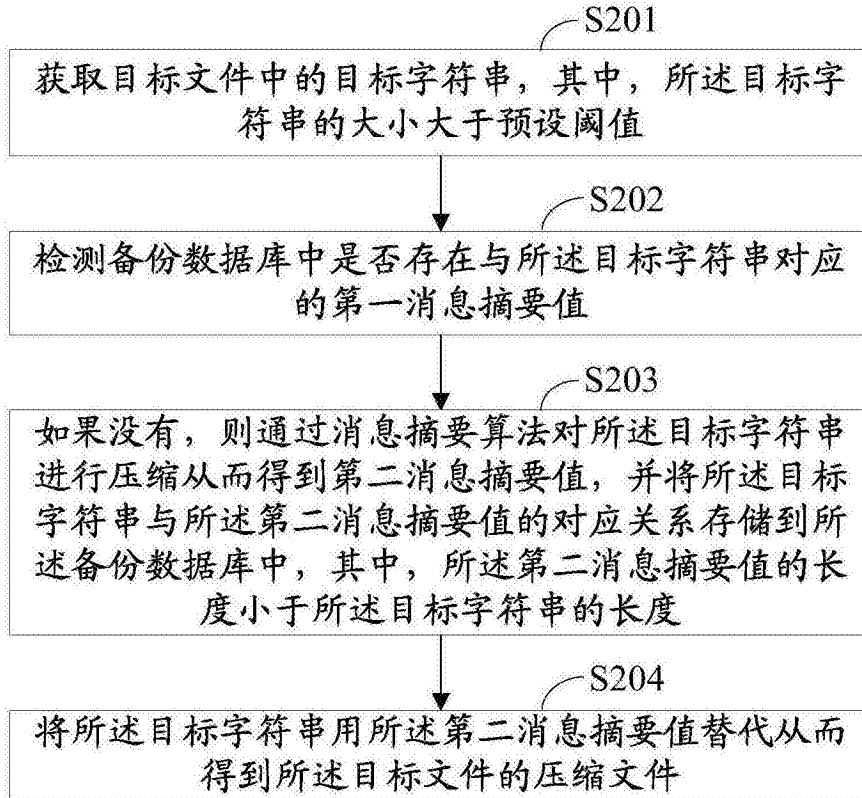


图2

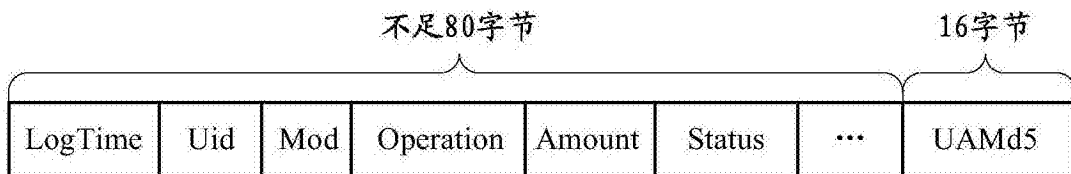


图3

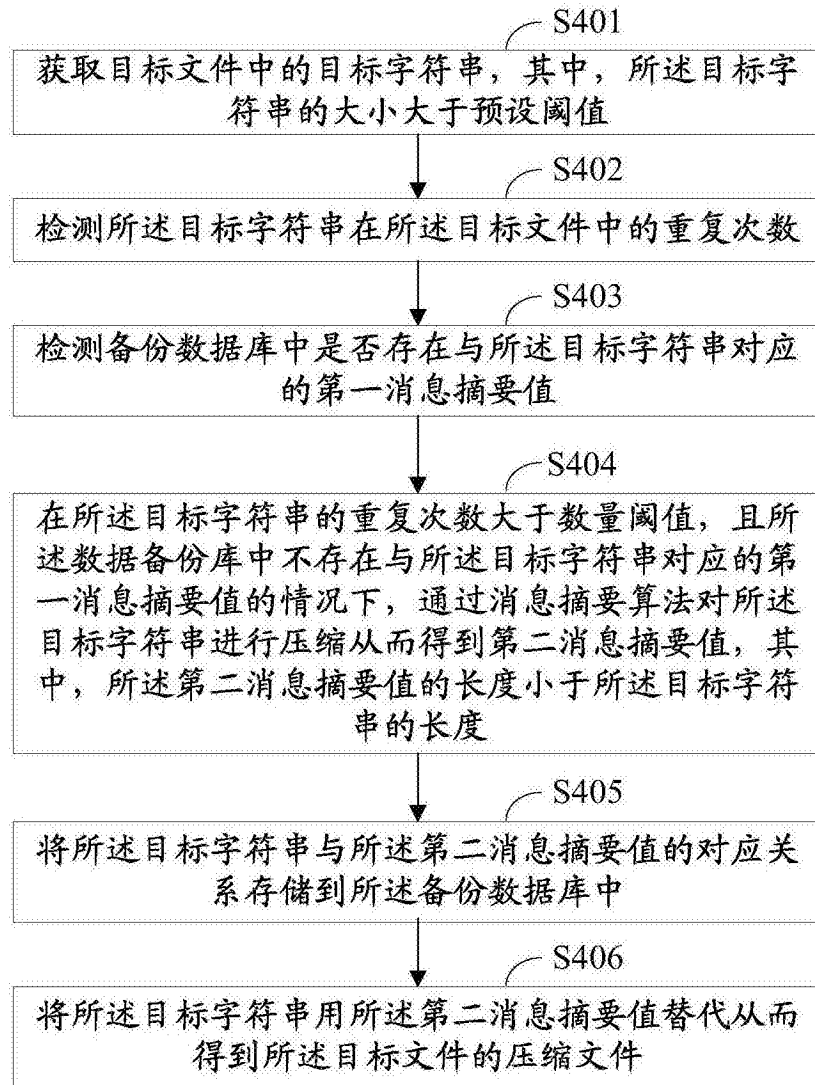


图4

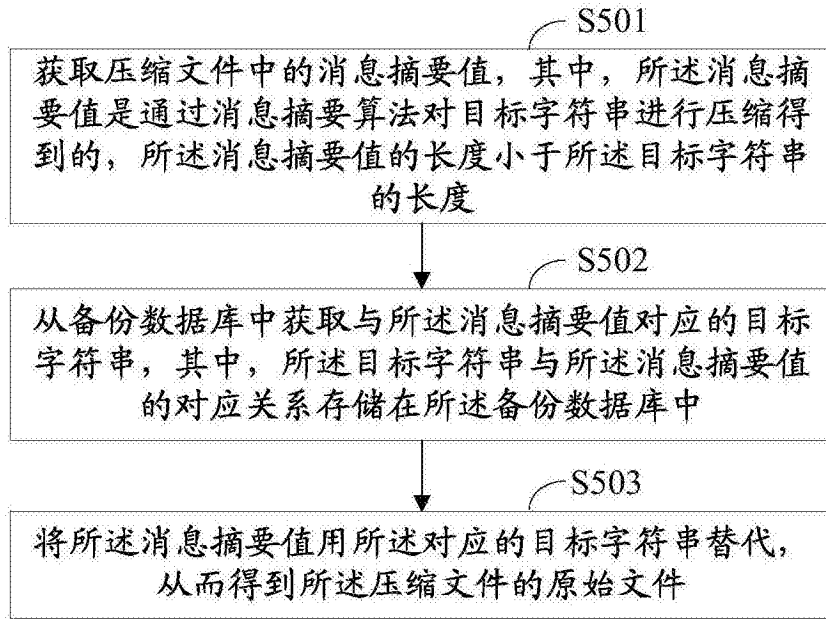


图5

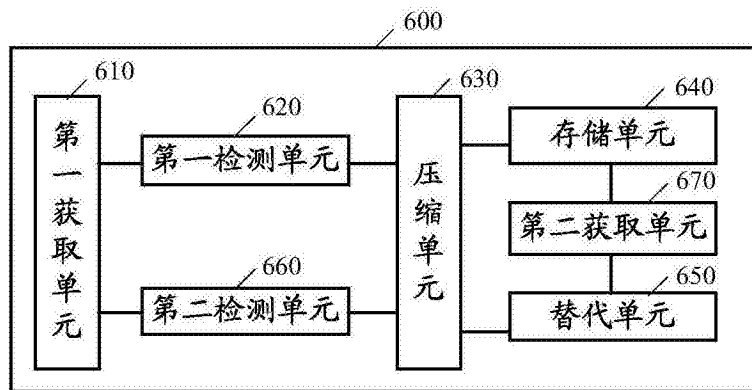


图6

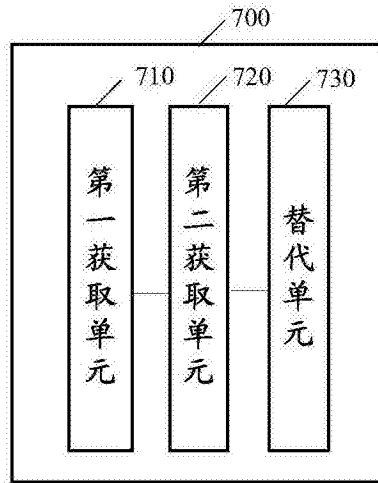


图7

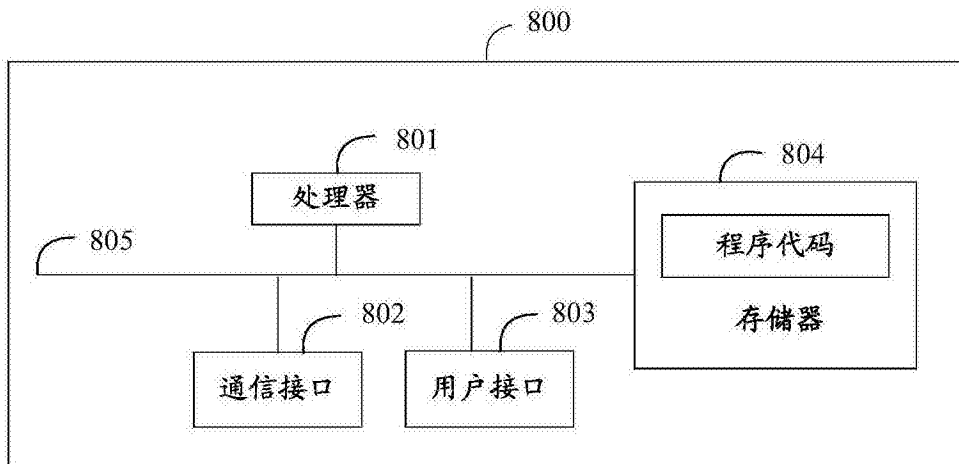


图8

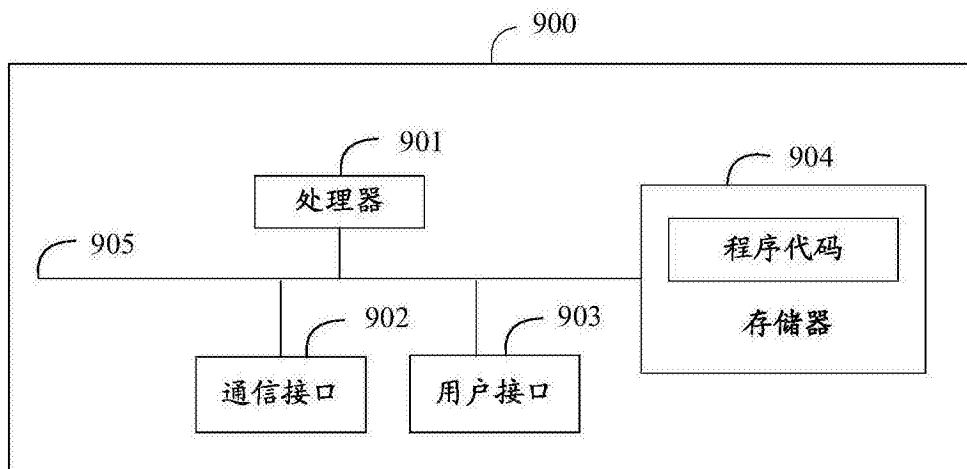


图9