

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 7 月 30 日 (30.07.2020)



(10) 国际公布号
WO 2020/151332 A1

(51) 国际专利分类号:
G06F 3/06 (2006.01)

(21) 国际申请号: PCT/CN2019/118317

(22) 国际申请日: 2019 年 11 月 14 日 (14.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910064084.8 2019 年 1 月 23 日 (23.01.2019) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

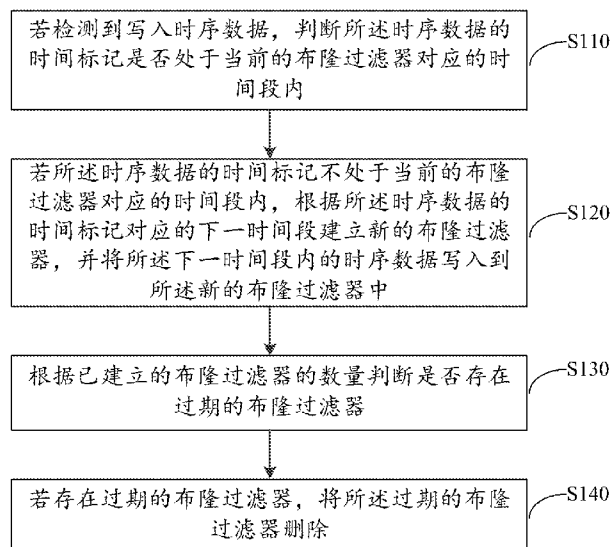
(72) 发明人: 李红伟(LI, Hongwei); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市精英专利事务所(SHENZHEN TALENT PATENT SERVICE); 中国广东省深圳市福田区深南中路 6009 号绿景广场 B 栋 20 层 B, Guangdong 518000 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: DATA DELETION METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 数据删除方法、装置、计算机设备及存储介质



S110 If write-in time sequence data is detected, determine whether a time mark of the time sequence data is within a time period corresponding to a current Bloom filter
S120 If the time mark of the time sequence data is not within the time period corresponding to the current Bloom filter, establish a new Bloom filter according to the next time period corresponding to the time mark of the time sequence data, and write the time sequence data within the next time period into the new Bloom filter
S130 Determine whether an expired Bloom filter exists according to the number of the established Bloom filters
S140 If yes, delete the expired Bloom filter

图 1

(57) Abstract: Disclosed in embodiments of the present application are a data deletion method and apparatus, a computer device, and a storage medium. The present application is applied to the field of data query in big data. The method comprises: if write-in time sequence data is detected, determining whether a time mark of the time sequence data is within a time period corresponding to a current Bloom filter; if the time mark of the time sequence data is not within the time period corresponding to the current Bloom filter, establishing a new Bloom filter according to the next time period corresponding to the time mark of the time sequence data, and writing the time sequence data within the next time period into the new Bloom filter; determining whether an expired Bloom filter exists according to the number of the established Bloom filters; and if yes, deleting the expired Bloom filter.

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 本申请实施例公开了一种数据删除方法、装置、计算机设备及存储介质。本申请应用于大数据中的数据查询领域。所述方法包括: 若检测到写入时序数据, 判断时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内; 若时序数据的时间标记不处于当前的布隆过滤器对应的时间段内, 根据时序数据的时间标记对应的下一时间段建立新的布隆过滤器, 并将下一时间段内的时序数据写入到新的布隆过滤器中; 根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器; 若存在过期的布隆过滤器, 将过期的布隆过滤器删除。

数据删除方法、装置、计算机设备及存储介质

本申请要求于2019年01月23日提交中国专利局、申请号为201910064084.8、申请名称为“数据删除方法、装置、计算机设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及数据处理技术领域，尤其涉及一种数据删除方法、装置、计算机设备及存储介质。

背景技术

随着信息与科技的发展，人们的日常生活中每时每刻不停地产生着巨量的数据，数据的价值越来越被重视，数据的处理也愈发重要。传统的数据查询方法，通常是将数据存储到磁盘中，然后通过链表、树以及哈希表等进行查询，然而随着时间的推移数据量的不断增长，需要的存储空间越来越大，查询的时间越来越长，导致空间消耗大，查询效率低。现有的布隆过滤器（Bloom Filter）是一种快速判断某个元素是否存在于集合中的算法，具有空间效率高和查询效率高的优点，但是在时序数据的处理过程中，时序数据的去重通常需要将过期数据删除，而布隆过滤器不支持删除操作，从而导致布隆过滤器的使用受限。

发明内容

本申请实施例提供了一种数据删除方法、装置、计算机设备及存储介质，旨在解决布隆过滤器不支持删除操作而导致使用受限的技术问题。

第一方面，本申请实施例提供了一种数据删除方法，其包括：若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中；根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；若存在过期的布隆过

滤波器，将所述过期的布隆过滤器删除。

第二方面，本申请实施例还提供了一种数据删除装置，其包括：第一判断单元，用于若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；建立单元，用于若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中；第二判断单元，用于根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；删除单元，用于若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

第三方面，本申请实施例还提供了一种计算机设备，其包括存储器以及与所述存储器相连的处理器；所述存储器用于存储计算机程序；所述处理器用于运行所述存储器中存储的计算机程序，以执行如下步骤：若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中；根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

第四方面，本申请实施例还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使所述处理器执行以下步骤：若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中；根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

附图说明

为了更清楚地说明本申请实施例技术方案，下面将对实施例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申请的一些实

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。

图1为本申请实施例提供的删除数据方法的流程图；

图2为本申请实施例提供的删除数据的子流程图；

图3为本申请另一实施例提供的删除数据方法的流程图；

图4为本申请实施例提供的删除数据的子流程图；

图5为本申请另一实施例提供的删除数据方法的流程图；

图6为本申请实施例提供的删除数据装置的示意性框图；

图7为本申请实施例提供的删除数据装置的具体单元的示意性框图；

图8为本申请另一实施例提供的删除数据装置的示意性框图；以及

图9为本申请另一实施例提供的删除数据装置的示意性框图；以及

图10为本申请实施例提供的计算机设备的示意性框图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请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申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

应当理解，当在本说明书和所附权利要求书中使用时，术语“包括”和“包含”指示所描述特征、整体、步骤、操作、元素和/或组件的存在，但并不排除一个或多个其它特征、整体、步骤、操作、元素、组件和/或其集合的存在或添加。

还应当理解，在此本申请说明书中所使用的术语仅仅是出于描述特定实施例的目的而并不意在限制本申请。如在本申请说明书和所附权利要求书中所使用的那样，除非上下文清楚地指明其它情况，否则单数形式的“一”、“一个”及“该”意在包括复数形式。

还应当进一步理解，在本申请说明书和所附权利要求书中使用的术语“和/或”是指相关联列出的项中的一个或多个的任何组合以及所有可能组合，并且包括这些组合。

请参阅图1，图1为本申请实施例提供的删除数据的示意性流程图。该删除数据方法应用于服务器中，基于布隆过滤器实现，布隆过滤器是一种快速

判断某个元素是否存在于集合中的算法，具有空间效率高和查询效率高的优点。然而由于布隆过滤器本身的特性不支持删除操作，通过布隆过滤器处理时序数据会导致过期数据无法删除，降低了布隆过滤器空间效率和查询效率。因此，通过本方案所提出的一种数据删除方法以实现时序数据在布隆过滤器中的删除操作。

图1是本申请实施例提供的数据删除方法的流程示意图。如图所示，该方法包括以下步骤S110-S140。

S110、若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内。

在一实施例中，时序数据指的是按照时间顺序记录的数据，时间段指的是根据用户需求设置的时间区间，例如，一个月，一天，一小时以及一分钟等。由于时序数据的特性，其数据本身即存在有时间标记，时间标记指的是数据生成时标记的时刻，在时序数据写入时将时间标记与当前的布隆过滤器对应的时间段进行对比，判断当前时间是否处于当前的布隆过滤器对应的时间段内。具体地，将时序数据的时间标记与当前的布隆过滤器对应的时间段的结束时间进行对比，若时间标记早于结束时间，则判定时间标记处于时间段内；若时间标记晚于结束时间，则判定时间标记不处于时间段内。例如，若时序数据的时间标记为8月29日9:00，当前的布隆过滤器对应的时间段为8月份，那么该时序数据的时间标记处于当前的布隆过滤器对应的时间段内；若时序数据的时间标记为9月1日1:00，当前的布隆过滤器对应的时间段为8月份，那么该时序数据的时间标记就不处于当前的布隆过滤器对应的时间段内。

S120、若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中。

在一实施例中，布隆过滤器根据时间段建立，每个时间段建立一个布隆过滤器，时序数据根据时间段写入到对应的布隆过滤器中。当时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，说明时序数据的时间标记已来到了下一个时间段，那么则根据时间标记对应的下一个时间段建立新的布隆过滤器，下一个时间段内的时序数据均写入到新的布隆过滤器中。例如，当前时间段为8月份，而时序数据的时间标记为9月1日，那么则根据该时间标记对应

的9月份建立一个新的布隆过滤器,9月份的时序数据均写入到新的布隆过滤器中。

在一实施例中,如图2所示,所述步骤S120可包括步骤S121-S123。

S121、新建位数组并将所述位数组中的所有位置均设置为零。

在一实施例中,位数组指的是一种通过二进制存储数据的数据存储方式,在位数组中,若数据存在,则其相应的二进制位就为1;若数据不存在,则其相应的二进制位就为0。在本实施例中,新建的位数组初始状态时,将位数组中的所有位置均设为零,即表明当前位数组是新建的,还未存储有数据。

S122、将下一个时间段内的时序数据通过预设数量的哈希函数映射到所述位数组的位置上。

在一实施例中,哈希函数是一种可将任意内容的数据映射到一个固定长度的字符串中的函数,相同内容的数据映射得到相同的字符串,因此,通过哈希函数构成的内容-索引的生成关系可以快速查询数据。基于单个哈希函数对数据的查询,当映射后的值限制在一定范围内时哈希冲突的概率会增大,而存储空间的大小又不能无限扩张,从而造成查询效率低下。为了提高空间利用率,采用多个不同的哈希函数来进行映射,可以较少的空间来表示较大集合的存在关系,大大提高了空间利用率。具体地,预设数量指的是哈希函数的数量,一个哈希函数可将时序数据映射到位数组的一个位置上,那么多个不同的哈希函数则可将相同的时序数据映射到位数组的多个位置上,哈希函数的数量对应映射到位数组上位置的数量。其中,哈希函数最优的预设数量由哈希函数的错误率计算得出,在此不再赘述。例如,预设时间段为一个月,预设数量为3,那么将一个月内的时序数据通过3个不同的哈希函数映射到新建的位数组的三个位置上。

S123、将所述位数组中被映射到的位置均设置为一。

在一实施例中,在得到时序数据映射到位数组的位置后,将位数组上的这些位置全部由0变为1,从而将时序数据写入到布隆过滤器中,这些位数组上的位置即为时序数据的索引,根据索引即可实现快速查询所写入的时序数据。

在一实施例中,如图3所示,所述步骤S120之后,还包括步骤S124-S126。

S124、若检测到查询时序数据,将所述时序数据通过预设数量的哈希函数映射到所述位数组的位置上。

S125、判断所述位数组中被映射到的位置是否均为一。

S126、若所述位数组中被映射到的位置均为一，判定存在所述时序数据。

在一实施例中，当检测到查询时序数据，从布隆过滤器中查询，与写入时序数据到布隆过滤器相类似，首先需要通过预设数量的哈希函数对时序数据进行映射得到位数组上预设数量的位置，需要注意的是，该步骤的预设数量需与步骤 S122 步骤中的预设数量相同，然后将这些位置与位数组上对应的位置进行比较，判断位数组中对应的位置是否全部都是一，若全部都是一，判定该时序数据存在；若非全部都是一，判定该时序数据不存在。例如，查询 A 客户的 8 月份的登录记录，那么则通过预设数量的哈希函数将 8 月份的登录记录映射到位数组的位置上，一个哈希函数对应映射到位数组上的一个位置，然后将这些位置与 8 月份的布隆过滤器的位数组进行对比，若位数组上所映射的这些位置全部都是 1，说明 A 客户最近一个月曾经登录过；若位数组上这些位置并非全部都是 1，说明 A 客户最近一个月未曾登录过。

S130、根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器。

在一实施例中，如图 4 所示，所述步骤 S130 可包括步骤 S131-S132。

S131、判断已建立的布隆过滤器的数量是否超过预设数量。

在一实施例中，预设数量是根据用户的需求设定的，例如，用户需要的数据实效性为 3 个月，那么 3 个月内的数据则为有用的数据，3 个月前的数据则为无用数据没有存在的价值。具体地，根据已建立的布隆过滤器的数量与预设数量进行比较来判断。

S132、若已建立的布隆过滤器的数量超过预设数量，判定存在过期的布隆过滤器。

在一实施例中，已建立的布隆过滤器的数量大于预设数量，说明当前存储了没有存在价值的无用数据，判定当前的布隆过滤器中存在过期的布隆过滤器。例如，预设数量为 3，已建立的布隆过滤器的数量有 4 个，那么已建立的布隆过滤器的数量大于预设数量，判定当前的布隆过滤器中存在过期的布隆过滤器。

S140、若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

在一实施例中，在得到已建立的布隆过滤器中存在过期的布隆过滤器后，需要将过期的布隆过滤器删除，过期的布隆过滤器具体是最早建立的布隆过滤

器。例如，预设数量为3，已经建立了8月份、9月份以及10月份的布隆过滤器，当11月份建立了新的布隆过滤器，导致已建立的布隆过滤器的数量4大于预设数量3，那么最早建立的布隆过滤器即8月份的布隆过滤器则为过期的布隆过滤器。具体地，将过期的布隆过滤器即8月份的布隆过滤器进行删除，即将位数组所有位置全部置0，从而将无用的时序数据全部删去。

在一实施例中，如图5所示，所述步骤S110之后，还包括步骤S111。

S111、若所述时序数据的时间标记处于当前的布隆过滤器对应的时间段内，将所述时序数据写入到当前的布隆过滤器中。

在一实施例中，当时序数据的时间标记处于当前的布隆过滤器对应的时间段内，那么则直接将时序数据写入到当前的布隆过滤器中。具体地，首先通过预设数量的哈希函数将时序数据映射到当前的位数组的位置上，每一个哈希函数映射一个位数组的位置，然后将位数组位置中被映射到的位置均置为1，从而将时序数据写入到当前的布隆过滤器中。相对于现有的布隆过滤器应用，将所有的时序数据均写入到一个布隆过滤器中导致不能删除时序数据，空间效率和查询效率低。本实施例通过根据时间段建立布隆过滤器写入时序数据，当时序数据属于当前的布隆过滤器对应的时间段，则写入当前的布隆过滤器；当时序数据不属于当前的布隆过滤器对应的时间段，则根据时序数据的时间标记对应下一个时间段建立新的布隆过滤器，写入到新的布隆过滤器中。当时序数据变为无用数据没有存在的价值时，将整个过期的布隆过滤器删除，从而实现了时序数据在布隆过滤器中的删除操作，且提高了空间效率和查询效率。

本申请实施例展示了一种数据删除方法，通过若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中；根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；若存在过期的布隆过滤器，将所述过期的布隆过滤器删除，可实现布隆过滤器的删除操作，扩展布隆过滤器应用到时序数据，提高空间效率和查询效率。

图6是本申请实施例提供的一种数据删除装置200的示意性框图。如图6所示，对应于以上数据删除方法，本申请还提供一种数据删除装置200。该数据

删除装置 200 包括用于执行上述数据删除方法的单元，该装置可以被配置于服务器中。具体地，请参阅图 6，该数据删除装置 200 包括第一判断单元 210、建立单元 220、第二判断单元 230 以及删除单元 240。

第一判断单元 210，用于若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；

建立单元 220，用于若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中。

在一实施例中，如图 7 所示，所述建立单元 220 包括：新建单元 221、映射单元 222 以及设置单元 223。

新建单元 221，用于新建位数组并将所述位数组中的所有位置均设置为零；

映射单元 222，用于将下一个时间段内的时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

设置单元 223，用于将所述位数组中被映射到的位置均设置为一。

在一实施例中，如图 8 所示，所述数据删除装置 200 还包括：

映射子单元 224，用于若检测到查询时序数据，将所述时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

判断子单元 225，用于判断所述位数组中被映射到的位置是否均为一；

判定子单元 226，用于若所述位数组中被映射到的位置均为一，判定存在所述时序数据。

第二判断单元 230，用于根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器。

在一实施例中，如图 7 所示，所述第二判断单元 230：第二判断子单元 231 以及判定单元 232。

第二判断子单元 231，用于判断已建立的布隆过滤器的数量是否超过预设数量；

判定单元 232，用于若已建立的布隆过滤器的数量超过预设数量，判定存在过期的布隆过滤器。

删除单元 240，用于若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

在一实施例中，如图 9 所示，所述数据删除装置 200 还包括：写入单元 211。

写入单元 211，用于若所述时序数据的时间标记处于当前的布隆过滤器对应的时间段内，将所述时序数据写入到当前的布隆过滤器中。

需要说明的是，所属领域的技术人员可以清楚地了解到，上述数据删除装置 200 和各单元的具体实现过程，可以参考前述方法实施例中的相应描述，为了描述的方便和简洁，在此不再赘述。

上述数据删除装置可以实现为一种计算机程序的形式，该计算机程序可以在如图 10 所示的计算机设备上运行。

请参阅图 10，图 10 是本申请实施例提供的一种计算机设备的示意性框图。该计算机设备 500 可以是服务器，服务器可以是独立的服务器，也可以是多个服务器组成的服务器集群。

参阅图 10，该计算机设备 500 包括通过系统总线 501 连接的处理器 502、存储器和网络接口 505，其中，存储器可以包括非易失性存储介质 503 和内存储器 504。

该非易失性存储介质 503 可存储操作系统 5031 和计算机程序 5032。该计算机程序 5032 包括程序指令，该程序指令被执行时，可使得处理器 502 执行一种数据删除方法。

该处理器 502 用于提供计算和控制能力，以支撑整个计算机设备 500 的运行。

该内存储器 504 为非易失性存储介质 503 中的计算机程序 5032 的运行提供环境，该计算机程序 5032 被处理器 502 执行时，可使得处理器 502 执行一种数据删除方法。

该网络接口 505 用于与其它设备进行网络通信。本领域技术人员可以理解，图 10 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备 500 的限定，具体的计算机设备 500 可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

其中，所述处理器 502 用于运行存储在存储器中的计算机程序 5032，以实现本申请实施例的数据删除方法

应当理解，在本申请实施例中，处理器 502 可以是中央处理单元 (Central

Processing Unit, CPU), 该处理器 502 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。其中, 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。

本领域普通技术人员可以理解的是实现上述实施例的方法中的全部或部分流程, 是可以通过计算机程序来指令相关的硬件来完成。该计算机程序包括程序指令, 计算机程序可存储于一存储介质中, 该存储介质为计算机可读存储介质。该程序指令被该计算机系统中的至少一个处理器执行, 以实现上述方法的实施例的流程步骤。

因此, 本申请还提供一种存储介质。该存储介质可以为计算机可读存储介质。该存储介质存储有计算机程序, 该计算机程序被处理器执行时使处理器执行以上各实施例中所描述的数据删除方法的步骤。

所述存储介质可以是 U 盘、移动硬盘、只读存储器 (Read-Only Memory, ROM)、磁碟或者光盘等各种可以存储程序代码的计算机可读存储介质。

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申请的范围。

在本申请所提供的几个实施例中, 应该理解到, 所揭露的装置和方法, 可以通过其它的方式实现。例如, 以上所描述的装置实施例仅仅是示意性的。例如, 各个单元的划分, 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。

本申请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。本申请实施例装置中的单元可以根据实际需要进行合并、划分和删减。另外, 在本申请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。

该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，可以存储在一个存储介质中。基于这样的理解，本申请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，终端，或者网络设备等等）执行本申请各个实施例所述方法的全部或部分步骤。

以上所述，仅为本申请的具体实施方式，但本申请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申请揭露的技术范围内，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申请的保护范围之内。因此，本申请的保护范围应以权利要求的保护范围为准。

权 利 要 求 书

1. 一种数据删除方法，包括：

若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；

若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的时序数据写入到所述新的布隆过滤器中；

根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；

若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

2. 根据权利要求 1 所述的数据删除方法，其中，所述根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中，包括：

新建位数组并将所述位数组中的所有位置均设置为零；

将下一个时间段内的时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

将所述位数组中被映射到的位置均设置为一。

3. 根据权利要求 2 所述的数据删除方法，其中，所述根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中之后，还包括：

若检测到查询时序数据，将所述时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

判断所述位数组中被映射到的位置是否均为一；

若所述位数组中被映射到的位置均为一，判定存在所述时序数据。

4. 根据权利要求 1 所述的数据删除方法，其中，所述根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器，包括：

判断已建立的布隆过滤器的数量是否超过预设数量；

若已建立的布隆过滤器的数量超过预设数量，判定存在过期的布隆过滤器。

5. 根据权利要求 1 所述的数据删除方法，其中，所述判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内之后，还包括：

若所述时序数据的时间标记处于当前的布隆过滤器对应的时间段内，将所述时序数据写入到当前的布隆过滤器中。

6. 根据权利要求3所述的数据删除方法，其中，所述预设数量为所述哈希函数的数量，所述哈希函数最优的预设数量由哈希函数的错误率计算得出。

7. 根据权利要求6所述的数据删除方法，其中，所述预设数量为3。

8. 一种数据删除装置，包括：

第一判断单元，用于若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；

建立单元，用于若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中；

第二判断单元，用于根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；

删除单元，用于若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

9. 根据权利要求8所述的数据删除装置，其中，所述建立单元包括：

新建单元，用于新建位数组并将所述位数组中的所有位置均设置为零；

映射单元，用于将下一个时间段内的时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

设置单元，用于将所述位数组中被映射到的位置均设置为一。

10. 根据权利要求8所述的数据删除装置，其中，所述第二判断单元，包括：

第二判断子单元，用于判断已建立的布隆过滤器的数量是否超过预设数量；

判定单元，用于若已建立的布隆过滤器的数量超过预设数量，判定存在过期的布隆过滤器。

11. 一种计算机设备，包括存储器以及与所述存储器相连的处理器；所述存储器用于存储计算机程序；所述处理器用于运行所述存储器中存储的计算机程序，以执行如下步骤：

若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；

若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述

下一时间段内的时序数据写入到所述新的布隆过滤器中；

根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；

若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

12.根据权利要求 11 所述的计算机设备，其中，所述根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中，包括：

新建位数组并将所述位数组中的所有位置均设置为零；

将下一个时间段内的时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

将所述位数组中被映射到的位置均设置为一。

13.根据权利要求 12 所述的计算机设备，其中，所述根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中之后，还包括：

若检测到查询时序数据，将所述时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

判断所述位数组中被映射到的位置是否均为一；

若所述位数组中被映射到的位置均为一，判定存在所述时序数据。

14.根据权利要求 11 所述的计算机设备，其中，所述根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器，包括：

判断已建立的布隆过滤器的数量是否超过预设数量；

若已建立的布隆过滤器的数量超过预设数量，判定存在过期的布隆过滤器。

15.根据权利要求 11 所述的计算机设备，其中，所述判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内之后，还包括：

若所述时序数据的时间标记处于当前的布隆过滤器对应的时间段内，将所述时序数据写入到当前的布隆过滤器中。

16.根据权利要求 13 所述的计算机设备，其中，所述预设数量为所述哈希函数的数量，所述哈希函数最优的预设数量由哈希函数的错误率计算得出。

17.根据权利要求 16 所述的计算机设备，其中，所述预设数量为 3。

18. 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使所述处理器执行以下步骤：

若检测到写入时序数据，判断所述时序数据的时间标记是否处于当前的布隆过滤器对应的时间段内；

若所述时序数据的时间标记不处于当前的布隆过滤器对应的时间段内，根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的时序数据写入到所述新的布隆过滤器中；

根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器；

若存在过期的布隆过滤器，将所述过期的布隆过滤器删除。

19.根据权利要求 18 所述的计算机可读存储介质，其中，所述根据所述时序数据的时间标记对应的下一时间段建立新的布隆过滤器，并将所述下一时间段内的所述时序数据写入到所述新的布隆过滤器中的步骤，包括：

新建位数组并将所述位数组中的所有位置均设置为零；

将下一个时间段内的时序数据通过预设数量的哈希函数映射到所述位数组的位置上；

将所述位数组中被映射到的位置均设置为一。

20.根据权利要求 18 所述的计算机可读存储介质，其中，所述根据已建立的布隆过滤器的数量判断是否存在过期的布隆过滤器的步骤，包括：

判断已建立的布隆过滤器的数量是否超过预设数量；

若已建立的布隆过滤器的数量超过预设数量，判定存在过期的布隆过滤器。

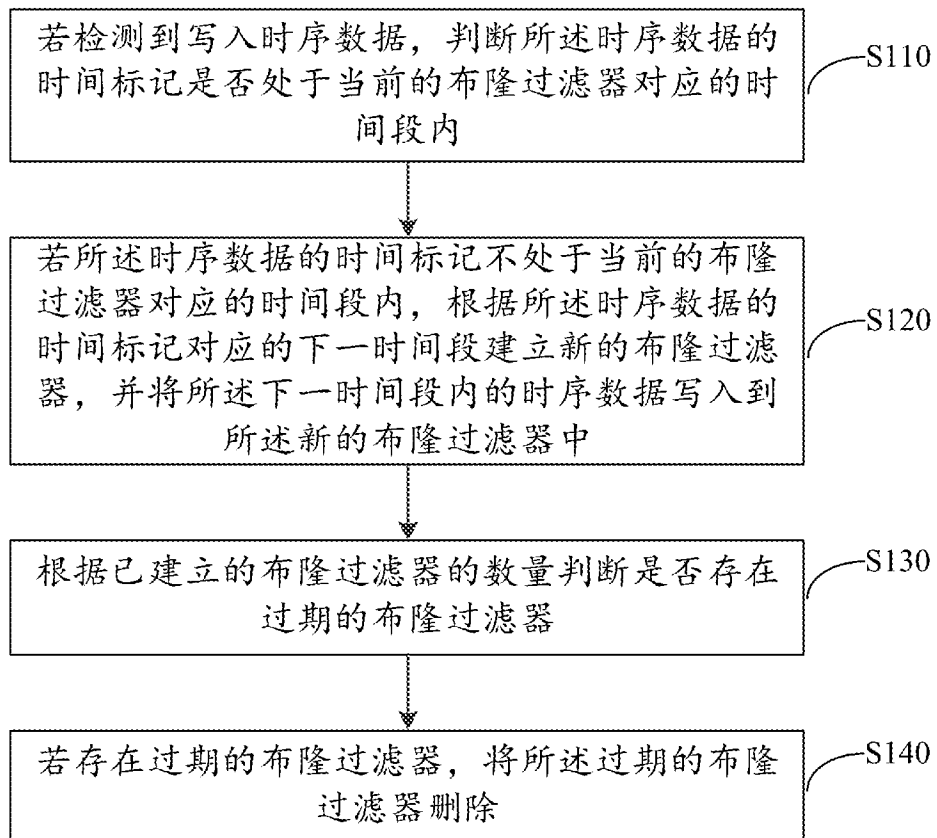


图 1

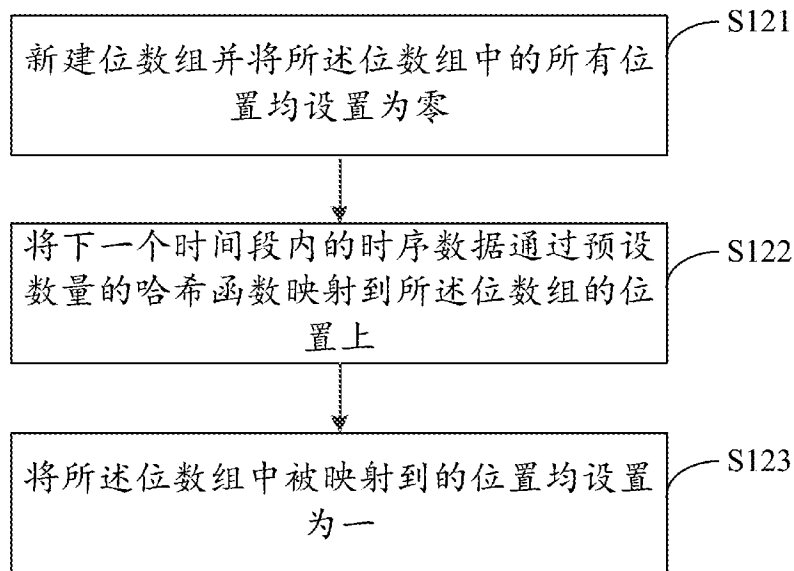


图 2

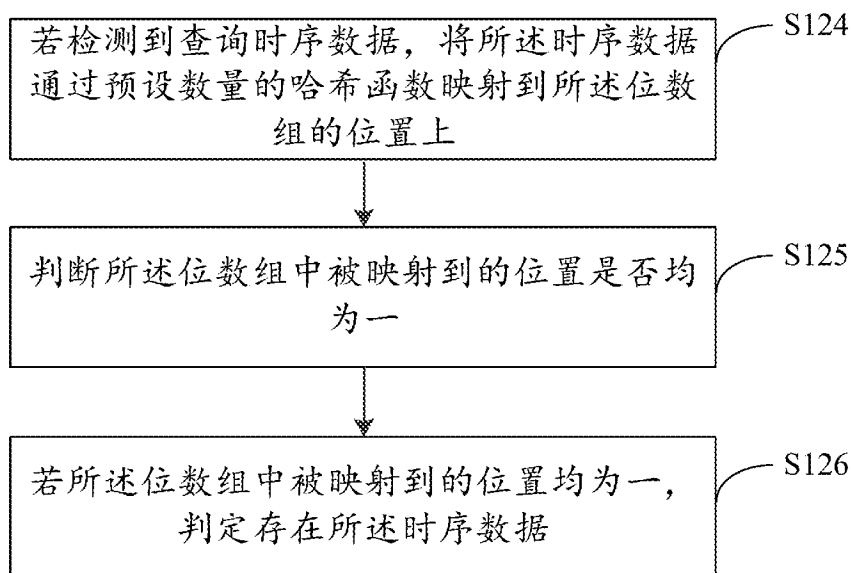


图 3

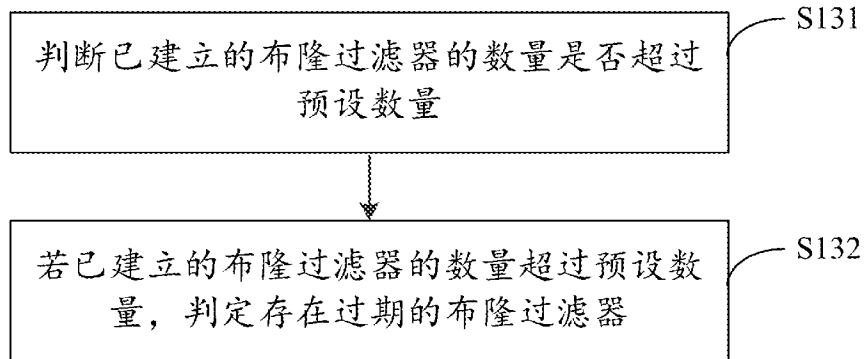


图 4

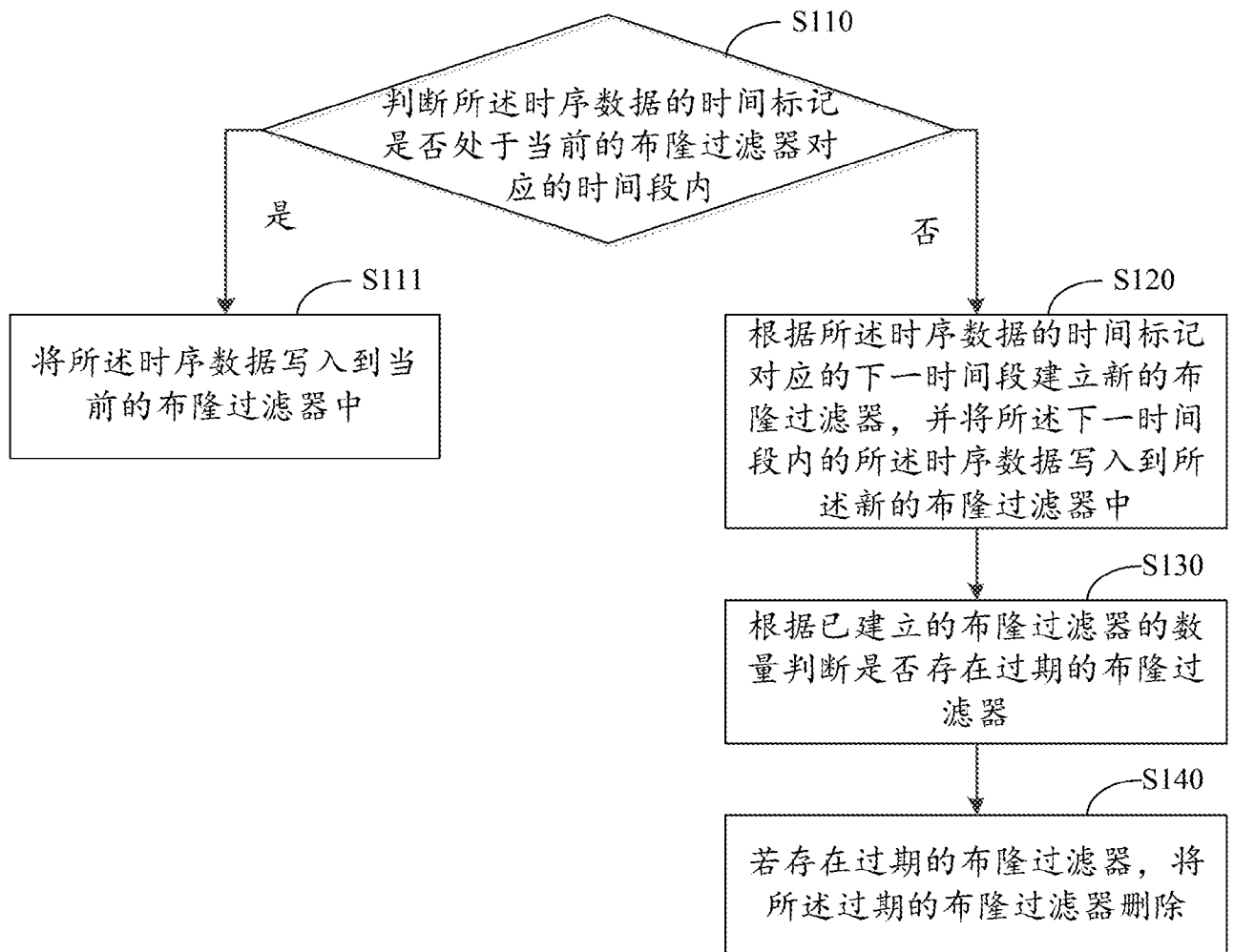


图 5



图 6

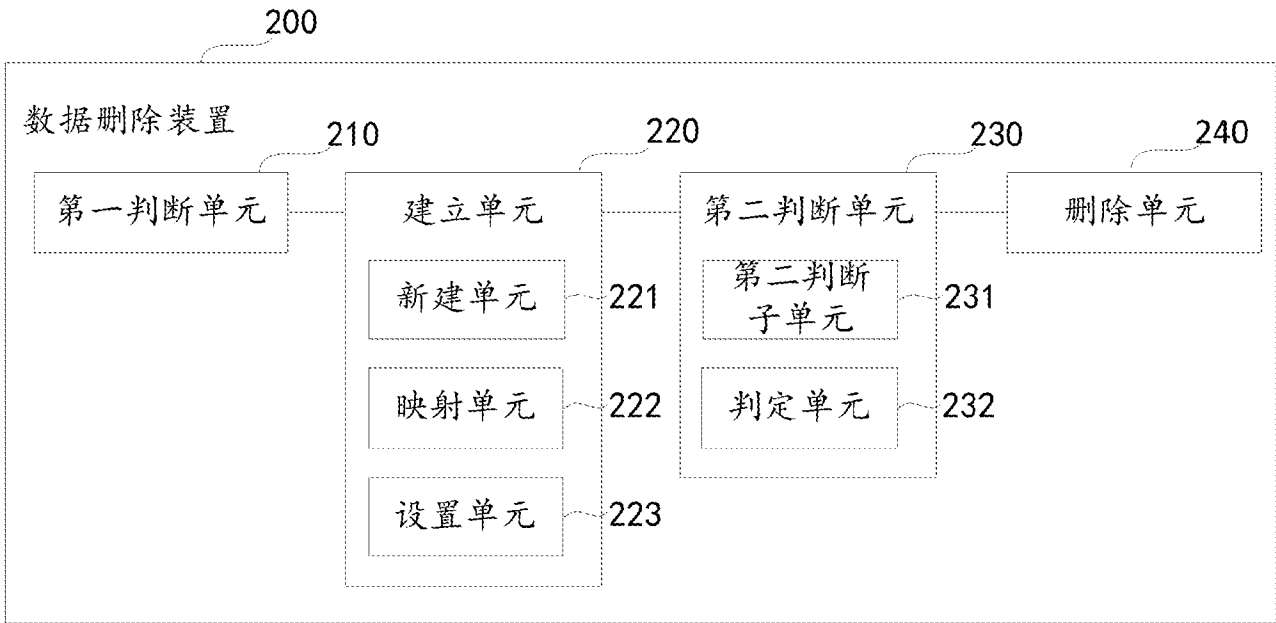


图 7

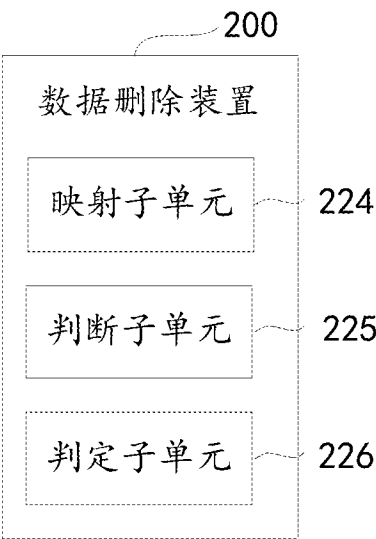


图 8

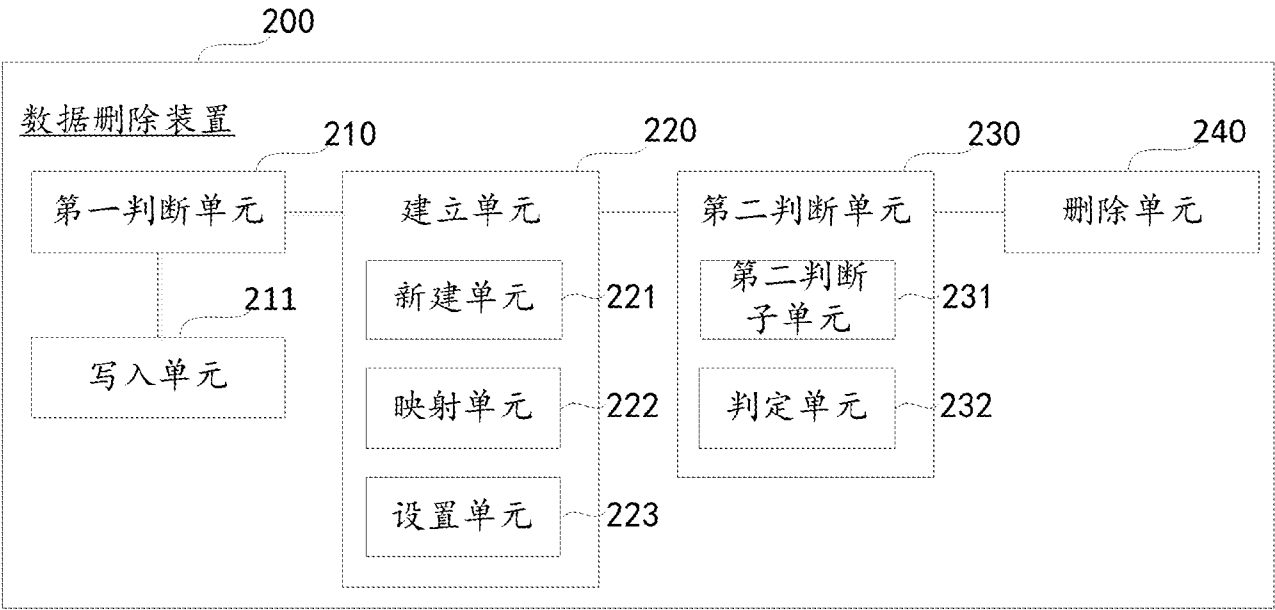


图 9

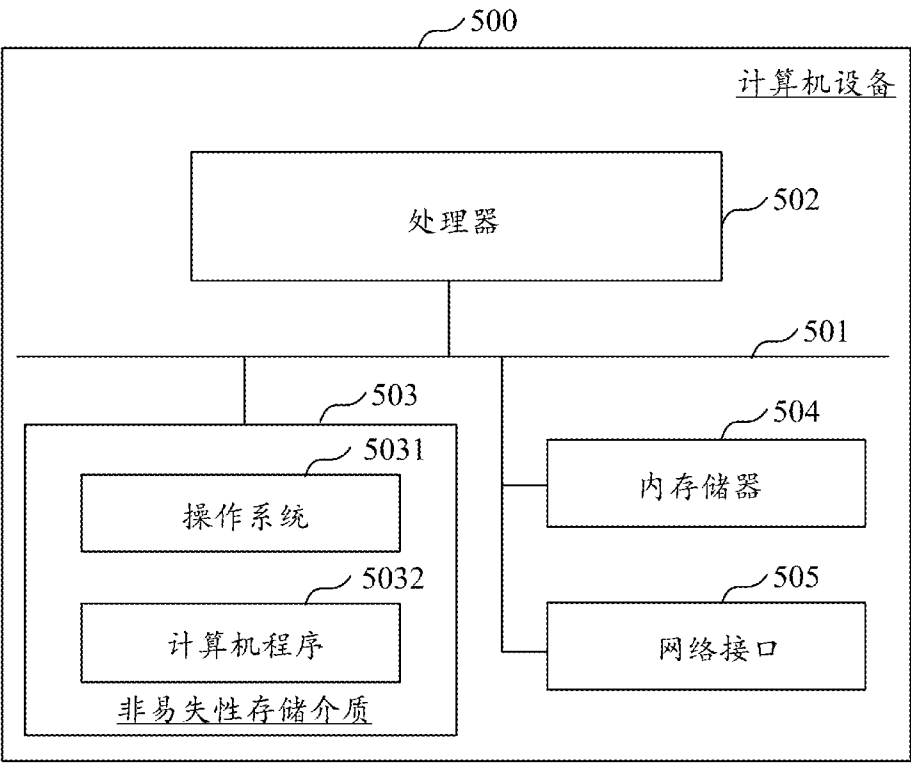


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/118317

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/06(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT; CNABS; SIPOABS; DWPI; CNKI: 布隆过滤器, 时序, 时间, 去重, 过期, 哈希, 删除, bloom, filter, timing, sequence, time, duplicate, removal, overdue, hash, delete

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109828721 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 31 May 2019 (2019-05-31) claims 1-10, and entire description	1-20
X	CN 109255001 A (ALIBABA GROUP HOLDING LIMITED) 22 January 2019 (2019-01-22) description, paragraphs [0070]-[0113]	1-20
A	CN 106649346 A (BEIJING GRIDSUM TECHNOLOGY CO., LTD.) 10 May 2017 (2017-05-10) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

12 February 2020

Date of mailing of the international search report

21 February 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/118317

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109828721	A	31 May 2019	None	
CN	109255001	A	22 January 2019	None	
CN	106649346	A	10 May 2017	None	

国际检索报告

国际申请号

PCT/CN2019/118317

A. 主题的分类

G06F 3/06(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNTXT;CNABS;SIPOABS;DWPI;CNKI:布隆过滤器, 时序, 时间, 去重, 过期, 哈希, 删除, bloom, filter, timing, sequence, time, duplicate, removal, overdue, hash, delete

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 109828721 A (平安科技深圳有限公司) 2019年 5月 31日 (2019 - 05 - 31) 权利要求1-10及说明书全文	1-20
X	CN 109255001 A (阿里巴巴集团控股有限公司) 2019年 1月 22日 (2019 - 01 - 22) 说明书第[0070]-[0113]段	1-20
A	CN 106649346 A (北京国双科技有限公司) 2017年 5月 10日 (2017 - 05 - 10) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2020年 2月 12日

国际检索报告邮寄日期

2020年 2月 21日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

刘芳

电话号码 86-(10)-62089108

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/118317

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109828721	A	2019年 5月 31日	无	
CN	109255001	A	2019年 1月 22日	无	
CN	106649346	A	2017年 5月 10日	无	