

(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
2023 年 2 月 2 日 (02.02.2023)

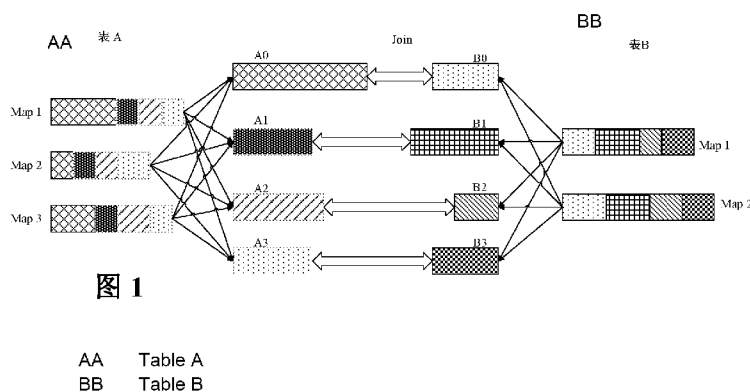


(10) 国际公布号
WO 2023/005264 A1

- (51) 国际专利分类号:
G06F 9/50 (2006.01)
- (21) 国际申请号: PCT/CN2022/087297
- (22) 国际申请日: 2022 年 4 月 18 日 (18.04.2022)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110871021.0 2021年7月30日 (30.07.2021) CN
- (71) 申请人: 北京沃东天骏信息技术有限公司 (**BEIJING WODONG TIANJUN INFORMATION TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市北京经济技术开发区科创十一街18号院2号楼4层A402室, Beijing 100176 (CN)。北京京东世纪贸易有限公司 (**BEIJING JINGDONG CENTURY TRADING CO., LTD.**) [CN/CN]; 中国北
- 京市北京经济技术开发区科创十一街18号C座2层201室, Beijing 100176 (CN)。
- (72) 发明人: 郑瑞峰 (**ZHENG, Ruifeng**); 中国北京市北京经济技术开发区科创十一街18号院2号楼4层A402室, Beijing 100176 (CN)。
- (74) 代理人: 中原信达知识产权代理有限责任公司 (**CHINA SINDA INTELLECTUAL PROPERTY LIMITED**); 中国北京市西城区金融街19号富凯大厦B座11层, Beijing 100033 (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, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW,

(54) Title: DATA PROCESSING METHOD AND APPARATUS

(54) 发明名称: 一种数据处理方法和装置



(57) Abstract: The present disclosure relates to the technical field of big data and discloses a data processing method and an apparatus. A specific embodiment of the present method comprises: utilizing an adaptive means and performing sampling on data to be processed in a data table, and obtaining a sample data distribution of the data table; dividing the data to be processed into a plurality of slices, and separately calculating amounts of output data corresponding to join operations of the plurality of slices according to the sample data distribution of the data table; selecting a target slice from the plurality of slices according to the amounts of output data corresponding to the join operations of the plurality of slices, and performing data slicing on the target slice. The present embodiment is able to solve the technical problem where data inflation cannot be dealt with.

(57) 摘要: 本公开公开了一种数据处理方法和装置, 涉及大数据技术领域。该方法的一具体实施方式包括: 采用自适应方式对数据表中的待处理数据进行采样, 得到所述数据表的样本数据分布; 将所述待处理数据划分为多个分片, 根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量; 根据所述多个分片的连接操作对应的输出数据量, 从所述多个分片中筛选出目标分片, 对所述目标分片进行数据分片。该实施方式能够解决无法处理数据膨胀的技术问题。

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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, 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))。

一种数据处理方法和装置

相关申请的交叉引用

本申请要求享有 2021 年 7 月 30 日提交的题为“一种数据处理方法和装置”的中国专利申请 No.202110871021.0 的中国专利申请的优先权，在此全文引用上述中国专利申请公开的内容以作为本申请的一部分或全部。

技术领域

本公开涉及大数据技术领域，尤其涉及一种数据处理方法和装置。

背景技术

数据倾斜一直是大数据领域里面的一个主要痛点，它是指在分布式的数据并行处理过程中，由于数据分片不均匀，导致大量数据集中到一个或少数计算节点上，导致少数任务的处理速度远低于平均速度，拖慢整个计算环节。其中一种常见的数据倾斜情况被称为数据膨胀，其表现为：输入分片大小并没有显著大于平均水平，但其对应数据只含有少数不同的 key，导致 Join 连接执行后，对应任务输出的数据量远大于其他任务。

在实现本公开过程中，发明人发现现有技术中至少存在如下问题：现有计算引擎基于输入分片的大小进行划分，不能处理数据膨胀的场景。而且，在实际生产环境中，情况更为复杂，多种数据倾斜场景（包含数据膨胀）往往一同出现，而现有技术方案均不能处理包含数据膨胀的情况。

发明内容

有鉴于此，本发明实施例提供一种数据处理方法和装置，以解决无法处理数据膨胀的技术问题。

为实现上述目的，根据本公开实施例的一个方面，提供了一种数据处理方法，包括：

采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布；

5 将所述待处理数据划分为多个分片，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量；

根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片。

10 可选地，采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布，包括：

从数据表中获取待处理数据，以获取所述待处理数据中的 key；

判断累计处理条数是否大于等于第一数量阈值；若否，则将所述待处理数据中的 key 写入计数表，将所述计数表中的 key 作为采样结果；

15 若是，则采用蓄水池采样算法对所述待处理数据进行采样；

根据所述待处理数据的采样结果得到所述数据表的样本数据分布。

可选地，采用蓄水池采样算法对所述待处理数据进行采样，包括：

20 判断累计处理条数是否大于等于第二数量阈值；若否，则将所述待处理数据中的 key 写入采样数组；若是，则采用蓄水池采样算法将所述采样数组中的一个 key 置换为所述待处理数据中的 key；

将所述采样数组中的 key 作为采样结果。

25 可选地，根据所述待处理数据的采样结果得到所述数据表的样本数据分布，包括：

分别计算每个样本的累计采样量；其中，所述待处理数据的采样结果包括多个样本；

30 对于每个数据表，将所述待处理数据的数据量除以所述每个样本的累计采样量，得到所述数据表中每个样本的权重，从而得到所述数据表的样本数据分布。

可选地，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量，包括：

5 对于所述多个分片中的任意一个分片，从采样结果中筛选出被分配至所述分片的目标样本；其中，每个所述数据表包含所述目标样本；

根据所述数据表中各个目标样本的权重，计算所述分片的连接操作对应的输出数据量。

10 可选地，根据所述数据表中各个目标样本的权重，计算所述分片的连接操作对应的输出数据量，包括：

将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量。

15 可选地，将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量，包括：

若连接操作的类型为内连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量；

20 若连接操作的类型为左连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量；

若连接操作的类型为右连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在右连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量；

25 若连接操作的类型为全连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表和右连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量。

30 可选地，根据所述多个分片的连接操作对应的输出数据量，从所

述多个分片中筛选出目标分片，对所述目标分片进行数据分片，包括：

对所述多个分片的连接操作对应的输出数据量进行排序，从所述多个分片中筛选出中位数的中位分片；

5 从所述多个分片中筛选出输出数据量大于等于所述中位分片的连接操作对应的输出数据量预设倍数的目标分片；

将每个所述目标分片划分为多个分片。

另外，根据本公开实施例的另一个方面，提供了一种数据处理装置，包括：

10 采样模块，用于采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布；

计算模块，用于将所述待处理数据划分为多个分片，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量；

15 分片模块，用于根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片。

根据本公开实施例的另一个方面，还提供了一种电子设备，包括：一个或多个处理器；

20 存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行时，所述一个或多个处理器实现上述任一实施例所述的方法。

25 根据本公开实施例的另一个方面，还提供了一种计算机可读介质，其上存储有计算机程序，所述程序被处理器执行时实现上述任一实施例所述的方法。

上述的非惯用的可选方式所具有的进一步效果将在下文中结合具体实施方式加以说明。

30

附图说明

附图用于更好地理解本公开，不构成对本公开的不当限定。其中：

图 1 是 Spark AQE 的基本原理示意图；

图 2 是 Spark AQE 中的动态优化策略基本原理示意图；

5 图 3 是根据本公开实施例的数据处理方法的主要流程的示意图；

图 4 是根据本公开一个可参考实施例的数据处理方法的主要流程的示意图；

图 5 是根据本公开另一个可参考实施例的数据处理方法的主要流程的示意图；

10 图 6a 和 6b 是采用本公开实施例的数据处理方法前后的对比示意图；

图 7 是根据本公开实施例的数据处理装置的主要模块的示意图；

图 8 是本公开实施例可以应用于其中的示例性系统架构图；

图 9 是适于用来实现本公开实施例的终端设备或服务器的计算机系统的结构示意图。

15

具体实施方式

以下结合附图对本公开的示范性实施例做出说明，其中包括本公开实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。

20 因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本公开的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

25 以 Spark、Hive 为首的分布式计算引擎，都针对数据倾斜提供了一定的处理能力。以 Spark 为例，其自从 3.0 版本开始，在运行时中引入了自适应查询执行优化（AQE：Adaptive Query Execution）功能，可以基于已完成环节（Stage）的统计信息，来自适应的、动态的调整和优化后续计算环节。

30 AQE 中一个很重要的策略为：动态优化倾斜的 Join 连接。如图 1 所示，以表 A 连接（Join）表 B 为例介绍其基本原理，假定表 A 的分

片 A0 数据量明显大于其他分片，在执行哈希数据分发（hash shuffle）后，处理 A0、B0 的计算任务的数据量明显大于其他分片，导致整个计算环节被拖慢。shuffle：分布式计算中，数据在网络上分发的过程；一般会基于 Hash 值进行 shuffle，对应同一 Hash 值的数据会被放在同一分片中。

Spark AQE 中的动态优化策略，会统计每个分片的大小，在发现 A0 显著的大于各个分片的中位数时，会触发对 A0 的划分，同时为了保证计算的顺利进行，会对 B0 进行复制，如图 2 所示。经过这样的处理后，虽然总任务数增多（由 4 个变成了 5 个），但每个任务的数据更加均匀，计算耗时差不多，为整个查询计算带来了更好的性能。

实际生产环境中，存在多种多样的数据倾斜场景。其中一种常见的数据倾斜情况被称为数据膨胀，其表现为：输入分片大小并没有显著大于平均水平，但其对应数据只含有少数不同的 key，导致 Join 连接执行后，对应任务输出的数据量远大于其他任务。现有计算引擎基于输入分片的大小进行划分，不能处理这种数据膨胀的场景。例如，输入的分片 A3 和 B3 的大小不存在倾斜情况，但是 A3 和 B3 中都存在一个频繁 key=Hot_Sku。该 key 在 A3 侧和 B3 侧的频次分别为 N_A 和 N_B ，则在 Join 连接后的结果中，该 key 的频次为 $N_A * N_B$ 。当 N_A 和 N_B 较输入分片尺寸正常，但 $N_A * N_B$ 较输出分片尺寸较大时，则会发生数据膨胀。

为了解决现有技术中存在的问题，本公开实施例针对在分布式环境下，通过自适应的数据统计和分析，对 Join 连接的输出数据量进行预估，并据此对数据分片进行划分，以处理数据膨胀。

图 3 是根据本公开实施例的数据处理方法的主要流程的示意图。作为本公开的一个实施例，如图 3 所示，所述数据处理方法可以包括：
步骤 301，采用自适应方式对数据表中的待处理数据进行采样，得

到所述数据表的样本数据分布。

首先采用自适应方式对数据表中的待处理数据进行采样，采样逻辑以分布式的形式执行在执行端（Executor）。通过自适应数据采样可以从数据表中采集到样本数据，从而计算出数据表的样本数据分布。

Executor: Spark 计算引擎中，实际执行具体计算任务的节点。

可选地，步骤 301 可以包括：从数据表中获取待处理数据，以获取所述待处理数据中的 key；判断累计处理条数是否大于等于第一数量阈值；若否，则将所述待处理数据中的 key 写入计数表，将所述计数表中的 key 作为采样结果；若是，则采用蓄水池采样算法对所述待处理数据进行采样；根据所述待处理数据的采样结果得到所述数据表的样本数据分布。可以预先设定第一数量阈值（MapSize），之后以流的形式进行增量更新。每次从数据表中获取一条待处理数据时，从该条待处理数据中提取出 key，并以累计方式记录累计处理条数，如果累计处理条数小于第一数量阈值，则进入计数模式（Counting），在该模式下，将该 key 写入计数表（CountMap），并将计数表中的所有 key 作为采样结果；如果累计处理条数大于等于第一数量阈值，即计数表的大小达到预设的第一数量阈值，则进入采样模式（Sampling），即采用蓄水池采样算法对所述待处理数据进行采样，保证整个流中的每个 key 以等概率被采样保留。蓄水池抽样(Reservoir Sampling): 一种流式采样算法，能够在 $O(n)$ 时间内对 n 个数据进行等概率随机抽取。

需要说明的是，待处理数据中包含 key 及其对应的 value，本公开实施例只需要 key，因此只需要从待处理数据中提取出 key。

可选地，采用蓄水池采样算法对所述待处理数据进行采样，包括：判断累计处理条数是否大于等于第二数量阈值；若否，则将所述待处理数据中的 key 写入采样数组；若是，则采用蓄水池采样算法将所述采样数组中的一个 key 置换为所述待处理数据中的 key；将所述采样数组

中的 key 作为采样结果。在采样模式下，需要预先设定第二数量阈值（ArraySize），第二数量阈值大于第一数量阈值。进入采样模式后，如果累计处理条数小于第二数量阈值，则将从待处理数据中提取出的 key 写入采样数组，如果累计处理条数大于等于第二数量阈值，则采用蓄水池采样算法将所述采样数组中的一个 key 替换为从待处理数据中提取出的 key，最后将采样数组中的所有 key 作为采样结果。

在本公开的实施例中，当进入采样模式时，将计数表中的所有 key 全部写入采样数组，以便于直接从采样数组中获取采样结果，采样数组中的所有 key 即为采样结果。首先尝试进行精确采样，在累计处理条数大于等于第一数量阈值的情况下，会回退到蓄水池采样。

可选地，根据所述待处理数据的采样结果得到所述数据表的样本数据分布，包括：分别计算每个样本的累计采样量；其中，所述待处理数据的采样结果包括多个样本；对于每个数据表，将所述待处理数据的数据量除以所述每个样本的累计采样量，得到所述数据表中每个样本的权重，从而得到所述数据表的样本数据分布。在本公开的实施例中，各个执行端（Executor）的采样执行完毕后，会将采样结果、分片数据量 C、采样量 S 发送回调度端（Driver），调度端为每个样本赋予权重 C/S ，并将每个执行端对应的样本权重进行汇总累加处理（每个样本的权重分别汇总累加）。Driver: Spark 计算引擎中，负责调度、汇总、分配任务的节点。

如图 1 所示，以表 A 为例，将每个 map 任务中各个样本的权重分别汇总累加，从而得到表 A 的样本数据分布。同理，也可以得到表 B 的样本数据分布。

步骤 302，将所述待处理数据划分为多个分片，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量。

可以采用哈希算法将待处理数据划分为多个分片，如图 1 所示，根据哈希算法将表 A 划分为四个分片 A0、A1、A2 和 A3，将表 B 划分为两个分片 B0 和 B1；然后根据步骤 301 计算出的表 A 和表 B 的样本数据分布，分别计算出每个分片的 Join 对应的输出数据量。

5

Join 连接：SQL 中用于把来自两个或多个表的行结合起来，获取匹配关系的计算逻辑。按照语义可以分为：InnerJoin（内连接）、LeftJoin（左连接）、RightJoin（右连接）、FullOuterJoin（全连接）等。

10

可选地，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量，包括：对于所述多个分片中的任意一个分片，从采样结果中筛选出被分配至所述分片的目标样本；其中，每个所述数据表包含所述目标样本；根据所述数据表中各个目标样本的权重，计算所述分片的连接操作对应的输出数据量。通过自适应采样，获得了表 A 和表 B 的样本数据分布，对于某个分片 X，从表 A 和表 B 的采样结果中筛选出落在分片 X 中的目标样本 A_X 和 B_X，然后根据各个目标样本 A_X 和 B_X 的权重，计算分片 X 的 Join 对应的输出数据量。需要指出的是，表 A 和表 B 中均包含目标样本，目标样本可以是多个。

20

可选地，根据所述数据表中各个目标样本的权重，计算所述分片的连接操作对应的输出数据量，包括：将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量。具体地，可以采用如下公式计算各个分片的 Join 对应的输出数据量：

25

$$\text{Estimated Size}(X) = \sum_{x \in A_X \cap B_X} w_A(x) * w_B(x)$$

其中， $w_A(x)$ 和 $w_B(x)$ 分别为表 A 和表 B 的数据中，样本 x 的权重。

可选地，将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量，包括：若连接

操作的类型为内连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量；若连接操作的类型为左连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量；若连接操作的类型为右连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在右连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量；若连接操作的类型为全连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表和右连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量。为了准确地计算分片的 Join 对应的输出数据量，不同类型的 Join，采用不同的计算公式：

如果 JoinType 为内连接（InnerJoin）：

$$\text{Estimated Size}(X) = \sum_{x \in A_X \cap B_X} w_A(x) * w_B(x);$$

如果 JoinType 为左连接（LeftJoin）：

$$\text{Estimated Size}(X) = \sum_{x \in A_X \cap B_X} w_A(x) * w_B(x) + \sum_{x \in A_X} w_A(x);$$

如果 JoinType 为右连接（RightJoin）：

$$\text{Estimated Size}(X) = \sum_{x \in A_X \cap B_X} w_A(x) * w_B(x) + \sum_{x \in B_X} w_B(x);$$

如果 JoinType 为全连接（FullOuterJoin）：

$$\text{Estimated Size}(X) = \sum_{x \in A_X \cap B_X} w_A(x) * w_B(x) + \sum_{x \in A_X} w_A(x) + \sum_{x \in B_X} w_B(x);$$

其中， $w_A(x)$ 和 $w_B(x)$ 分别为表 A 和表 B 的数据中，样本 x 的权重。

步骤 303，根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片。

由于计算出了每个分片的 Join 对应的输出数据量，因此可以基于输出数据量对各个分片进行排序，从而筛选出目标分片，并对目标分片进行数据分片，以控制分片的输出数据量的大小。

可选地，步骤 303 可以包括：对所述多个分片的连接操作对应的输出数据量进行排序，从所述多个分片中筛选出中位数的中位分片；从所述多个分片中筛选出输出数据量大于等于所述中位分片的连接操作对应的输出数据量预设倍数的目标分片；将每个所述目标分片划分为多个分片。在本公开的实施例中，先基于输出数据量对各个分片进行排序，然后从这些分片中筛选出中位数的中位分片，最后筛选出输出数据量大于等于所述中位分片的连接操作对应的输出数据量 W 倍数（预设参数）的目标分片。

在本公开的实施例中，筛选出目标分片之后，可以采用哈希算法对目标分片进行数据分片，以控制分片的输出数据量的大小。

根据上面所述的各种实施例，可以看出本公开实施例通过采用自适应方式对数据表中的待处理数据进行采样，得到数据表的样本数据分布，根据数据表的样本数据分布分别计算多个分片的连接操作对应的输出数据量，从而筛选出目标分片并对其进行数据分片的技术手段，解决了现有技术中无法处理数据膨胀的技术问题。本公开实施例通过自适应数据采用得到数据表的样本数据分布，以获得数据的整体分布，然后根据样本数据分布计算出各个分配的输出数据量，准确地筛选出需要再次进行数据分片的目标分片，从而有限地处理数据膨胀，避免发生数据膨胀。

图 4 是根据本公开一个可参考实施例的数据处理方法的主要流程的示意图。作为本公开的又一个实施例，如图 4 所示，采用自适应方式对数据表中的待处理数据进行采样的步骤可以包括：

步骤 401，从数据表中获取待处理数据，以获取所述待处理数据中的 key。

步骤 402，累计处理条数是否大于等于第一数量阈值；若是，则执行步骤 403；若否，则执行步骤 406。

步骤 403，累计处理条数是否大于等于第二数量阈值；若是，则执行步骤 404；若否，则执行步骤 405。

5 步骤 404，采用蓄水池采样算法将所述采样数组中的一个 key 置换为所述待处理数据中的 key，将所述采样数组中的 key 作为采样结果。

 步骤 405，将所述待处理数据中的 key 写入采样数组，将所述采样数组中的 key 作为采样结果。

10

 步骤 406，将所述待处理数据中的 key 写入计数表，将所述计数表中的 key 作为采样结果。

15

 另外，在本公开一个可参考实施例中数据处理方法的具体实施内容，在上面所述数据处理方法中已经详细说明了，故在此重复内容不再说明。

20

 图 5 是根据本公开另一个可参考实施例的数据处理方法的主要流程的示意图。作为本公开的另一个实施例，如图 5 所示，采用自适应方式对数据表中的待处理数据进行采样的步骤可以包括：

 步骤 501，采用自适应方式对数据表中的待处理数据进行采样。

25

 可选地，步骤 501 可以包括：从数据表中获取待处理数据，以获取所述待处理数据中的 key；判断累计处理条数是否大于等于第一数量阈值；若否，则将所述待处理数据中的 key 写入计数表，将所述计数表中的 key 作为采样结果；若是，则采用蓄水池采样算法对所述待处理数据进行采样；根据所述待处理数据的采样结果得到所述数据表的样本数据分布。通过预先设定第一数量阈值，可以自适应地采用计数模式或者采用采样模式，采样逻辑以分布式的形式执行在执行端。

30

在本公开的实施例中，每次从数据表中获取一条待处理数据时，从该条待处理数据中提取出 **key**，并以累计方式记录累计处理条数，如果累计处理条数小于第一数量阈值，则进入计数模式，在该模式下，将该 **key** 写入计数表（**CountMap**），并将计数表中的所有 **key** 作为采
5 样结果；如果累计处理条数大于等于第一数量阈值，即计数表的大小达到预设的第一数量阈值，则进入采样模式，即采用蓄水池采样算法对所述待处理数据进行采样，保证整个流中的每个 **key** 以等概率被采样保留。

10 步骤 502，分别计算每个样本的累计采样量。

其中，所述待处理数据的采样结果包括多个样本，从采用结果中统计出每个样本的累计采样量。

15 步骤 503，对于每个数据表，将所述待处理数据的数据量除以所述每个样本的累计采样量，得到所述数据表中每个样本的权重，从而得到所述数据表的样本数据分布。

各个执行端（**Executor**）的采样执行完毕后，会将采样结果、分片数据量 **C**、采样量 **S** 发送回调度端（**Driver**），调度端为每个样本赋予
20 权重 C/S ，并将每个执行端对应的样本权重进行汇总累加处理（每个样本的权重分别汇总累加），从而得到每个数据表的样本数据分布。

步骤 504，将所述待处理数据划分为多个分片。

25 可以采用哈希算法将待处理数据划分为多个分片。

步骤 505，对于所述多个分片中的任意一个分片，从采样结果中筛选出被分配至所述分片的目标样本。

30 每个所述数据表包含所述目标样本，目标样本被分配至所述分片。

步骤 506, 根据所述数据表中各个目标样本的权重, 计算所述分片的连接操作对应的输出数据量。

5 具体地, 将各个目标样本在各个数据表中的权重相乘并将乘积累加, 从而得到所述分片的连接操作对应的输出数据量。具体地, 可以采用如下公式计算各个分片的 Join 对应的输出数据量:

$$\text{Estimated Size}(X) = \sum_{x \in A_X \cap B_X} w_A(x) * w_B(x)$$

其中, $w_A(x)$ 和 $w_B(x)$ 分别为表 A 和表 B 的数据中, 样本 x 的权重。

10 为了准确地计算分片的 Join 对应的输出数据量, 不同类型的 Join, 采用不同的计算公式, 不再赘述。

步骤 507, 对所述多个分片的连接操作对应的输出数据量进行排序, 从所述多个分片中筛选出中位数的中位分片。

15

步骤 508, 从所述多个分片中筛选出输出数据量大于等于所述中位分片的连接操作对应的输出数据量预设倍数的目标分片。

步骤 509, 将每个所述目标分片划分为多个分片。

20

筛选出目标分片之后, 可以采用哈希算法对目标分片进行数据分片, 以控制分片的输出数据量的大小。

25 另外, 在本公开另一个可参考实施例中数据处理方法的具体实施内容, 在上面所述数据处理方法中已经详细说明了, 故在此重复内容不再说明。

在实践中, 本公开实施例提供的方法可以有效地处理数据膨胀的问题, 以下为例:

如图 6a 所示，分片 5 的输入数据（Shuffle Read Size）为 1.8MiB，与其他分片差别不大；但是其输出数据量（Output Size）为 75.5MiB，远超其他分片（中位数为 853.4KiB）。本例中共包含 10 个分片，数据膨胀的分片运行时长 17 秒，远超其他分片的运行时长 2 秒。由于现有技术方案仅仅考虑输入数据量的大小，故不能解决数据膨胀问题。

如图 6b 所示，采用本公开实施例提供的方法后，依据预估的输出数据量大小，对输入数据进行切分，从原有的 10 个分片，切为 72 个分片；切分之后，所有分片的输出更加均匀，且所有分片的执行时间都不超过 2 秒，使得总运行时长从 17 秒优化为 2 秒，效率获得极大提升。

图 7 是根据本公开实施例的数据处理装置的主要模块的示意图，如图 7 所示，所述数据处理装置 700 包括采样模块 701、计算模块 702 和分片模块 703；其中，采样模块 701 用于采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布；计算模块 702 用于将所述待处理数据划分为多个分片，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量；分片模块 703 用于根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片。

可选地，所述采样模块 701 还用于：

从数据表中获取待处理数据，以获取所述待处理数据中的 key；

判断累计处理条数是否大于等于第一数量阈值；若否，则将所述待处理数据中的 key 写入计数表，将所述计数表中的 key 作为采样结果；若是，则采用蓄水池采样算法对所述待处理数据进行采样；

根据所述待处理数据的采样结果得到所述数据表的样本数据分布。

可选地，所述采样模块 701 还用于：

判断累计处理条数是否大于等于第二数量阈值；若否，则将所述

待处理数据中的 key 写入采样数组；若是，则采用蓄水池采样算法将所述采样数组中的一个 key 置换为所述待处理数据中的 key；

将所述采样数组中的 key 作为采样结果。

5 可选地，所述采样模块 701 还用于：

分别计算每个样本的累计采样量；其中，所述待处理数据的采样结果包括多个样本；

对于每个数据表，将所述待处理数据的数据量除以所述每个样本的累计采样量，得到所述数据表中每个样本的权重，从而得到所述数据表的样本数据分布。

10

可选地，所述计算模块 702 还用于：

对于所述多个分片中的任意一个分片，从采样结果中筛选出被分配至所述分片的目标样本；其中，每个所述数据表包含所述目标样本；

15 根据所述数据表中各个目标样本的权重，计算所述分片的连接操作对应的输出数据量。

可选地，所述计算模块 702 还用于：

将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量。

20

可选地，所述计算模块 702 还用于：

若连接操作的类型为内连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量；

25

若连接操作的类型为左连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量；

若连接操作的类型为右连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在右连接数据表中的

30

权重，从而得到所述分片的连接操作对应的输出数据量；

若连接操作的类型为全连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表和右连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量。

可选地，所述分片模块 703 还用于：

对所述多个分片的连接操作对应的输出数据量进行排序，从所述多个分片中筛选出中位数的中位分片；

从所述多个分片中筛选出输出数据量大于等于所述中位分片的连接操作对应的输出数据量预设倍数的目标分片；

将每个所述目标分片划分为多个分片。

根据上面所述的各种实施例，可以看出本公开实施例通过采用自适应方式对数据表中的待处理数据进行采样，得到数据表的样本数据分布，根据数据表的样本数据分布分别计算多个分片的连接操作对应的输出数据量，从而筛选出目标分片并对其进行数据分片的技术手段，解决了现有技术中无法处理数据膨胀的技术问题。本公开实施例通过自适应数据采用得到数据表的样本数据分布，以获得数据的整体分布，然后根据样本数据分布计算出各个分配的输出数据量，准确地筛选出需要再次进行数据分片的目标分片，从而有限地处理数据膨胀，避免发生数据膨胀。

需要说明的是，在本公开所述数据处理装置的具体实施内容，在上面所述数据处理方法中已经详细说明了，故在此重复内容不再说明。

图 8 示出了可以应用本公开实施例的数据处理方法或数据处理装置的示例性系统架构 800。

如图 8 所示，系统架构 800 可以包括终端设备 801、802、803，网

络 804 和服务器 805。网络 804 用以在终端设备 801、802、803 和服务器 805 之间提供通信链路的介质。网络 804 可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

5 用户可以使用终端设备 801、802、803 通过网络 804 与服务器 805 交互，以接收或发送消息等。终端设备 801、802、803 上可以安装有各种通讯客户端应用，例如购物类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等（仅为示例）。

10 终端设备 801、802、803 可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。

15 服务器 805 可以是提供各种服务的服务器，例如对用户利用终端设备 801、802、803 所浏览的购物类网站提供支持的后台管理服务器（仅为示例）。后台管理服务器可以对接收到的物品信息查询请求等数据进行分析等处理，并将处理结果反馈给终端设备。

20 需要说明的是，本公开实施例所提供的数据处理方法一般由服务器 805 执行，相应地，所述数据处理装置一般设置在服务器 805 中。

 应该理解，图 8 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

25 下面参考图 9，其示出了适于用来实现本公开实施例的终端设备的计算机系统 900 的结构示意图。图 9 示出的终端设备仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

30 如图 9 所示，计算机系统 900 包括中央处理单元（CPU）901，其可以根据存储在只读存储器（ROM）902 中的程序或者从存储部分 908

加载到随机访问存储器（RAM）903 中的程序而执行各种适当的动作和处理。在 RAM 903 中，还存储有系统 900 操作所需的各种程序和数据。CPU 901、ROM 902 以及 RAM903 通过总线 904 彼此相连。输入/输出（I/O）接口 905 也连接至总线 904。

5

以下部件连接至 I/O 接口 905：包括键盘、鼠标等的输入部分 906；包括诸如阴极射线管（CRT）、液晶显示器（LCD）等以及扬声器等的输出部分 907；包括硬盘等的存储部分 908；以及包括诸如 LAN 卡、调制解调器等的网络接口卡的通信部分 909。通信部分 909 经由诸如因特网的网络执行通信处理。驱动器 910 也根据需要连接至 I/O 接口 905。可拆卸介质 911，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 910 上，以便于从其上读出的计算机程序根据需要被安装入存储部分 908。

15

特别地，根据本公开公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开公开的实施例包括一种计算机程序，其包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信部分 909 从网络上被下载和安装，和/或从可拆卸介质 911 被安装。在该计算机程序被中央处理单元（CPU）901 执行时，执行本公开的系统中限定的上述功能。

20

需要说明的是，本公开所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存

25

30

储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：无线、电线、光缆、RF 等等，或者上述的任意合适的组合。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的模块可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的模块也可以设置在处理器中，例如，可以描述为：一种处理器包括采样模块、计算模块和分片模块，其中，这些模块的名称在某种情况下并不构成对该模块本身的限定。

作为另一方面，本公开还提供了一种计算机可读介质，该计算机

可读介质可以是上述实施例中描述的设备中所包含的；也可以是单独存在，而未装配入该设备中。上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被一个该设备执行时，该设备实现如下方法：采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布；将所述待处理数据划分为多个分片，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量；根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片。

根据本公开实施例的技术方案，因为通过采用自适应方式对数据表中的待处理数据进行采样，得到数据表的样本数据分布，根据数据表的样本数据分布分别计算多个分片的连接操作对应的输出数据量，从而筛选出目标分片并对其进行数据分片的技术手段，所以克服了现有技术中无法处理数据膨胀的技术问题。本公开实施例通过自适应数据采用得到数据表的样本数据分布，以获得数据的整体分布，然后根据样本数据分布计算出各个分配的输出数据量，准确地筛选出需要再次进行数据分片的目标分片，从而有限地处理数据膨胀，避免发生数据膨胀。

上述具体实施方式，并不构成对本公开保护范围的限制。本领域技术人员应该明白的是，取决于设计要求和因素，可以发生各种各样的修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等，均应包含在本公开保护范围之内。

权 利 要 求 书

1. 一种数据处理方法，包括：

采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布；

将所述待处理数据划分为多个分片，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量；

根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片。

2. 根据权利要求 1 所述的方法，其中，采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布，包括：

从数据表中获取待处理数据，以获取所述待处理数据中的 key；

判断累计处理条数是否大于等于第一数量阈值；若否，则将所述待处理数据中的 key 写入计数表，将所述计数表中的 key 作为采样结果；若是，则采用蓄水池采样算法对所述待处理数据进行采样；

根据所述待处理数据的采样结果得到所述数据表的样本数据分布。

3. 根据权利要求 2 所述的方法，其中，采用蓄水池采样算法对所述待处理数据进行采样，包括：

判断累计处理条数是否大于等于第二数量阈值；若否，则将所述待处理数据中的 key 写入采样数组；若是，则采用蓄水池采样算法将所述采样数组中的一个 key 替换为所述待处理数据中的 key；

将所述采样数组中的 key 作为采样结果。

4. 根据权利要求 2 所述的方法，其中，根据所述待处理数据的采样结果得到所述数据表的样本数据分布，包括：

分别计算每个样本的累计采样量；其中，所述待处理数据的采样结果包括多个样本；

对于每个数据表，将所述待处理数据的数据量除以所述每个样本

的累计采样量，得到所述数据表中每个样本的权重，从而得到所述数据表的样本数据分布。

5. 根据权利要求 4 所述的方法，其中，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量，包括：

对于所述多个分片中的任意一个分片，从采样结果中筛选出被分配至所述分片的目标样本；其中，每个所述数据表包含所述目标样本；

根据所述数据表中各个目标样本的权重，计算所述分片的连接操作对应的输出数据量。

6. 根据权利要求 5 所述的方法，其中，根据所述数据表中各个目标样本的权重，计算所述分片的连接操作对应的输出数据量，包括：

将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量。

7. 根据权利要求 6 所述的方法，其中，将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量，包括：

若连接操作的类型为内连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，从而得到所述分片的连接操作对应的输出数据量；

若连接操作的类型为左连接，则将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量；

若连接操作的类型为右连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在右连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据量；

若连接操作的类型为全连接，将各个目标样本在各个数据表中的权重相乘并将乘积累加，然后加上各个目标样本在左连接数据表和右连接数据表中的权重，从而得到所述分片的连接操作对应的输出数据

量。

8. 根据权利要求 1 所述的方法，其中，根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片，包括：

对所述多个分片的连接操作对应的输出数据量进行排序，从所述多个分片中筛选出中位数的中位分片；

从所述多个分片中筛选出输出数据量大于等于所述中位分片的连接操作对应的输出数据量预设倍数的目标分片；

将每个所述目标分片划分为多个分片。

9. 一种数据处理装置，包括：

采样模块，用于采用自适应方式对数据表中的待处理数据进行采样，得到所述数据表的样本数据分布；

计算模块，用于将所述待处理数据划分为多个分片，根据所述数据表的样本数据分布分别计算所述多个分片的连接操作对应的输出数据量；

分片模块，用于根据所述多个分片的连接操作对应的输出数据量，从所述多个分片中筛选出目标分片，对所述目标分片进行数据分片。

10. 一种电子设备，包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行时，所述一个或多个处理器实现如权利要求 1-8 中任一所述的方法。

11. 一种计算机可读介质，其上存储有计算机程序，所述程序被处理器执行时实现如权利要求 1-8 中任一所述的方法。

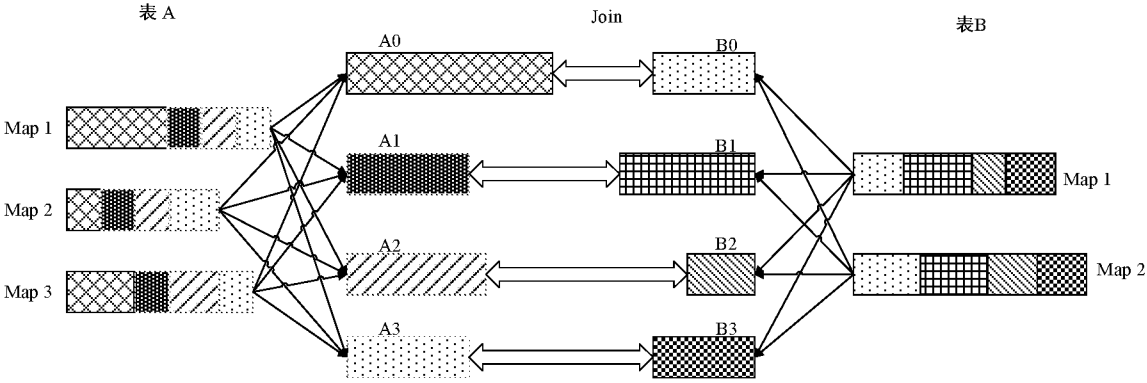


图 1

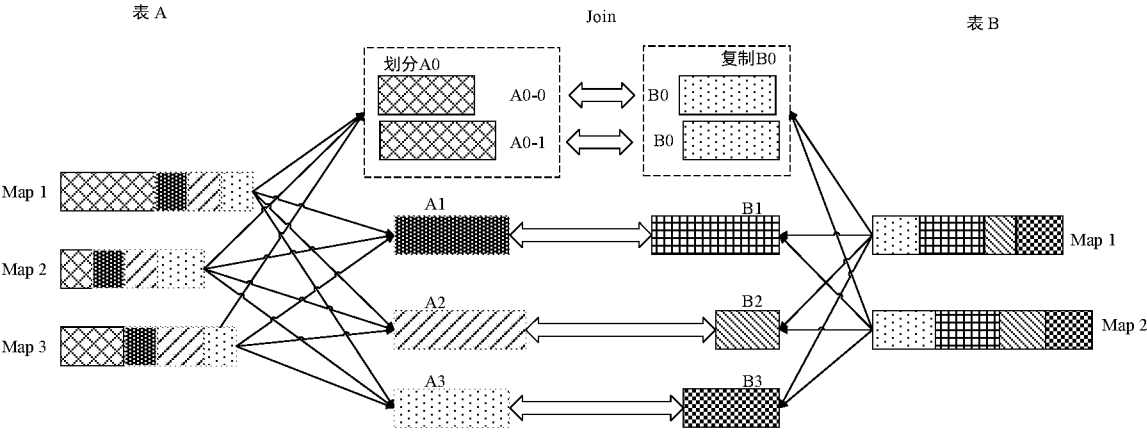


图 2

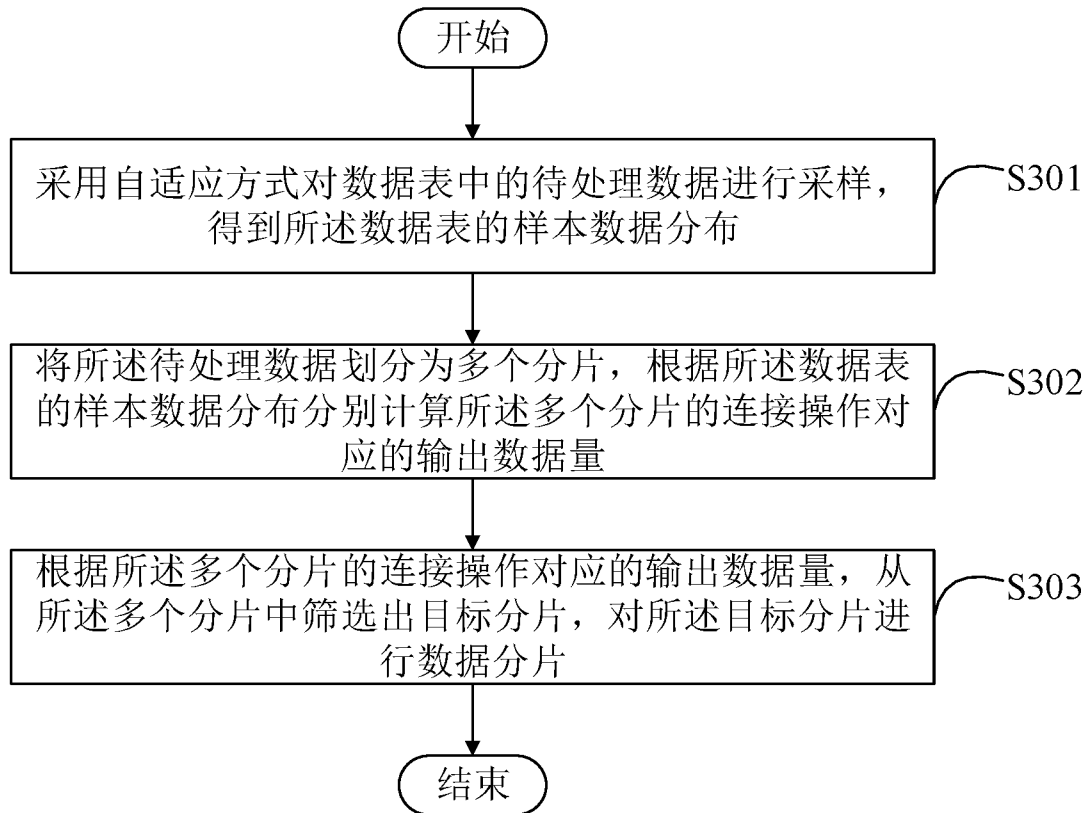


图 3

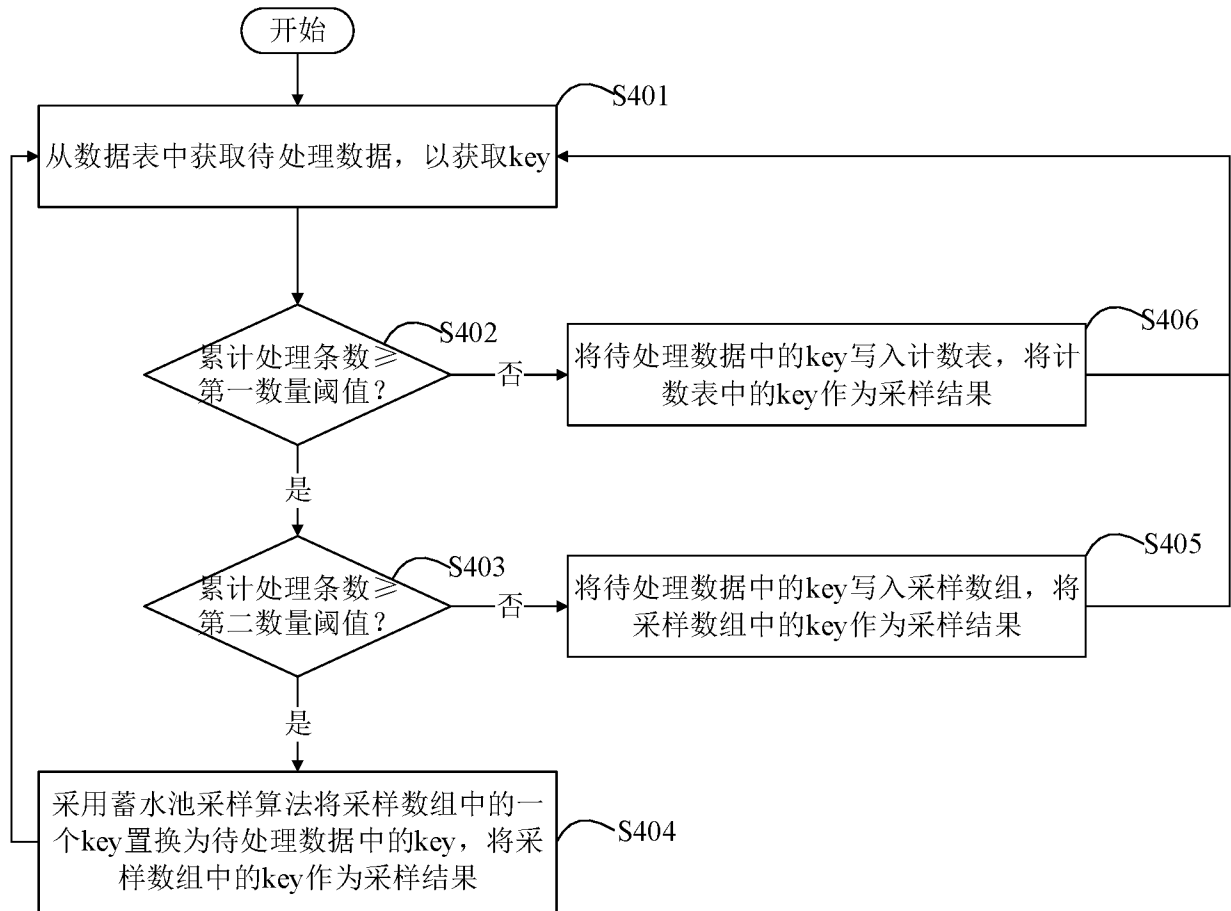


图 4

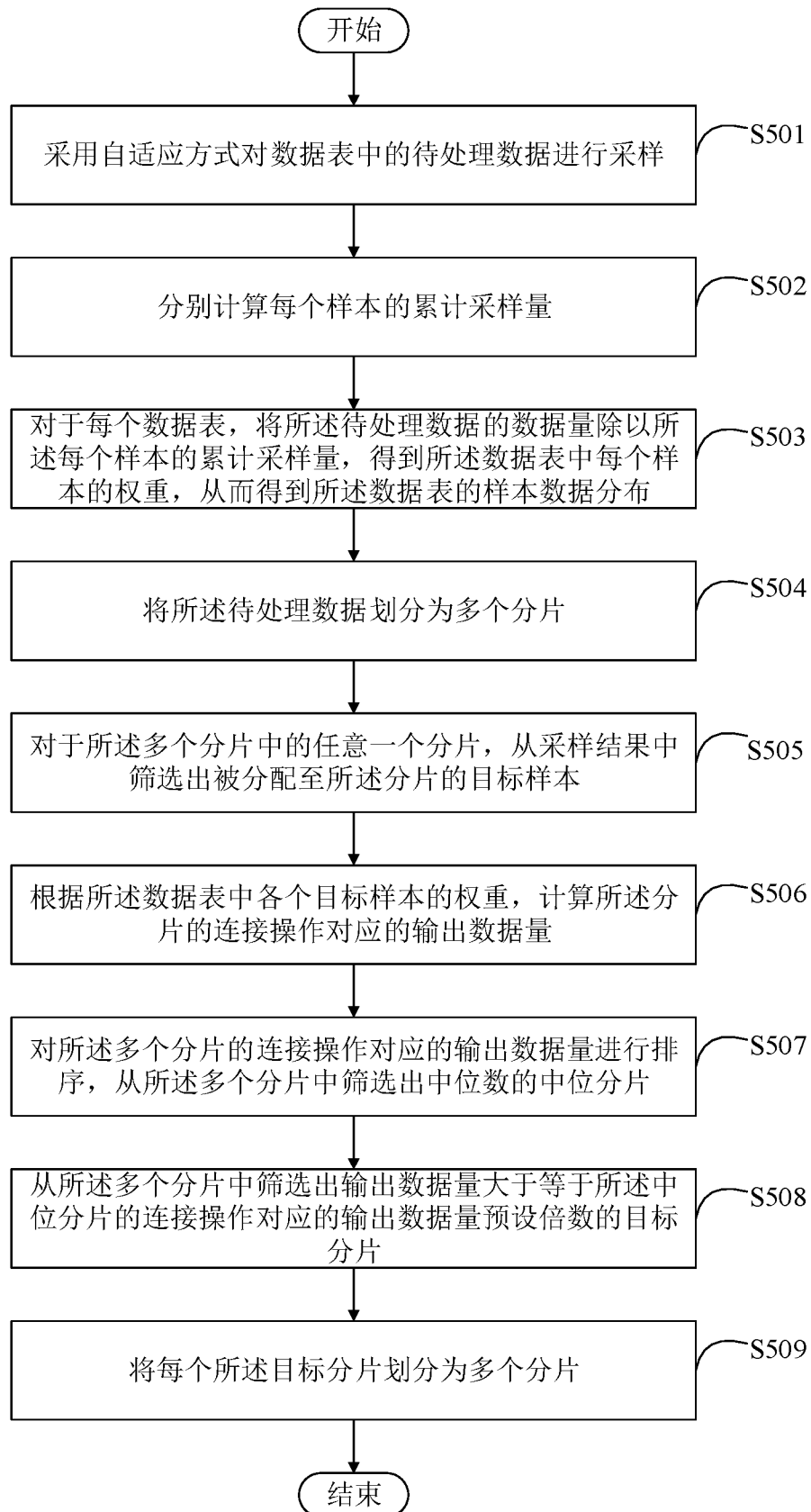


图 5

[illegible]

Index	Year	Month	Category	Country	Subcategory	Value	Label	Value	Unit	Year	Month	Category	Country	Subcategory	Value	Label	Value	Unit
1	2019	01	01	USA	01	1000000	1000000	1000000	1000000	2019	01	01	USA	01	1000000	1000000	1000000	1000000
2	2019	02	02	USA	02	1000000	1000000	1000000	1000000	2019	02	02	USA	02	1000000	1000000	1000000	1000000
3	2019	03	03	USA	03	1000000	1000000	1000000	1000000	2019	03	03	USA	03	1000000	1000000	1000000	1000000
4	2019	04	04	USA	04	1000000	1000000	1000000	1000000	2019	04	04	USA	04	1000000	1000000	1000000	1000000
5	2019	05	05	USA	05	1000000	1000000	1000000	1000000	2019	05	05	USA	05	1000000	1000000	1000000	1000000
6	2019	06	06	USA	06	1000000	1000000	1000000	1000000	2019	06	06	USA	06	1000000	1000000	1000000	1000000
7	2019	07	07	USA	07	1000000	1000000	1000000	1000000	2019	07	07	USA	07	1000000	1000000	1000000	1000000
8	2019	08	08	USA	08	1000000	1000000	1000000	1000000	2019	08	08	USA	08	1000000	1000000	1000000	1000000
9	2019	09	09	USA	09	1000000	1000000	1000000	1000000	2019	09	09	USA	09	1000000	1000000	1000000	1000000
10	2019	10	10	USA	10	1000000	1000000	1000000	1000000	2019	10	10	USA	10	1000000	1000000	1000000	1000000
11	2019	11	11	USA	11	1000000	1000000	1000000	1000000	2019	11	11	USA	11	1000000	1000000	1000000	1000000
12	2019	12	12	USA	12	1000000	1000000	1000000	1000000	2019	12	12	USA	12	1000000	1000000	1000000	1000000
13	2020	01	01	USA	01	1000000	1000000	1000000	1000000	2020	01	01	USA	01	1000000	1000000	1000000	1000000
14	2020	02	02	USA	02	1000000	1000000	1000000	1000000	2020	02	02	USA	02	1000000	1000000	1000000	1000000
15	2020	03	03	USA	03	1000000	1000000	1000000	1000000	2020	03	03	USA	03	1000000	1000000	1000000	1000000
16	2020	04	04	USA	04	1000000	1000000	1000000	1000000	2020	04	04	USA	04	1000000	1000000	1000000	1000000
17	2020	05	05	USA	05	1000000	1000000	1000000	1000000	2020	05	05	USA	05	1000000	1000000	1000000	1000000
18	2020	06	06	USA	06	1000000	1000000	1000000	1000000	2020	06	06	USA	06	1000000	1000000	1000000	1000000
19	2020	07	07	USA	07	1000000	1000000	1000000	1000000	2020	07	07	USA	07	1000000	1000000	1000000	1000000
20	2020	08	08	USA	08	1000000	1000000	1000000	1000000	2020	08	08	USA	08	1000000	1000000	1000000	1000000
21	2020	09	09	USA	09	1000000	1000000	1000000	1000000	2020	09	09	USA	09	1000000	1000000	1000000	1000000
22	2020	10	10	USA	10	1000000	1000000	1000000	1000000	2020	10	10	USA	10	1000000	1000000	1000000	

5/7

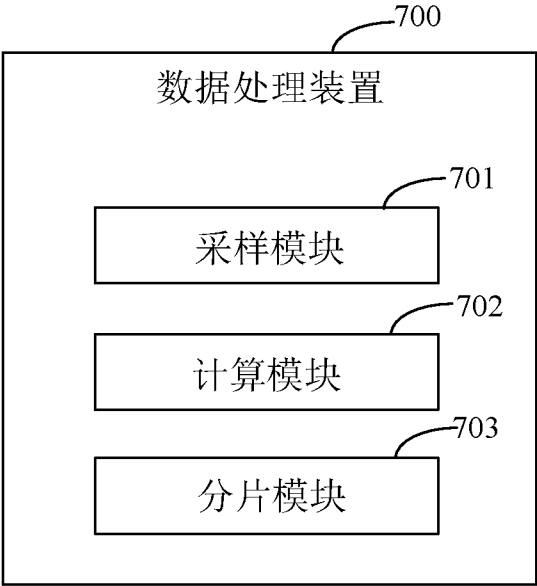


图 7

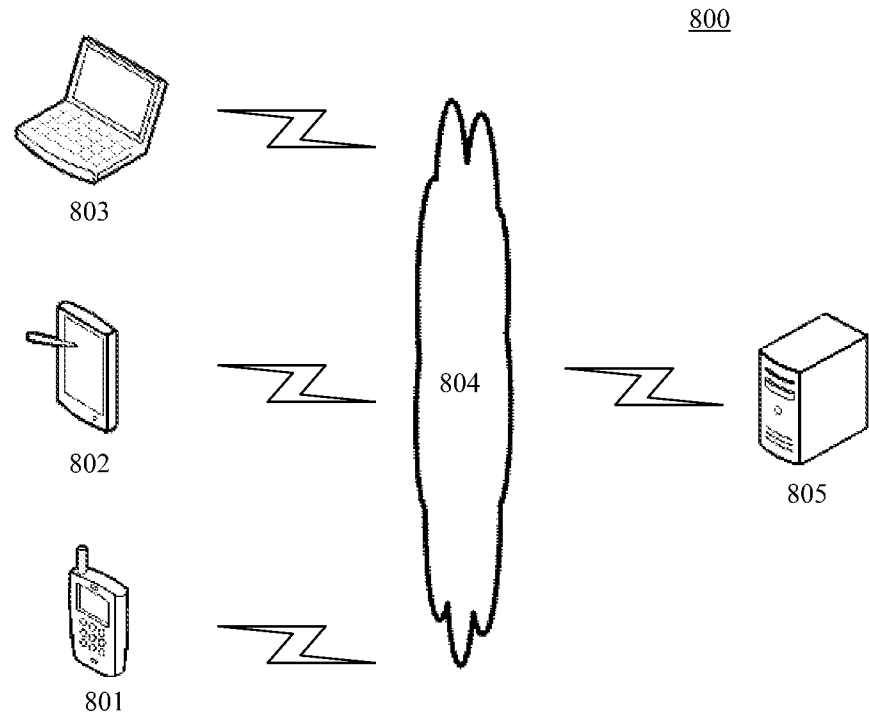


图 8

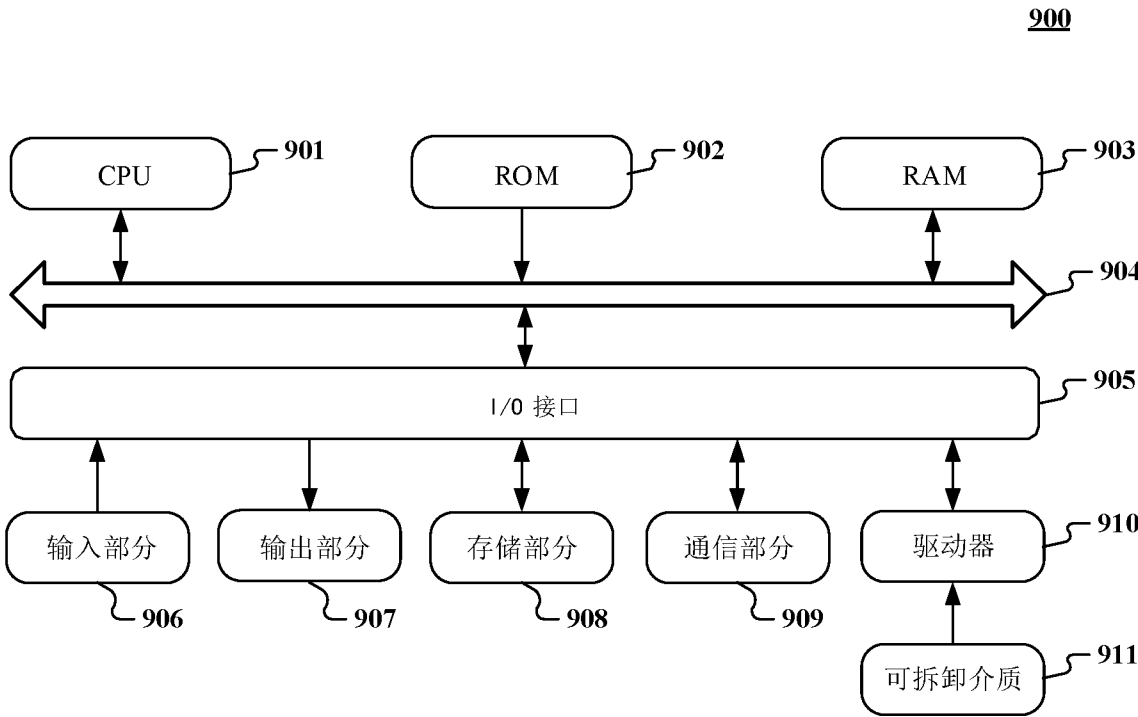


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/087297

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/50(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)

USTXT; VEN; CNABS; WPABS; CNTXT; EPTXT; WOTXT; CNKI; IEEE: 数据, 倾斜, 偏置, 膨胀, 分布, 分散, 平均, 采样, 抽样, 阈值, 累计, 连接, 筛选, 权重, data, inclination, tilt, skew, bias, expand, distribute, disperse, average, sample, threshold, accumulate, join, filter, weight

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113590322 A (BEIJING WODONG TIANJUN INFORMATION TECHNOLOGY CO., LTD. et al.) 02 November 2021 (2021-11-02) claims 1-11, and description, paragraphs 85-187	1-11
Y	CN 112905596 A (BEIJING ZHONGJING HUIZHONG TECHNOLOGY CO., LTD.) 04 June 2021 (2021-06-04) description, paragraphs 22-102	1-11
Y	胡忠奎 等 (HU, Zhongkui et al.). "一种基于虚拟处理区间划分的负载均衡等值连接算法 (A Load Balanced Equi-Join Algorithm Based on Virtual Processor Range Partition)" <i>现代计算机 (Modern Computer)</i> , No. 03, 25 January 2016 (2016-01-25), ISSN: 1007-1423, pages 3-7	1-11
Y	张占峰 等 (ZHANG, Zhanfeng et al.). "Spark数据倾斜问题研究 (Research on Data Skew of Spark)" <i>河北省科学院学报 (Journal of the Hebei Academy of Sciences)</i> , Vol. 37, No. 1, 31 March 2020 (2020-03-31), ISSN: 1001-9383, pages 1-7	1-11

☒ 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

10 June 2022

Date of mailing of the international search report

23 June 2022

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/087297

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	黄子纯 (HUANG, Zichun). "基于Spark数据倾斜的分区负载均衡研究 (Research on Partition Loading Balance Based on Spark Data Skew)" 中国优秀硕士学位论文全文数据库(硕士)信息科技辑 (月刊) (<i>China Doctoral Dissertations/Master's Theses full-text database (Monthly)</i>), No. 2, 15 February 2021 (2021-02-15), ISSN: 1674-0246, text, pages 12-43	2-7, 10-11
Y	CN 112000467 A (GUANGDONG POLYTECHNIC NORMAL UNIVERSITY et al.) 27 November 2020 (2020-11-27) description, paragraphs 67-117	2-7, 10-11
A	CN 107145394 A (NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY OF PLA) 08 September 2017 (2017-09-08) entire document	1-11
A	US 2015331724 A1 (TERADATA US INC.) 19 November 2015 (2015-11-19) entire document	1-11

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/087297

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113590322	A	02 November 2021	None			
CN	112905596	A	04 June 2021	None			
CN	112000467	A	27 November 2020	None			
CN	107145394	A	08 September 2017	CN	107145394	B	08 May 2020
US	2015331724	A1	19 November 2015	US	9405592	B2	02 August 2016

国际检索报告

国际申请号

PCT/CN2022/087297

A. 主题的分类 G06F 9/50 (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)) USTXT; VEN; CNABS; WPABS; CNTXT; EPTXT; WOTXT; CNKI; IEEE: 数据, 倾斜, 偏置, 膨胀, 分布, 分散, 平均, 采样, 抽样, 阈值, 累计, 连接, 筛选, 权重, data, inclination, tilt, skew, bias, expand, distribute, disperse, average, sample, threshold, accumulate, join, filter, weight																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 113590322 A (北京沃东天骏信息技术有限公司 等) 2021年11月2日 (2021 - 11 - 02) 权利要求1-11、说明书第85-187段</td> <td>1-11</td> </tr> <tr> <td>Y</td> <td>CN 112905596 A (北京中经惠众科技有限公司) 2021年6月4日 (2021 - 06 - 04) 说明书第22-102段</td> <td>1-11</td> </tr> <tr> <td>Y</td> <td>胡忠奎 等. "一种基于虚拟处理区间划分的负载均衡等值连接算法" 现代计算机, 第03期, 2016年1月25日 (2016 - 01 - 25), ISSN: 1007-1423, 第3-7页</td> <td>1-11</td> </tr> <tr> <td>Y</td> <td>张占峰 等. "Spark数据倾斜问题研究" 河北省科学院学报, 第37卷, 第1期, 2020年3月31日 (2020 - 03 - 31), ISSN: 1001-9383, 第1-7页</td> <td>1-11</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 113590322 A (北京沃东天骏信息技术有限公司 等) 2021年11月2日 (2021 - 11 - 02) 权利要求1-11、说明书第85-187段	1-11	Y	CN 112905596 A (北京中经惠众科技有限公司) 2021年6月4日 (2021 - 06 - 04) 说明书第22-102段	1-11	Y	胡忠奎 等. "一种基于虚拟处理区间划分的负载均衡等值连接算法" 现代计算机, 第03期, 2016年1月25日 (2016 - 01 - 25), ISSN: 1007-1423, 第3-7页	1-11	Y	张占峰 等. "Spark数据倾斜问题研究" 河北省科学院学报, 第37卷, 第1期, 2020年3月31日 (2020 - 03 - 31), ISSN: 1001-9383, 第1-7页	1-11
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
PX	CN 113590322 A (北京沃东天骏信息技术有限公司 等) 2021年11月2日 (2021 - 11 - 02) 权利要求1-11、说明书第85-187段	1-11															
Y	CN 112905596 A (北京中经惠众科技有限公司) 2021年6月4日 (2021 - 06 - 04) 说明书第22-102段	1-11															
Y	胡忠奎 等. "一种基于虚拟处理区间划分的负载均衡等值连接算法" 现代计算机, 第03期, 2016年1月25日 (2016 - 01 - 25), ISSN: 1007-1423, 第3-7页	1-11															
Y	张占峰 等. "Spark数据倾斜问题研究" 河北省科学院学报, 第37卷, 第1期, 2020年3月31日 (2020 - 03 - 31), ISSN: 1001-9383, 第1-7页	1-11															
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
<table border="0"> <tr> <td> * 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 </td> <td> "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 </td> </tr> </table>			* 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													
* 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		国际检索报告邮寄日期															
2022年6月10日		2022年6月23日															
ISA/CN的名称和邮寄地址		受权官员															
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		赵静 电话号码 86-(20)-28958139															

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
Y	黄子纯. “基于Spark数据倾斜的分区负载均衡研究” 中国优秀硕士学位论文全文数据库(硕士)信息科技辑 (月刊), 第2期, 2021年2月15日 (2021 - 02 - 15), ISSN: 1674-0246, 正文第12-43页	2-7、10-11
Y	CN 112000467 A (广东技术师范大学 等) 2020年11月27日 (2020 - 11 - 27) 说明书第67-117段	2-7、10-11
A	CN 107145394 A (中国人民解放军国防科学技术大学) 2017年9月8日 (2017 - 09 - 08) 全文	1-11
A	US 2015331724 A1 (TERADATA US INC) 2015年11月19日 (2015 - 11 - 19) 全文	1-11

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/087297

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	113590322	A	2021年11月2日	无	
CN	112905596	A	2021年6月4日	无	
CN	112000467	A	2020年11月27日	无	
CN	107145394	A	2017年9月8日	CN	107145394 B 2020年5月8日
US	2015331724	A1	2015年11月19日	US	9405592 B2 2016年8月2日