



(12) 发明专利

(10) 授权公告号 CN 112818183 B

(45) 授权公告日 2024.05.17

(21) 申请号 202110152682.8

(22) 申请日 2021.02.03

(65) 同一申请的已公布的文献号
申请公布号 CN 112818183 A

(43) 申请公布日 2021.05.18

(73) 专利权人 恒安嘉新(北京)科技股份有限公司
地址 100098 北京市海淀区北三环西路25号27号楼五层5002室

(72) 发明人 张建浩 盛磊 武林红 陈浩
侯立冬 孟宝权 傅强 蔡琳
梁彧 田野 王杰 杨满智 金红
陈晓光

(74) 专利代理机构 北京品源专利代理有限公司
11332
专利代理师 孟金喆

(51) Int. Cl.
G06F 16/901 (2019.01)

(56) 对比文件

- CN 105264535 A, 2016.01.20
- CN 106933204 A, 2017.07.07
- CN 107181776 A, 2017.09.19
- CN 107704590 A, 2018.02.16
- CN 107748797 A, 2018.03.02
- CN 108833346 A, 2018.11.16
- CN 109299121 A, 2019.02.01
- CN 109412867 A, 2019.03.01
- CN 109491928 A, 2019.03.19
- CN 111078147 A, 2020.04.28
- CN 111737271 A, 2020.10.02
- CN 112000703 A, 2020.11.27
- CN 112231405 A, 2021.01.15
- US 2009196458 A1, 2009.08.06
- US 6999968 B1, 2006.02.14

姚庆华等. 面向综合智能交通系统的多源异构数据集成框架研究.《云南大学学报(自然科学版)》.2017,第41-45页.

审查员 唐文俊

权利要求书2页 说明书7页 附图3页

(54) 发明名称

一种数据合成方法、装置、计算机设备和存储介质

(57) 摘要

本发明公开了一种数据合成方法、装置、计算机设备和存储介质。该方法包括：实时获取多个数据源接入的待合成数据，并获取各待合成数据对应的数据生成时间；根据各待合成数据对应的数据生成时间，确定目标时间窗口对应的目标待合成数据，并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据；对目标待合成数据和/或缓存数据进行数据合成，获得至少一个合成数据并存储到数据库中。使用本发明的技术方案，可以实现对不同数据源和不同到达时间的数据进行实时合成处理。



1. 一种数据合成方法,其特征在于,包括:

实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间;

所述待合成数据包括主数据源对应的待合成数据,以及至少一个从数据源对应的待合成数据;

根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据;

其中,根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,包括:

在主数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内的待合成数据作为目标待合成数据;

在至少一个从数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内以及延迟等待时间之前的待合成数据作为目标待合成数据;

对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中;

其中,对目标待合成数据和/或缓存数据进行数据合成,包括:

如果确定当前时间为与目标时间窗口匹配的延迟等待时间,则对目标待合成数据和/或缓存数据进行数据合成。

2. 根据权利要求1所述的方法,其特征在于,在实时获取多个数据源接入的待合成数据之前,还包括:

响应于用户的目标数据源接入指令,通过与所述目标数据源接入指令匹配的目标数据源接入组件,接入目标数据源的待合成数据。

3. 根据权利要求1所述的方法,其特征在于,确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据,包括:

在至少一个从数据源对应的各待合成数据中,获取数据生成时间位于与目标时间窗口匹配的缓存时间区间内的待合成数据作为缓存数据。

4. 根据权利要求3所述的方法,其特征在于,对目标待合成数据和/或缓存数据进行数据合成,包括:

通过Join连接方法,对目标待合成数据和/或缓存数据进行数据合成。

5. 根据权利要求1-4任一项所述的方法,其特征在于,所述待合成数据为流式数据。

6. 一种数据合成装置,其特征在于,包括:

数据生成时间获取模块,用于实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间;

所述待合成数据包括主数据源对应的待合成数据,以及至少一个从数据源对应的待合成数据;

数据确定模块,用于根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据;

所述数据确定模块,包括:

第一目标待合成数据获取单元,用于在主数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内的待合成数据作为目标待合成数据;

第二目标待合成数据获取单元,用于在至少一个从数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内以及延迟等待时间之前的待合成数据作为目标待合成数据;

数据合成模块,用于对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中;

所述数据合成模块,包括:

延迟等待时间判断单元,用于如果确定当前时间为与目标时间窗口匹配的延迟等待时间,则对目标待合成数据和/或缓存数据进行数据合成。

7.一种计算机设备,包括存储器、处理器及存储在存储器上并可在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程序时实现如权利要求1-5中任一所述的数据合成方法。

8.一种包含计算机可执行指令的存储介质,其特征在于,所述计算机可执行指令在由计算机处理器执行时用于执行如权利要求1-5中任一所述的数据合成方法。

一种数据合成方法、装置、计算机设备和存储介质

技术领域

[0001] 本发明实施例涉及数据处理技术,尤其涉及一种数据合成方法、装置、计算机设备和存储介质。

背景技术

[0002] 数据合成也即数据的关联合并,是指从大规模数据中发现数据之间隐含的关系与规律,将多种数据信息进行提取计算,合并生成一类数据的过程。

[0003] 传统的数据关联分析为离线分析,将多种数据存储到数据库中进行关联分析,在海量数据的关联分析场景下,传统的离线关联分析无法满足数据的实时计算需求,并且离线分析需要耗费的计算成本较高,效率较低。而在大数据领域,如果要对数据进行实时合成,在多个数据源接入的数据到达时间不同的情况下,可能会存在部分数据无法被合成的情况。

发明内容

[0004] 本发明实施例提供了一种数据合成方法、装置、计算机设备和存储介质,以实现对不同数据源和不同到达时间的数据进行实时合成处理。

[0005] 第一方面,本发明实施例提供了一种数据合成方法,该方法包括:

[0006] 实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间;

[0007] 根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据;

[0008] 对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中。

[0009] 第二方面,本发明实施例还提供了一种数据合成装置,该装置包括:

[0010] 数据生成时间获取模块,用于实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间;

[0011] 数据确定模块,用于根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据;

[0012] 数据合成模块,用于对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中。

[0013] 第三方面,本发明实施例还提供了一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现如本发明实施例中任一所述的数据合成方法。

[0014] 第四方面,本发明实施例还提供了一种包含计算机可执行指令的存储介质,所述计算机可执行指令在由计算机处理器执行时用于执行如本发明实施例中任一所述的数据

合成方法。

[0015] 本发明实施例通过接入多个数据源,实时获取待合成数据,根据各待合成数据的数据生成时间,确定目标时间窗口对应的目标待合成数据和缓存时间区间对应的缓存数据,对目标待合成数据和/或缓存数据进行数据合成,得到合成数据。解决了现有技术中传统的离线关联分析无法满足数据的实时计算需求,并且需要耗费的计算成本较高,效率较低的问题,以及实时合成时数据到达时间不一导致部分数据无法被合成的问题,实现了对不同数据源和不同到达时间的数据的实时合成处理。

附图说明

- [0016] 图1是本发明实施例一中的一种数据合成方法的流程图;
- [0017] 图2a是本发明实施例二中的一种数据合成方法的流程图;
- [0018] 图2b是适用于本发明实施例中的一种数据合成原理的示意图;
- [0019] 图3是本发明实施例三中的一种数据合成装置的结构示意图;
- [0020] 图4是本发明实施例四中的一种计算机设备的结构示意图。

具体实施方式

[0021] 下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是,此处所描述的具体实施例仅仅用于解释本发明,而非对本发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与本发明相关的部分而非全部结构。

[0022] 实施例一

[0023] 图1是本发明实施例一提供一种数据合成方法的流程图,本实施例可适用于对流式数据进行实时的数据合成的情况,该方法可以由数据合成装置来执行,该装置可以由软件和/或硬件来实现,并一般集成在计算机设备中。

[0024] 如图1所示,本发明实施例的技术方案,具体包括如下步骤:

[0025] S110、实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间。

[0026] 其中,数据源也即数据的来源,是提供某种所需要数据的器件或原始媒体。示例性的,本实施例中可以基于Flink框架进行数据源的接入和待合成数据的合成,在Flink框架下,数据源可以包括内置数据源和外部数据源,内置数据源包括文件数据源或者集合数据源等,外部数据源包括Kafka、RabbitMQ等,采用基于Flink的流式数据框架,可以实现大规模分布式的数据合成,提高数据实时合成的效率。但本实施例对采用的框架类型不进行限制。

[0027] 待合成数据是根据接入的多个数据源获取的数据。可选的,所述待合成数据可以为流式数据。也即,数据源接入的待合成数据以DataStream数据流的形式存在。对不同的数据源,将其数据均转化为统一的流式数据,对于批处理的数据,也将其转换为流式数据进行后续处理。

[0028] 可选的,所述待合成数据可以包括主数据源对应的待合成数据,以及至少一个从数据源对应的待合成数据。

[0029] 在本发明实施例中,对来自多个数据源的待合成数据进行数据合成时,需要将不

同的数据源分为主数据源和从数据源。具体的,可以将接入的流式数据较为连续、无延时的数据源作为主数据源,其他数据源作为从数据源,但本实施例对主数据源的选择依据不进行限制。将数据源划分为主数据源和从数据源之后,就可以以主数据源接入的数据和时间窗口为依据,确定是否要对从数据源接入的数据进行缓存。

[0030] 数据生成时间可以根据待合成数据的时间戳确定,时间戳表示在特定时间点已经存在完整的可验证的数据,可以用来表示数据的产生时间。可选的,可以解析各待合成数据中的时间字段,从而获取各待合成数据的数据生成时间。

[0031] S120、根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据。

[0032] 时间窗口用于将待合成数据按照时间进行分组,从而将每一组作为一个整体进行数据合成。在本发明实施例中,各时间窗口之间是不重叠的,示例性的,可以设置时间窗口长度为2h,则每隔2h进行一次数据合成。确定待合成数据的数据生成时间之后,即可确定待合成数据对应的时间窗口,例如,当待合成数据的数据生成时间为12:00时,该待合成数据对应12:00-14:00这一时间窗口。

[0033] 缓存时间区间是与时间窗口对应的一个时间区间。不同数据源的待合成数据通过键名进行关联,但不同数据源对应的接口不同,不同数据源的待合成数据通过接口接入的时间也不同,则对应相同键名的、来自不同数据源的待合成数据可能存在对应不同的时间窗口的情况,这种情况下这些待合成数据将无法被合成。因此,本发明实施例在时间窗口之前设置缓存时间区间,对于缓存时间区间内接入的待合成数据进行缓存,与时间窗口内接入的待合成数据一起进行数据合成。

[0034] 具体的,可以通过MapState方法,将缓存时间区间内接入的待合成数据存储到HDFS(Hadoop Distributed File System,Hadoop分布式文件系统)中。这样设置的好处在于,HDFS可以实现大规模的状态存储,可以避免数据合成的过程中,程序内存溢出。但本实施例对缓存时间区间内接入的待合成数据存储的位置,以及具体存储方式不进行限制。

[0035] S130、对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中。

[0036] 在本发明实施例中,对目标待合成数据和/或缓存数据,如果存在对应同一键名的数据,则对其进行数据合成生成至少一个合成数据,将各合成数据保存到数据库中。

[0037] 本实施例的技术方案,通过接入多个数据源,实时获取待合成数据,根据各待合成数据的数据生成时间,确定目标时间窗口对应的目标待合成数据和缓存时间区间对应的缓存数据,对目标待合成数据和/或缓存数据进行数据合成,得到合成数据。解决了现有技术中传统的离线关联分析无法满足数据的实时计算需求,并且需要耗费的计算成本较高,效率较低的问题,以及实时合成时数据到达时间不一导致部分数据无法被合成的问题,实现了对不同数据源和不同到达时间的数据的实时合成处理。

[0038] 实施例二

[0039] 图2a是本发明实施例二提供了一种数据合成方法的流程图,本发明实施例在上述实施例的基础上,对确定目标待合成数据的过程、确定缓存数据的过程以及进行数据合成的过程进行了进一步的具体化,并加入了接入数据源的步骤。

[0040] 相应的,如图2a所示,本发明实施例的技术方案,具体包括如下步骤:

[0041] S210、响应于用户的目标数据源接入指令,通过与所述目标数据源接入指令匹配的目标数据源接入组件,接入目标数据源的待合成数据。

[0042] 在本发明实施例中,用户可以通过可视化操作界面,设置数据源信息、待合成数据对应的键名信息以及进行数据合成时的延迟等待时间等。

[0043] 当接收到用户的目标数据源接入指令时,通过目标数据源接入组件,即可对目标数据源的数据进行抽取,获得待合成数据。

[0044] S220、实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间。

[0045] 接入多路数据源之后,实时获取各数据源接入的各待合成数据,通过提取各带合成数据的时间字段,获取各待合成数据对应的数据生成时间。

[0046] S230、在主数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内的待合成数据作为目标待合成数据。

[0047] 主数据源的数据接入较为连续,延迟较小,对主数据源接入的各待合成数据,根据数据生成时间确定其对应的时间窗口,当数据生成时间位于目标时间窗口时,待合成数据为目标待合成数据。

[0048] 图2b提供了一种数据合成原理的示意图,如图2b所示,数据源A为主数据源,数据源B为从数据源,时间窗口为12:00-14:00,数据源A对应的待合成数据为a1、a2、a3和a4,a1、a2、a3和a4的数据生成时间都在时间窗口内,则a1、a2、a3和a4都是目标待合成数据。

[0049] S240、在至少一个从数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内以及延迟等待时间之前的待合成数据作为目标待合成数据。

[0050] 延迟等待时间可以由用户在可视化操作界面进行自定义设置,对应同一时间窗口的待合成数据需要等待到延迟等待时间时,再触发进行数据合成。设置延迟等待时间的原因在于,对于各从数据源,如果待合成数据在时间窗口之后但延迟等待时间之前到达,仍能参与本轮时间窗口的数据合成,可以提高数据合成的效率。

[0051] 如图2b所示,延迟等待时间为10min,时间窗口为12:00-14:00时,14:10进行数据合成,对于从数据源来说,数据生成时间位于12:00-14:10的都作为目标待合成数据。数据源B对应的b2和b3都是目标待合成数据。

[0052] S250、在至少一个从数据源对应的各待合成数据中,获取数据生成时间位于与目标时间窗口匹配的缓存时间区间内的待合成数据作为缓存数据。

[0053] 对于从数据源,数据生成时间位于缓存时间区间内的待合成数据将被保存至缓冲区,等待与目标时间窗口匹配的目标待合成数据共同进行数据合成。

[0054] 示例性的,图2b中时间窗口为12:00-14:00,缓存时间区间为11:00-12:00,在此期间内的从数据源的待合成数据进行缓存,等待在14:10时进行数据合成。数据源B对应的b1数据生成时间是在11:00-12:00之间,则对b1进行缓存,在14:10时,b1即可与a1进行合成。数据源B对应的b4数据生成时间是在14:10之后,无法在本轮目标时间窗口对应的延迟等待时间进行合成,则将b4进行缓存,在下一个时间窗口对应的延迟等待时间等待进行合成。

[0055] S260、判断当前时间是否为与目标时间窗口匹配的延迟等待时间,如果是,则执行S270,否则返回执行S260。

[0056] 在本发明实施例中,在延迟等待时间对目标时间窗口匹配的数据进行数据合成。

[0057] S270、通过Join连接方法,对目标待合成数据和/或缓存数据进行数据合成。

[0058] Join连接方法用于基于表之间的共同字段,将两个或多个表的行结合起来,在本发明实施例中,Join连接方法基于相同的键名,将来自不同数据源的数据进行结合。

[0059] 示例性的,如图2b所示,目标待合成数据包括:a1、a2、a3、a4、b2和b3,缓存数据包括b1,在14:10时,a1和b1、a2和b2、a3和b3将进行数据合成。

[0060] S280、获得至少一个合成数据存储到数据库中。

[0061] 数据合成之后生成新的合成数据,将合成数据存储到数据库中。在本轮目标时间窗口的数据合成之后,未进行合成的数据将被丢弃,不再参与下一轮时间窗口的数据合成。

[0062] 示例性的,在图2b中,由于b4的数据生成时间是在本轮延迟等待时间之后,则本轮a4在延迟等待时间时不存在对应的数据,a4将被丢弃。

[0063] 本实施例的技术方案,通过响应于用户的数据源接入指令,接入多个数据源,实时获取待合成数据,根据各待合成数据的数据生成时间,将数据生成时间位于目标时间窗口内的主数据源的待合成数据作为目标待合成数据,将数据生成时间位于目标时间窗口内和延迟等待时间之前的从数据源的待合成数据作为目标待合成数据,将数据生成时间位于缓存时间区间内的从数据源的待合成数据作为缓存数据,在当前时间为延迟等待时间时,通过Join连接方法,对目标待合成数据和/或缓存数据进行数据合成。解决了现有技术中传统的离线关联分析无法满足数据的实时计算需求,并且需要耗费的计算成本较高,效率较低的问题,以及实时合成时数据到达时间不一导致部分数据无法被合成的问题,实现了对不同数据源和不同到达时间的数据的实时合成处理。

[0064] 实施例三

[0065] 图3是本发明实施例三中的一种数据合成装置的结构示意图,该装置可以由软件和/或硬件来实现,并一般集成在计算机设备中。该装置包括:数据生成时间获取模块310、数据确定模块320以及数据合成模块330。其中:

[0066] 数据生成时间获取模块310,用于实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间;

[0067] 数据确定模块320,用于根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据;

[0068] 数据合成模块330,用于对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中。

[0069] 本实施例的技术方案,通过接入多个数据源,实时获取待合成数据,根据各待合成数据的数据生成时间,确定目标时间窗口对应的目标待合成数据和缓存时间区间对应的缓存数据,对目标待合成数据和/或缓存数据进行数据合成,得到合成数据。解决了现有技术中传统的离线关联分析无法满足数据的实时计算需求,并且需要耗费的计算成本较高,效率较低的问题,以及实时合成时数据到达时间不一导致部分数据无法被合成的问题,实现了对不同数据源和不同到达时间的数据的实时合成处理。

[0070] 在上述实施例的基础上,所述装置,还包括:

[0071] 数据源接入模块,用于响应于用户的目标数据源接入指令,通过与所述目标数据源接入指令匹配的目标数据源接入组件,接入目标数据源的待合成数据。

[0072] 在上述实施例的基础上,所述数据合成模块330,包括:

[0073] 延迟等待时间判断单元,用于如果确定当前时间为与目标时间窗口匹配的延迟等待时间,则对目标待合成数据和/或缓存数据进行数据合成。

[0074] 在上述实施例的基础上,所述待合成数据包括主数据源对应的待合成数据,以及至少一个从数据源对应的待合成数据;

[0075] 所述数据确定模块320,包括:

[0076] 第一目标待合成数据获取单元,用于在主数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内的待合成数据作为目标待合成数据;

[0077] 第二目标待合成数据获取单元,用于在至少一个从数据源对应的各待合成数据中,获取数据生成时间位于目标时间窗口内以及延迟等待时间之前的待合成数据作为目标待合成数据。

[0078] 在上述实施例的基础上,所述数据确定模块320,包括:

[0079] 缓存数据获取单元,用于在至少一个从数据源对应的各待合成数据中,获取数据生成时间位于与目标时间窗口匹配的缓存时间区间内的待合成数据作为缓存数据。

[0080] 在上述实施例的基础上,所述数据合成模块330,包括:

[0081] 数据合成单元,用于通过Join连接方法,对目标待合成数据和/或缓存数据进行数据合成。

[0082] 在上述实施例的基础上,所述待合成数据为流式数据。

[0083] 本发明实施例所提供的数据合成装置可执行本发明任意实施例所提供的数据合成方法,具备执行方法相应的功能模块和有益效果。

[0084] 实施例四

[0085] 图4为本发明实施例四提供的一种计算机设备的结构示意图,如图4所示,该计算机设备包括处理器70、存储器71、输入装置72和输出装置73;计算机设备中处理器70的数量可以是一个或多个,图4中以一个处理器70为例;计算机设备中的处理器70、存储器71、输入装置72和输出装置73可以通过总线或其他方式连接,图4中以通过总线连接为例。

[0086] 存储器71作为一种计算机可读存储介质,可用于存储软件程序、计算机可执行程序以及模块,如本发明实施例中的数据合成方法对应的模块(例如,数据合成装置中的数据生成时间获取模块310、数据确定模块320以及数据合成模块330)。处理器70通过运行存储在存储器71中的软件程序、指令以及模块,从而执行计算机设备的各种功能应用以及数据处理,即实现上述的数据合成方法。该方法包括:

[0087] 实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间;

[0088] 根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据;

[0089] 对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中。

[0090] 存储器71可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序;存储数据区可存储根据终端的使用所创建的数据等。此外,存储器71可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁

盘存储器件、闪存器件、或其他非易失性固态存储器件。在一些实例中,存储器71可进一步包括相对于处理器70远程设置的存储器,这些远程存储器可以通过网络连接至计算机设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0091] 输入装置72可用于接收输入的数字或字符信息,以及产生与计算机设备的用户设置以及功能控制有关的键信号输入。输出装置73可包括显示屏等显示设备。

[0092] 实施例五

[0093] 本发明实施例五还提供一种包含计算机可执行指令的存储介质,所述计算机可执行指令在由计算机处理器执行时用于执行一种数据合成方法,该方法包括:

[0094] 实时获取多个数据源接入的待合成数据,并获取各待合成数据对应的数据生成时间;

[0095] 根据各待合成数据对应的数据生成时间,确定目标时间窗口对应的目标待合成数据,并确定与所述目标时间窗口匹配的缓存时间区间对应的缓存数据;

[0096] 对目标待合成数据和/或缓存数据进行数据合成,获得至少一个合成数据存储到数据库中。

[0097] 当然,本发明实施例所提供的一种包含计算机可执行指令的存储介质,其计算机可执行指令不限于如上所述的方法操作,还可以执行本发明任意实施例所提供的数据合成方法中的相关操作。

[0098] 通过以上关于实施方式的描述,所属领域的技术人员可以清楚地了解到,本发明可借助软件及必需的通用硬件来实现,当然也可以通过硬件实现,但很多情况下前者是更佳的实施方式。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可以以软件产品的形式体现出来,该计算机软件产品可以存储在计算机可读存储介质中,如计算机的软盘、只读存储器(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory, RAM)、闪存(FLASH)、硬盘或光盘等,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的方法。

[0099] 值得注意的是,上述数据合成装置的实施例中,所包括的各个单元和模块只是按照功能逻辑进行划分的,但并不局限于上述的划分,只要能够实现相应的功能即可;另外,各功能单元的具体名称也只是为了便于相互区分,并不用于限制本发明的保护范围。

[0100] 注意,上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解,本发明不限于这里所述的特定实施例,对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本发明的保护范围。因此,虽然通过以上实施例对本发明进行了较为详细的说明,但是本发明不仅仅限于以上实施例,在不脱离本发明构思的情况下,还可以包括更多其他等效实施例,而本发明的范围由所附的权利要求范围决定。

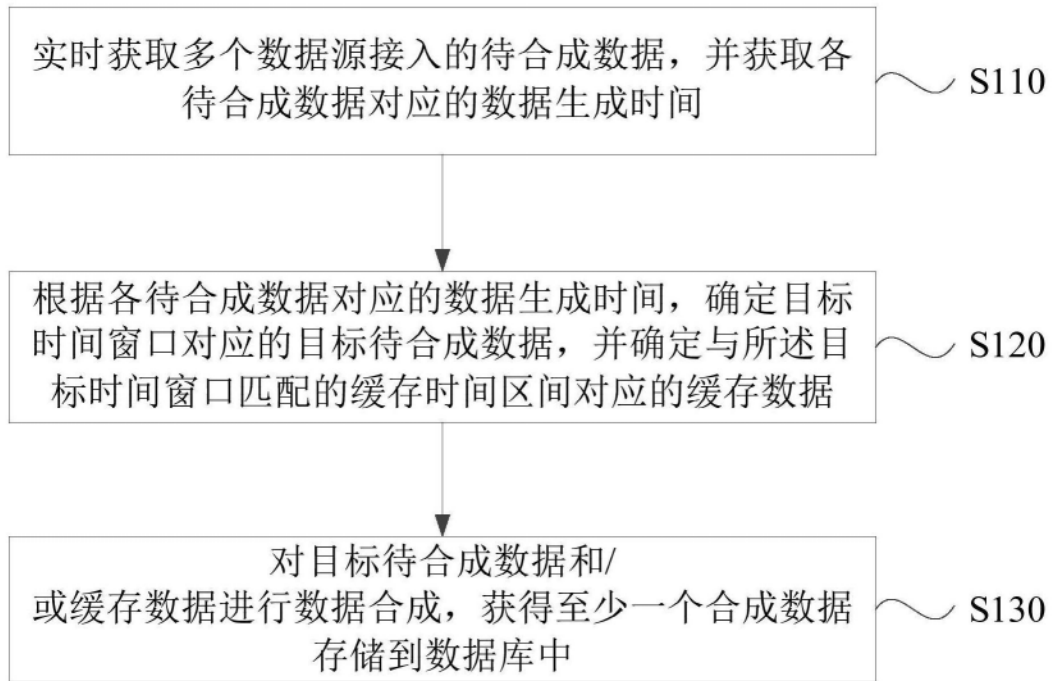


图1

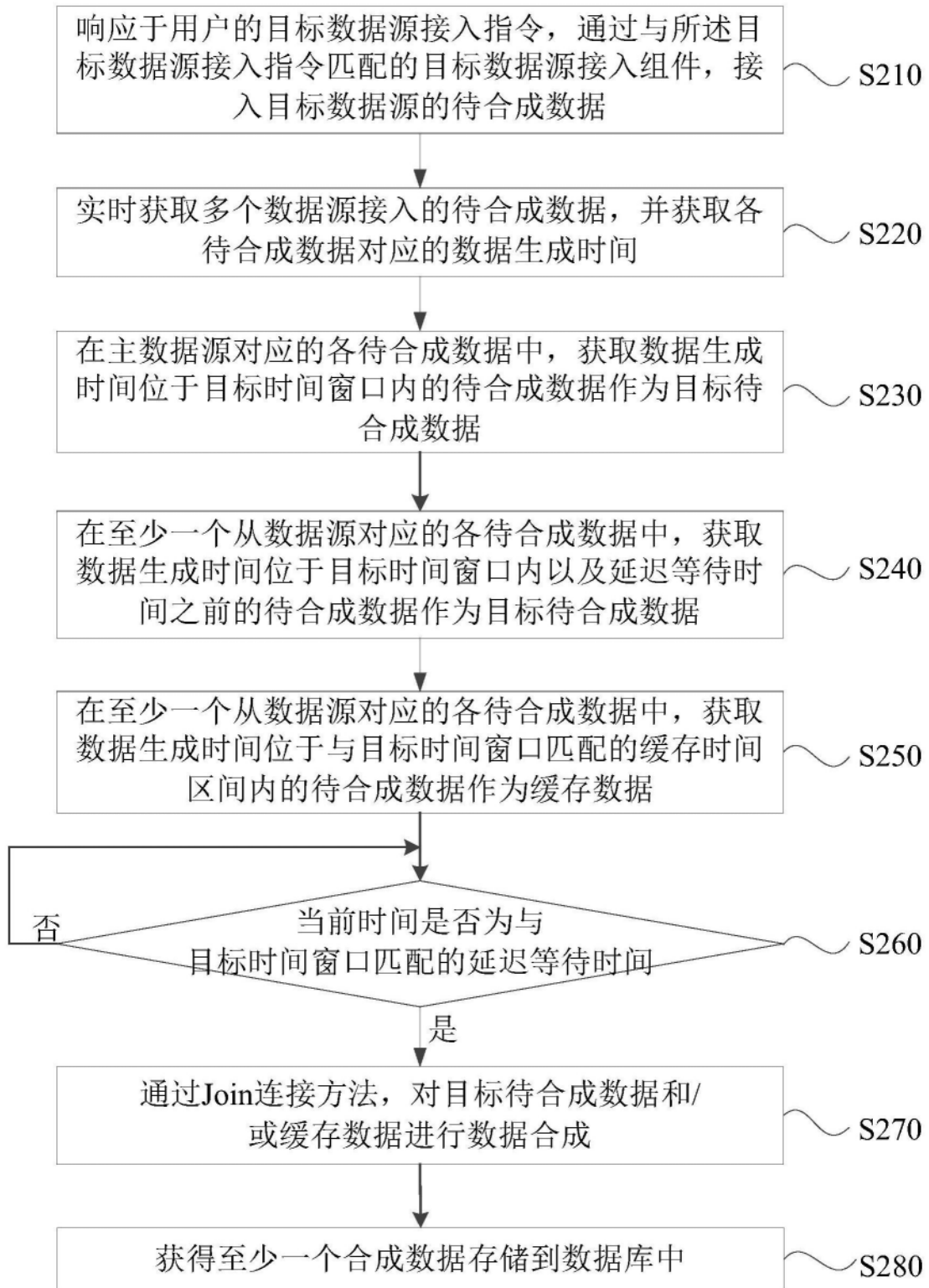


图2a

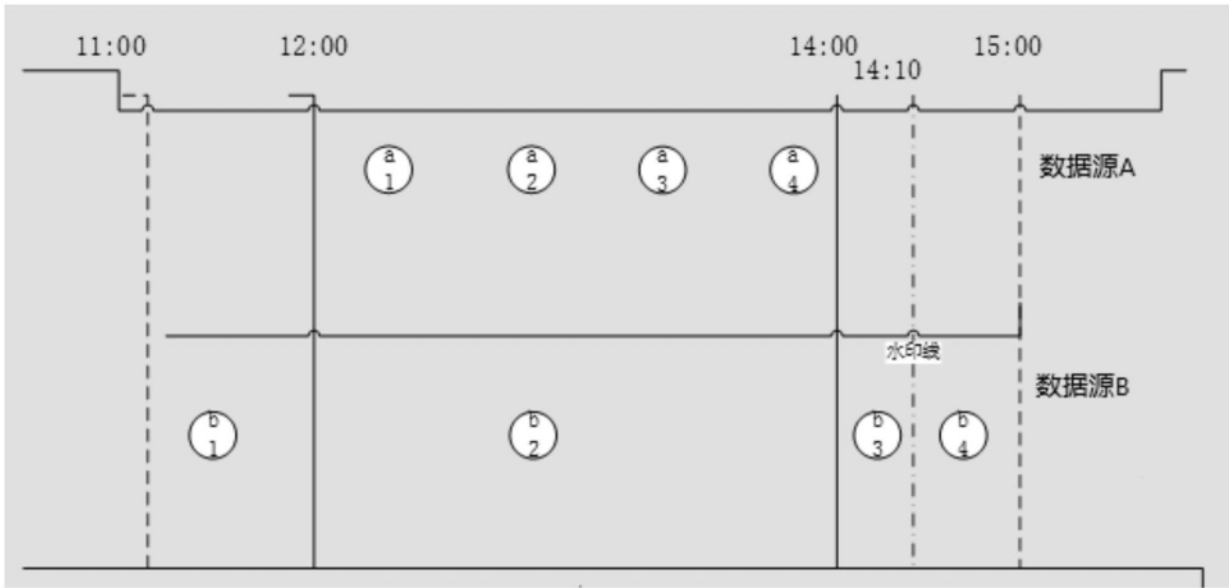


图2b

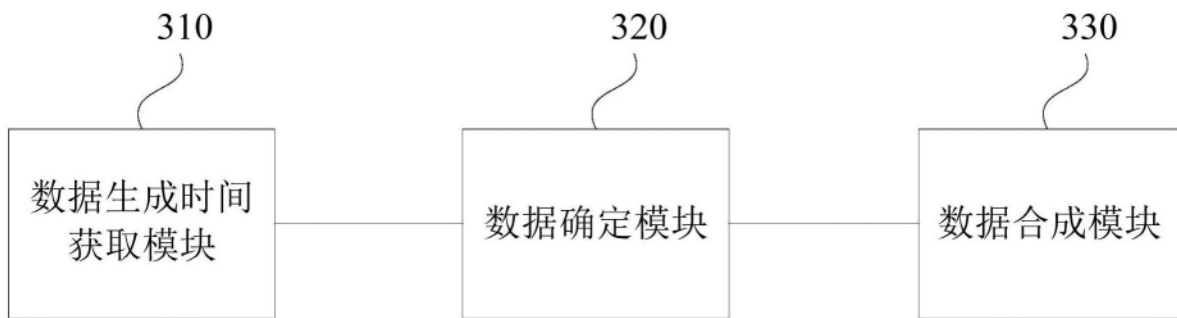


图3

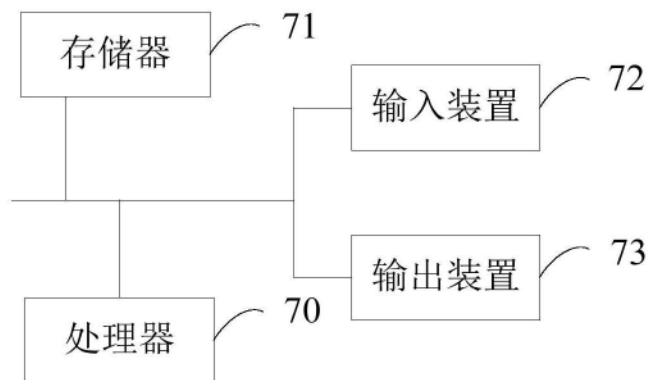


图4