



(12) 发明专利申请

(10) 申请公布号 CN 112699142 A

(43) 申请公布日 2021. 04. 23

(21) 申请号 202011603604.7

(22) 申请日 2020.12.29

(71) 申请人 平安普惠企业管理有限公司
地址 518000 广东省深圳市前海深港合作区前湾一路1号A栋201室(入驻深圳市前海商务秘书有限公司)

(72) 发明人 高玉环

(74) 专利代理机构 深圳市沃德知识产权代理事务所(普通合伙) 44347
代理人 高杰 于志光

(51) Int. Cl.
G06F 16/242 (2019.01)
G06F 16/2455 (2019.01)
G06F 16/23 (2019.01)
G06F 16/28 (2019.01)

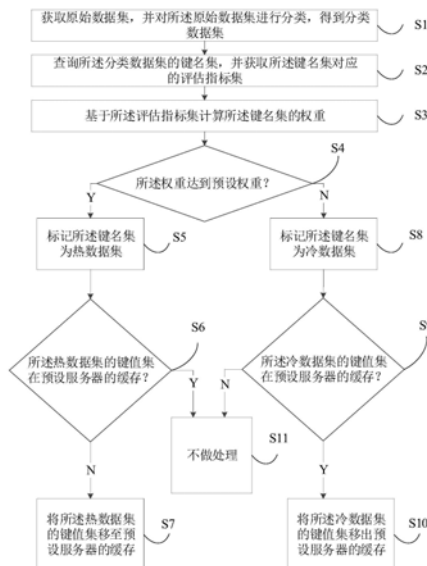
权利要求书2页 说明书10页 附图2页

(54) 发明名称

冷热数据处理方法、装置、电子设备及存储介质

(57) 摘要

本发明涉及数据存储领域,揭露一种冷热数据处理方法,包括:获取原始数据集,对所述原始数据集进行分类,得到分类数据集,查询所述分类数据集的键名集,并获取所述键名集对应的评估指标集,基于所述评估指标集计算所述键名集的权重,并评估所述权重是否达到预设权重,若所述权重达到预设权重,则标记所述键名集为热数据集,并将所述热数据集的键值集移至缓存,若所述权重未达到预设权重,则标记所述键名集为冷数据集,并将所述冷数据集的键值集移出缓存。本发明还提出一种冷热数据处理装置、电子设备以及存储介质。本发明还涉及区块链技术,所述热数据集可以存储于区块链节点中。本发明可以实现高效的数据查询,同时降低数据存储成本。



1. 一种冷热数据处理方法,其特征在于,所述方法包括:
 - 获取原始数据集,对所述原始数据集进行分类,得到分类数据集;
 - 查询所述分类数据集的键名集,并获取所述键名集对应的评估指标集;
 - 基于所述评估指标集计算所述键名集的权重,并评估所述权重是否达到预设权重;
 - 若所述权重达到所述预设权重,则标记所述键名集为热数据集,并判断所述热数据集的键值集不在预设服务器的缓存内时,将所述热数据集的键值集移至预设服务器的缓存;
 - 若所述权重未达到所述预设权重,则标记所述键名集为冷数据集,并判断所述冷数据集的键值集在预设服务器的缓存内时,将所述冷数据集的键值集移出预设服务器的缓存。
2. 如权利要求1所述的冷热数据处理方法,其特征在于,所述对所述原始数据集进行分类,得到分类数据集,包括:
 - 获取所述原始数据集对应的业务场景类别;
 - 根据所述业务场景类别,将所述原始数据集进行分类,得到分类数据集。
3. 如权利要求1所述的冷热数据处理方法,其特征在于,所述获取所述键名集对应的评估指标集,包括:
 - 根据所述键名集在数据库实例中预设时间段内的访问结果,得到访问日志;
 - 从所述访问日志中的访问信息中提取所述键名集对应的评估指标集。
4. 如权利要求1所述的冷热数据处理方法,其特征在于,所述基于所述评估指标集计算所述键名集的权重,包括:
 - 根据预设的时间阈值定时基于所述评估指标集计算所述键名集的权重,并根据重新计算的结果更新所述键名集的权重。
5. 如权利要求4所述的冷热数据处理方法,其特征在于,所述评估指标集包括访问数据长度、访问频次、更新频次和最近访问时间,所述根据预设的时间阈值定时基于所述评估指标集计算所述键名集的权重,包括:
 - 对定时更新的所述访问频次、所述访问数据长度、所述最近访问时间、所述更新频次执行聚合运算,得到所述键名集的权重。
6. 如权利要求1至5中任意一项所述的冷热数据处理方法,其特征在于,所述判断所述热数据集的键值集不在预设服务器的缓存内时,将所述热数据集的键值集移至缓存,包括:
 - 若所述缓存中不包含所述热数据集的键值集,则从数据库中提取对应的键值集,并将所述键值集保存至缓存中。
7. 如权利要求1至5中任意一项所述的冷热数据处理方法,其特征在于,所述标记所述键名集为热数据集之后,所述方法还包括:
 - 将所述热数据集备份在区块链节点中;
 - 当检测到存储所述热数据集的服务器出现宕机或者重启时,在确认服务器重启成功后,从所述区块链节点中获取所述热数据集,并将所述热数据集重新移至所述服务器的缓存中。
8. 一种冷热数据处理装置,其特征在于,所述装置包括:
 - 获取模块,用于获取原始数据集,对所述原始数据集进行分类,得到分类数据集;
 - 查询模块,用于查询所述分类数据集的键名集,并获取所述键名集对应的评估指标集;
 - 计算模块,用于基于所述评估指标集计算所述键名集的权重,并评估所述权重是否达

到预设权重；

处理模块,用于若所述权重达到所述预设权重,则标记所述键名集为热数据集,并判断所述热数据集的键值集不在预设服务器的缓存内时,将所述热数据集的键值集移至预设服务器的缓存,若所述权重未达到所述预设权重,则标记所述键名集为冷数据集,并判断所述冷数据集的键值集在预设服务器的缓存内时,将所述冷数据集的键值集移出预设服务器的缓存。

9. 一种电子设备,其特征在于,所述电子设备包括:

至少一个处理器;以及,

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如权利要求1至7中任一项所述的冷热数据处理方法。

10. 一种计算机可读存储介质,包括存储数据区和存储程序区,所述存储数据区存储创建的数据,所述存储程序区存储有计算机程序,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1至7中任一项所述的冷热数据处理方法。

冷热数据处理方法、装置、电子设备及存储介质

技术领域

[0001] 本发明涉及数据存储领域,尤其涉及一种冷热数据处理方法、装置、电子设备及计算机可读存储介质。

背景技术

[0002] 在企业业务不断变化的过程中,企业配置系统内的业务配置数据呈现爆炸式增长,随着数据量的不断激增,给存储系统带来巨大压力。数据会随着访问热度划分为冷数据和热数据,且冷热数据在一定时间内具有逻辑空间上的相对独立性。如果将冷数据和热数据都按照相同的存储方式进行存储,会影响系统的服务性能同时大幅度提升存储成本。如何根据数据的冷热程度,在保证系统服务性能稳定的前提下降低数据存储成本是一个亟需解决的问题。

发明内容

[0003] 本发明提供一种冷热数据处理方法、装置、电子设备及计算机可读存储介质,其主要目的在于实现高效的数据查询,同时降低数据存储成本。

[0004] 为实现上述目的,本发明提供的一种冷热数据处理方法,包括:

[0005] 获取原始数据集,对所述原始数据集进行分类,得到分类数据集;

[0006] 查询所述分类数据集的键名集,并获取所述键名集对应的评估指标集;

[0007] 基于所述评估指标集计算所述键名集的权重,并评估所述权重是否达到预设权重;

[0008] 若所述权重达到所述预设权重,则标记所述键名集为热数据集,并判断所述热数据集的键值集不在预设服务器的缓存内时,将所述热数据集的键值集移至预设服务器的缓存;

[0009] 若所述权重未达到所述预设权重,则标记所述键名集为冷数据集,并判断所述冷数据集的键值集在预设服务器的缓存内时,将所述冷数据集的键值集移出预设服务器的缓存。

[0010] 可选地,所述对所述原始数据集进行分类,得到分类数据集,包括:

[0011] 获取所述原始数据集对应的业务场景类别;

[0012] 根据所述业务场景的场景类别,将所述原始数据集进行分类,得到分类数据集。

[0013] 可选地,所述获取键名集对应的评估指标集,包括:

[0014] 根据所述键名集在数据库实例中预设时间段内的访问结果,得到访问日志;

[0015] 从所述访问日志中的访问信息中提取所述键名集的评估指标集。

[0016] 可选地,所述基于所述评估指标集计算所述键名集的权重,包括:

[0017] 根据预设的时间阈值定时基于所述评估指标集计算所述键名集的权重,并根据重新计算的结果更新所述键名集的权重。

[0018] 可选地,所述评估指标集包括访问数据长度、访问频次、更新频次和最近访问时

间,所述根据预设的时间阈值定时基于所述评估指标集计算所述键名集的权重,包括:

[0019] 对定时更新的所述访问频次、所述访问数据长度、所述最近访问时间、所述更新频次执行聚合运算,得到所述键名集的权重。

[0020] 可选地,所述判断所述热数据集的键值集不在预设服务器的缓存内时,将所述热数据集的键值集移至预设服务器的缓存,包括:

[0021] 若所述缓存中不包含所述热数据集的键值集,则从数据库中提取对应的键值集,并将所述键值集保存至缓存中。

[0022] 可选地,所述标记所述键名集为热数据集之后,所述方法还包括:

[0023] 将所述热数据集备份在区块链节点中;

[0024] 当存储所述热数据集的服务器出现宕机或者重启时,在确认服务器重启成功后,从所述区块链节点中获取所述热数据集,并将所述热数据集重新移至所述服务器的缓存中。

[0025] 为了解决上述问题,本发明还提供一种冷热数据处理装置,所述装置包括:

[0026] 获取模块,用于获取原始数据集,对所述原始数据集进行分类,得到分类数据集;

[0027] 查询模块,用于查询所述分类数据集的键名集,并获取所述键名集对应的评估指标集;

[0028] 计算模块,用于基于所述评估指标集计算所述键名集的权重,并评估所述权重是否达到预设权重;

[0029] 处理模块,用于若所述权重达到所述预设权重,则标记所述键名集为热数据集,并判断所述热数据集的键值集不在预设服务器的缓存内时,将所述热数据集的键值集移至预设服务器的缓存,若所述权重未达到所述预设权重,则标记所述键名集为冷数据集,并判断所述冷数据集的键值集在预设服务器的缓存内时,将所述冷数据集的键值集移出预设服务器的缓存。

[0030] 为了解决上述问题,本发明还提供一种电子设备,所述电子设备包括:

[0031] 至少一个处理器;以及,

[0032] 与所述至少一个处理器通信连接的存储器;其中,

[0033] 所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够实现上述所述的冷热数据处理方法。

[0034] 为了解决上述问题,本发明还提供一种计算机可读存储介质,包括存储数据区和存储程序区,所述存储数据区存储创建的数据,所述存储程序区存储有计算机程序,其中,所述计算机程序被处理器执行时实现上述所述的冷热数据处理方法。

[0035] 本发明实施例提供的冷热数据处理方法通过获取数据的键名,并基于所述键名对应的评估指标集计算所述键名的权重,根据键名的权重判断结果定义数据的冷热程度,并按照数据的冷热程度进行相应的操作,本发明实施例通过对冷热数据进行不同的存储处理,将热数据存储于缓存中,实现数据更高效的访问,保证了系统的服务性能,将冷数据移出缓存,减少缓存资源的浪费,降低数据存储的成本。

附图说明

[0036] 图1为本发明一实施例提供的冷热数据处理方法的流程示意图;

- [0037] 图2为本发明一实施例提供的冷热数据处理装置的模块示意图；
- [0038] 图3为本发明一实施例提供的实现冷热数据处理方法的电子设备的内部结构示意图；
- [0039] 本发明目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

- [0040] 应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。
- [0041] 本申请实施例提供一种冷热数据处理方法。所述冷热数据处理方法的执行主体包括但不限于服务端、终端等能够被配置为执行本申请实施例提供的该方法的电子设备中的至少一种。换言之，所述冷热数据处理方法可以由安装在终端设备或服务端设备的软件或硬件来执行，所述软件可以是区块链平台。所述服务端包括但不限于：单台服务器、服务器集群、云端服务器或云端服务器集群等。
- [0042] 参照图1所示，为本发明一实施例提供的冷热数据处理方法的流程示意图。在本实施例中，所述冷热数据处理方法包括：
- [0043] S1、获取原始数据集，并对所述原始数据集进行分类，得到分类数据集。
- [0044] 本发明较佳实施例中，所述原始数据集可以是某一具体应用场景下的配置系统内存在的所有业务数据的集合，可以通过与各个业务系统交互接入数据源，从而获取业务数据。其中，所述具体应用场景包括银行、保险等，如银行的配置系统的业务数据包括负债业务数据、中间业务数据、资产业务数据等。
- [0045] 进一步地，所述对所述原始数据集进行分类，得到分类数据集，包括：
- [0046] 获取所述原始数据集对应的业务场景类别；
- [0047] 根据所述业务场景类别，将所述原始数据集进行分类，得到分类数据集。
- [0048] 本发明实施例可以通过采集所述原始数据集中的业务参数，将所述业务参数进行解耦，得到所述原始数据集反映的业务场景的信息交互流程及逻辑框架信息，将所述信息交互流程及逻辑框架信息加载到Storm中进行计算，还原所述原始数据集对应的业务场景类别。
- [0049] 进一步地，本发明实施例根据所述业务场景类别，将所述原始数据集进行分类，得到分类数据集。如对负债业务数据进行业务场景还原后对应的业务场景包括存款业务场景、借款业务场景、同业业务场景等，则将所述负债业务数据划分为存款业务数据、借款业务数据、同业业务数据等。
- [0050] S2、查询所述分类数据集的键名集，并获取所述键名集对应的评估指标集。
- [0051] 本发明其中一个实施例，可以通过js脚本识别所述分类数据集的数据属性，根据所述数据属性，利用基于SQL语句的select查询命令查询所述分类数据集在数据库表中对应的键名。
- [0052] 本发明较佳实施例通过获取所述键名集在数据库实例中预设时间段内的访问日志，利用python脚本从所述访问日志中的访问信息中提取所述键名集的评估指标集。其中，所述评估指标集包括访问数据长度、访问频次、更新频次、最近访问时间访问数据长度等评估指标。本发明其中一个实施例，所述预设时间段可以为30天，对于所述预设时间段，不限于本实施例列举出的值，还可以按照实际需要进行设定。

[0053] S3、基于所述评估指标集计算所述键名集的权重。

[0054] 具体地,本发明实施例对所述访问频次、访问数据长度、最近访问时间、更新频次执行聚合运算,得到所述键名集的权重。较佳地,本发明实施例中,所述聚合运算是利用预设的权重聚合函数依次对所述键名集的评估指标集进行计算,得到所述键名集的权重。本发明其中一个实施例中,所述权重聚合函数可以为:

$$[0055] \quad w = \sum_{i=1}^n \lambda_i f_i$$

[0056] 其中,所述w表示权重,n表示评估指标集中评估指标的个数, λ 为超参数,表示每个评估指标的影响能力,f表示评估指标,在本发明实施例中所述 λ 可以为0.1,本发明其他实施例不限于上述列举出的值,还可以按照实际需要进行设定。

[0057] 进一步地,本发明实施例中所述键名集的评估指标集会随着所述键名集在数据库实例中的访问进行实时变化的,所以所述键名集的权重需要根据实时变化的评估指标集进行实时计算。但是为了节约计算损耗及实时计算带来的计算延迟,本发明其中一个实施例可以通过预设一个时间阈值,定时基于所述评估指标集计算所述键名集的权重,并根据重新计算的结果更新所述键名集的权重,便于后续对分类数据集进行对应的处理。本发明其中一个实施例,所述时间阈值可以为7天。对于所述时间阈值的设定,不限于本实施例列举出的值,还可以按照实际需要进行设定。

[0058] S4、评估所述权重是否达到预设权重。

[0059] 本发明较佳实施例中,所述预设权重(Fix Weight)可以是根据Redis的内存阈值综合设定的。本发明其中一个实施例,所述预设权重可以为0.7,当所述键名集的权重不小于0.7时,则所述键名集的权重达到所述预设权重,当所述键名集的权重小于0.7时,则所述键名集的权重未达到所述预设权重。

[0060] 本发明其中一个实施例中,若所述权重达到所述预设权重,则执行S5、标记所述键名集为热数据集。

[0061] 本发明较佳实施例中,当所述键名集的权重评估结果为达到所述预设权重,说明所述键名集在数据库实例上的访问热度高,此时将所述键名集标记为热数据集。

[0062] 本发明较佳实施例中,步骤S5之后还包括:将所述热数据集备份在区块链节点中,当存储所述热数据集的服务器出现宕机或者重启时,在确认服务器重启成功后,从所述区块链节点中获取所述热数据集,并重新移至所述服务器的缓存中。

[0063] 本发明较佳实施例中,所述缓存是指基于Redis做的一个数据缓存层。

[0064] S6、判断所述热数据集的键值集是否在预设服务器的缓存。

[0065] 本发明较佳实施例中,利用所述Redis的set命令判断所述热数据集的键值集是否在预设服务器的缓存。

[0066] 当所述热数据集的键值集在预设服务器的缓存时,执行S11,对所述热数据集的键值集不做处理。

[0067] 当所述热数据集的键值集不在预设服务器的缓存时,执行S7、将所述热数据集的键值集移至预设服务器的缓存。

[0068] 本发明较佳实施例中,从数据库中提取所述热数据集对应的键值集,并保存至所述缓存。

[0069] 进一步地,本发明另一个实施例中,若所述权重未达到预设权重,则执行S8、标记所述键名集为冷数据集。

[0070] 本发明较佳实施例中,当所述键名集的权重评估结果为未达到所述预设权重,则说明所述键名集在数据库实例上的访问热度低,此时将所述键名集标记为冷数据集。

[0071] 本发明较佳实施例中,步骤S8之后,还包括,根据预设的时间阈值定时基于所述评估指标集计算所述键名集的权重,并根据重新计算的结果更新所述键名集的权重。按照所述时间阈值对键名集的权重进行更新,其更新依据为基于更新后的访问频次、访问数据长度、最近访问时间及更新频次,执行聚合运算,得到所述键名集的权重,便于后续对分类数据集进行对应的处理。

[0072] S9、判断所述冷数据集的键值集是否在预设服务器的缓存。

[0073] 本发明较佳实施例中,利用所述Redis的set命令判断所述冷数据集的键值集是否在预设服务器的缓存。

[0074] 当所述冷数据集的键值集不在预设服务器的缓存时,执行S11,对所述冷数据集的键值集不做处理。

[0075] 当所述冷数据集的键值集在预设服务器的缓存时,执行S10、将所述冷数据集的键值集移出预设服务器的缓存。

[0076] 本发明较佳实施例中,删除所述缓存中包含的所述冷数据集对应的键值集。

[0077] 本发明实施例提供的冷热数据处理方法通过获取数据的键名,并基于所述键名对应的评估指标集计算所述键名的权重,根据键名的权重判断结果定义数据的冷热程度,并按照数据的冷热程度进行相应的操作,本发明实施例通过对冷热数据进行不同的存储处理,将热数据存储于缓存中,实现数据更高效的访问,保证了系统的服务性能,将冷数据移出缓存,减少缓存资源的浪费,降低数据存储的成本。

[0078] 如图2所示,是本发明冷热数据处理装置的模块示意图。

[0079] 本发明所述冷热数据处理装置100可以安装于电子设备中。根据实现的功能,所述冷热数据处理装置可以包括获取模块101、查询模块102、计算模块103及处理模块104。本发明所述模块也可以称之为单元,是指一种能够被电子设备处理器所执行,并且能够完成固定功能的一系列计算机程序段,其存储在电子设备的存储器中。

[0080] 在本实施例中,关于各模块/单元的功能如下:

[0081] 所述获取模块101,用于获取原始数据集,对所述原始数据集进行分类,得到分类数据集;

[0082] 本发明较佳实施例中,所述原始数据集可以是某一具体应用场景下的配置系统内存在的所有业务数据的集合,可以通过与各个业务系统交互接入数据源,从而获取业务数据。其中,所述具体应用场景包括银行、保险等,如银行的配置系统的业务数据包括负债业务数据、中间业务数据、资产业务数据等。

[0083] 所述获取模块101,用于对所述原始数据集进行分类,得到分类数据集,包括获取所述原始数据集对应的业务场景类别,根据所述业务场景类别,将所述原始数据集进行分类,得到分类数据集。

[0084] 本发明实施例可以通过采集所述原始数据集中的业务参数,将所述业务参数进行解耦,得到所述原始数据集反映的业务场景的信息交互流程及逻辑框架信息,将所述信息

交互流程及逻辑框架信息加载到Storm中进行计算,还原所述原始数据集对应的业务场景类别。

[0085] 进一步地,本发明实施例根据所述业务场景类别,将所述原始数据集进行分类,得到分类数据集。如对负债业务数据进行业务场景还原后对应的业务场景包括存款业务场景、借款业务场景、同业业务场景等,则将所述负债业务数据划分为存款业务数据、借款业务数据、同业业务数据等。

[0086] 所述查询模块102,用于询所述分类数据集的键名集,并获取所述键名集对应的评估指标集;

[0087] 所述查询模块102,可以通过js脚本识别所述分类数据集的数据属性,根据所述数据属性,利用基于SQL语句的select查询命令查询所述分类数据集在数据库表中对应的键名。

[0088] 本发明较佳实施例通过获取所述键名集在数据库实例中预设时间段内的访问日志,利用python脚本从所述访问日志中的访问信息中提取所述键名集的评估指标集。其中,所述评估指标集包括访问数据长度、访问频次、更新频次、最近访问时间访问数据长度等评估指标。本发明其中一个实施例,所述预设时间段可以为30天,对于所述预设时间段,不限于本实施例列举出的值,还可以按照实际需要进行设定。

[0089] 所述计算模块103,用于基于所述评估指标集计算所述键名集的权重,并评估所述权重是否达到预设权重;

[0090] 所述计算模块103,对所述访问频次、访问数据长度、最近访问时间、更新频次执行聚合运算,得到所述键名集的权重。较佳地,本发明实施例中,所述聚合运算是利用预设的权重聚合函数依次对所述键名集的评估指标集进行计算,得到所述键名集的权重。所述计算模块103利用下述方法,计算所述键名集的权重:

$$[0091] \quad w = \sum_{i=1}^n \lambda_i f_i$$

[0092] 其中,所述w表示权重,n表示评估指标集中评估指标的个数, λ 为超参数,表示每个评估指标的影响能力,f表示评估指标,在本发明实施例中所述 λ 可以为0.1,本发明其他实施例不限于上述列举出的值,还可以按照实际需要进行设定。

[0093] 进一步地,本发明实施例中所述键名集的评估指标集会随着所述键名集在数据库实例中的访问进行实时变化的,所以所述键名集的权重需要根据实时变化的评估指标集进行实时计算。但是为了节约计算损耗及实时计算带来的计算延迟,本发明其中一个实施例可以通过预设一个时间阈值,定时基于所述评估指标集计算所述键名集的权重,并根据重新计算的结果更新所述键名集的权重,便于后续对分类数据集进行对应的处理。本发明其中一个实施例,所述时间阈值可以为7天。对于所述时间阈值的设定,不限于本实施例列举出的值,还可以按照实际需要进行设定。

[0094] 本发明较佳实施例中,所述预设权重(Fix Weight)可以是根据Redis的内存阈值综合设定的。本发明其中一个实施例,所述预设权重可以为0.7,当所述键名集的权重不小于0.7时,则所述键名集的权重达到所述预设权重,当所述键名集的权重小于0.7时,则所述键名集的权重未达到所述预设权重。

[0095] 所述处理模块104,用于若所述权重达到所述预设权重,则标记所述键名集为热数

据集,并判断所述热数据集的键值集不在预设服务器的缓存内时,将所述热数据集的键值集移至预设服务器的缓存,若所述权重未达到所述预设权重,则标记所述键名集为冷数据集,并判断所述冷数据集的键值集在预设服务器的缓存内时,将所述冷数据集的键值集移出预设服务器的缓存。

[0096] 所述处理模块104,用于若所述权重达到所述预设权重,则标记所述键名集为热数据集。

[0097] 本发明较佳实施例中,当所述键名集的权重评估结果为达到所述预设权重,说明所述键名集在数据库实例上的访问热度高,此时将所述键名集标记为热数据集。

[0098] 本发明较佳实施例中,所述标记所述键名集为热数据集之后还包括:将所述热数据集备份在区块链节点中,当存储所述热数据集的服务器出现宕机或者重启时,在确认服务器重启成功后,从所述区块链节点中获取所述热数据集,并重新移至所述服务器的缓存中。

[0099] 本发明较佳实施例中,所述缓存是指基于Redis做的一个数据缓存层。

[0100] 判断所述热数据集的键值集是否在预设服务器的缓存,本发明较佳实施例中,利用所述Redis的set命令判断所述热数据集的键值集是否在预设服务器的缓存。

[0101] 当所述热数据集的键值集在预设服务器的缓存时,对所述热数据集的键值集不做处理。

[0102] 当所述热数据集的键值集不在预设服务器的缓存时,将所述热数据集的键值集移至预设服务器的缓存,本发明较佳实施例中,从数据库中提取所述热数据集对应的键值集,并保存至所述缓存。

[0103] 所述处理模块104,还用于若所述权重未达到预设权重,则标记所述键名集为冷数据集。

[0104] 本发明较佳实施例中,当所述键名集的权重评估结果为未达到所述预设权重,则说明所述键名集在数据库实例上的访问热度低,此时将所述键名集标记为冷数据集。

[0105] 本发明较佳实施例中,所述标记所述键名集为冷数据集之后,还包括,根据预设的时间阈值定时基于所述评估指标集计算所述键名集的权重,并根据重新计算的结果更新所述键名集的权重。按照所述时间阈值对键名集的权重进行更新,其更新依据为基于更新后的访问频次、访问数据长度、最近访问时间及更新频次,执行聚合运算,得到所述键名集的权重,便于后续对分类数据集进行对应的处理。

[0106] 判断所述冷数据集的键值集是否在预设服务器的缓存,本发明较佳实施例中,利用所述Redis的set命令判断所述冷数据集的键值集是否在预设服务器的缓存。

[0107] 当所述冷数据集的键值集不在预设服务器的缓存时,对所述冷数据集的键值集不做处理。

[0108] 当所述冷数据集的键值集在预设服务器的缓存时,将所述冷数据集的键值集移出预设服务器的缓存,本发明较佳实施例中,删除所述缓存中包含的所述冷数据集对应的键值集。

[0109] 如图3所示,是本发明实现冷热数据处理方法的电子设备的结构示意图。

[0110] 所述电子设备1可以包括处理器10、存储器11和总线,还可以包括存储在所述存储器11中并可在所述处理器10上运行的计算机程序,如冷热数据处理程序12。

[0111] 其中,所述存储器11至少包括一种类型的计算机可读存储介质,所述计算机可读存储介质包括闪存、移动硬盘、多媒体卡、卡型存储器(例如:SD或DX存储器等)、磁性存储器、磁盘、光盘等。所述存储器11在一些实施例中可以是电子设备1的内部存储单元,例如该电子设备1的移动硬盘。所述存储器11在另一些实施例中也可以是电子设备1的外部存储设备,例如电子设备1上配备的插接式移动硬盘、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11还可以既包括电子设备1的内部存储单元也包括外部存储设备。所述存储器11不仅可以用于存储安装于电子设备1的应用软件及各类数据,例如冷热数据处理程序12的代码等,还可以用于暂时地存储已经输出或者将要输出的数据。

[0112] 所述处理器10在一些实施例中可以由集成电路组成,例如可以由单个封装的集成电路所组成,也可以是由多个相同功能或不同功能封装的集成电路所组成,包括一个或者多个中央处理器(Central Processing unit,CPU)、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。所述处理器10是所述电子设备的控制核心(Control Unit),利用各种接口和线路连接整个电子设备的各个部件,通过运行或执行存储在所述存储器11内的程序或者模块(例如执行冷热数据处理程序等),以及调用存储在所述存储器11内的数据,以执行电子设备1的各种功能和处理数据。

[0113] 所述总线可以是外设部件互连标准(peripheral component interconnect,简称PCI)总线或扩展工业标准结构(extended industry standard architecture,简称EISA)总线等。该总线可以分为地址总线、数据总线、控制总线等。所述总线被设置为实现所述存储器11以及至少一个处理器10等之间的连接通信。

[0114] 图3仅示出了具有部件的电子设备,本领域技术人员可以理解的是,图3示出的结构并不构成对所述电子设备1的限定,可以包括比图示更少或者更多的部件,或者组合某些部件,或者不同的部件布置。

[0115] 例如,尽管未示出,所述电子设备1还可以包括给各个部件供电的电源(比如电池),优选地,电源可以通过电源管理装置与所述至少一个处理器10逻辑相连,从而通过电源管理装置实现充电管理、放电管理、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述电子设备1还可以包括多种传感器、蓝牙模块、Wi-Fi模块等,在此不再赘述。

[0116] 进一步地,所述电子设备1还可以包括网络接口,可选地,所述网络接口可以包括有线接口和/或无线接口(如WI-FI接口、蓝牙接口等),通常用于在该电子设备1与其他电子设备之间建立通信连接。

[0117] 可选地,该电子设备1还可以包括用户接口,用户接口可以是显示器(Display)、输入单元(比如键盘(Keyboard)),可选地,用户接口还可以是标准的有线接口、无线接口。可选地,在一些实施例中,显示器可以是LED显示器、液晶显示器、触控式液晶显示器以及OLED(Organic Light-Emitting Diode,有机发光二极管)触摸器等。其中,显示器也可以适当的称为显示屏或显示单元,用于显示在电子设备1中处理的信息以及用于显示可视化的用户界面。

[0118] 应该了解,所述实施例仅为说明之用,在专利申请范围上并不受此结构的限制。

[0119] 所述电子设备1中的所述存储器11存储的冷热数据处理程序12是多个指令的组合,在所述处理器10中运行时,可以实现:

[0120] 获取原始数据集,对所述原始数据集进行分类,得到分类数据集;

[0121] 查询所述分类数据集的键名集,并获取所述键名集对应的评估指标集;

[0122] 基于所述评估指标集计算所述键名集的权重,并评估所述权重是否达到预设权重;

[0123] 若所述权重达到所述预设权重,则标记所述键名集为热数据集,并将所述热数据集的键值集移至预设服务器的缓存;

[0124] 若所述权重未达到所述预设权重,则标记所述键名集为冷数据集,并将所述冷数据集的键值集移出预设服务器的缓存。

[0125] 进一步地,所述电子设备1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,可以存储在一个计算机可读取存储介质中。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)。

[0126] 进一步地,所述计算机可用存储介质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等;存储数据区可存储根据区块链节点的使用所创建的数据等。

[0127] 在本发明所提供的几个实施例中,应该理解到,所揭露的设备,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,所述模块的划分,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。

[0128] 所述作为分离部件说明的模块可以是或者也可以不是物理上分开的,作为模块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模块来实现本实施例方案的目的。

[0129] 另外,在本发明各个实施例中的各功能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,也可以采用硬件加软件功能模块的形式实现。

[0130] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。

[0131] 因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附关联图表记视为限制所涉及的权利要求。

[0132] 本发明所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain),本质上是一个去中心化的数据库,是一串使用密码学方法相关联产生的数据块,每一个数据块中包含了一批网络交易的信息,用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

[0133] 此外,显然“包括”一词不排除其他单元或步骤,单数不排除复数。系统权利要求中

陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第二等词语用来表示名称,而并不表示任何特定的顺序。

[0134] 最后应说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管参照较佳实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或等同替换,而不脱离本发明技术方案的精神和范围。

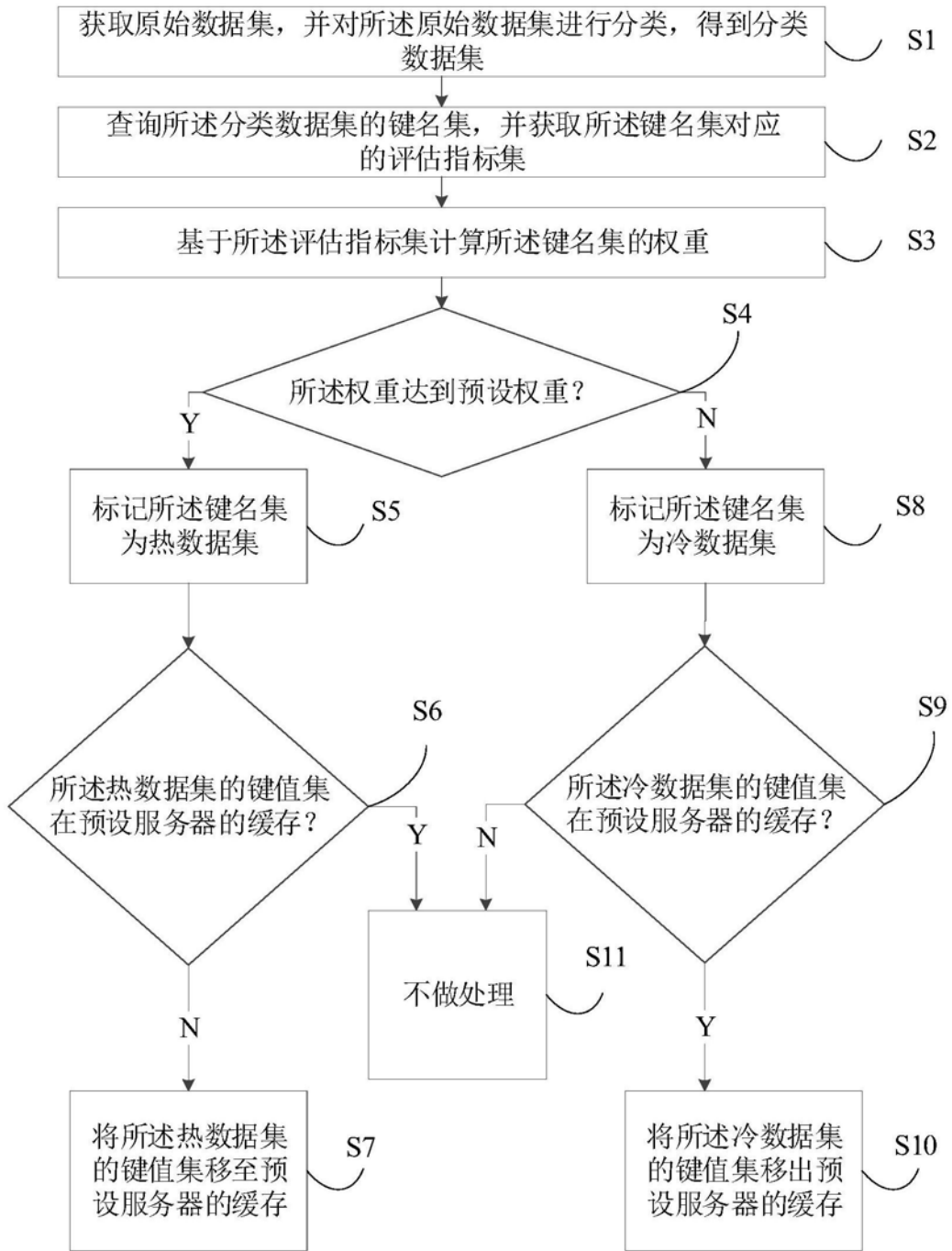


图1



图2

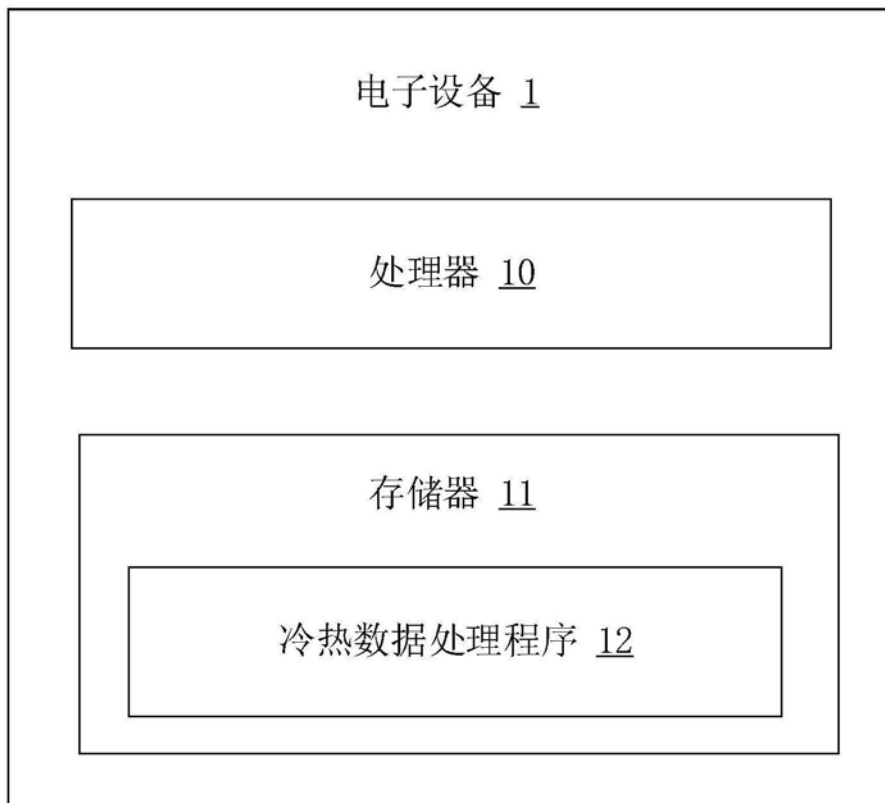


图3