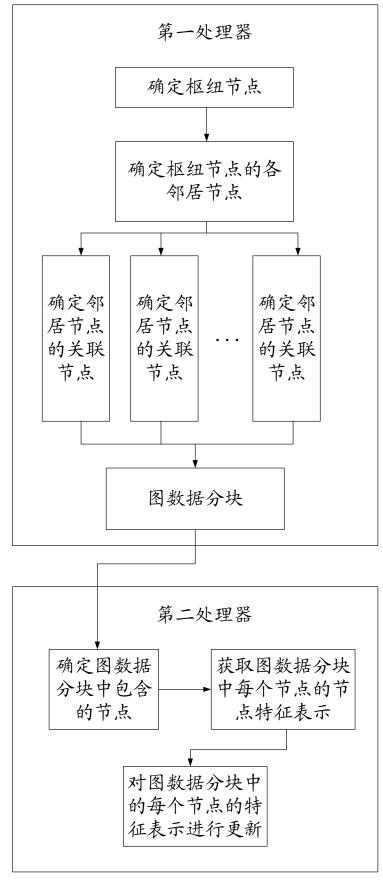
## 说明书摘要

本说明书公开了一种数据处理系统、方法、设备及存储介质,可以从原始图数据中包含的每个节点中筛选出度数高于预设阈值的节点,作为用于连接各社区的枢纽节点,进而可以将枢纽节点的每个邻居节点作为起始节点,通过访问与该邻居节点之间存在连接关系的其他节点,作为该邻居节点的关联节点,从而可以将该邻居节点和该邻居节点的关联节点划分出来,作为原始图数据的一个图数据分块,并针对每个图数据分块进行节点更新处理,进而可以避免由于直接使用原始图数据中包含的全部节点对应的特征矩阵进行计算而产生的冗余计算,以提升数据处理效率。

5

# 摘要附图



1

### 权利要求书

1、一种数据处理系统, 其特征在于, 所述数据处理系统包括: 第一处理器、 第二处理器;

所述第一处理器用于针对原始图数据中包含的每个节点,判断该节点的度数是否超过预设阈值,若是,则确定该节点为枢纽节点,并针对所述枢纽节点的每个邻居节点,通过多轮节点查询,确定所述原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点,并根据该邻居节点以及该邻居节点的关联节点,确定所述原始图数据分块;

所述第二处理器用于针对所述图数据分块中包含的每个节点,根据该节点 在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示 进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后 特征表示,对所述原始图数据进行更新处理。

10

15

20

25

2、如权利要求 1 所述的数据处理系统, 其特征在于, 所述第一处理器用于针对所述枢纽节点的每个邻居节点, 将该邻居节点作为起始节点, 通过多轮节点查询, 确定该邻居节点的关联节点; 其中,

针对每轮节点查询,确定该轮节点查询中的各目标节点,判断所述各目标节点的各邻居节点中是否存在未访问节点,若是,则将所述未访问节点作为所述起始节点的关联节点,并将所述为访问节点设置为已访问节点,以及,将所述未访问节点作为下一轮节点查询的目标节点,所述各目标节点是将所述起始节点迭代至上一轮后得到的;

在确定满足预设的第一终止条件后,得到所述起始节点的各关联节点。

3、如权利要求 2 所述的数据处理系统,其特征在于,所述第一处理器用于针对每轮节点查询中的每个目标节点,判断在预设的全局已访问节点集中是否包含该目标节点,若否,则确定该目标节点为所述起始节点对应的已访问邻居节点,并将该目标节点添加到所述全局已访问节点集中,若是,则确定所述起始节点不存在关联节点,并将所述起始节点对应的已访问邻居节点从所述全

局已访问节点集中移除。

15

20

4、如权利要求 1 所述的数据处理系统, 其特征在于, 所述第二处理器包括: 至少一个处理单元;

各处理单元中的至少部分处理单元用于针对所述图数据分块中包含的每个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示;以及

其他处理单元用于根据该节点的各子聚合特征表示,对该节点的节点特征 表示进行更新,得到该节点的更新后特征表示。

5、如权利要求 4 所述的数据处理系统, 其特征在于, 所述数据处理系统 10 还包括: 任务调度器;

所述任务调度器用于生成针对所述图数据分块的第一数据处理任务,并根据所述各处理单元的负载值,将所述第一数据处理任务分配给所述处理单元, 所述第一数据处理任务用于将部分节点的节点特征表示进行聚合;以及

所述任务调度器用于生成针对所述图数据分块的第二数据处理任务,并根据所述各处理单元的负载值,将所述第二数据处理任务分配给所述处理单元,所述第二数据处理任务用于对所述图数据分块中的至少一个节点的节点特征表示进行更新。

- 6、如权利要求 5 所述的数据处理系统, 其特征在于, 所述任务调取器用于针对每个处理单元, 判断该处理单元的负载值是否超过预设的第一负载阈值, 若是, 则确定在各处理单元中与该处理单元的位置相邻的各处理单元, 作为各候选处理单元, 判断所述各候选处理单元中负载值最低的候选处理单元, 作为目标处理单元, 将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给所述目标处理单元。
- 7、如权利要求 5 所述的数据处理系统, 其特征在于, 所述任务调取器用 25 于针对所述各处理单元中的每个处理单元, 判断该处理单元的负载值与其他处 理单元的负载值之间的差值是否超过预设的第二负载阈值, 若是, 则从各处理

单元中选择负载值最小的处理单元作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给所述目标处理单元。

- 8、如权利要求 5 所述的数据处理系统, 其特征在于, 所述任务调取器用于针对所述各处理单元中的每个处理单元, 判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值, 若是, 则对该处理单元正在处理的第一数据处理任务或第二数据处理任务进行拆分, 得到各子数据处理任务, 将所述各子数据处理任务分配给各处理单元中选择负载值最小的处理单元。
- 9、如权利要求 1 所述的数据处理系统, 其特征在于, 所述第二处理器用 10 于针对所述图数据分块中包含的每个节点, 通过多轮迭代, 确定该节点的更新 后特征表示; 其中,

针对每轮迭代,确定该节点的待更新节点特征表示,根据上一轮迭代中确定出的该节点对应的聚合特征表示,对所述待更新节点特征表示进行更新,得到该节点在该轮迭代中对应的更新后特征表示,并将该节点在该轮迭代中对应的更新后特征表示,作为下一轮迭代中的待更新节点特征表示,所述聚合特征表示用于表征该节点在所述图数据分块中的各邻居节点的聚合结果,所述待更新节点特征表示是将该节点的节点特征表示迭代至上一轮后得到的;

15

20

25

当确定满足预设的第二终止条件时,得到该节点的更新后特征表示。

10、一种数据处理方法,其特征在于,所述数据处理方法应用于数据处理 系统中,所述数据处理系统包括:第一处理器、第二处理器,所述方法包括:

所述第一处理器针对原始图数据中包含的每个节点,判断该节点的度数是 否超过预设阈值;

若是,则确定该节点为枢纽节点,并针对所述枢纽节点的每个邻居节点,通过多轮节点查询,确定所述原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点;

根据该邻居节点以及该邻居节点的关联节点,确定所述原始图数据的图数

据分块,以通过所述第二处理器针对所述图数据分块中包含的每个节点,根据该节点在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后特征表示,对所述原始图数据进行更新处理。

11、如权利要求 10 所述的方法, 其特征在于, 针对所述枢纽节点的每个邻居节点, 通过多轮节点查询, 确定所述原始图数据中与该邻居节点存在连接关系的其他节点, 作为该邻居节点的关联节点, 具体包括:

5

10

15

针对所述枢纽节点的每个邻居节点,将该邻居节点作为起始节点,通过多轮节点查询,确定该邻居节点的关联节点;其中,

针对每轮节点查询,确定该轮节点查询中的各目标节点,判断所述各目标节点是否存在邻居节点,若是,则将所述各目标节点作为所述起始节点的关联节点,以及,将所述各目标节点的邻居节点作为下一轮节点查询的目标节点,所述各目标节点是将所述起始节点迭代至上一轮后得到的;

在确定满足预设的第一终止条件后,得到所述起始节点的各关联节点。

12、如权利要求 11 所述的方法, 其特征在于, 判断所述各目标节点是否存在邻居节点, 具体包括:

针对每轮节点查询中的每个目标节点,判断在预设的全局已访问节点集中是否包含该目标节点;

若是,则确定所述起始节点不存在关联节点,并将所述起始节点对应的已 20 访问邻居节点从所述全局已访问节点集中移除;

若否,则将确定该目标节点为所述起始节点对应已访问邻居节点,并将该目标节点添加到所述全局已访问节点集中,并判断所述各目标节点是否存在邻居节点。

13、如权利要求 10 所述的方法, 其特征在于, 所述第二处理器包括: 至少 25 一个处理单元;

根据该节点在所述图数据分块中的各邻居节点的节点特征表示, 对该节点

的节点特征表示进行更新,得到该节点的更新后特征表示,具体包括:

通过各处理单元中的至少部分处理单元针对所述图数据分块中包含的每个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示;以及

5 通过其他处理单元根据该节点的各子聚合特征表示,对该节点的节点特征 表示进行更新,得到该节点的更新后特征表示。

14、如权利要求 13 所述的方法, 其特征在于, 所述数据处理系统还包括: 任务调度器;

通过各处理单元中的至少部分处理单元针对所述图数据分块中包含的每 10 个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进 行聚合,得到该节点的子聚合特征表示,具体包括:

通过所述任务调度器生成针对所述图数据分块的第一数据处理任务,并根据所述各处理单元的负载值,将所述第一数据处理任务分配给所述处理单元,以使所述处理单元针对所述图数据分块中包含的每个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示:

15

20

25

通过其他处理单元根据该节点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,具体包括:

通过所述任务调度器生成针对所述图数据分块的第二数据处理任务,并根据所述各处理单元的负载值,将所述第二数据处理任务分配给所述处理单元,以使所述处理单元根据该节点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示。

15、如权利要求 14 所述的方法, 其特征在于, 所述方法还包括:

通过所述任务调度器,针对每个处理单元,判断该处理单元的负载值是否 超过预设的第一负载阈值,若是,则确定在各处理单元中与该处理单元的位置 相邻的各处理单元,作为各候选处理单元,判断所述各候选处理单元中负载值 最低的候选处理单元,作为目标处理单元,将该处理单元正在处理的第一数据 处理任务或第二数据处理任务分配给所述目标处理单元。

16、如权利要求 14 所述的方法, 其特征在于, 所述方法还包括:

通过所述任务调取器,针对所述各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则从各处理单元中选择负载值最小的处理单元作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给所述目标处理单元。

17、如权利要求 14 所述的方法, 其特征在于, 所述方法还包括:

10

15

20

25

通过所述任务调取器,针对所述各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则对该处理单元正在处理的第一数据处理任务或第二数据处理任务进行拆分,得到各子数据处理任务,将所述各子数据处理任务分配给各处理单元中选择负载值最小的处理单元。

18、如权利要求 10 所述的方法, 其特征在于, 针对所述图数据分块中包含的每个节点, 根据该节点在所述图数据分块中的各邻居节点的节点特征表示, 对该节点的节点特征表示进行更新, 得到该节点的更新后特征表示, 具体包括:

针对所述图数据分块中包含的每个节点,通过多轮迭代,确定该节点的更新后特征表示;其中,

针对每轮迭代,确定该节点的待更新节点特征表示,根据上一轮迭代中确定出的该节点对应的聚合特征表示,对所述待更新节点特征表示进行更新,得到该节点在该轮迭代中对应的更新后特征表示,并将该节点在该轮迭代中对应的更新后特征表示,作为下一轮迭代中的待更新节点特征表示,所述聚合特征表示用于表征该节点在所述图数据分块中的各邻居节点的聚合结果,所述待更新节点特征表示是将该节点的节点特征表示迭代至上一轮后得到的;

当确定满足预设的第二终止条件时,得到该节点的更新后特征表示。

- 19、一种计算机可读存储介质,其特征在于,所述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实现上述权利要求 10~18 任一项所述的方法。
- 20、一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器 5 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上述权利 要求 10~18 任一项所述的方法。

### 说明书

一种数据处理系统、方法、设备及存储介质

### 技术领域

本说明书涉及图计算技术领域,尤其涉及一种数据处理系统、方法、设 5 备及存储介质。

### 背景技术

目前,在通过图卷积神经网络对图数据进行处理的过程中,由于需要处理 的图数据具有尺寸大、数据分布较为稀疏等特点,使得在使用图数据对应的特 征矩阵进行计算时,存在较多的无效计算操作,进而使得处理的效率较低。

因此,如何提升数据处理效率,则是一个亟待解决的问题。

#### 发明内容

10

15

20

本说明书提供一种数据处理方法、装置、设备及存储介质,以部分的解决现有技术存在的上述问题。

本说明书采用下述技术方案:

本说明书提供了一种数据处理系统,所述数据处理系统包括:第一处理器、 第二处理器;

所述第一处理器用于针对原始图数据中包含的每个节点,判断该节点的度数是否超过预设阈值,若是,则确定该节点为枢纽节点,并针对所述枢纽节点的每个邻居节点,通过多轮节点查询,确定所述原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点,并根据该邻居节点以及该邻居节点的关联节点,确定所述原始图数据分块;

所述第二处理器用于针对所述图数据分块中包含的每个节点,根据该节点在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后

特征表示,对所述原始图数据进行更新处理。

10

15

20

25

可选地,所述第一处理器用于针对所述枢纽节点的每个邻居节点,将该邻居节点作为起始节点,通过多轮节点查询,确定该邻居节点的关联节点;其中,

针对每轮节点查询,确定该轮节点查询中的各目标节点,判断所述各目标节点是否存在邻居节点,若是,则将所述各目标节点作为所述起始节点的关联节点,以及,将所述各目标节点的邻居节点作为下一轮节点查询的目标节点,所述各目标节点是将所述起始节点迭代至上一轮后得到的;

在确定满足预设的第一终止条件后,得到所述起始节点的各关联节点。

可选地,所述第一处理器用于针对每轮节点查询中的每个目标节点,判断 在预设的全局已访问节点集中是否包含该目标节点,若否,则确定该目标节点 为所述起始节点对应的已访问邻居节点,并将该目标节点添加到所述全局已访问节点集中,若是,则确定所述起始节点不存在关联节点,并将所述起始节点 对应的已访问邻居节点从所述全局已访问节点集中移除。

可选地,所述第二处理器包括: 至少一个处理单元;

各处理单元中的至少部分处理单元用于针对所述图数据分块中包含的每个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示;以及

其他处理单元用于根据该节点的各子聚合特征表示,对该节点的节点特征 表示进行更新,得到该节点的更新后特征表示。

可选地, 所述数据处理系统还包括: 任务调度器;

所述任务调度器用于生成针对所述图数据分块的第一数据处理任务,并根据所述各处理单元的负载值,将所述第一数据处理任务分配给所述处理单元,所述第一数据处理任务用于将部分节点的节点特征表示进行聚合;以及

所述任务调度器用于生成针对所述图数据分块的第二数据处理任务,并根据所述各处理单元的负载值,将所述第二数据处理任务分配给所述处理单元, 所述第二数据处理任务用于对所述图数据分块中的至少一个节点的节点特征 表示进行更新。

5

10

15

20

25

可选地,所述任务调取器用于针对每个处理单元,判断该处理单元的负载值是否超过预设的第一负载阈值,若是,则确定在各处理单元中与该处理单元的位置相邻的各处理单元,作为各候选处理单元,判断所述各候选处理单元中负载值最低的候选处理单元,作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给所述目标处理单元。

可选地,所述任务调取器用于针对所述各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则从各处理单元中选择负载值最小的处理单元作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给所述目标处理单元。

可选地,所述任务调取器用于针对所述各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则对该处理单元正在处理的第一数据处理任务或第二数据处理任务进行拆分,得到各子数据处理任务,将所述各子数据处理任务分配给各处理单元中选择负载值最小的处理单元。

可选地,所述第二处理器用于针对所述图数据分块中包含的每个节点,通过多轮迭代,确定该节点的更新后特征表示;其中,

针对每轮迭代,确定该节点的待更新节点特征表示,根据上一轮迭代中确定出的该节点对应的聚合特征表示,对所述待更新节点特征表示进行更新,得到该节点在该轮迭代中对应的更新后特征表示,并将该节点在该轮迭代中对应的更新后特征表示,作为下一轮迭代中的待更新节点特征表示,所述聚合特征表示用于表征该节点在所述图数据分块中的各邻居节点的聚合结果,所述待更新节点特征表示是将该节点的节点特征表示迭代至上一轮后得到的;

当确定满足预设的第二终止条件时,得到该节点的更新后特征表示。

本说明书提供了一种数据处理方法, 所述数据处理方法应用于数据处理系

统中,所述数据处理系统包括:第一处理器、第二处理器,所述方法包括:

所述第一处理器针对原始图数据中包含的每个节点,判断该节点的度数是 否超过预设阈值;

若是,则确定该节点为枢纽节点,并针对所述枢纽节点的每个邻居节点,通过多轮节点查询,确定所述原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点;

根据该邻居节点以及该邻居节点的关联节点,确定所述原始图数据的图数据分块,以通过所述第二处理器针对所述图数据分块中包含的每个节点,根据该节点在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后特征表示,对所述原始图数据进行更新处理。

10

20

可选地,针对所述枢纽节点的每个邻居节点,通过多轮节点查询,确定所述原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点,具体包括:

15 针对所述枢纽节点的每个邻居节点,将该邻居节点作为起始节点,通过多 轮节点查询,确定该邻居节点的关联节点;其中,

针对每轮节点查询,确定该轮节点查询中的各目标节点,判断所述各目标节点是否存在邻居节点,若是,则将所述各目标节点作为所述起始节点的关联节点,以及,将所述各目标节点的邻居节点作为下一轮节点查询的目标节点,所述各目标节点是将所述起始节点迭代至上一轮后得到的;

在确定满足预设的第一终止条件后,得到所述起始节点的各关联节点。可选地,判断所述各目标节点是否存在邻居节点,具体包括:

针对每轮节点查询中的每个目标节点,判断在预设的全局已访问节点集中是否包含该目标节点;

25 若是,则确定所述起始节点不存在关联节点,并将所述起始节点对应的已 访问邻居节点从所述全局已访问节点集中移除; 若否,则将确定该目标节点为所述起始节点对应已访问邻居节点,并将该目标节点添加到所述全局已访问节点集中,并判断所述各目标节点是否存在邻居节点。

可选地,所述第二处理器包括: 至少一个处理单元;

5 根据该节点在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,具体包括:

通过各处理单元中的至少部分处理单元针对所述图数据分块中包含的每个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示;以及

通过其他处理单元根据该节点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示。

可选地, 所述数据处理系统还包括: 任务调度器;

10

15

20

25

通过各处理单元中的至少部分处理单元针对所述图数据分块中包含的每个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示,具体包括:

通过所述任务调度器生成针对所述图数据分块的第一数据处理任务,并根据所述各处理单元的负载值,将所述第一数据处理任务分配给所述处理单元,以使所述处理单元针对所述图数据分块中包含的每个节点,将该节点在所述图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示;

通过其他处理单元根据该节点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,具体包括:

通过所述任务调度器生成针对所述图数据分块的第二数据处理任务,并根据所述各处理单元的负载值,将所述第二数据处理任务分配给所述处理单元,以使所述处理单元根据该节点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示。

可选地,所述方法还包括:

通过所述任务调度器,针对每个处理单元,判断该处理单元的负载值是否超过预设的第一负载阈值,若是,则确定在各处理单元中与该处理单元的位置相邻的各处理单元,作为各候选处理单元,判断所述各候选处理单元中负载值最低的候选处理单元,作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给所述目标处理单元。

可选地,所述方法还包括:

5

10

通过所述任务调取器,针对所述各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则从各处理单元中选择负载值最小的处理单元作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给所述目标处理单元。

可选地,所述方法还包括:

通过所述任务调取器,针对所述各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则对该处理单元正在处理的第一数据处理任务或第二数据处理任务进行拆分,得到各子数据处理任务,将所述各子数据处理任务分配给各处理单元中选择负载值最小的处理单元。

可选地,针对所述图数据分块中包含的每个节点,根据该节点在所述图数 20 据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新, 得到该节点的更新后特征表示,具体包括:

针对所述图数据分块中包含的每个节点,通过多轮迭代,确定该节点的更新后特征表示;其中,

针对每轮迭代,确定该节点的待更新节点特征表示,根据上一轮迭代中确 25 定出的该节点对应的聚合特征表示,对所述待更新节点特征表示进行更新,得 到该节点在该轮迭代中对应的更新后特征表示,并将该节点在该轮迭代中对应 的更新后特征表示,作为下一轮迭代中的待更新节点特征表示,所述聚合特征表示用于表征该节点在所述图数据分块中的各邻居节点的聚合结果,所述待更新节点特征表示是将该节点的节点特征表示迭代至上一轮后得到的;

当确定满足预设的第二终止条件时,得到该节点的更新后特征表示。

本说明书提供了一种计算机可读存储介质,所述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实现上述数据处理方法。

本说明书提供了一种电子设备,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上述数据处理方法。

本说明书采用的上述至少一个技术方案能够达到以下有益效果:

在本说明书提供的数据处理方法,第一处理器针对原始图数据中包含的每个节点,判断该节点的度数是否超过预设阈值,若是,则确定该节点为枢纽节点,并针对枢纽节点的每个邻居节点,确定原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点,根据该邻居节点以及该邻居节点的关联节点,确定原始图数据的图数据分块,以通过第二处理器针对图数据分块中包含的每个节点,根据该节点在图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后特征表示,对原始图数据进行更新处理。

从上述方法中可以看出,可以从原始图数据中包含的每个节点中筛选出度数高于预设阈值的节点,作为用于连接各社区的枢纽节点,进而可以将枢纽节点的每个邻居节点作为起始节点,通过访问与该邻居节点之间存在连接关系的其他节点,作为该邻居节点的关联节点,从而可以将该邻居节点和该邻居节点的关联节点划分出来,作为原始图数据的一个图数据分块,并针对每个图数据分块进行节点更新处理,进而可以避免由于直接使用原始图数据中包含的全部节点对应的特征矩阵进行计算而产生的冗余计算,以提升数据处理效率。

### 附图说明

5

10

15

20

25

此处所说明的附图用来提供对本说明书的进一步理解,构成本说明书的一部分,本说明书的示意性实施例及其说明用于解释本说明书,并不构成对本说明书的不当限定。在附图中:

- 图 1 为本说明书中提供的一种数据处理系统的示意图;
- 图 2 为本说明书中提供的一种数据处理方法的流程示意图;
- 图 3 为本说明书提供的一种对应于图 2 的电子设备的示意图。

#### 具体实施方式

5

10

15

20

25

为使本说明书的目的、技术方案和优点更加清楚,下面将结合本说明书具体实施例及相应的附图对本说明书技术方案进行清楚、完整地描述。显然,所描述的实施例仅是本说明书一部分实施例,而不是全部的实施例。基于本说明书中的实施例,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,都属于本说明书保护的范围。

以下结合附图,详细说明本说明书各实施例提供的技术方案。

本说明书中提供了一种数据处理系统,如图1所示:

图1为本说明书中提供的一种数据处理系统的示意图。

从图 1 中可以看出,数据处理系统包括:第一处理器、第二处理器。

其中,第一处理器用于针对原始图数据中包含的每个节点,判断该节点的 度数是否超过预设阈值,若是,则确定该节点为枢纽节点,并针对枢纽节点的 每个邻居节点,通过多轮节点查询,确定原始图数据中与该邻居节点存在连接 关系的其他节点,作为该邻居节点的关联节点,并根据该邻居节点以及该邻居 节点的关联节点,确定原始图数据的图数据分块。

上述的节点的度数可以是指该节点的入度和出度的和,节点入度表示图数据中进入该节点的边的条数,节点的出度是指从该节点出发的边的条数。

具体地,第一处理器可以针对枢纽节点的每个邻居节点,将该邻居节点作 为起始节点,通过多轮节点查询,确定原始图数据中与该邻居节点存在连接关 系的其他节点,作为该邻居节点的关联节点,其中,针对每轮节点查询,确定 该轮节点查询中的各目标节点,判断各目标节点的各邻居节点中是否存在未访问节点,若是,则将各目标节点的各邻居节点中的未访问节点作为起始节点的关联节点,并将各目标节点的各邻居节点中的未访问节点设为已访问节点,以及,将各目标节点的各邻居节点中的未访问节点作为下一轮节点查询的目标节点,各目标节点是将起始节点迭代至上一轮后得到的,在确定满足预设的第一终止条件后,得到起始节点的各关联节点。

其中,这里的第一终止条件可以根据实际需求设置,例如:该轮节点查询中的各目标节点的各邻居节点均为已访问节点。再例如:已经访问的节点数超过预设阈值。

需要说明的是,第一处理器查询每个目标节点邻居节点的过程,可以通过 多个线程,并行执行。

10

15

20

25

另外,从上述内容中可以看出,第一处理器在以枢纽节点的每个邻居节点作为起始节点,通过多轮节点查询,以确定每个邻居节点的关联节点的过程中,可能会出现一个节点被多次重复访问的情况出现,即,在当前起始节点的多轮节点查询过程中未访问的节点可能在其他起始节点的多轮节点查询过程中已访问。

基于此,第一处理器在判断各目标节点是否存在邻居节点之前,可以针对每轮节点查询中的每个目标节点,判断在预设的全局已访问节点集中是否包含该目标节点,若否,则确定该目标节点为起始节点对应的已访问邻居节点,并将该目标节点添加到全局已访问节点集中,若是,则确定起始节点不存在关联节点,并将起始节点对应的已访问邻居节点从全局已访问节点集中移除。

进一步地,第二处理器用于针对图数据分块中包含的每个节点,根据该节点在图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后特征表示,对原始图数据进行更新处理。

进一步地, 在第二处理器中包含有至少一个处理单元, 各处理单元中的至

少部分处理单元用于针对图数据分块中包含的每个节点,将该节点在图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示,以及,各处理单元中的其他处理单元用于根据该节点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示。

需要说明的是,数据处理系统还包括:任务调度器,可以通过任务调度器 生成针对图数据分块的第一数据处理任务,并根据各处理单元的负载值,将第 一数据处理任务分配给处理单元,这里的第一数据处理任务用于将部分节点的 节点特征表示进行聚合。以及可以通过任务调度器生成针对图数据分块的第二 数据处理任务,并根据各处理单元的负载值,将第二数据处理任务分配给处理 单元,第二数据处理任务用于对图数据分块中的至少一个节点的节点特征表示 进行更新。

5

10

15

20

25

进一步地,任务调取器还可以针对每个处理单元,判断该处理单元的负载值是否超过预设的第一负载阈值,若是,则确定在各处理单元中与该处理单元的位置相邻的各处理单元,作为各候选处理单元,判断各候选处理单元中负载值最低的候选处理单元,作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给目标处理单元。

需要说明的是,之所以从各处理单元中与该处理单元的位置相邻的各处理 单元中筛选出目标处理单元,是因为相邻的处理单元之间的数据传输效率更高, 更容易实现第一数据处理任务或第二数据处理任务的转移。

除此之外,任务调取器还可以针对各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则从各处理单元中选择负载值最小的处理单元作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给目标处理单元。

在实际应用场景中,可能存在部分第一数据处理任务或第二数据处理任务 所需处理的数据量较大,从而使得任务调度器在将该任务分配给任意一个处理

单元均使该处理单元的负载较高,此时,任务调取器可以针对各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则对该处理单元正在处理的第一数据处理任务或第二数据处理任务进行拆分,得到各子数据处理任务,将各子数据处理任务分配给各处理单元中选择负载值最小的处理单元。

5

10

15

20

25

上述内容中,第二处理器根据该节点在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示的方法可以是针对图数据分块中包含的每个节点,通过多轮迭代,确定该节点的更新后特征表示。

其中,针对每轮迭代,第二处理器可以确定该节点的待更新节点特征表示,根据上一轮迭代中确定出的待更新节点对应的聚合特征表示,对待更新节点特征表示进行更新,得到该节点在该轮迭代中对应的更新后特征表示,并将该节点在该轮迭代中对应的更新后特征表示,作为下一轮迭代中的待更新节点特征表示,聚合特征表示用于表征该节点在图数据分块中的各邻居节点的聚合结果,待更新节点特征表示是将该节点的节点特征表示迭代至上一轮后得到的,当确定满足预设的第二终止条件时,得到该节点的更新后特征表示。

上述的第二终止条件可以根据实际需要设置,例如: 迭代的轮数达到预设轮数。上述的根据待更新节点对应的聚合特征表示,对待更新节点特征表示进行更新,得到待更新节点在该轮迭代中对应的更新后特征表示的方法可以是,将待更新节点对应的聚合特征表示与待更新节点的节点特征表示进行加权融合,以得到待更新节点的更新后节点特征表示。

从上述内容中可以看出,在针对图数据分块的每个节点进行节点更新的过程中,主要分为两个阶段,即,将该节点在图数据分块中的各邻居节点进行聚合,以得到该节点对应的聚合特征表示的聚合阶段,以及,根据该节点对应的聚合特征表示,对待更新节点特征表示进行更新的更新阶段。

而在实际应用场景中, 可能存在部分节点在所述图数据分块中的各邻居节

点的数量较多,进而使得如果在每轮迭代中使用本轮迭代得到的各邻居节点的聚合特征表示,对待更新节点进行更新时,可能会由于将待更新节点的各邻居节点进行聚合,以得到聚合特征表示所需的时间较长,而影响对待更新节点进行更新的效率,因此,可以在每轮迭代中使用上一轮迭代得到的各邻居节点的聚合特征表示,这样可以在每轮迭代中同时进行上述的聚合阶段以及更新阶段,从而提升了对图数据分块中包含的节点进行节点更新的效率。

5

10

15

20

25

进一步地,还可以在每轮迭代中,针对该轮迭代所需聚合的各邻居节点,将这些邻居节点划分为各邻居节点分组,分别针对每个邻居节点分组进行聚合,以得到各子聚合特征表示,最终可以将各子聚合特征表示进行聚合,以得到各邻居节点的聚合特征表示。

上述内容中可以看出,可以从原始图数据中包含的每个节点中筛选出度数高于预设阈值的节点,作为用于连接各社区的枢纽节点,进而可以将枢纽节点的每个邻居节点作为起始节点,通过访问与该邻居节点之间存在连接关系的其他节点,作为该邻居节点的关联节点,从而可以将该邻居节点和该邻居节点的关联节点划分出来,作为原始图数据的一个图数据分块,并针对每个图数据分块进行节点更新处理,进而可以避免由于直接使用原始图数据中包含的全部节点对应的特征矩阵进行计算而产生的冗余计算,以提升数据处理效率。

为了进一步地对上述数据处理系统进行详细说明,本说明书还提供了通过上述的数据处理系统进行数据处理的方法,具体如图 2 所示。

图 2 为本说明书中提供的一种数据处理方法的示意图,包括以下步骤:

S201: 所述第一处理器针对原始图数据中包含的每个节点,判断该节点的度数是否超过预设阈值;

S202: 若是,则确定该节点为枢纽节点,并针对所述枢纽节点的每个邻居节点,通过多轮节点查询,确定所述原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点;

S203: 根据该邻居节点以及该邻居节点的关联节点,确定所述原始图数据

的图数据分块,以通过所述第二处理器针对所述图数据分块中包含的每个节点, 根据该节点在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后特征表示,对所述原始图数据进行更新处理。

5

10

15

20

25

第一处理器针对原始图数据中包含的每个节点,判断该节点的度数是否超过预设阈值,若是,则确定该节点为枢纽节点,并针对枢纽节点的每个邻居节点,确定原始图数据中与该邻居节点存在连接关系的其他节点,作为该邻居节点的关联节点,根据该邻居节点以及该邻居节点的关联节点,确定原始图数据的图数据分块,以通过第二处理器针对图数据分块中包含的每个节点,根据该节点在图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示,根据该节点以及其他节点的更新后特征表示,对原始图数据进行更新处理。

针对枢纽节点的每个邻居节点,将该邻居节点作为起始节点,通过多轮节点查询,确定该邻居节点的关联节点。

其中,针对每轮节点查询,确定该轮节点查询中的各目标节点,判断各目标节点是否存在邻居节点,若是,则将各目标节点作为起始节点的关联节点,以及,将各目标节点的邻居节点作为下一轮节点查询的目标节点,各目标节点是将起始节点迭代至上一轮后得到的,在确定满足预设的第一终止条件后,得到起始节点的各关联节点。

针对每轮节点查询中的每个目标节点,判断在预设的全局已访问节点集中是否包含该目标节点,若是,则确定起始节点不存在关联节点,并将起始节点对应的已访问邻居节点从全局已访问节点集中移除,若否,则将确定该目标节点为起始节点对应已访问邻居节点,并将该目标节点添加到全局已访问节点集中,并判断各目标节点是否存在邻居节点。

通过各处理单元中的至少部分处理单元针对图数据分块中包含的每个节点,将该节点在图数据分块中的至少部分邻居节点的节点特征表示进行聚合,

得到该节点的子聚合特征表示。以及,通过其他处理单元根据该节点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更新后特征表示。

通过任务调度器生成针对图数据分块的第一数据处理任务,并根据各处理单元的负载值,将第一数据处理任务分配给处理单元,以使处理单元针对图数据分块中包含的每个节点,将该节点在图数据分块中的至少部分邻居节点的节点特征表示进行聚合,得到该节点的子聚合特征表示。

5

10

15

20

25

通过任务调度器生成针对图数据分块的第二数据处理任务,并根据各处理 单元的负载值,将第二数据处理任务分配给处理单元,以使处理单元根据该节 点的各子聚合特征表示,对该节点的节点特征表示进行更新,得到该节点的更 新后特征表示。

通过任务调度器,针对每个处理单元,判断该处理单元的负载值是否超过 预设的第一负载阈值,若是,则确定在各处理单元中与该处理单元的位置相邻 的各处理单元,作为各候选处理单元,判断各候选处理单元中负载值最低的候 选处理单元,作为目标处理单元,将该处理单元正在处理的第一数据处理任务 或第二数据处理任务分配给目标处理单元。

通过任务调取器,针对各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则从各处理单元中选择负载值最小的处理单元作为目标处理单元,将该处理单元正在处理的第一数据处理任务或第二数据处理任务分配给目标处理单元。

通过任务调取器,针对各处理单元中的每个处理单元,判断该处理单元的负载值与其他处理单元的负载值之间的差值是否超过预设的第二负载阈值,若是,则对该处理单元正在处理的第一数据处理任务或第二数据处理任务进行拆分,得到各子数据处理任务,将各子数据处理任务分配给各处理单元中选择负载值最小的处理单元。

针对图数据分块中包含的每个节点,通过多轮迭代,确定该节点的更新后特征表示。

其中,针对每轮迭代,确定该节点的待更新节点特征表示,根据上一轮迭代中确定出的该节点对应的聚合特征表示,对待更新节点特征表示进行更新,得到该节点在该轮迭代中对应的更新后特征表示,并将该节点在该轮迭代中对应的更新后特征表示,作为下一轮迭代中的待更新节点特征表示,聚合特征表示用于表征该节点在图数据分块中的各邻居节点的聚合结果,待更新节点特征表示是将该节点的节点特征表示迭代至上一轮后得到的,当确定满足预设的第二终止条件时,得到该节点的更新后特征表示。

5

10

15

20

从上述内容中可以看出,可以从原始图数据中包含的每个节点中筛选出度数高于预设阈值的节点,作为用于连接各社区的枢纽节点,进而可以将枢纽节点的每个邻居节点作为起始节点,通过访问与该邻居节点之间存在连接关系的其他节点,作为该邻居节点的关联节点,从而可以将该邻居节点和该邻居节点的关联节点划分出来,作为原始图数据的一个图数据分块,并针对每个图数据分块进行节点更新处理,进而可以避免由于直接使用原始图数据中包含的全部节点对应的特征矩阵进行计算而产生的冗余计算,以提升数据处理效率。

本说明书还提供了一种计算机可读存储介质,该存储介质存储有计算机程序,计算机程序可用于执行上述图 1 提供的一种的方法。

本说明书还提供了图 5 所示的一种对应于图 1 的电子设备的示意结构图。如图 5 所示,在硬件层面,该电子设备包括处理器、内部总线、网络接口、内存以及非易失性存储器,当然还可能包括其他业务所需要的硬件。处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行,以实现上述图 1 的方法。

当然,除了软件实现方式之外,本说明书并不排除其他实现方式,比 25 如逻辑器件抑或软硬件结合的方式等等,也就是说以下处理流程的执行主 体并不限定于各个逻辑单元,也可以是硬件或逻辑器件。

在20世纪90年代,对于一个技术的改进可以很明显地区分是硬件上的改 进(例如,对二极管、晶体管、开关等电路结构的改进)还是软件上的改进(对 于方法流程的改进)。然而,随着技术的发展,当今的很多方法流程的改进已 经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程 编程到硬件电路中来得到相应的硬件电路结构。因此,不能说一个方法流程的 改进就不能用硬件实体模块来实现。例如,可编程逻辑器件(Programmable Logic Device, PLD)(例如现场可编程门阵列(Field Programmable Gate Array, FPGA))就是这样一种集成电路,其逻辑功能由用户对器件编程来确定。由设 计人员自行编程来把一个数字系统"集成"在一片 PLD 上,而不需要请芯片制 造厂商来设计和制作专用的集成电路芯片。而且,如今,取代手工地制作集成 电路芯片,这种编程也多半改用"逻辑编译器(logic compiler)"软件来实现, 它与程序开发撰写时所用的软件编译器相类似,而要编译之前的原始代码也得 用特定的编程语言来撰写,此称之为硬件描述语言(Hardware Description Language, HDL), 而 HDL 也并非仅有一种, 而是有许多种, 如 ABEL( Advanced Boolean Expression Language), AHDL( Altera Hardware Description Language), Confluence, CUPL (Cornell University Programming Language), HDCal, JHDL (Java Hardware Description Language), Lava, Lola, MyHDL, PALASM, RHDL (Ruby Hardware Description Language)等,目前最普遍使用的是 VHDL(Very-High-Speed Integrated Circuit Hardware Description Language)与 Verilog。本领 域技术人员也应该清楚, 只需要将方法流程用上述几种硬件描述语言稍作逻辑 编程并编程到集成电路中,就可以很容易得到实现该逻辑方法流程的硬件电路。

5

10

15

20

25

控制器可以按任何适当的方式实现,例如,控制器可以采取例如微处理器或处理器以及存储可由该(微)处理器执行的计算机可读程序代码(例如软件或固件)的计算机可读介质、逻辑门、开关、专用集成电路(Application Specific Integrated Circuit,ASIC)、可编程逻辑控制器和嵌入微控制器的形式,控制器的例子包括但不限于以下微控制器: ARC 625D、Atmel AT91SAM、Microchip

PIC18F26K20 以及 Silicone Labs C8051F320,存储器控制器还可以被实现为存储器的控制逻辑的一部分。本领域技术人员也知道,除了以纯计算机可读程序代码方式实现控制器以外,完全可以通过将方法步骤进行逻辑编程来使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件,而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至,可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

5

10

15

20

25

上述实施例阐明的系统、装置、模块或单元,具体可以由计算机芯片或实体实现,或者由具有某种功能的产品来实现。一种典型的实现设备为计算机。具体的,计算机例如可以为个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任何设备的组合。

为了描述的方便,描述以上装置时以功能分为各种单元分别描述。当然, 在实施本说明书时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

本领域内的技术人员应明白,本说明书的实施例可提供为方法、系统、或计算机程序产品。因此,本说明书可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本说明书可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

本说明书是参照根据本说明书实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程

图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

5

10

15

20

25

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使 得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处 理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个 流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

在一个典型的配置中,计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器 (RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带、磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

还需要说明的是,术语"包括"、"包含"或者其任何其他变体意在涵盖非排

他性的包含,从而使得包括一系列要素的过程、方法、商品或者设备不仅包括 那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、 方法、商品或者设备所固有的要素。在没有更多限制的情况下,由语句"包括一 个……"限定的要素,并不排除在包括所述要素的过程、方法、商品或者设备中 还存在另外的相同要素。

5

10

15

20

本领域技术人员应明白,本说明书的实施例可提供为方法、系统或计算机程序产品。因此,本说明书可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本说明书可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

本说明书可以在由计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本说明书,在这些分布式计算环境中,由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中,程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。 尤其,对于系统实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

以上所述仅为本说明书的实施例而已,并不用于限制本说明书。对于本领域技术人员来说,本说明书可以有各种更改和变化。凡在本说明书的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本说明书的权利要求范围之内。

# 说明书附图

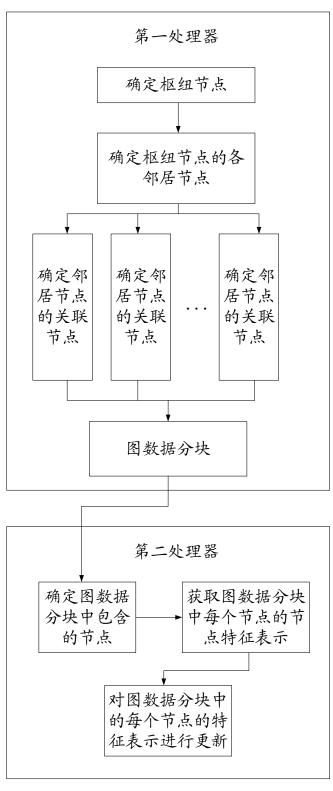


图 1

S201

所述第一处理器针对原始图数据中包含的 每个节点,判断该节点的度数是否超过预 设阈值

S202

若是,则确定该节点为枢纽节点,并针对 所述枢纽节点的每个邻居节点,通过多轮 节点查询,确定所述原始图数据中与该邻 居节点存在连接关系的其他节点,作为该 邻居节点的关联节点

S203

根据该邻居节点以及该邻居节点的关联节点,确定所述原始图数据的图数据分块,以通过所述第二处理器针对所述图数据分块中包含的每个节点,根据该节点在所述图数据分块中的各邻居节点的节点特征表示,对该节点的节点特征表示,根据该节点以及其他节点的更新后特征表示,对所述原始图数据进行更新处理

图 2

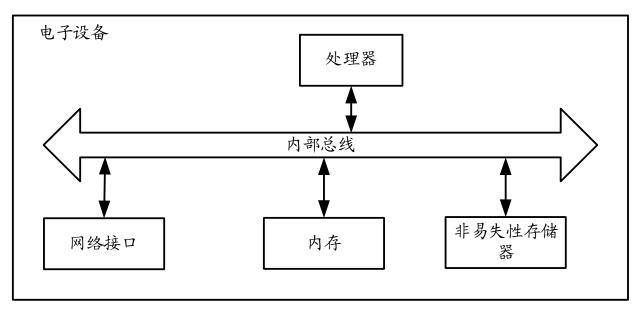


图 3