

权利要求书

- 1、一种基于集群的随机游走方法，包括：
所述集群获取图数据包含的各节点的信息；
- 5 根据所述各节点的信息，生成哈希表，所述哈希表的键包括所述节点的标识，值包括所述节点的相邻节点的标识；
根据所述哈希表，生成随机序列，所述随机序列反映在所述图数据中的随机游走。
- 2、如权利要求 1 所述的方法，所述集群包括服务器集群和工作机集群；
所述集群获取图数据包含的各节点的信息，具体包括：
- 10 所述工作机集群从数据库读取图数据包含的各节点的相邻节点的标识，其中，每个工作机读取一部分节点的相邻节点的标识。
- 3、如权利要求 2 所述的方法，所述根据所述各节点的信息，生成哈希表，具体包括：
各所述工作机分别根据自己读取标识的相邻节点及其对应节点的标识，生成非全量
- 15 的哈希表；
所述工作机集群将各所述非全量的哈希表向所述服务器集群同步；
所述服务器集群根据各所述非全量的哈希表，得到全量的哈希表。
- 4、如权利要求 3 所述的方法，所述根据所述哈希表，生成随机序列前，所述方法还包括：
- 20 所述服务器集群将所述全量的哈希表向各所述工作机同步，以便各所述工作机根据所述全量的哈希表，生成随机序列。
- 5、如权利要求 2 所述的方法，所述根据所述哈希表，生成随机序列，具体包括：
所述工作机在所述各节点的标识中，随机确定一个标识，作为目标节点的标识；
以所述目标节点的标识为键，在所述哈希表中索引得到对应的值，所述对应的值包
- 25 括所述目标节点的相邻节点的标识；
确定所述对应的值包括的相邻节点的标识的数量；
随机确定一个小于所述数量的非负整数，并获取所述对应的值包括的第所述非负整数个相邻节点的标识；
通过将该第所述非负整数个相邻节点重新作为目标节点进行迭代计算，生成由依次
- 30 得到的各目标节点的标识构成的随机序列。
- 6、如权利要求 5 所述的方法，所述对应的值存储于数组，所述目标节点的第 n 个

相邻节点的标识为该数组的第 n 个元素, n 从 0 开始计数;

所述非负整数记作 j , 所述获取所述对应的值包括的第所述非负整数个相邻节点的标识, 具体包括:

通过读取该数组的第 j 个元素, 获取所述目标节点的第 j 个相邻节点的标识。

- 5 7、如权利要求 5 所述的方法, 所述生成由依次得到的各目标节点的标识构成的随机序列, 具体包括:

当依次得到的各目标节点总数量达到预设的随机游走步数时, 生成由所述依次得到的各目标节点的标识构成的随机序列。

8、如权利要求 2 所述的方法, 所述生成随机序列, 具体包括:

- 10 各所述工作机分别生成随机序列, 直至生成的随机序列总数量达到设定阈值。

9、如权利要求 5 所述的方法, 所述方法还包括:

所述工作机若本地已有的所述哈希表丢失, 则重新从所述服务器集群获取。

10、一种随机游走方法, 包括:

获取根据图数据包含的各节点的信息生成的哈希表, 所述哈希表的键包括所述节点

- 15 的标识, 值包括所述节点的相邻节点的标识;

根据所述哈希表, 生成随机序列, 所述随机序列反映在所述图数据中的随机游走。

11、一种基于集群的随机游走装置, 所述装置属于所述集群, 包括:

获取模块, 获取图数据包含的各节点的信息;

第一生成模块, 根据所述各节点的信息, 生成哈希表, 所述哈希表的键包括所述节

- 20 点的标识, 值包括所述节点的相邻节点的标识;

第二生成模块, 根据所述哈希表, 生成随机序列, 所述随机序列反映在所述图数据中的随机游走。

12、如权利要求 11 所述的装置, 所述集群包括服务器集群和工作机集群;

所述获取模块获取图数据包含的各节点的信息, 具体包括:

- 25 所述工作机集群从数据库读取图数据包含的各节点的相邻节点的标识, 其中, 每个工作机读取一部分节点的相邻节点的标识。

13、如权利要求 12 所述的装置, 所述第一生成模块根据所述各节点的信息, 生成哈希表, 具体包括:

- 30 各所述工作机分别根据自己读取标识的相邻节点及其对应节点的标识, 生成非全量的哈希表;

所述工作机集群将各所述非全量的哈希表向所述服务器集群同步;

所述服务器集群根据各所述非全量的哈希表，得到全量的哈希表。

14、如权利要求 13 所述的装置，所述第二生成模块根据所述哈希表，生成随机序列前，所述服务器集群将所述全量的哈希表向各所述工作机同步，以便各所述工作机根据所述全量的哈希表，生成随机序列。

5 15、如权利要求 12 所述的装置，所述第二生成模块根据所述哈希表，生成随机序列，具体包括：

所述工作机在所述各节点的标识中，随机确定一个标识，作为目标节点的标识；

以所述目标节点的标识为键，在所述哈希表中索引得到对应的值，所述对应的值包括所述目标节点的相邻节点的标识；

10 确定所述对应的值包括的相邻节点的标识的数量；

随机确定一个小于所述数量的非负整数，并获取所述对应的值包括的第所述非负整数个相邻节点的标识；

通过将该第所述非负整数个相邻节点重新作为目标节点进行迭代计算，生成由依次得到的各目标节点的标识构成的随机序列。

15 16、如权利要求 15 所述的装置，所述对应的值存储于数组，所述目标节点的第 n 个相邻节点的标识为该数组的第 n 个元素， n 从 0 开始计数；

所述非负整数记作 j ，所述第二生成模块获取所述对应的值包括的第所述非负整数个相邻节点的标识，具体包括：

通过读取该数组的第 j 个元素，获取所述目标节点的第 j 个相邻节点的标识。

20 17、如权利要求 15 所述的装置，所述工作机生成由依次得到的各目标节点的标识构成的随机序列，具体包括：

所述工作机当依次得到的各目标节点总数量达到预设的随机游走步数时，生成由所述依次得到的各目标节点的标识构成的随机序列。

18、如权利要求 12 所述的装置，所述第二生成模块生成随机序列，具体包括：

25 各所述工作机分别生成随机序列，直至生成的随机序列总数量达到设定阈值。

19、如权利要求 15 所述的装置，所述工作机若本地已有的所述哈希表丢失，则重新从所述服务器集群获取。

20、一种随机游走装置，包括：

30 获取模块，获取根据图数据包含的各节点的信息生成的哈希表，所述哈希表的键包括所述节点的标识，值包括所述节点的相邻节点的标识；

生成模块，根据所述哈希表，生成随机序列，所述随机序列反映在所述图数据中的

随机游走。

- 21、一种基于集群的随机游走设备，所述设备属于所述集群，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，
- 5 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：
- 获取图数据包含的各节点的信息；
- 根据所述各节点的信息，生成哈希表，所述哈希表的键包括所述节点的标识，值包括所述节点的相邻节点的标识；
- 10 根据所述哈希表，生成随机序列，所述随机序列反映在所述图数据中的随机游走。
- 22、一种随机游走设备，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
- 15 处理器执行，以使所述至少一个处理器能够：
- 获取根据图数据包含的各节点的信息生成的哈希表，所述哈希表的键包括所述节点的标识，值包括所述节点的相邻节点的标识；
- 根据所述哈希表，生成随机序列，所述随机序列反映在所述图数据中的随机游走。