

分布式图嵌入方法、装置、设备及系统

技术领域

[01] 本说明书涉及数据处理领域，尤其涉及分布式图嵌入方法、装置、设备及系统。

背景技术

5 [02] 图数据结构是一种抽象程度高，表达能力强的数据结构：顶点（又可以称为节点）可用于表示现实世界的实体，而实体之间的关系则通过顶点之间的边来体现。常用的图数据结构有社交网络、商品网络、知识图谱、用户-商品网络等等。Graph embedding（图嵌入）是一种图计算的方法，是学习顶点的一种连续、稠密、低维的向量表示，可以通过向量的关系来衡量顶点之间的关系。例如，采用 Deep Walk、node2vec 等图嵌入学习
10 算法，可以生成由图数据结构中各顶点的向量构成的图嵌入矩阵。图嵌入技术因为其结果的通用性，经常被用作特征提取，进一步将结果输入给下游任务/项目。如社交媒体通过提取用户的图嵌入特征进行反作弊检测；电商通过商品的图嵌入特征优化商品推荐任务；内容分发商通过文章的图嵌入特征优化内容推荐任务等。

[03] 然而，对图数据结构进行大规模图嵌入计算时，由于会将图嵌入矩阵加载到内存进行计算，而图嵌入矩阵的大小由图数据结构中顶点数目和每个向量大小决定，在顶点数目较多、向量维度较大时，图嵌入矩阵需要占用大量内存，常常导致内存溢出。
15

发明内容

[04] 为克服相关技术中存在的问题，本说明书提供了分布式图嵌入方法、装置、设备及系统。

20 [05] 根据本说明书实施例的第一方面，提供一种分布式图嵌入方法，图数据结构至少分布于两个计算节点，所述方法包括：

[06] 在利用预设采样算法对图数据结构进行路径采样时，将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构，所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得；

25 [07] 在消息数据结构的数据长度达到词向量算法所定义的窗口长度时，利用所述词向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度；

[08]基于更新后的消息数据结构中的标识, 将各更新梯度分别发送至相应顶点, 并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新;

[09]达到预设收敛条件后, 将图数据结构中各顶点的最终嵌入向量, 构成图嵌入矩阵。

[10]在一个实施例中, 图数据结构中顶点的当前嵌入向量存储于顶点所在计算节点的图
5 顶点数据中。

[11]在一个实施例中, 所述方法还包括:

[12]在消息数据结构的数据长度未达到词向量算法所定义的窗口长度时, 利用预设采样算法对与所述当前采样顶点相邻且连接的顶点进行采样, 将更新后的消息数据结构发送至所确定的采样顶点所在计算节点。

10 [13]在一个实施例中, 所述消息数据结构的数据长度上限值为窗口长度, 所述方法还包括:

[14]在消息数据结构的数据长度达到或大于词向量算法所定义的窗口长度时, 删除消息数据结构中加入时间最早的采样顶点的标识和当前嵌入向量。

[15]在一个实施例中, 所述方法还包括:

15 [16]基于预设任务利用图嵌入矩阵构建训练样本, 并利用所构建的训练样本对预设模型进行训练, 获得用于处理预设任务的模型。

[17]根据本说明书实施例的第二方面, 提供一种分布式图嵌入装置, 图数据结构至少分布于两个计算节点, 所述装置设于其中一个计算节点, 包括:

[18]消息更新模块, 用于在利用预设采样算法对图数据结构进行路径采样时, 将当前采
20 样顶点的标识和当前嵌入向量加入预定义的消息数据结构, 所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得;

[19]梯度计算模块, 用于在消息数据结构的数据长度达到词向量算法所定义的窗口长度
25 度; 时, 利用所述词向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯

[20]向量更新模块, 用于基于更新后的消息数据结构中的标识, 将各更新梯度分别发送至相应顶点, 并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新;

[21]矩阵生成模块,用于达到预设收敛条件后,将图数据结构中各顶点的最终嵌入向量,构成图嵌入矩阵。

[22]在一个实施例中,图数据结构中顶点的当前嵌入向量存储于顶点所在计算节点的图顶点数据中。

5 [23]在一个实施例中,所述装置还包括:

[24]顶点采样模块,用于在消息数据结构的数据长度未达到词向量算法所定义的窗口长度时,利用预设采样算法对与所述当前采样顶点相邻且连接的顶点进行采样;

[25]信息传输模块,用于将更新后的消息数据结构发送至所确定的采样顶点所在计算节点。

10 [26]在一个实施例中,所述消息数据结构的数据长度上限值为窗口长度,所述消息更新模块还用于:

[27]在消息数据结构的数据长度达到或大于词向量算法所定义的窗口长度时,删除消息数据结构中加入时间最早的采样顶点的标识和当前嵌入向量。

[28]根据本说明书实施例的第三方面,提供一种分布式图嵌入系统,所述系统至少包括

15 两个计算节点,图数据结构分布式存储在所述系统中,任意一个计算节点执行以下方法:

[29]在利用预设采样算法对图数据结构进行路径采样时,将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构,所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得;

[30]在消息数据结构的数据长度达到词向量算法所定义的窗口长度时,利用所述词向量

20 算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度;

[31]基于更新后的消息数据结构中的标识,将各更新梯度分别发送至相应顶点,并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新;

[32]达到预设收敛条件后,将图数据结构中各顶点的最终嵌入向量,构成图嵌入矩阵。

[33]根据本说明书实施例的第四方面,提供一种计算机设备,包括存储器、处理器及存

25 储在存储器上并可在处理器上运行的计算机程序,其中,图数据结构至少分布于两个计算节点,所述处理器执行所述程序时实现如下方法:

[34]在利用预设采样算法对图数据结构进行路径采样时,将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构,所述消息数据结构中记录有采样过程中采样顶点

的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得；

[35]在消息数据结构的数据长度达到词向量算法所定义的窗口长度时，利用所述词向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度；

[36]基于更新后的消息数据结构中的标识，将各更新梯度分别发送至相应顶点，并利用
5 所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新；

[37]达到预设收敛条件后，将图数据结构中各顶点的最终嵌入向量，构成图嵌入矩阵。

[38]本说明书的实施例提供的技术方案可以包括以下有益效果：

[39]应用本说明书实施例，通过将图数据结构至少分布于两个计算节点，由多个计算节点来存储和计算顶点的嵌入向量，无需在计算时由同一个计算节点存储所有顶点的嵌入
10 向量，可以最大程度的分散内存消耗，同时，采用消息传播的方式使分布于不同计算节点的顶点，也能进行相应的向量计算。

[40]应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本说明书。

附图说明

15 [41]此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本说明书的实施例，并与说明书一起用于解释本说明书的原理。

[42]图 1 为本说明书一个或多个实施例提供的网络图的一个示例。

[43]图 2A 为本说明书根据一示例性实施例示出的一种分布式图嵌入方法的流程图。

[44]图 2B 为本说明书根据一示例性实施例示出的一种图数据结构分布示意图。

20 [45]图 3 为本说明书根据一示例性实施例示出的另一种分布式图嵌入方法的示意图。

[46]图 4 为本说明书根据一示例性实施例示出的一种分布式图嵌入装置所在计算机设备的一种硬件结构图。

[47]图 5 为本说明书根据一示例性实施例示出的一种分布式图嵌入装置的框图。

具体实施方式

25 [48]这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附

图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中描述的实施方式并不代表与本发明相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

[49]在本说明书使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本发明。

- 5 在本说明书和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

- [50]应当理解，尽管在本说明书可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本发明范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。
- 10

- [51]图数据结构是一种抽象程度高，表达能力强的数据结构。通常，在图计算中，基本的数据结构表达可以是： $G=(V, E, D)$ ， $V = \text{vertex}$ （顶点或者节点）， $E = \text{edge}$ （边）， $D = \text{data}$ （权重或其他属性信息）。对于许多物理世界的图问题，都可以利用图数据结构来抽象表达：比如社交网络、商品网络、知识图谱、用户-商品网络等等。图数据结构很好的表达了数据之间的关联性(relationships between data)，关联性计算是大数据计算的核心——通过获得数据的关联性，可以从噪音很多的海量数据中抽取有用的信息。比如，通过为购物者之间的关系建模，就能很快找到兴趣相似的用户，并为之推荐商品；
- 15
- 20 或者在社交网络中，通过传播关系发现意见领袖。

- [52]例如，对于凑单业务，可以基于用户购买行为构建 graph，节点：商品，边：商品间同时购买的行为，权重：同时购买的比重，可以是购买次数、购买时间、金额等特征。图 1 为本说明书一个或多个实施例提供的网络图的一个示例。在该示例中，顶点 A 至顶点 J 用于表示商品，顶点间的连线可以用于表示商品间同时购买的行为。可以理解的是，
- 25
- 图 1 仅是为了示例，而实际网络图可以根据业务需求确定。在基于用户购买行为构建网络图后，可以基于权重采样作为正样本的候选，负样本从用户非购买行为中随机抽样，利用样本对有监督模型进行训练，获得用于对用户进行凑单商品推荐的模型。

- [53]Graph embedding 是一种图计算的方法，能够将图中的顶点（节点）进行嵌入变成可计算的向量，各顶点向量可以构成图向量矩阵。然而，相关技术中对图数据结构进行大规模图嵌入计算时，图嵌入矩阵需要占用大量内存，常常导致内存溢出。
- 30

[54] 本说明书提供一种最大程度在分布式计算节点中分散内存消耗的方案，将大幅度降低计算中由于内存消耗导致的计算瓶颈。以下结合附图对本说明书实施例进行示例说明。

[55] 如图 2A 所示，是本说明书根据一示例性实施例示出的一种分布式图嵌入方法的流程图，图数据结构至少分布于两个计算节点，所述方法包括：

- 5 [56] 在步骤 202 中，在利用预设采样算法对图数据结构进行路径采样时，将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构，所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得；

[57] 在步骤 204 中，在消息数据结构的数据长度达到词向量算法所定义的窗口长度时，

- 10 利用所述词向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度；

[58] 在步骤 206 中，基于更新后的消息数据结构中的标识，将各更新梯度分别发送至相应顶点，并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新；

[59] 在步骤 208 中，达到预设收敛条件后，将图数据结构中各顶点的最终嵌入向量，构成图嵌入矩阵。

- 15 [60] 本实施例分布式图嵌入方法可以应用于至少包括两个计算节点的分布式系统中。例如如图 2B 所示，为本说明书根据一示例性实施例示出的一种图数据结构分布示意图。顶点 A、B、C、D 可以存储在计算节点 22 上，顶点 E、H、I、J 可以存储在计算节点 24 上，顶点 G、F 可以存储在计算节点 26 上。通过将图数据结构至少分布于两个计算节点，由多个计算节点来存储和计算顶点的嵌入向量，无需在计算时由同一个计算节点存储所有顶点的嵌入向量，可以最大程度的分散内存消耗，同时，采用消息传播的方式使分布于不同计算节点的顶点，也能进行相应的向量计算。

[61] 其中，图数据结构可以基于不同的业务需求而创建，例如，可以是社交网络、商品网络、知识图谱、用户-商品网络等等。预设采样算法可以是挖掘顶点与顶点之间关系的算法，通过预设采样算法可以采样获得顶点序列（又可以称为节点序列）。例如，预

- 25 设采样算法可以是 Deep Walk 形式的随机采样，也可以是 node2vec 形式的有倾向性采样，或者其他采样方法，该方案可适用于不同顶点/路径采样算法。

[62] 在利用预设采样算法对图数据结构进行路径采样过程中，在确定当前采样顶点时，可以将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构。其中，可以理解的是，为了区分，可以将当前采样所确定的顶点称为当前采样顶点。标识可以是用于

唯一标识当前采样顶点的标识，例如，可以是顶点 ID。当前嵌入向量可以是当前采样顶点此刻所存储的嵌入向量。例如，当前采样顶点首次被采样时，当前嵌入向量可以是初始嵌入向量。每个顶点的初始嵌入向量由初始化图数据结构获得。在训练过程中会更新顶点的嵌入向量，因此，当前采样顶点非首次被采样时，当前嵌入向量可以是上一次更新获得的嵌入向量。嵌入向量可以是向量化特征表示。图数据结构中顶点的当前嵌入向量存储于顶点所在计算节点的图顶点数据中，实现将图嵌入向量直接存储与图顶点数据中，并分布于全部计算节点，减少同一计算节点的内存消耗。

[63]关于消息数据结构，消息数据结构是用于在嵌入向量计算过程中传递数据的特定数据结构，也可以称为路径窗口消息。例如，消息数据结构可以是链表或列表等数据结构。

- 10 由于图数据结构分布于至少两个计算节点，不同计算节点分别存储不同顶点的顶点数据，而同一顶点序列可能分布于不同计算节点，因此，采用消息数据结构传播的方式，可以实现不同计算节点的数据传递，以便进行顶点嵌入向量更新。

[64]消息数据结构中至少记录有采样顶点的标识以及当前嵌入向量，消息数据结构可以在采样顶点之间进行传递。例如，消息数据结构可以由上一采样顶点所在计算节点发送而获得。若上一采样顶点和当前采样顶点在同一计算节点，则该计算节点可以直接将当前采样顶点的标识和当前嵌入向量加入消息数据结构；若上一采样顶点和当前采样顶点不在同一计算节点，则上一采样顶点所在计算节点将消息数据结构发送至当前采样顶点所在计算节点，当前采样顶点所在计算节点可以将当前采样顶点的标识和当前嵌入向量加入消息数据结构。

- 20 [65]例如，在一个实施例中，在消息数据结构的数据长度未达到词向量算法所定义的窗口长度时，利用预设采样算法对与所述当前采样顶点相邻且连接的顶点进行采样，将更新后的消息数据结构发送至所确定的采样顶点所在计算节点。

[66]可见，消息数据结构可以在顶点之间进行传递，避免同一顶点序列分布在不同计算节点而导致无法计算顶点嵌入向量的情况。

- 25 [67]消息数据结构的作用之一是用于在计算同一顶点路径中顶点的嵌入向量的更新梯度，因此，消息数据结构中采样顶点数目的上限值远远小于图数据结构中所有顶点数目。将采样顶点的标识和当前嵌入向量作为采样顶点的一组数据，所谓消息数据结构中采样顶点数目，可以是数据的组数。

[68]在一个实施例中，消息数据结构的数据长度上限值可以根据词向量算法所定义的窗

口长度而确定。其中，消息数据结构的数据长度可以是消息数据结构中当前嵌入向量的个数。在一个例子中，消息数据结构的数据长度上限值可以是窗口长度；在另一个例子中，消息数据结构的数据长度上限值可以稍大于窗口长度。

[69]消息数据结构可以是一种动态链表或列表，可以随着消息传递执行添加或删除操作，

5 以保证消息数据结构的当前长度不超过长度上限值。

[70]在一个实施例中，所述消息数据结构的数据长度上限值为窗口长度，所述方法还包括：

[71]在消息数据结构的数据长度达到或大于词向量算法所定义的窗口长度时，删除消息数据结构中加入时间最早的采样顶点的标识和当前嵌入向量。

10 [72]在该实施例中，可以采用先进先删除的原则，在消息数据结构的数据长度达到或大于词向量算法所定义的窗口长度时，可以根据数据加入时间对消息数据结构中的数据进行删除操作。例如，删除消息数据结构中加入时间最早的采样顶点的标识和当前嵌入向量，以实现只保留并通讯窗口长度内的顶点数据，并对其进行动态更新，降低存储训练样本造成的不必要的内存消耗。

15 [73]进一步的，为了提高删除效率，可以按加入顺序存储采样顶点的标识和当前嵌入向量，例如，将首次加入的标识和当前嵌入向量加入消息数据结构的头部，然后将后续加入的标识和当前嵌入向量加入当前消息数据结构的尾部，从而实现按加入顺序对采样顶点的标识和当前嵌入向量进行排列。

[74]在消息数据结构的数据长度达到词向量算法所定义的窗口长度时，可以利用词向量

20 算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度。

[75]其中，词向量算法可以是用于进行词向量计算的工​​具。例如，可以采用 SkipGram 或者 CBoW 算法计算嵌入向量的更新梯度，还可以采用其他词向量算法，也可支持自定义的损失函数计算嵌入向量的更新梯度，由于词向量算法是常规的算法，对其具体的计算方式不再详述。

25 [76]基于更新后的消息数据结构中的标识，将各更新梯度分别发送至相应顶点，并通过顶点所在计算节点利用更新梯度对采样顶点的当前嵌入向量进行更新。

[77]每次训练是对顶点嵌入向量的一次更新，经过多轮的训练，可以将顶点的嵌入向量更新到相对稳定的状态，因此，每次训练需要将顶点的嵌入向量进行调整。通过计算嵌入向量的梯度（可以称为更新梯度），而梯度也是一个向量，将当前嵌入向量加上或减

去该梯度，从而实现对当前嵌入向量的更新。

[78]其中，预设收敛条件可以是使嵌入向量达到稳定的条件。例如，可以是在所有顶点序列中滑动的窗口达到预设的训练步数（迭代步数）。

[79]进一步的，不同顶点序列可以同时进行嵌入向量的更新。不同顶点序列可以是同一

5 计算节点中的不同顶点序列，也可以是不同计算节点的不同顶点序列。

[80]由上述实施例可见，本说明书实施例可以将图嵌入向量作为顶点数据的一部分，分布存储于各计算节点，避免由单一计算节点造成的内存瓶颈。同时，由于只保留并通讯窗口长度内的顶点数据，并对其进行动态更新，降低存储训练样本造成的不必要的内存消耗。

10 [81]进一步的，本说明书还提供图嵌入矩阵的应用。在一个实施例中，可以基于预设任务利用图嵌入矩阵构建训练样本，并利用所构建的训练样本对预设模型进行训练，获得用于处理预设任务的模型。例如，如果图数据结构是商品网络图，则预设任务可以是商品推荐任务等。

[82]以上实施方式中的各种技术特征可以任意进行组合，只要特征之间的组合不存在冲突或矛盾，但是限于篇幅，未进行一一描述，因此上述实施方式中的各种技术特征的任意进行组合也属于本说明书公开的范围。

[83]以下以一个应用实例进行示例说明。

[84]如图3所示，是本说明书根据一示例性实施例示出的另一种分布式图嵌入方法的示意图。图3中示出部分顶点，顶点1至顶点N+1可以分布存储于不同计算节点。

20 [85]初始化图数据结构，获得各顶点的初始嵌入向量，并将初始嵌入向量存储于顶点数据结构中（步骤1）。

[86]利用预设采用算法对图数据结构进行路径采样，假设所确定的采样顶点为顶点1，则顶点1所在计算节点将顶点1的顶点ID和初始嵌入向量加入消息数据结构的头部。在消息数据结构的数据长度未达到Skip Gram/CboW算法所定义的窗口长度时，继续对
25 邻接顶点进行采样，并将更新后的消息数据结构发送至所确定的采样顶点。假设采样顶点为顶点2，则顶点1所在计算节点将更新后的消息数据结构发送至顶点2所在计算节点（步骤2）。

[87]顶点2所在计算节点将顶点2的顶点ID和当前嵌入向量加入消息数据结构的尾部，

例如消息数据结构为：顶点 1 的数组（ID 和当前嵌入向量）、顶点 2 的数组（ID 和当前嵌入向量）。在消息数据结构的数据长度未达到 Skip Gram/CboW 算法所定义的窗口长度时，继续对邻接顶点进行采样，并将更新后的消息数据结构发送至所确定的采样顶点（步骤 3）。

- 5 [88]重复消息数据结构添加操作以及数据长度判断操作，直到消息数据结构的数据长度达到 Skip Gram/CboW 算法所定义的窗口长度（步骤 4）。

[89]假设顶点 N 所在计算节点将顶点 N 的顶点 ID 和当前嵌入向量加入消息数据结构的尾部后，消息数据结构的数据长度达到 Skip Gram/CboW 算法所定义的窗口长度，则根据 Skip Gram/CBoW 算法计算消息数据结构中各顶点嵌入向量的更新梯度（步骤 5）。

- 10 [90]顶点 N 所在计算节点将更新梯度发送至所对应顶点（步骤 6）。

[91]所对应顶点所在计算节点收到更新梯度后，利用更新梯度对本顶点的嵌入向量进行更新（步骤 7）。

[92]将消息数据结构中头部的顶点数据删除，并将消息数据结构发送至采样所确定的顶点，以便该顶点所在计算节点将该顶点的标识和当前嵌入向量加入消息数据结构（步骤

15 8）。

[93]重复上述步骤 3 至 8，直到规定训练步数截止（步骤 9）。

[94]由此可见，将图嵌入向量直接存储于图顶点数据中，并分布于全部计算节点，在图上进行路径采样的同时，动态生成训练样本进行训练，从而降低内存消耗。

[95]与前述分布式图嵌入方法的实施例相对应，本说明书还提供了分布式图嵌入装置、

- 20 系统及装置所应用的电子设备的实施例。

[96]本说明书分布式图嵌入装置的实施例可以应用在计算机设备。装置实施例可以通过软件实现，也可以通过硬件或者软硬件结合的方式实现。以软件实现为例，作为一个逻辑意义上的装置，是通过其所在计算机设备的处理器将非易失性存储器中对应的计算机程序指令读取到内存中运行形成的。从硬件层面而言，如图 4 所示，为本说明书分布式图嵌入装置所在计算机设备的一种硬件结构图，除了图 4 所示的处理器 410、网络接口 420、内存 430、以及非易失性存储器 440 之外，实施例中分布式图嵌入装置 431 所在的计算机设备通常根据该设备的实际功能，还可以包括其他硬件，对此不再赘述。

[97]如图 5 所示，是本说明书根据一示例性实施例示出的一种分布式图嵌入装置的框图，

图数据结构至少分布于两个计算节点，所述装置设于其中一个计算节点，包括：

[98]消息更新模块 52，用于在利用预设采样算法对图数据结构进行路径采样时，将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构，所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所
5 在计算节点传递而获得；

[99]梯度计算模块 54，用于在消息数据结构的数据长度达到词向量算法所定义的窗口长度时，利用所述词向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度；

[100] 向量更新模块 56，用于基于更新后的消息数据结构中的标识，将各更新梯度分
10 别发送至相应顶点，并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新；

[101] 矩阵生成模块 58，用于达到预设收敛条件后，将图数据结构中各顶点的最终嵌入向量，构成图嵌入矩阵。

[102] 在一个实施例中，图数据结构中顶点的当前嵌入向量存储于顶点所在计算节点
15 的图顶点数据中。

[103] 在一个实施例中，所述装置还包括：

[104] 顶点采样模块，用于在消息数据结构的数据长度未达到词向量算法所定义的窗口长度时，利用预设采样算法对与所述当前采样顶点相邻且连接的顶点进行采样；

[105] 信息传输模块，用于将更新后的消息数据结构发送至所确定的采样顶点所在计
20 算节点。

[106] 在一个实施例中，所述消息数据结构的数据长度上限值为窗口长度，所述消息更新模块还用于：

[107] 在消息数据结构的数据长度达到或大于词向量算法所定义的窗口长度时，删除消息数据结构中加入时间最早的采样顶点的标识和当前嵌入向量。

25 [108] 在一个实施例中，所述装置还包括任务模型确定模块，用于：

[109] 基于预设任务利用图嵌入矩阵构建训练样本，并利用所构建的训练样本对预设模型进行训练，获得用于处理预设任务的模型。

[110] 对于装置实施例而言，由于其基本对应于方法实施例，所以相关之处参见方法

实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中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模块，即可以位于一个地方，或者也可以分布到多个网络模块上。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说明书方案的目的。本领域普通

5 技术人员在不付出创造性劳动的情况下，即可以理解并实施。

[111] 相应的，本说明书实施例还提供一种计算机设备，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，其中，图数据结构至少分布于两个计算节点，所述处理器执行所述程序时实现如下方法：

[112] 在利用预设采样算法对图数据结构进行路径采样时，将当前采样顶点的标识和
10 当前嵌入向量加入预定义的消息数据结构，所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得；

[113] 在消息数据结构的数据长度达到词向量算法所定义的窗口长度时，利用所述词向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度；

15 [114] 基于更新后的消息数据结构中的标识，将各更新梯度分别发送至相应顶点，并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新；

[115] 达到预设收敛条件后，将图数据结构中各顶点的最终嵌入向量，构成图嵌入矩阵。

[116] 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，对于设备实施例而言，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。

[117] 相应的，本说明书实施例还提供一种计算机存储介质，所述存储介质中存储有程序指令，所述程序指令包括：

25 [118] 在利用预设采样算法对图数据结构进行路径采样时，将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构，所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得；

[119] 在消息数据结构的数据长度达到词向量算法所定义的窗口长度时，利用所述词

向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度;

[120] 基于更新后的消息数据结构中的标识, 将各更新梯度分别发送至相应顶点, 并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新;

[121] 达到预设收敛条件后, 将图数据结构中各顶点的最终嵌入向量, 构成图嵌入矩阵。

[122] 相应的, 本说明书实施例还提供一种分布式图嵌入系统, 所述系统至少包括两个计算节点, 图数据结构分布式存储在所述系统中, 任意一个计算节点可以执行以下方法:

[123] 在利用预设采样算法对图数据结构进行路径采样时, 将当前采样顶点的标识和当前嵌入向量加入预定义的消息数据结构, 所述消息数据结构中记录有采样过程中采样顶点的标识和当前嵌入向量、且所述消息数据结构由上一采样顶点所在计算节点传递而获得;

[124] 在消息数据结构的数据长度达到词向量算法所定义的窗口长度时, 利用所述词向量算法计算更新后的消息数据结构中各采样顶点的嵌入向量的更新梯度;

[125] 基于更新后的消息数据结构中的标识, 将各更新梯度分别发送至相应顶点, 并利用所述更新梯度对归属当前计算节点的采样顶点的当前嵌入向量进行更新;

[126] 达到预设收敛条件后, 将图数据结构中各顶点的最终嵌入向量, 构成图嵌入矩阵。

[127] 本说明书实施例可采用在一个或多个其中包含有程序代码的存储介质(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。计算机可用存储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。

[128] 本领域技术人员在考虑说明书及实践这里申请的发明后, 将容易想到本说明书

的其它实施方案。本说明书旨在涵盖本说明书的任何变型、用途或者适应性变化，这些变型、用途或者适应性变化遵循本说明书的一般性原理并包括本说明书未申请的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的，本说明书的真正范围和精神由下面的权利要求指出。

- 5 [129] 应当理解的是，本说明书并不局限于上面已经描述并在附图中示出的精确结构，并且可以在不脱离其范围进行各种修改和改变。本说明书的范围仅由所附的权利要求来限制。

[130] 以上所述仅为本说明书的较佳实施例而已，并不用以限制本说明书，凡在本说明书的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本说明书保

10 护的范围之内。