(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 114841296 A
(43)申请公布日 2022.08.02

(21)申请号 202210776495.1

(22)申请日 2022.07.04

(71)申请人 北京六方云信息技术有限公司
地址 100085 北京市海淀区上地信息路12号1幢2层C202室
申请人 北京六方云科技有限公司

(72)发明人 卯路宁

(74)专利代理机构 北京恒程知识产权代理有限公司 11914
专利代理师 张婷

(51)Int.Cl.
G06K 9/62(2022.01)
G06N 3/04(2006.01)
G06N 3/08(2006.01)

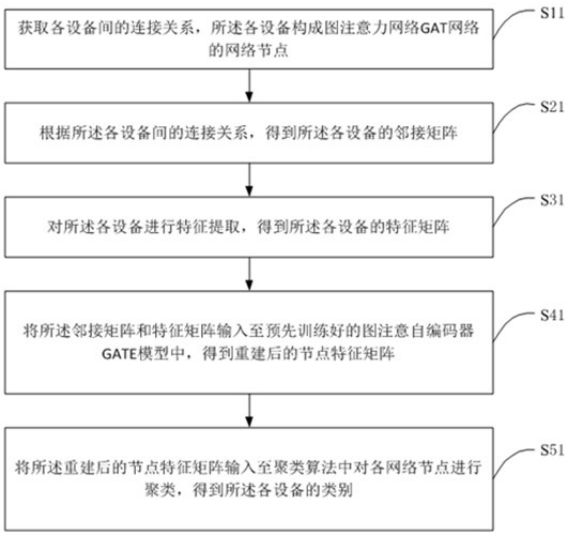
权利要求书3页 说明书19页 附图5页

(54)发明名称

设备聚类方法、终端设备以及存储介质

(57)摘要

本申请公开了一种设备聚类方法、终端设备以及存储介质，其设备聚类方法包括：获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；对所述各设备进行特征提取，得到所述各设备的特征矩阵；将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。本申请提供了在设备节点增加的情况下，无需更新全图的节点特征并对全图进行重新计算的设备聚类方案，节省大量计算空间。



CN 114841296 A

1.一种设备聚类方法,其特征在于,所述设备聚类方法包括:

获取各设备间的连接关系,所述各设备构成图注意力网络GAT网络的网络节点;

根据所述各设备间的连接关系,得到所述各设备的邻接矩阵;

对所述各设备进行特征提取,得到所述各设备的特征矩阵;

将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中,得到重建后的节点特征矩阵;

将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类,得到所述各设备的类别。

2.根据权利要求1所述的设备聚类方法,其特征在于,所述获取各设备间的连接关系的步骤包括:

统计GAT网络在预设时间段内的网络流量;

获取有网络流量流入和/或流出的设备IP;

将各设备IP的设备编号作为网络节点;

根据设备间的网络流量交互关系给各网络节点添加边;

基于各网络节点和边,构建GAT网络图结构,得到图结构的网络设备关系。

3.根据权利要求2所述的设备聚类方法,其特征在于,所述对所述各设备进行特征提取,得到所述各设备的特征矩阵的步骤包括:

获取每个设备的网络进出流量信息;

根据每个设备的网络进出流量信息,提取每个设备的节点特征;

根据每个设备的节点特征,生成对应的特征矩阵。

4.根据权利要求1、2或3所述的设备聚类方法,其特征在于,所述将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中,得到重建后的节点特征矩阵的步骤之前,还包括:

基于堆叠的编码器层和解码器层训练得到所述GATE模型,具体包括:

获取预先采集的数据集,所述数据集包括若干样本节点;

通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性,得到各样本节点的第一相关系数;

基于编码器层得到的第一相关系数,通过编码器层采用邻居的表示来生成各样本节点的表示;

通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性,得到各样本节点的第二相关系数;

基于解码器层得到的第二相关系数,通过解码器层采用邻居的表示来重建各样本节点的表示;

结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练,最小化节点特征和图结构的重建损失,输出重建后的节点特征矩阵,得到训练后的GATE模型。

5.根据权利要求4所述的设备聚类方法,其特征在于,所述通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性,得到各样本节点的第一相关系数的步骤包括:

采用具有节点间共享参数的自我关注机制,通过编码器层计算得到各样本节点的第一

相关系数；

对各样本节点的第一相关系数进行归一化处理；

所述基于编码器层得到的第一相关系数，通过编码器层采用邻居的表示来生成各样本节点的表示的步骤包括：

结合归一化处理后的第一相关系数，通过将各样本节点的节点特征作为初始节点表示，采用对应的编码器层生成各样本节点的表示。

6.根据权利要求5所述的设备聚类方法，其特征在于，所述通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第二相关系数的步骤包括：

采用具有节点间共享参数的自我关注机制，通过解码器层计算得到各样本节点的第二相关系数；

对各样本节点的第二相关系数进行归一化处理；

所述基于解码器层得到的第二相关系数，通过解码器层采用邻居的表示来重建各样本节点的表示的步骤包括：

结合归一化处理后的第二相关系数，通过将对应编码器层输出的样本节点的表示作为解码器层的输入，采用对应的解码器层重建各样本节点的表示。

7.根据权利要求6所述的设备聚类方法，其特征在于，所述结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型的步骤包括：

计算各样本节点重建后的节点表示与初始节点表示之间的差值；

计算重建后的各样本节点与对应的相邻节点之间的表示相似；

将所述差值与所述表示相似代入预设的目标损失函数，计算得到节点特征和图结构的重建损失；

将所述节点特征和图结构的重建损失回传到GATE模型，对编码器层和解码器层的可训练参数进行更新；并返回执行步骤；通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数；

以此循环，进行参数迭代，最小化节点特征和图结构的重建损失，直到所述GATE模型收敛，终止训练，输出重建后的节点特征矩阵，得到训练后的GATE模型。

8.根据权利要求7所述的设备聚类方法，其特征在于，所述输出重建后的节点特征矩阵的步骤包括：

根据编码器层得到的第一相关系数，获取所述编码器层的第一注意力矩阵；

根据所述第一注意力矩阵，通过所述编码器层输出各样本节点的节点表示矩阵；

根据解码器层得到的第二相关系数，获取所述解码器层的第二注意力矩阵；

结合各样本节点的节点表示矩阵和解码器层的第二注意力矩阵，通过解码器层输出重建后的节点特征矩阵。

9.一种终端设备，其特征在于，所述终端设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的设备聚类程序，所述设备聚类程序被所述处理器执行时实现如权利要求1-8中任一项所述的设备聚类方法的步骤。

10.一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有设备聚

类程序,所述设备聚类程序被处理器执行时实现如权利要求1-8中任一项所述的设备聚类方法的步骤。

设备聚类方法、终端设备以及存储介质

技术领域

[0001] 本申请涉及网络安全检测领域，尤其涉及一种设备聚类方法、终端设备以及存储介质。

背景技术

[0002] 随着计算机技术的发展，全球信息化已经成为人类发展的大趋势。但由于计算机网络具有连接形式多样性、终端分布不均匀和网络的开放性、互联性等特征，致使网络容易受到黑客、恶意软件等的攻击。网络设备作为互联网中重要的元素，通过对设备之间的通信关系进行分析，普遍认为有相似通信行为的设备间具有较高的相似度。通过对有相似通信行为的设备进行聚类，可将设备按照相似性分类，这对于研究设备间关系及设备的异常行为至关重要。目前已有的设备聚类技术大多基于机器学习，通过提取各个设备的行为作为特征，然后利用决策树分类、无监督聚类等方法，对设备进行分组画像。然而在实现本申请过程中，发明人发现采用现有设备聚类技术进行设备聚类时，由于其学习的参数与图结构有关，因此每当有设备节点增加时，就需要在每一次计算时更新全图的节点特征，致使其需要耗费大量的计算资源，因而难以完成动态变化的设备聚类任务。

[0003] 因此，有必要提出一种无需更新全图的节点特征并对全图进行重新计算的设备聚类解决方案。

发明内容

[0004] 本申请的主要目的在于提供一种设备聚类方法、终端设备以及存储介质，旨在提供一种在设备节点增加的情况下，无需更新全图的节点特征并对全图进行重新计算的设备聚类方案，节省大量计算空间。

[0005] 为实现上述目的，本申请提供一种设备聚类方法，所述设备聚类方法包括：

获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；

根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；

对所述各设备进行特征提取，得到所述各设备的特征矩阵；

将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；

将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。

[0006] 可选地，所述获取各设备间的连接关系的步骤包括：

统计GAT网络在预设时间段内的网络流量；

获取有网络流量流入和/或流出的设备IP；

将各设备IP的设备编号作为网络节点；

根据设备间的网络流量交互关系给各网络节点添加边；

基于各网络节点和边，构建GAT网络图结构，得到图结构的网络设备关系。

[0007] 可选地，所述对所述各设备进行特征提取，得到所述各设备的特征矩阵的步骤包括：

获取每个设备的网络进出流量信息；

根据每个设备的网络进出流量信息，提取每个设备的节点特征；

根据每个设备的节点特征，生成对应的特征矩阵。

[0008] 可选地，所述将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵的步骤之前，还包括：

基于堆叠的编码器层和解码器层训练得到所述GATE模型，具体包括：

获取预先采集的数据集，所述数据集包括若干样本节点；

通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数；

基于编码器层得到的第一相关系数，通过编码器层采用邻居的表示来生成各样本节点的表示；

通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第二相关系数；

基于解码器层得到的第二相关系数，通过解码器层采用邻居的表示来重建各样本节点的表示；

结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型。

[0009] 可选地，所述通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数的步骤包括：

采用具有节点间共享参数的自我关注机制，通过编码器层计算得到各样本节点的第一相关系数；

对各样本节点的第一相关系数进行归一化处理；

所述基于编码器层得到的第一相关系数，通过编码器层采用邻居的表示来生成各样本节点的表示的步骤包括：

结合归一化处理后的第一相关系数，通过将各样本节点的节点特征作为初始节点表示，采用对应的编码器层生成各样本节点的表示。

[0010] 可选地，所述通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第二相关系数的步骤包括：

采用具有节点间共享参数的自我关注机制，通过解码器层计算得到各样本节点的第二相关系数；

对各样本节点的第二相关系数进行归一化处理；

所述基于解码器层得到的第二相关系数，通过解码器层采用邻居的表示来重建各样本节点的表示的步骤包括：

结合归一化处理后的第二相关系数，通过将对应编码器层输出的样本节点的表示作为解码器层的输入，采用对应的解码器层重建各样本节点的表示。

[0011] 可选地，所述结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE

模型的步骤包括：

计算各样本节点重建后的节点表示与初始节点表示之间的差值；

计算重建后的各样本节点与对应的相邻节点之间的表示相似；

将所述差值与所述表示相似代入预设的目标损失函数，计算得到节点特征和图结构的重建损失；

将所述节点特征和图结构的重建损失回传到GATE模型，对编码器层和解码器层的可训练参数进行更新；并返回执行步骤；通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数；

以此循环，进行参数迭代，最小化节点特征和图结构的重建损失，直到所述GATE模型收敛，终止训练，输出重建后的节点特征矩阵，得到训练后的GATE模型。

[0012] 可选地，所述输出重建后的节点特征矩阵的步骤包括：

根据编码器层得到的第一相关系数，获取所述编码器层的第一注意力矩阵；

根据所述第一注意力矩阵，通过所述编码器层输出各样本节点的节点表示矩阵；

根据解码器层得到的第二相关系数，获取所述解码器层的第二注意力矩阵；

结合各样本节点的节点表示矩阵和解码器层的第二注意力矩阵，通过解码器层输出重建后的节点特征矩阵。

[0013] 本申请实施例还提出一种设备聚类装置，所述设备聚类装置包括：

关系获取模块，用于获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；

邻接矩阵模块，用于根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；

特征矩阵模块，用于对所述各设备进行特征提取，得到所述各设备的特征矩阵；

图注意力模块，用于将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；

设备聚类模块，用于将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。

[0014] 本申请实施例还提出一种终端设备，所述终端设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的设备聚类程序，所述设备聚类程序被所述处理器执行时实现如上所述的设备聚类方法的步骤。

[0015] 本申请实施例还提出一种计算机可读存储介质，所述计算机可读存储介质上存储有设备聚类程序，所述设备聚类程序被处理器执行时实现如上所述的设备聚类方法的步骤。

[0016] 本申请实施例提出的设备聚类方法、终端设备以及存储介质，通过获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；对所述各设备进行特征提取，得到所述各设备的特征矩阵；将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；将所述重建后的节点特征矩阵输入至聚类算法中对各节点进行聚类，得到所述各设备的类别。基于本申请方案，从相似通信行为的设备间具备较高相似度的规则出发，构建了一个能将设备按照相似性进行分类的设备聚类模型，基于本申请构建的设备聚类模型，在设备增加的情况下，无需更新全图的节点特征并对全图进行重新计算，最

后经过本申请方案实现设备聚类的方法节省了大量计算空间。

附图说明

[0017] 图1为本申请设备聚类装置所属终端设备的功能模块示意图；

图2为本申请设备聚类方法第一示例性实施例的流程示意图；

图3为本申请设备聚类方法第二示例性实施例的流程示意图；

图4为本申请设备聚类方法第三示例性实施例的流程示意图；

图5为本申请设备聚类方法实施例涉及的GATE模型的一种训练流程示意图；

图6为本申请实施例中结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型的具体流程示意图；

图7为本申请实施例中输出重建后的节点特征矩阵的具体流程示意图；

图8为本申请设备聚类方法第五示例性实施例的整体流程示意图；

图9为本申请实施例中将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别的具体流程示意图。

[0018] 本申请目的的实现、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

[0019] 应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

[0020] 本申请实施例的主要解决方案是：获取各设备间的连接关系，所述各设备构成图注意力网络(GAT网络)的网络节点；根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；对所述各设备进行特征提取，得到所述各设备的特征矩阵。通过在GATE模型中配置堆叠的编码器层和解码器层，获取预先采集的数据集，所述数据集包括若干样本节点；通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数；基于编码器层得到的第一相关系数，通过编码器层采用邻居的表示来生成各样本节点的表示；通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第二相关系数；基于解码器层得到的第二相关系数，通过解码器层采用邻居的表示来重建各样本节点的表示；结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型。通过训练后的GATE模型对所述各网络节点的邻接矩阵和特征矩阵进行处理后，得到重建后的节点特征矩阵，将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。基于本申请方案，从相似通信行为的设备间具备较高相似度的规则出发，构建了一个基于GATE模型并将设备按照相似性进行分类的设备聚类模型，基于本申请构建的设备聚类模型，在设备节点增加的情况下，无需更新全图的节点特征并对全图进行重新计算，最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0021] 本申请实施例涉及的技术术语：

注意力机制，Attention；

自注意力机制，Self Attention；

图卷积神经网络,GCN,Graph Convolutional Network;

图注意力网络,GAT,Graph Attention Network;

图注意自编码器,GATE,Graph Attention Auto-encoder。

[0022] 其中,注意力机制(Attention)是聚焦于局部信息的机制。而随着任务的变化,注意力区域往往会发生变化,而注意力机制便是要找到任务所需的最有用的信息。近年来,注意力机制在图像、自然语言处理等领域都取得了重要的突破,被证明有益于提高模型的性能。注意力机制本质上就是定位到感兴趣的信息,抑制无用信息,结果通常都是以概率图或者概率特征向量的形式展示。深度学习中的注意力机制从本质上来讲和人类的选择性视觉注意力机制类似,核心目标也是从众多信息中选择出对当前任务目标更为关键的信息。所以,注意力机制的目的是根据我们的目标,去关注部分细节,而不是基于全局进行分析,所以其核心就是如何基于目标确定我们要关注我部分,以及在找到这部分细节之后做进一步地分析。基于注意力机制构建的注意力模型(AM,Attention Model)通过允许模型动态地关注有助于执行手头任务的输入的某些部分,将这种相关性概念结合起来。

[0023] 自注意力机制(也叫自我关注机制,Self Attention)是注意力机制中的一种,也是转换器(Transformer)中的重要组成部分。自注意力机制要解决的问题是:当神经网络的输入是多个大小不一样的向量,并且可能因为不同向量之间有一定的关系,而在训练时却无法充分发挥这些关系,导致模型训练结果较差。

[0024] 图卷积神经网络(GCN,Graph Convolutional Network)是一种特征提取器,其对象为图数据,即图卷积神经网络就是一种处理图数据的深度学习方法。GCN精妙地设计了一种从图数据中提取特征的方法,从而可以使用这些特征去对图数据进行节点分类(Node Classification)、图分类(Graph Classification)、边预测(Link Prediction),还可以顺便得到图的嵌入表示(Graph Embedding),用途广泛。然而,由于GCN模型在训练期间需要整个图里面所有的节点参与,即需要将整个具体的图作为输出,所以改变图或者节点时,GCN模型就需要从头开始训练。

[0025] 图注意力网络(也叫图注意力机制,GAT,Graph Attention Network)是一种基于图结构数据的新型神经网络架构,利用隐藏的自注意力层来解决之前基于图卷积或者近似方法的不足。GAT引入了注意力机制(Attention)来实现更好的邻居聚合,通过学习邻居的权重,GAT可以实现对邻居的加权聚合。GAT通过堆叠层,使节点能够参与邻居的特征,可以(隐式地)为邻域中的不同节点指定不同的权值,而不需要任何代价高昂的矩阵操作(如反转),也不需要预先知道图的结构。通过这种方法,该模型克服了基于频谱的神经网络的几个关键挑战,并使得模型适用于归纳和推理问题。

[0026] 对比图卷积神经网络(GCN,Graph Convolutional Network)和图注意力网络(GAT,Graph Attention Network),GAT由自注意力层构建而成,自注意力层能够解决先前使用神经网络对图结构数据建模方法中存在的问题,包括:计算更高效,自注意力层的操作可以在所有边之间并行进行,同时所有节点的输出特征的计算也可以并行,不需要特征分解以及类似代价较高的矩阵操作;GAT模型与GCN模型相反,GAT模型允许(隐式地)将不同的重要性分配给同一邻居的节点,从而实现了模型容量的提升,此外,分析学习到的注意力权重可能会带来可解释性方面的好处;图注意力机制以共享的方式应用于图中的所有边,因此它不依赖于对全局图结构或者其所有节点(特性)的预先访问,即,图不需要是无向的,且

它使所提技术可以直接用于归纳学习,包括在训练过程中完全看不见的图上评估模型的任务。

[0027] 图注意自编码器(GATE,Graph Attention Auto-encoder)是一个基于图结构数据进行无监督学习的神经网络架构。此神经网络架构能够通过结合自注意力机制(Self Attention)的堆叠的编码器层/解码器层重构图结构化输入,包括节点属性和图结构。在编码器层中,通过考虑节点属性作为初始节点表示,每一层通过关注其邻居的表产生新的节点表示。在解码器层中,通过反转编码过程来重构节点属性。此外通过正则化节点表示来重构图结构。此外,神经网络架构不需要预先知道图结构,因此可以应用于归纳学习。

[0028] 归一化,Normalization;Sigmoid函数;Softmax函数。

[0029] 其中,归一化(Normalization)也称数据规范化。在机器学习领域中,不同评价指标(即特征向量中的不同特征就是所述的不同评价指标)往往具有不同的量纲和量纲单位,这样的情况会影响到数据分析的结果。为了消除指标之间的量纲影响,需要进行数据归一化处理,将数据按照相应比例进行缩放,使之落入一个特定的区域(映射于指定区域),以解决数据指标之间的可比性。原始数据经过数据归一化处理后,各指标处于同一数量级,适合进行综合对比评价。归一化的目的就是使得预处理的数据被限定在一定的范围内,比如[0,1]或者[-1,1],从而消除奇异样本数据导致的不良影响。其中奇异样本数据是指相对于其他输入样本特别大或者特别小的样本矢量(即特征向量)。通常神经网络需要归一化处理,一般变量的取值在-1到1之间,这样做是为了弱化某些变量的值较大而对模型产生影响。

[0030] Sigmoid函数,也称S型生长曲线,是神经元的非线性作用函数。Sigmoid函数在(0,0.5)处中心对称,在(0, 0.5)附近有比较大的斜率,而当数据趋向于正无穷和负无穷的时候,映射出来的值就会无限趋向于1和0。在深度学习中,由于其单增以及反函数单增等性质,Sigmoid函数常被用作神经网络的激活函数,将变量映射到[0,1]之间。激活函数给神经元引入了非线性因素,当加入多层神经元网络时,可以让神经网络拟合任何线性函数及非线性函数,从而使得神经网络可以适用于更多的非线性问题,而不仅仅是线性问题。

[0031] Softmax函数,又称归一化指数函数,它是二分类函数-sigmoid函数在多分类上的推广,目的是将实数范围内的分类结果转化为0-1之间的概率,通过利用指数的特性,将实数映射到0-正无穷(非负),并利用归一化方法,将实数映射的结果转化为0-1之间的概率。通常单个输出节点的二分类问题一般在输出节点上使用Sigmoid函数,拥有两个及其以上的输出节点的二分类或者多分类问题一般在输出节点上使用Softmax函数。

[0032] K-means算法、肘部法则:

其中,K-means算法又叫K均值算法。K-means算法中的K表示的是聚类为K个簇,means表示取每一个聚类中数据值的均值作为该簇的中心,或称为质心,即用每一个类的质心对该簇进行描述。K-means算法的思想大致为:先从样本集中随机选取K个样本作为簇中心,并计算所有样本与这K个簇中心的距离,对于每一个样本,将其划分到与其距离最近的簇中心所在的簇中,对于新的簇计算各个簇的新的簇中心,直到簇中心没有移动。

[0033] 肘部法则,一种K-means聚类的K值选择规则。肘部法则的核心思想是分类数K越大,样本划分会更加精细,每个类的聚合程度会逐渐提高,那么误差平方和SEE自然会逐渐变小。肘部法则的计算原理是成本函数,成本函数是类别畸变程度之和,每个类的畸变程度等于每个变量点到其类别中心的位置距离平方和(类内部的成员彼此越紧凑则类的畸变程

度越小,越分散越大)。在选择类别数量上,肘部法则会把不同值的成本函数值画出来。随着值的增大,每个类包含的样本数会减少,于是样本离其重心会更近平均畸变程度会减小。随着值继续增大,平均畸变程度的改善效果会不断减低。值增大过程中,畸变程度的改善效果下降幅度最大的位置对应的值就是肘部。而这个值就可以考虑为聚类性能较好的点。因此,肘部法则对于K-means算法的K值确定起到指导作用。

[0034] 本申请实施例方案,从相似通信行为的设备间具备较高相似度的规则出发,构建了一个基于GATE模型并将设备按照相似性进行分类的设备聚类模型,基于本申请构建的设备聚类模型,在设备节点增加的情况下,无需更新全图的节点特征并对全图进行重新计算,最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0035] 具体地,参照图1,图1为本申请设备聚类装置所属终端设备的功能模块示意图。该设备聚类装置可以为独立于终端设备的、能够进行设备聚类、网络模型训练的装置,其可以通过硬件或软件的形式承载于终端设备上。该终端设备可以为手机、平板电脑等具有数据处理功能的智能移动终端,还可以为具有数据处理功能的固定终端设备或服务器等。

[0036] 在本实施例中,该设备聚类装置所属终端设备至少包括输出模块110、处理器120、存储器130以及通信模块140。

[0037] 存储器130中存储有操作系统以及设备聚类程序,设备聚类装置可以将获取的各设备间的连接关系,通过各设备间的连接关系得到的各设备的邻接矩阵,通过对设备进行特征提取得到的特征矩阵,以及获取的预先采集的包括节点和相邻节点的数据集,通过编码器层对节点和相邻节点的相关性进行计算得到节点的第一相关系数,通过编码器层采用邻居的表示来生成的节点的表示,通过解码器层对节点和相邻节点的相关性进行计算得到节点的第二相关系数,通过解码器层采用邻居的表示来重建节点的表示,通过结合所述重建后的节点表示与预设的目标损失函数训练所得的最小化节点特征和图结构的重建损失,通过训练好的GATE模型输出重建后的节点特征矩阵,通过将所述重建后的节点特征矩阵输入至聚类算法中对各节点进行聚类得到的各设备的类别等信息存储于该存储器130中;输出模块110可为显示屏等。

[0038] 通信模块140可以包括WIFI模块、移动通信模块以及蓝牙模块等,通过通信模块140与外部设备或服务器进行通信。

[0039] 其中,存储器130中的设备聚类程序被处理器执行时实现以下步骤:

获取各设备间的连接关系,所述各设备构成图注意力网络GAT网络的网络节点;

根据所述各设备间的连接关系,得到所述各设备的邻接矩阵;

对所述各设备进行特征提取,得到所述各设备的特征矩阵;

将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中,得到重建后的节点特征矩阵;

将所述重建后的节点特征矩阵输入至聚类算法中对各节点进行聚类,得到所述各设备的类别。

[0040] 基于上述终端设备架构但不限于上述架构,提出本申请方法实施例。

[0041] 参照图2,图2为本申请设备聚类方法第一示例性实施例的流程示意图。所述设备聚类方法包括:

步骤S11,获取各设备间的连接关系,所述各设备构成图注意力网络GAT网络的网

络节点。

[0042] 具体地，获取各设备间的连接关系，其中，各设备间的连接关系是指设备间因存在相互通信行为而相互连接的关系，包括但不限于设备间的网络流量交互关系。其中，所述各设备构成GAT网络的网络节点，在本实施例中表示为可将所有设备看做是图G=(E,V)中的节点，节点之间的连接关系可看做是边，由节点和边共同构成了GAT网络。更进一步地，采用GAT网络的原因在于：GAT网络中重要的学习参数是W和α(.)，这两个参数仅与节点特征相关，与图的结构毫无关系，所以改变图的结构时，对于GAT网络的影响不大，因此更适合于归纳式(inductive)任务。而其他神经网络算法，如GCN，是一种全图的计算方式，由于其学习的参数很大程度上与图结构相关，每一次计算都要更新全图的节点特征，使得GCN在inductive任务上遇到困难。

[0043] 步骤S21，根据所述各设备间的连接关系，得到所述各设备的邻接矩阵。

[0044] 具体地，根据所述各设备间的连接关系，得到所述各设备的邻接矩阵，其中，为各设备添加设备编号作为节点，为各设备的连接关系添加边，由节点和边共同构成GAT网络的图结构，通过所得图结构计算得到各设备节点的邻接矩阵。更进一步地，邻接矩阵是表示节点之间相邻关系的矩阵，对于一个具有N(N∈R)个节点的图来说，邻接矩阵A为一个大小为N*N的对称矩阵，主对角线为0，若两个节点i，j之间有连接，则$A_{ij}=A_{ji}=1$，否则为0。

[0045] 步骤S31，对所述各设备进行特征提取，得到所述各设备的特征矩阵。

[0046] 具体地，提取各设备的特征，其中，设备的特征是与设备间的相互通信行为相关的特征，包括但不限于流入/流出的流量大小、端口信息、协议类型等。将提取的各设备特征作为节点特征，形成对应排列的特征矩阵。

[0047] 步骤S41，将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵。

[0048] 具体地，将基于图结构得到的邻接矩阵和基于节点特征得到的特征矩阵输入到预先训练好的GATE模型中，其中，GATE模型采用深度学习自编码器结合引入了注意力机制(Attention)的GAT网络，通过预先采集的数据集完成训练。将邻接矩阵和特征矩阵输入至训练好的GATE模型中，根据节点与相邻节点的相关性利用其邻居的表示来实现节点特征和图结构的重建，输出重建后的节点特征矩阵。

[0049] 步骤S51，将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。

[0050] 具体地，由于经GATE模型输出的网络节点不具备类别标签，因此需要将重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，根据聚类结果得到各设备的类别，其中，可采用的聚类算法包括但不限于K-means算法。

[0051] 本实施例方法的执行主体可以是一种设备聚类装置，也可以是一种设备聚类终端设备或服务器，本实施例以设备聚类装置进行举例，该设备聚类装置可以集成在具有数据处理功能的智能手机、平板电脑等终端设备上。

[0052] 本实施例通过上述方案，具体通过获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；对所述各设备进行特征提取，得到所述各设备的特征矩阵；将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；将所述重

建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。基于本申请方案，从相似通信行为的设备间具备较高相似度的规则出发，构建了一个能将设备按照相似性进行分类的设备聚类模型，基于本申请构建的设备聚类模型，在设备增加的情况下，无需更新全图的节点特征并对全图进行重新计算，最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0053] 参照图3，图3为本申请设备聚类方法第二示例性实施例的流程示意图。基于上述图2所示的实施例，在本实施例中，上述步骤S11，获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点，可以包括：

步骤S111，统计GAT网络在预设时间段内的网络流量。

[0054] 具体地，统计GAT网络中各资产在预设时间段内的网络流量，其中，资产包括但不限于系统程序、应用程序等软件产品，资产需承托在智能手机、电脑等设备上运行，产生网络流量。其中，网络流量的具体形式包括：流入该资产的网络流量、该资产向其他设备发送的网络流量。

[0055] 步骤S112，获取有网络流量流入和/或流出的设备IP。

[0056] 具体地，获取有网络流量流入和/或流出的设备IP，其中，每个设备都有其固定的IP，通过统计资产的网络流量流入和/或流出信息，获取承载该资产的设备的IP，具体形式包括：获取该资产的设备IP作为第一列表、获取有网络流量流入该资产的设备IP作为第二列表、获取接收了该资产流出的网络流量的其他设备IP作为第三列表，如表一所示：

资产IP	流入IP	流出IP
A	[C,E...]	[B,D...]

表一：网络流量流入和/或流出设备IP表

其中A、B、C、D、E均为设备，表示为设备A和B、C、D、E都有网络流量交互关系。

[0057] 步骤S113，将各设备IP的设备编号作为网络节点。

[0058] 具体地，将获取到的设备IP添加设备编码，并使用设备编码作为GAT网络的网络节点。

[0059] 步骤S114，根据设备间的网络流量交互关系给各网络节点添加边。

[0060] 具体地，基于获取的网络流量流入和/或流出设备IP表，得到设备间的网络流量交互关系。根据设备间的网络流量交互关系给各网络节点添加边。

[0061] 步骤S115，基于各网络节点和边，构建GAT网络图结构，得到图结构的网络设备关系。

[0062] 具体地，本实施例基于GAT网络中各节点间的网络流量流入和/或流出信息，获得节点间的网络流量交互关系，并根据网络流量交互关系给各节点添加边，构建GAT网络图结构，得到图结构的设备关系。

[0063] 进一步地，基于上述图2所示的实施例，在本实施例中，上述步骤S21，根据所述各设备间的连接关系，得到所述各设备的邻接矩阵具体为：根据所述图结构的网络设备关系，计算得到各网络节点的邻接矩阵。更为具体地，邻接矩阵是表示图结构的网络设备关系中各网络节点之间相邻关系的矩阵。

[0064] 本实施例通过上述方案，具体通过获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；根据所述各设备间的连接关系，得到所述各设备的邻接矩

阵;对所述各设备进行特征提取,得到所述各设备的特征矩阵;将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中,得到重建后的节点特征矩阵;将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类,得到所述各设备的类别。基于本申请方案,从相似通信行为的设备间具备较高相似度的规则出发,构建了一个能将设备按照相似性进行分类的设备聚类模型,基于本申请构建的设备聚类模型,在设备增加的情况下,无需更新全图的节点特征并对全图进行重新计算,最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0065] 进一步地,参照图4,图4为本申请设备聚类方法第三示例性实施例的流程示意图。基于上述图3所示的实施例,在本实施例中,上述步骤S31,对所述各设备进行特征提取,得到所述各设备的特征矩阵可以包括:

步骤S311,获取每个设备的网络进出流量信息。

[0066] 具体地,获取每个设备的网络进出流量信息,其中,网络进出流量信息包括但不限于每个设备使用的网络协议,以及每个设备下的每种网络协议的进出流量大小。在本实施例中,统计了常见的网络协议及其进出流量大小,其中,统计的网络协议包括但不限于ARP、DNS、FTP、IMAP、HHTPS、POP3、RDP、SIP、SMB、SMTP、SNMP、SSH、ICMP。

[0067] 步骤S312,根据每个设备的网络进出流量信息,提取每个设备的节点特征。

[0068] 具体地,根据每个设备的网络协议的进出流量信息,以每种协议的进出流量大小作为该设备的节点特征进行提取。

[0069] 步骤S313,根据每个设备的节点特征,生成对应的特征矩阵。

[0070] 具体地,根据提取到的设备的节点特征,生成对应排列的特征矩阵,其中,可以以节点数作为特征矩阵的行数,以网络协议数作为特征矩阵的列数,根据每个设备的节点特征生成特征矩阵的元素。

[0071] 本实施例通过上述方案,具体通过获取各设备间的连接关系,所述各设备构成图注意力网络GAT网络的网络节点;根据所述各设备间的连接关系,得到所述各设备的邻接矩阵;对所述各设备进行特征提取,得到所述各设备的特征矩阵;将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中,得到重建后的节点特征矩阵;将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类,得到所述各设备的类别。基于本申请方案,从相似通信行为的设备间具备较高相似度的规则出发,构建了一个能将设备按照相似性进行分类的设备聚类模型,基于本申请构建的设备聚类模型,在设备增加的情况下,无需更新全图的节点特征并对全图进行重新计算,最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0072] 参照图5,图5为本申请设备聚类方法实施例涉及的GATE模型的一种训练流程示意图。在本实施例中,在上述步骤S41,将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中,得到重建后的节点特征矩阵之前,还可以包括:

步骤S40,基于堆叠的编码器层和解码器层训练得到所述GATE模型。

[0073] 本实施例以步骤S40在步骤S41之前实施。具体地,GATE模型由GAT网络和自编码器结合而成,其框架使用了自编码器的模型框架,包括堆叠的编码器层和解码器层,其中,每一层的系数使用了注意力系数的计算方法。GATE模型采用预先采集的数据集,通过堆叠的编码器层和解码器层训练得到。

[0074] 相比上述图4所示的实施例，本实施例还包括训练GATE模型的方案。

[0075] 具体地，步骤S40，基于堆叠的编码器层和解码器层训练得到所述GATE模型，具体包括：

步骤S401，获取预先采集的数据集，所述数据集包括若干样本节点。

[0076] 更为具体地，本实施例预先收集了若干数量的样本节点，组成样本数据集；将样本数据集划分为训练集和测试集，所述训练集和测试集均包括若干样本节点，其中，训练集用于对GATE模型进行训练，测试集用于对训练过的GATE模型进行测试，以验证GATE模型的节点特征矩阵的生成效果。

[0077] 步骤S402，通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数。

[0078] 具体地，在训练GATE模型时，将预先采集的训练集输入GATE模型中进行训练，将各样本节点与对应的相邻节点输入至编码器层中，编码器层通过计算各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数。

[0079] 进一步地，步骤S402，通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数，可以包括：

采用具有节点间共享参数的自我关注机制，通过编码器层计算得到各样本节点的第一相关系数；

对各样本节点的第一相关系数进行归一化处理。

[0080] 更为具体地，为了确定数据集中各样本节点与其邻居(即相邻节点)之间的相关性，采用具有节点间共享参数的自我关注机制(又称自注意力机制)，通过编码器层计算该样本节点与其相邻节点之间的相关系数，并称作第一相关系数。在第k个编码器层中，相邻节点j和节点i的相关性计算如下公式1所示：

$$e_{ij}^{(k)} = sigmoid(v_s^{(k)^T}\sigma(W^{(k)}h_i^{(k-1)}) + v_r^{(k)^T}\sigma(W^{(k)}h_j^{(k-1)})) \quad (1)$$

其中，$e_{ij}^{(k)}$是第k个编码器层中相邻节点j与节点i的第一相关系数/注意系数，W、v_s、v_r是可训练参数，σ表示激活函数，Sigmoid表示Sigmoid函数，$h_i^{(k-1)}$是第k-1个编码器层生成的节点i的表示，$h_j^{(k-1)}$是第k-1个编码器层生成的节点j的表示。

[0081] 为了使节点i的邻域的相关系数具有可比性，采用Softmax函数进行归一化处理，如下公式2所示：

$$a_{ij}^{(k)} = \frac{exp(e_{ij}^{(k)})}{\sum_{l\in\mathcal{N}_i} exp(e_{il}^{(k)})} \quad (2)$$

其中，$\alpha_{ij}^{(k)}$是第k个编码器层中相邻节点j与节点i的归一化第一相关系数/注意系数，$e_{ij}^{(k)}$是第k个编码器层中相邻节点j与节点i的第一相关系数/注意系数，Ni是包括节点i的节点i的邻域，l是节点i的邻域内任一相邻节点。

[0082] 通过上述公式1和2可知，在本实施例中，采用了具有节点间共享参数的自我关注机制计算得到相邻节点j和节点i的第一相关系数，并为了使节点i的相关系数具有可比性，对第一相关系数进行了归一化处理。

[0083] 步骤S403,基于编码器层得到的第一相关系数,通过编码器层采用邻居的表示来生成各样本节点的表示,可以包括:

结合归一化处理后的第一相关系数,通过将各样本节点的节点特征作为初始节点表示,采用对应的编码器层生成各样本节点的表示。

[0084] 具体地,通过将该样本节点的节点特征作为初始节点表示(即$h_i^{(0)}=x_i$,$\forall i\in\{1,2,.,N\}$),编码器在第k层生成节点i的表示,具体计算过程如下公式3所示:

$$\hbar_i^{(k)}=\sum_{j\in\mathcal{N}_i}a_{ij}^{(k)}\sigma(W^{(k)}\hbar_j^{(k-1)}) \quad (3)$$

其中,$h_i^{(k)}$是第k个编码器层生成的节点i的表示,$h_j^{(k-1)}$是第k-1个编码器层生成的节点j的表示,$\alpha_{ij}^{(k)}$是第k个编码器层中相邻节点j与节点i的归一化第一相关系数/注意系数,W是可训练参数,σ表示激活函数,N_i是包括节点i的节点i的邻域。

[0085] 通过上述公式3可知,本实施例中,编码器在第k层生成的节点i的表示通过可训练参数、σ激活函数、第k-1个编码器层生成的节点j的表示以及第k个编码器层中相邻节点j与节点i的归一化第一相关系数/注意系数计算得到。

[0086] 步骤S404,通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性,得到各样本节点的第二相关系数。

[0087] 具体地,通过使用与编码器相同层数的解码器,每个解码器层试图逆转其相应编码器层的过程。换句话说,每个解码器层根据节点的相关性利用其邻居的表示来重建节点的表示。在训练GATE模型时,将预先采集的训练集输入GATE模型中进行训练,将各样本节点与对应的相邻节点输入至解码器层中,解码器层通过计算各样本节点与对应的相邻节点之间的相关性,得到各样本节点与对应的相邻节点之间的第二相关系数。

[0088] 进一步地,步骤S404,通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性,得到各样本节点的第二相关系数,可以包括:

采用具有节点间共享参数的自我关注机制,通过解码器层计算得到各样本节点的第二相关系数;

对各样本节点的第二相关系数进行归一化处理。

[0089] 更为具体地,采用具有节点间共享参数的自我关注机制,通过解码器层计算该样本节点与其相邻节点之间的相关系数,并称作第二相关系数。在第k个解码器层中,相邻节点j和节点i的相关性计算如下公式4所示,得到第k个解码器层中相邻节点j与节点i的第二相关系数/注意系数,然后采用Softmax函数对第k个解码器层生成的节点i的第二相关系数进行归一化处理,如下公式5所示:

$$\tilde{e}_{ij}^{(k)}=sigmoid(\tilde{v}_s^{(k)^T}\sigma(\widetilde{W}^{(k)}\widetilde{\hbar}_i^{(k)})+\tilde{v}_r^{(k)^T}\sigma(\widetilde{W}^{(k)}\widetilde{\hbar}_j^{(k)})) \quad (4)$$

$$\tilde{a}_{ij}^{(k)}=\frac{exp(\tilde{e}_{ij}^{(k)})}{\sum_{l\in\mathcal{N}_i}exp(\tilde{e}_{il}^{(k)})} \quad (5)$$

其中，$\tilde{e}_{ij}^{(k)}$ 是第k个解码器层中相邻节点j与节点i的第二相关系数/注意系数，$\tilde{W}$、$\tilde{v}_s$、$\tilde{v}_r$是可训练参数，σ表示激活函数，Sigmoid表示Sigmoid函数，$\tilde{h}_i^{(k)}$是第k个解码器层生成的节点i的表示，$\tilde{h}_j^{(k)}$是第k个解码器层生成的节点j的表示；$\tilde{a}_{ij}^{(k)}$ 是第k个解码器层中相邻节点j与节点i的归一化第二相关系数/注意系数，Ni是包括节点i的节点i的邻域，l是节点i的邻域内任一相邻节点。

[0090] 通过上述公式4和5可知，在本实施例中，解码器层采用了具有节点间共享参数的自我关注机制计算得到相邻节点j和节点i的第二相关系数，并为了使节点i的相关系数具有可比性，对第二相关系数进行了归一化处理。

[0091] 步骤S405，基于解码器层得到的第二相关系数，通过解码器层采用邻居的表示来重建节点的表示，可以包括：

结合归一化处理后的第二相关系数，通过将对应编码器层输出的样本节点的表示作为解码器层的输入，采用对应的解码器层重建各样本节点的表示。

[0092] 具体地，通过将对应编码器层输出的样本节点的表示作为解码器层的输入（即 $\tilde{\mathbf{h}}_i^{(L)}=\mathbf{h}_i^{(L)}$，∀i∈{1,2,.,N}），解码器在第k-1层重建节点i的表示，具体计算过程如下公式6所示：

$$\tilde{h}_i^{(k-1)}=\sum_{j\in\mathcal{N}_i}\tilde{a}_{ij}^{(k)}\sigma(\tilde{W}^{(k)}\tilde{h}_j^{(k)}) \quad (6)$$

其中，$\tilde{h}_i^{(k-1)}$ 是第k-1个解码器层重建的节点i的表示；$\tilde{h}_j^{(k)}$是第k个解码器层生成的节点j的表示，其中，可将编码器层的输出视为解码器层的输入（即 $\tilde{\mathbf{h}}_i^{(L)}=\mathbf{h}_i^{(L)}$，∀i∈{1,2,.,N}）；$\tilde{a}_{ij}^{(k)}$ 是第k个解码器层中相邻节点j与节点i的归一化第二相关系数/注意系数；$\tilde{\mathbf{W}}$ 是可训练参数；σ表示激活函数；Ni是包括节点i的节点i的邻域。

[0093] 通过上述公式6可知，本实施例中，解码器在第k-1层重建的节点i的表示通过可训练参数、σ激活函数、第k个解码器层生成的节点j的表示及第k个解码器层中相邻节点j与节点i的归一化第二相关系数/注意系数计算得到。

[0094] 步骤S406，结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型。

[0095] 具体地，结合重建后的各样本节点的表示与预设的目标损失函数进行可训练参数的循环迭代，以此最小化节点特征和图结构的重建损失，当节点特征和图结构的重建损失在预设损失值内，得到确定的可训练参数，即GATE模型收敛，通过收敛后的GATE模型输出重

建后的节点特征矩阵，并得到训练后的GATE模型。

[0096] 进一步地，参照图6，图6为本申请实施例中结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型的具体流程示意图。上述步骤S406，结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型，可以包括：

步骤S4061，计算各样本节点重建后的节点表示与初始节点表示之间的差值。

[0097] 具体地，为将节点特征的重建损失最小化，采用节点表示作为节点特征，计算各样本节点在重建后的节点表示与初始节点表示之间的差值，具体的计算如公式7所示：

$$\sum_{i=1}^{N} \|x_i - \hat{x}_i\|_2 \quad (7)$$

其中，N是图结构中的节点数，x_i是节点i的初始节点特征(其中，$h_i^{(0)} = x_i$)，$\hat{\mathbf{X}}_i$是节点i的重建节点特征(其中，$\hat{h}_i^{(0)} = \hat{x}_i$)。

[0098] 步骤S4062，计算重建后的各样本节点与对应的相邻节点之间的表示相似。

[0099] 具体地，由于特征相似的可能性，图中两个节点之间没有边不一定意味着不相似。换句话说，两个节点间没有联系/边，并不代表两个节点之间没有相似性。因此，可通过使用各样本节点与对应的相邻节点之间的表示相似来最小化图结构的重建损失。具体的该样本节点与对应的相邻节点之间的表示相似通过如下公式8所示来计算：

$$-\sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i} \log\left(\frac{1}{1+\exp(-h_i^T h_j)}\right) \quad (8)$$

其中，N是图结构中的节点数，Ni是包括节点i的节点i的邻域，h_i是节点i的表示，h_j是相邻节点j的表示。

[0100] 步骤S4063，将所述差值与所述表示相似代入预设的目标损失函数，计算得到节点特征和图结构的重建损失。

[0101] 具体地，结合各样本节点重建后的节点表示与初始节点表示之间的差值，和重建后的各样本节点与对应的相邻节点之间的表示相似，将所述差值和所述表示相似代入预设的目标损失函数中，具体目标损失函数的形式如公式9所示，计算得到节点特征和图结构的重建损失：

$$\text{Loss} = \sum_{i=1}^{N} \|x_i - \tilde{x}_i\|^2 - \lambda \sum_{j\in N_i} \log\left(\frac{1}{1+\exp(-h_i^T h_j)}\right) \quad (9)$$

其中，Loss为节点特征和图结构的重建损失，x_i是节点i的初始节点特征(其中，$h_i^{(0)} = x_i$)，$\hat{\mathbf{X}}_i$是节点i的重建节点特征(其中，$\hat{h}_i^{(0)} = \hat{x}_i$)，$h_i$是节点i的表示，$h_j$是相邻

节点j的表示，N是图结构中的节点数，Ni是包括节点i的节点i的邻域，λ是调整图结构重建损失的参数。

[0102] 步骤S4064，将所述节点特征和图结构的重建损失回传到GATE模型，对编码器层和解码器层的可训练参数进行更新；并返回执行步骤；通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数。

[0103] 具体地，将通过计算得到的节点特征和图结构的重建损失回传到GATE模型中，对编码器层和解码器层中的可训练参数进行更新，并返回执行步骤，重新按照步骤S402至步骤S406对GATE模型进行训练，输出新的节点特征和图结构的重建损失。

[0104] 步骤S4065，以此循环，进行参数迭代，最小化节点特征和图结构的重建损失，直到所述GATE模型收敛，终止训练，输出重建后的节点特征矩阵，得到训练后的GATE模型。

[0105] 具体地，对步骤S4064进行循环，不断迭代更新可训练参数，直到节点特征和图结构的重建损失在预设损失值范围内，即认为得到最小化的节点特征和图结构的重建损失，此时得到确定的可训练参数，即GATE模型收敛，可终止可训练参数的迭代。同时，通过训练好的GATE模型输出重建后的节点特征矩阵。

[0106] 通过上述公式7、8和9可知，本实施例中，节点特征和图结构的重建损失通过结合各样本节点重建后的节点表示与初始节点表示之间的差值，以及不同权重的重建后的各样本节点与对应的相邻节点之间的表示相似，来最小化节点特征和图结构的重建损失，以此训练GATE模型。

[0107] 进一步地，参照图7，图7为本申请实施例中输出重建后的节点特征矩阵的具体流程示意图。上述输出重建后的节点特征矩阵的步骤可以包括：

步骤S4071，根据编码器层得到的第一相关系数，获取所述编码器层的第一注意力矩阵；

具体地，在第k个编码器层中获得该层的第一注意力矩阵，其中，第一注意力矩阵向量通过该编码器层得到的归一化第一相关系数获得，即$C_{ij}^{(k)}=\alpha_{ij}^{(k)}$，如果节点i和节点j之间有边，则$C_{ij}^{(k)}=0$。具体编码器层的第一注意力矩阵的计算方式如下公式10、11和12所示：

$$C^{(k)} = softmax(sigmoid(M_s^{(k)} + M_r^{(k)})) \quad (10)$$

$$M_s^{(k)} = A \odot (v_s^{(k)^T} \sigma(W^{(k)} H^{(k-1)})) \quad (11)$$

$$M_r^{(k)} = A \odot (v_r^{(k)^T} \sigma(W^{(k)} H^{(k-1)}))^T \quad (12)$$

其中，$C^{(k)}$是第k个编码器层中的第一注意力矩阵；A是邻接矩阵；$H^{(k-1)}$是第k-1个编码器层生成的节点表示矩阵；W、v_s、v_r是可训练参数；⊙是同或运算符，表示为两个输入变量值相同时输出为1；σ表示激活函数；Sigmoid表示Sigmoid函数；Softmax表示Softmax函数；$M_s^{(k)}$和$M_r^{(k)}$分别为两个指代函数。

[0108] 步骤S4072，根据所述第一注意力矩阵，通过所述编码器层输出各样本节点的节点表示矩阵；

具体地，通过考虑$H^{(0)}=X$，其中，X是节点特征矩阵，在第k个编码器层中生成第k层的节点表示矩阵，其中，第k层的各样本节点的节点表示矩阵计算方式如下公式13所示：

$$H^{(k)}=\sigma(W^{(k)}H^{(k-1)})C^{(k)} \quad (13)$$

其中，$H^{(k)}$是第k个编码器层生成的各样本节点的节点表示矩阵，$H^{(k-1)}$是第k-1个编码器层生成的各样本节点的节点表示矩阵，$C^{(k)}$是第k个编码器层中的第一注意力矩阵，W是可训练参数，σ表示激活函数。

[0109] 步骤S4073，根据解码器层得到的第二相关系数，获取所述解码器层的第二注意力矩阵；

具体地，在第k个解码器层中获得该层的第二注意力矩阵，其中，第二注意力矩阵向量通过该解码器层得到的归一化第二相关系数获得，即$\hat{C}_{ij}^{(k)}=\hat{\alpha}_{ij}^{(k)}$，如果节点i和节点j之间有边，则$\hat{C}_{ij}^{(k)}=0$。具体解码器层的第二注意力矩阵的计算方式如下公式14、15和16所示：

$$\widetilde{C}^{(k)}=softmax(sigmoid(\widetilde{M}_s^{(k)}+\widetilde{M}_r^{(k)})) \quad (14)$$

$$\widetilde{M}_s^{(k)}=A\odot(\tilde{v}_s^{(k)^T}\sigma(\widetilde{W}^{(k)}\widetilde{H}^{(k)})) \quad (15)$$

$$\widetilde{M}_r^{(k)}=A\odot(\tilde{v}_r^{(k)^T}\sigma(\widetilde{W}^{(k)}\widetilde{H}^{(k)}))^T \quad (16)$$

其中，$\widetilde{C}^{(k)}$是第k个解码器层中的第二注意力矩阵；A是邻接矩阵；$\widetilde{H}^{(k)}$是第k个解码器层生成的各样本节点的节点特征矩阵；⊙是同或运算符，表示为两个输入变量值相同时输出为1；$\widetilde{W}$、$\tilde{v}_s$、$\tilde{v}_r$是可训练参数；σ表示激活函数；Sigmoid表示Sigmoid函数；Softmax表示Softmax函数，$\widetilde{M}_s^{(k)}$和$\widetilde{M}_r^{(k)}$分别为两个指代函数。

[0110] 步骤S4074，结合各样本节点的节点表示矩阵和解码器层的第二注意力矩阵，通过解码器层输出重建后的节点特征矩阵。

[0111] 具体地，通过考虑$\hat{H}^{(L)}=H^{(L)}$，其中，L是解码器层的层数，$H^{(L)}$是第L层编码器层生成的各样本节点的节点表示矩阵，即，将编码器层输出的各样本节点的节点表示矩阵作为对应解码器层的输入，在第k个解码器层中重建第k-1层的各样本节点的节点特征矩阵，具体的第k-1层的节点表示矩阵的计算方式如下公式17所示：

$$\widetilde{H}^{(k-1)}=\sigma(\widetilde{W}^{(k)}\widetilde{H}^{(k)})\widetilde{C}^{(k)} \quad (17)$$

其中，$\widetilde{H}^{(k-1)}$ 是由第k个解码器层重建的第k-1层的各样本节点的节点特征矩阵，$\widetilde{H}^{(k)}$ 是第k个解码器层生成的各样本节点的节点特征矩阵，$\widetilde{C}^{(k)}$ 是第k个解码器层中的第二注意力矩阵，$\widetilde{W}$ 是可训练参数，σ表示激活函数。

[0112] 通过上述公式10-17可知，本实施例中，由于邻接矩阵A在实践中通常是稀疏的，因此利用稀疏矩阵操作，如Softmax函数来处理图结构。GATE模型通过计算上述编码器层操作所对应的矩阵公式(如公式10-13)和解码器层操作所对应的矩阵公式(如公式14-17)，得到重建后的节点特征矩阵。

[0113] 相比现有技术，本实施例方案，通过获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；对所述各设备进行特征提取，得到所述各设备的特征矩阵；将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。基于本申请方案，从相似通信行为的设备间具备较高相似度的规则出发，构建了一个能将设备按照相似性进行分类的设备聚类模型，基于本申请构建的设备聚类模型，在设备增加的情况下，无需更新全图的节点特征并对全图进行重新计算，最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0114] 参考图8，图8为本申请设备聚类方法第五示例性实施例的整体流程示意图，本实施例的整体流程包括：通过获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；对所述各设备进行特征提取，得到所述各设备的特征矩阵；将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。

[0115] 进一步地，参照图9，图9为本申请实施例中将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别的具体流程示意图。基于上述图5所示的实施例，在本实施例中，上述步骤S51，将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别，可以包括：

步骤S511，根据所述重建后的节点特征矩阵，采用肘部法则自适应选取K值；

步骤S512，根据所述K值，采用K-means聚类算法对各网络节点进行聚类，得到每个所述设备的类别。

[0116] 具体地，本实施例中，由于经GATE模型输出的重建后的节点特征矩阵不具备类别标签，因此事先无法知晓网络节点的分类数。采用聚类算法对各网络节点进行聚类，首先需要确定聚类的簇中心数量，即K值。

[0117] 本实施例中采用K-means聚类算法对各网络节点进行聚类，其中K值的选取采用肘部法则来确定。肘部法则是一种K-means聚类的K值选择规则，通过在K值增大过程中，自适应选取畸变程度的改善效果下降幅度最大的位置所对应的值(即肘部)作为K-means聚类性能较好的K值。

[0118] 然后，根据肘部法则自适应选取的K值，将重建后的节点特征矩阵输入至K-means

聚类算法中对各网络节点进行聚类，根据聚类结果，得到每个所述设备的类别。

[0119] 通过肘部法则自适应选取K值，可以提高K-means聚类算法的聚类性能，应用此K-means聚类算法对GATE模型输出的重建后的节点特征矩阵进行网络节点间的聚类，能更好地找出设备间的聚类关系。

[0120] 本实施例通过上述方案，从相似通信行为的设备间具备较高相似度的规则出发，构建了一个能将设备按照相似性进行分类的设备聚类模型，基于本申请构建的设备聚类模型，在设备增加的情况下，无需更新全图的节点特征并对全图进行重新计算，最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0121] 此外，本申请实施例还提出一种设备聚类装置，所述设备聚类装置包括：

关系获取模块，用于获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；

邻接矩阵模块，用于根据所述各设备间的连接关系，得到所述各设备的邻接矩阵；

特征矩阵模块，用于对所述各设备进行特征提取，得到所述各设备的特征矩阵；

图注意力模块，用于将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；

设备聚类模块，用于将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。

[0122] 进一步地，所述设备聚类装置还包括：

模型训练模块，用于基于堆叠的编码器层和解码器层训练得到所述GATE模型。

[0123] 本实施例实现设备聚类的原理及实施过程，请参照上述各实施例，在此不再一一赘述。

[0124] 本实施例通过上述方案，从相似通信行为的设备间具备较高相似度的规则出发，构建了一个能将设备按照相似性进行分类的设备聚类模型，基于本申请构建的设备聚类模型，在设备增加的情况下，无需更新全图的节点特征并对全图进行重新计算，最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0125] 此外，本申请实施例还提出一种终端设备，所述终端设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的设备聚类程序，所述设备聚类程序被所述处理器执行时实现如上所述的设备聚类方法的步骤。

[0126] 由于本设备聚类程序被处理器执行时，采用了前述所有实施例的全部技术方案，因此至少具有前述所有实施例的全部技术方案所带来的所有有益效果，在此不再一一赘述。

[0127] 此外，本申请实施例还提出一种计算机可读存储介质，所述计算机可读存储介质上存储有设备聚类程序，所述设备聚类程序被处理器执行时实现如上所述的设备聚类方法的步骤。

[0128] 由于本设备聚类程序被处理器执行时，采用了前述所有实施例的全部技术方案，因此至少具有前述所有实施例的全部技术方案所带来的所有有益效果，在此不再一一赘述。

[0129] 相比现有技术，本申请实施例提出的设备聚类方法、装置、终端设备以及存储介质，通过获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点；根

据所述各设备间的连接关系，得到所述各设备的邻接矩阵；对所述各设备进行特征提取，得到所述各设备的特征矩阵；将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵；将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别。基于本申请方案，从相似通信行为的设备间具备较高相似度的规则出发，构建了一个能将设备按照相似性进行分类的设备聚类模型，基于本申请构建的设备聚类模型，在设备增加的情况下，无需更新全图的节点特征并对全图进行重新计算，最后经过本申请方案实现设备聚类的方法节省了大量计算空间。

[0130] 需要说明的是，在本文中，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、物品或者系统不仅包括那些要素，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、物品或者系统所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、方法、物品或者系统中还存在另外的相同要素。

[0131] 上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。

[0132] 通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请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，该计算机软件产品存储在如上的一个存储介质(如ROM/RAM、磁碟、光盘)中，包括若干指令用以使得一台终端设备(可以是手机，计算机，服务器，被控终端，或者网络设备等)执行本申请每个实施例的方法。

[0133] 以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

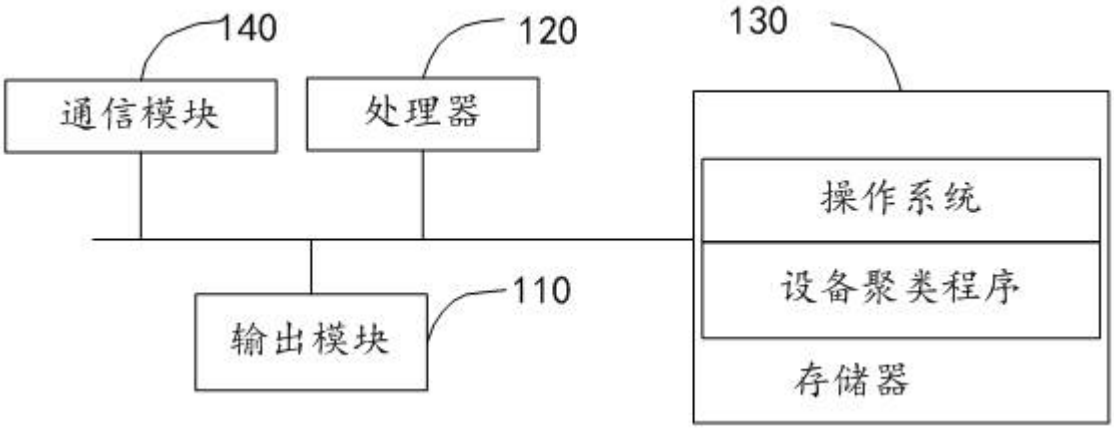

140
120
130
通信模块
处理器
操作系统
设备聚类程序
110
输出模块
存储器

图1

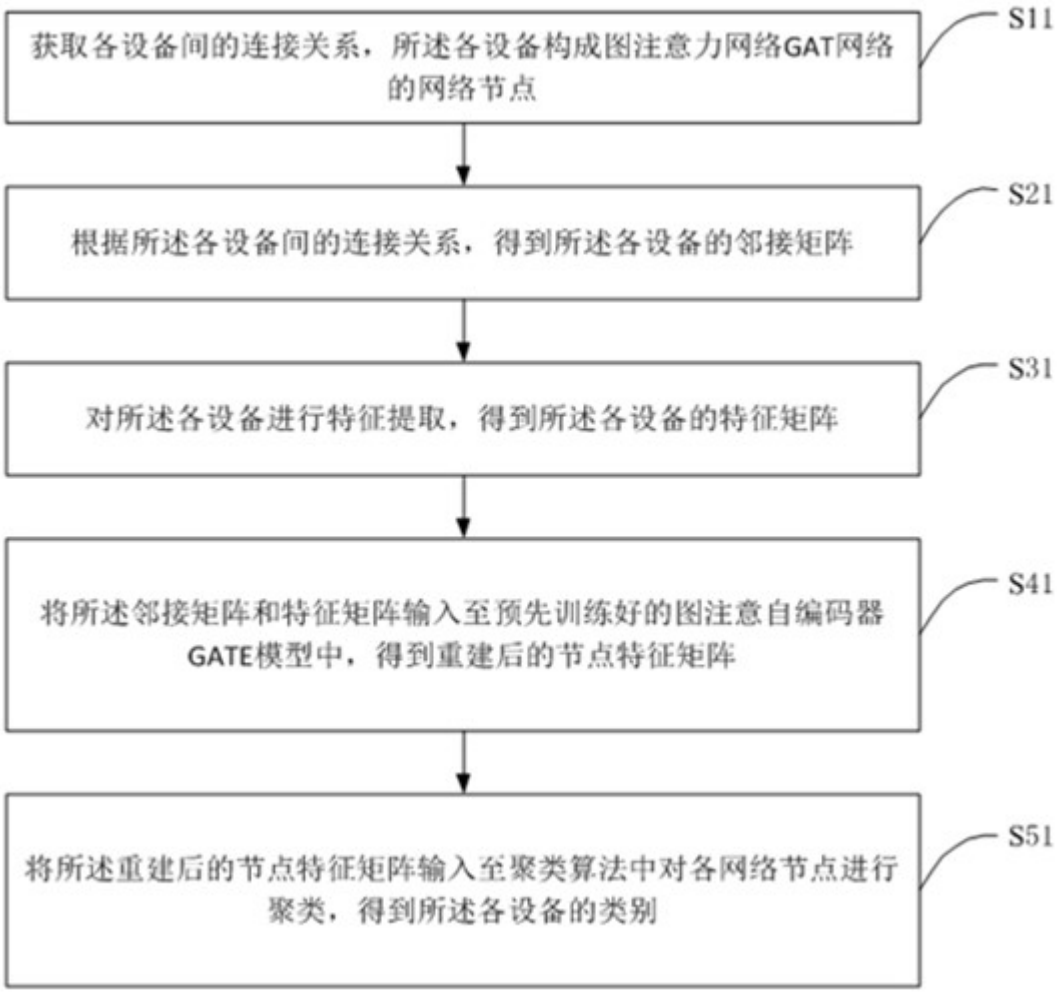

获取各设备间的连接关系，所述各设备构成图注意力网络GAT网络的网络节点
S11
根据所述各设备间的连接关系，得到所述各设备的邻接矩阵
S21
对所述各设备进行特征提取，得到所述各设备的特征矩阵
S31
将所述邻接矩阵和特征矩阵输入至预先训练好的图注意自编码器GATE模型中，得到重建后的节点特征矩阵
S41
将所述重建后的节点特征矩阵输入至聚类算法中对各网络节点进行聚类，得到所述各设备的类别
S51

图2

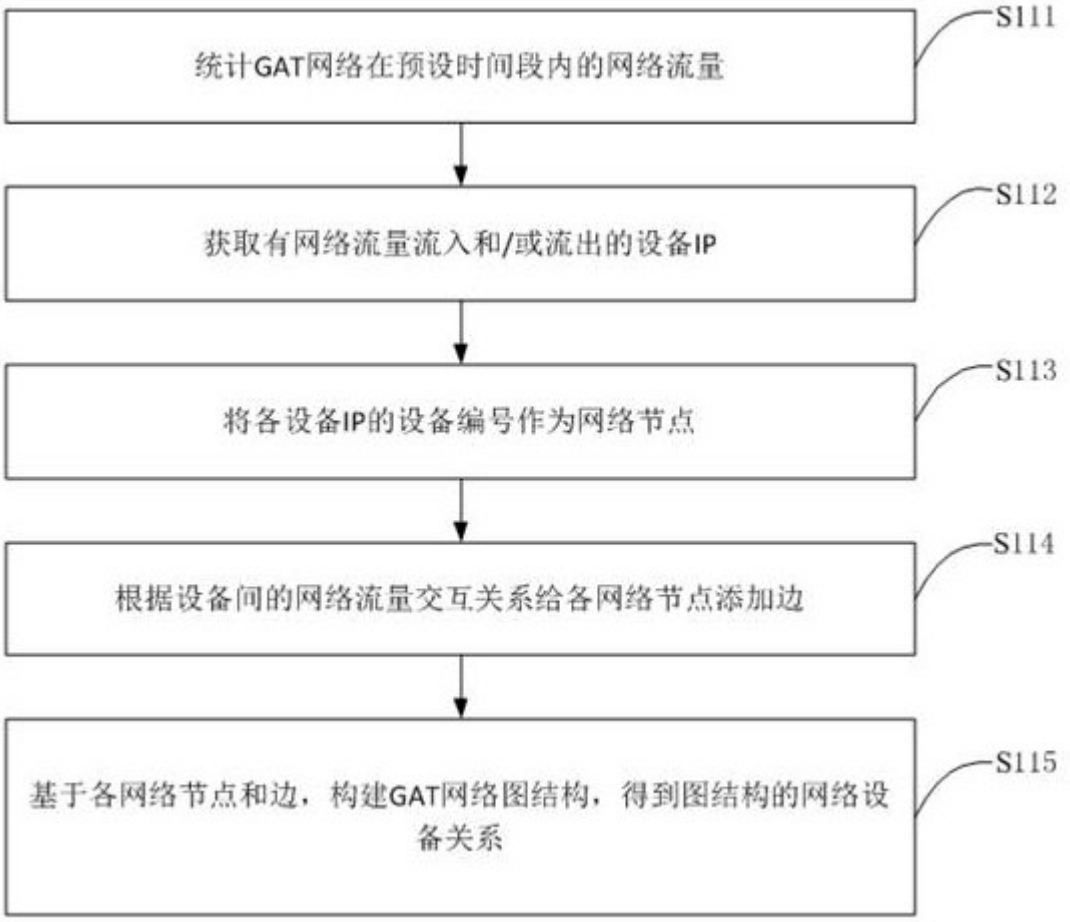

S111
统计GAT网络在预设时间段内的网络流量
S112
获取有网络流量流入和/或流出的设备IP
S113
将各设备IP的设备编号作为网络节点
S114
根据设备间的网络流量交互关系给各网络节点添加边
S115
基于各网络节点和边，构建GAT网络图结构，得到图结构的网络设备关系

图3

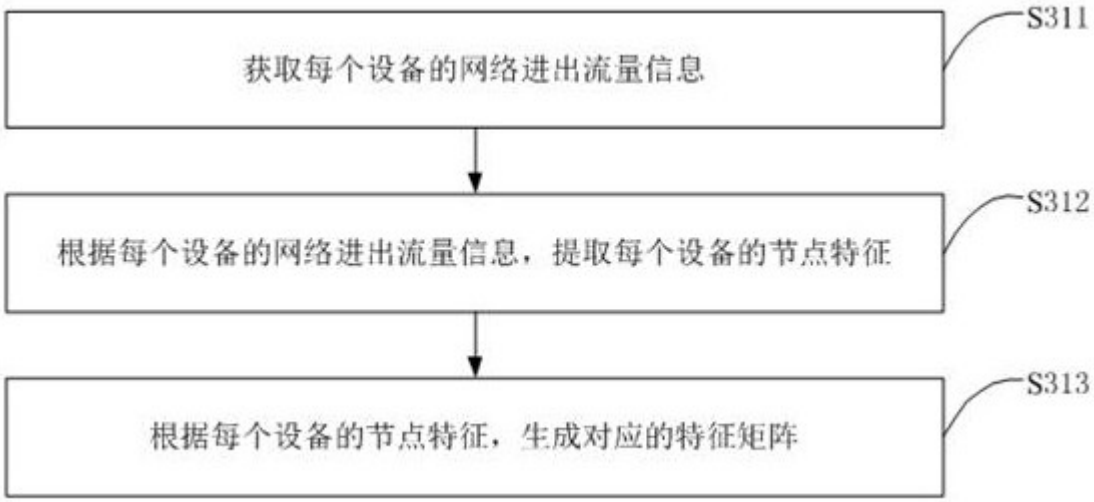

S311
获取每个设备的网络进出流量信息
S312
根据每个设备的网络进出流量信息，提取每个设备的节点特征
S313
根据每个设备的节点特征，生成对应的特征矩阵

图4

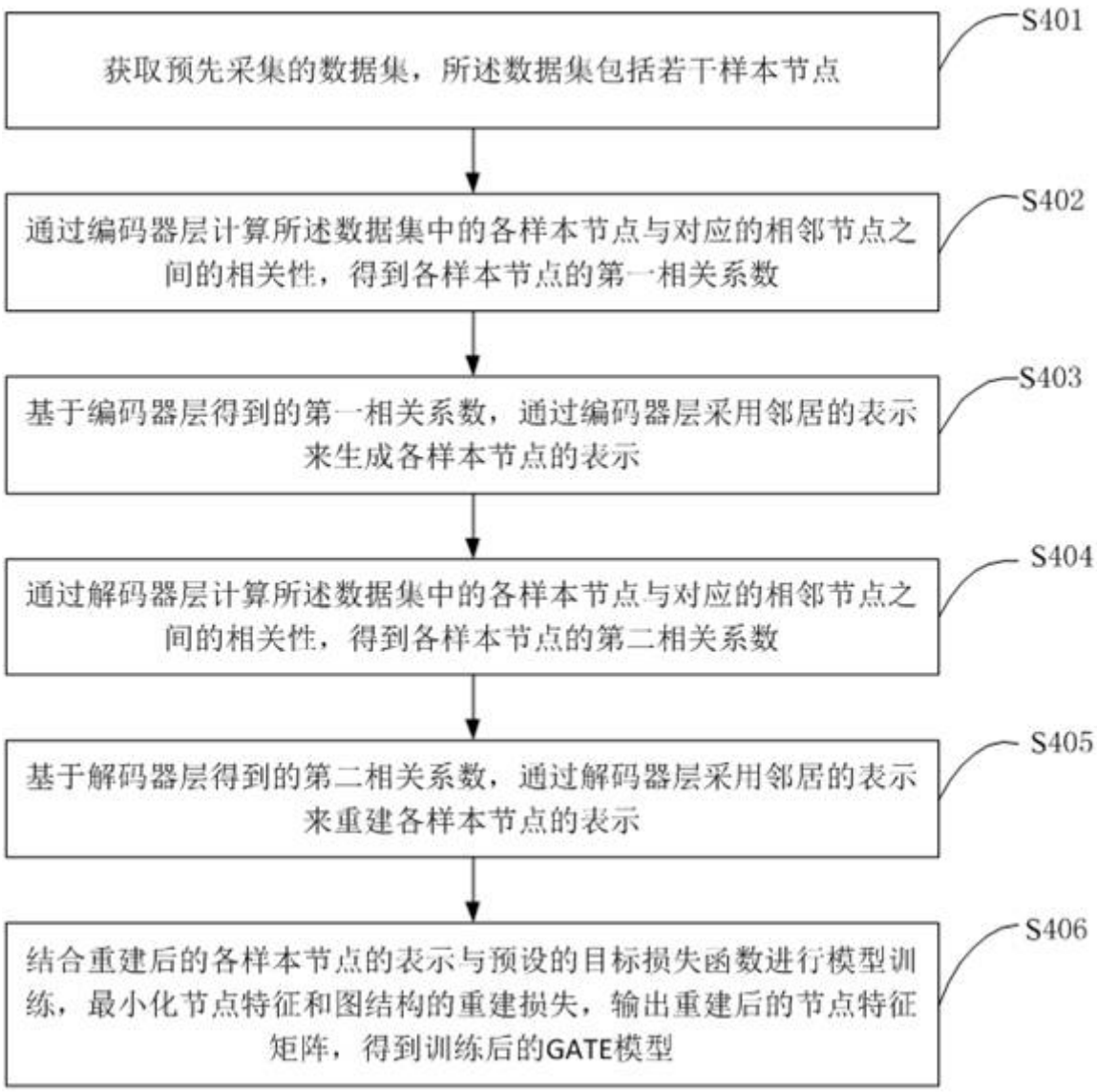

S401
获取预先采集的数据集，所述数据集包括若干样本节点
S402
通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数
S403
基于编码器层得到的第一相关系数，通过编码器层采用邻居的表示来生成各样本节点的表示
S404
通过解码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第二相关系数
S405
基于解码器层得到的第二相关系数，通过解码器层采用邻居的表示来重建各样本节点的表示
S406
结合重建后的各样本节点的表示与预设的目标损失函数进行模型训练，最小化节点特征和图结构的重建损失，输出重建后的节点特征矩阵，得到训练后的GATE模型

图5

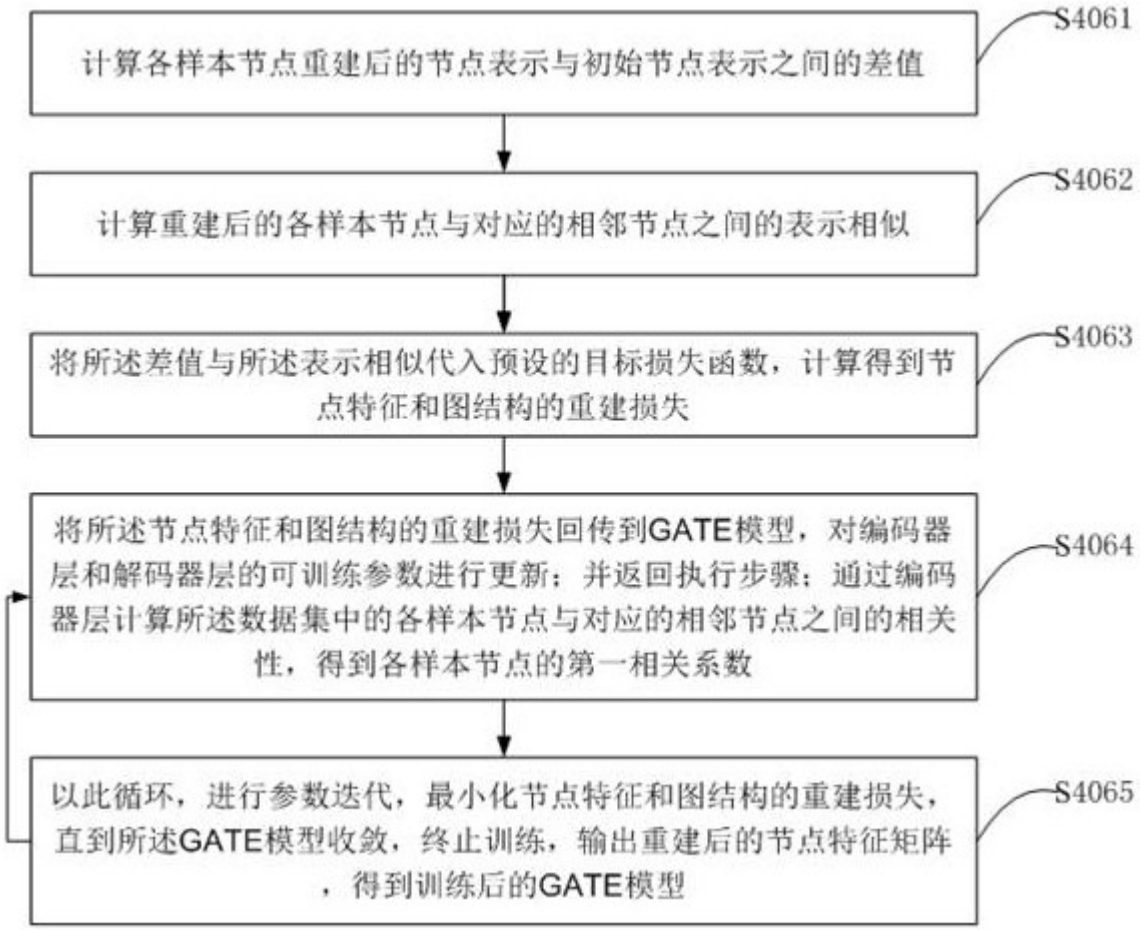

计算各样本节点重建后的节点表示与初始节点表示之间的差值
S4061
计算重建后的各样本节点与对应的相邻节点之间的表示相似
S4062
将所述差值与所述表示相似代入预设的目标损失函数，计算得到节点特征和图结构的重建损失
S4063
将所述节点特征和图结构的重建损失回传到GATE模型，对编码器层和解码器层的可训练参数进行更新；并返回执行步骤：通过编码器层计算所述数据集中的各样本节点与对应的相邻节点之间的相关性，得到各样本节点的第一相关系数
S4064
以此循环，进行参数迭代，最小化节点特征和图结构的重建损失，直到所述GATE模型收敛，终止训练，输出重建后的节点特征矩阵，得到训练后的GATE模型
S4065

图6

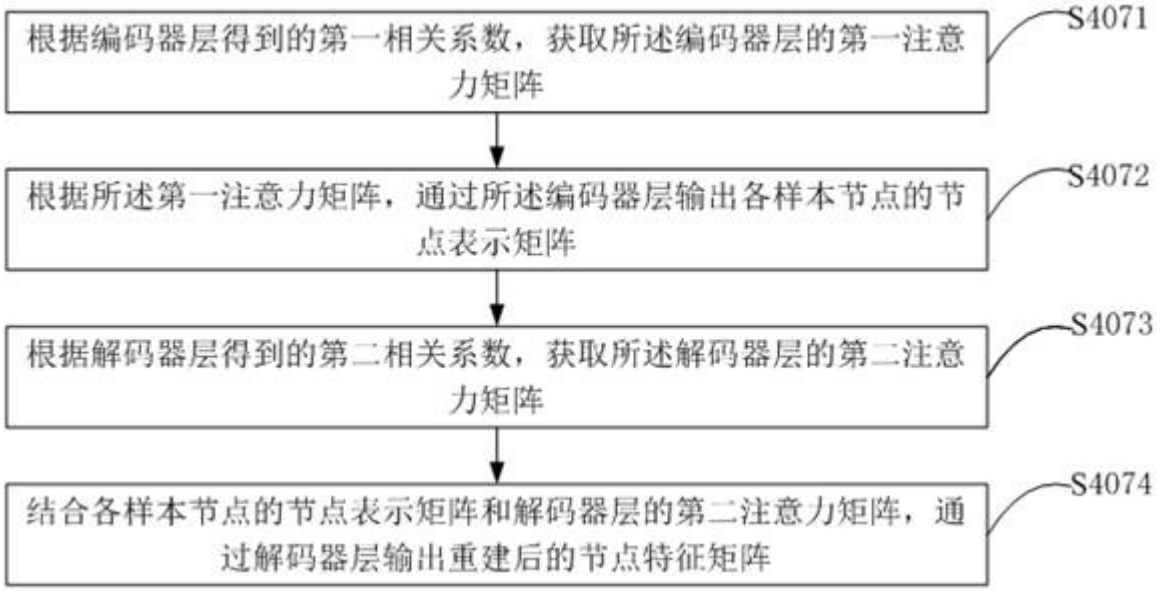

根据编码器层得到的第一相关系数，获取所述编码器层的第一注意力矩阵
S4071
根据所述第一注意力矩阵，通过所述编码器层输出各样本节点的节点表示矩阵
S4072
根据解码器层得到的第二相关系数，获取所述解码器层的第二注意力矩阵
S4073
结合各样本节点的节点表示矩阵和解码器层的第二注意力矩阵，通过解码器层输出重建后的节点特征矩阵
S4074

图7

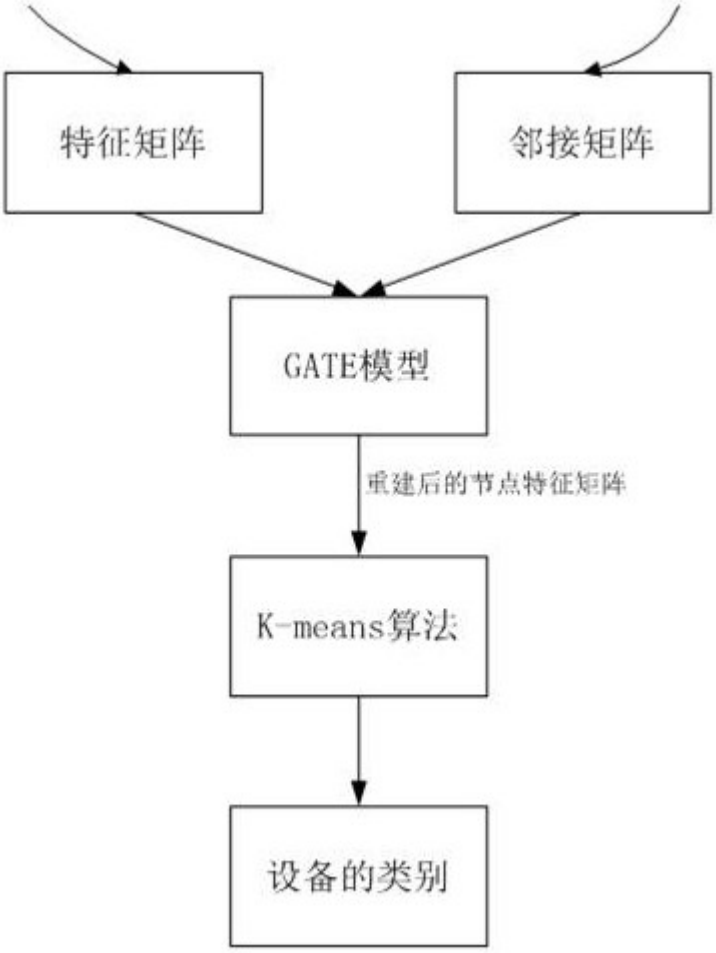

特征矩阵
邻接矩阵
GATE模型
重建后的节点特征矩阵
K-means算法
设备的类别

图8

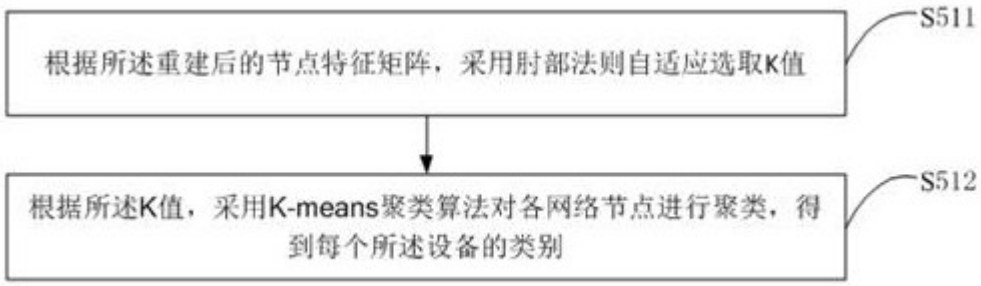

根据所述重建后的节点特征矩阵，采用肘部法则自适应选取K值
S511
根据所述K值，采用K-means聚类算法对各网络节点进行聚类，得到每个所述设备的类别
S512

图9