**专利名称：一种基于贝叶斯时空图变分自编码器的道路行驶时间预测模型**

**申请人 ：南京邮电大学**

**发明人：柯昌博、张永超**

**第一发明人身份证号：**

本发明公开了一种基于贝叶斯时空图和变分自编码器的道路行驶时间预测模型。该模型首先进行数据预处理，包括OSM文件解析以及图反转，以提高数据质量和准确性。接着，模型通过贝叶斯时空图和变分自编码器对预处理后的数据进行深度学习。模型首先将数据表达为时空图结构，通过时空图卷积神经网络，在这个图结构上进行信息的传播和融合，以捕捉道路网络的拓扑关系和空间依赖性。基于时空图卷积的输出，模型使用贝叶斯图变分自编码器进一步处理数据。通过引入贝叶斯理论，可以考虑行驶时间的不确定性，这使得预测结果不仅可以得到预测值，还可以得到预测值的不确定性，从而更好地反映行驶时间的分布特性。最后，模型使用生成对抗网络优化预测模型。在生成对抗网络中，贝叶斯时空图和变分自编码器作为生成器尝试产生看起来与真实行驶时间数据相似的数据，图注意力机制作为判别器的主要结构，其任务是尽可能区分生成的数据和真实数据。通过这种方式，可以进一步优化预测模型，使预测结果更接近真实的行驶时间数据。



1. 一种基于贝叶斯时空图和变分自编码器的道路行驶时间预测模型，包括以下步骤：

步骤 1）数据预处理，为了使模型更准确地学习和预测，需要对原始数据进行一系列预处理，如OSM文件解析以及图反转，然后提取相关特征并对提取的特征做编码处理。

步骤 2）时空图卷积， 在数据预处理后，将清洗后的数据表达为一个时空图结构，每条道路都被表示为图中的一个节点，相邻的道路之间则通过边连接。通过时空图卷积神经网络，可以在这个图结构上进行信息的传播和融合，以捕捉道路网络的拓扑关系和空间依赖性。

步骤 3）贝叶斯时空图变分自编码器，基于时空图卷积的输出，模型使用贝叶斯图变分自编码器进一步处理数据。通过引入贝叶斯理论，可以考虑行驶时间的不确定性，这使得预测结果不仅可以得到预测值，还可以得到预测值的不确定性，从而更好地反映行驶时间的分布特性。

步骤 4）生成对抗网络优化模型，在生成对抗网络中，基于贝叶斯时空图和变分自编码器的道路行驶时间预测模型作为生成器尝试产生看起来与真实行驶时间数据相似的数据，图注意力机制作为判别器的主要部分，任务是尽可能区分生成的数据和真实数据。通过这种方式，可以进一步优化预测模型，使预测结果更接近真实的行驶时间数据。

其中所述步骤1）具体如下：

步骤11）解析OSM文件，OSM文件记录了选定范围内的所有道路、交叉路口、节点、地理区域以及元数据等信息。由于本文只关注路网结构，因此只关注与道路相关的node标签和way标签。node标签的数据包括node唯一标识、经纬度、修改版本、以及更新时间戳等信息。way标签的数据包括way唯一标识、修改版本、更新时间戳路名以及形状等信息。本文将OSM数据解析成路网图结构 ，其中为节点的集合，表示路口， 为边集合，表示道路，每条道路有name、highway、length、lanes以及maxspeed等特征，节点除经纬度位置信息外无其他特征。

步骤 12）反转图，为更好地适应基于贝叶斯时空图变分自编码器的学习与表示，本文将上一节中得到的路网图 进行反转，即将原有的以道路为边、路口为节点的图结构反转为以道路为节点、道路之间连接关系为边的图结构。反转后的图结构可以使模型更好地学习道路特征并利用道路之间的时空关联进行推理。换言之， 更符合模型的学习机制，有助于模型学习交通网络的时空依赖关系。

步骤 13）特征选取与处理，经过分析，发现能够提取的有效特征包括道路长度、道路类型、限速和平均行驶速度等。使用MinMaxScaler将道路特征即道路长度、道路类型以及限速进行归一化，其中道路类型先使用OneHot进行编码，使用StandardScaler对道路之间连接关系的特征即一条道路行驶到另一条道路的时间特征进行归一化。对这些特征进行归一化处理的目的是以消除量纲的影响。特征选取与处理确保了模型能够有效地处理和学习从数据中提取的特征，从而提高预测的准确性。

所述步骤 2）具体如下：

步骤21）构建时空图，为了充分考虑时空关系对道路行驶时间预测的影响，文将数据集构建为时空图（STG）。时空图能够捕捉到道路网络中不同地点和不同时间的关联关系，通过将历史行驶时间与道路网络的拓扑结构相结合，时空图可以更准确地表示不同地点和不同时段之间的交通状况，更好地捕捉到交通流量的时空相关性。首先将所有道路的历史行驶速度数据按时间戳升序排列，每1800秒为一组，再聚合每一组中的所有数据到对应的路网中，如果一条道路有多条记录，则取平均值作为该条道路对应时间戳下的历史行驶速度，如果一条道路没有数据则使用0填充。即生成了多张时空图，其中节点表示道路，其特征分别为长度和限速；边表示道路之间的连接关系，边特征为一条道路到另一条道路的行驶时间。每个时空图的时间间隔为1800秒，即30分钟。这种构建方式既能够体现道路之间的拓扑关系，又能够包含道路的时空动态变化，为模型提供了丰富的信息。

步骤 22）时空图卷积（STGCN），STGCN用于在图结构中捕捉交通网络中的时空依赖关系，包括时序关系以及局部邻居信息。在交通网络行驶时间预测的背景下，本文采用了一种名为FeaStConv的图卷积操作，用于在图结构数据上捕获局部邻域信息。FeaStConv是一种特殊的图神经网络（GNN）操作，其基本思想是将每个节点的特征与其邻居的特征结合起来，以捕获局部结构信息。在交通网络中，FeaStConv可以提取道路网络中各个路段的时空特征。具体来说，对于节点，表示其邻居集合为，其中表示节点是节点的邻居。然后计算节点的聚合信息，由节点的特征以及其邻居的特征组成：

其中，表示第层节点的特征向量，表示节点的邻接节点集合，是节点和之间的归一化权重，是第层的权重矩阵，是激活函数（如ReLU）。最后，将节点的原始特征与新的节点特征结合起来，得到最终的节点特征表示：

FeaStConv的优点在于其具有良好的局部感知能力，在交通网络图中能够捕捉节点之间的空间相关性以及其随时间的动态变化。

所述步骤 3）具体如下：

步骤31）贝叶斯图卷积层（BayesianGCNConv），图卷积神经网络（GCN）通过在图结构上执行局部卷积操作来处理节点特征。GCN的基本操作定义为：

其中，表示第层的节点特征矩阵，是输入图的邻接矩阵加上自连接，是对应的度矩阵，是第层的权重矩阵，是激活函数。为了捕获节点特征的不确定性，引入贝叶斯神经网络的思想，将权重矩阵替换为随机变量，其中是权重矩阵的后验分布。这使得能够对每个权重参数进行概率建模，从而为节点特征提供更丰富的表示。

步骤32）编码器，编码器部分基于贝叶斯图卷积的神经网络，其目的是将输入的时空特征映射到一个潜在空间，从而捕获交通网络中的不确定性信息。编码器包括两个上一节中提出的贝叶斯图卷积层（BayesianGCNConv），贝叶斯图卷积是一种基于图卷积神经网络（GCN）的扩展，它可以在捕获图结构特征的同时，学习到不确定性信息。对于编码器中的第一层贝叶斯图卷积（BayesianGCNConv），其作用是将输入特征（FeaStConv的输出）映射到一个隐藏特征空间。给定一个图G=(V,E)，其中是节点集合，是边集合。对于每个节点i，其输入特征为，可以表示第一层贝叶斯图卷积的操作为：

其中，表示节点的邻居节点集合，和分别是该层的权重矩阵和偏置项，ReLU为激活函数。第二层贝叶斯图卷积的作用是将隐藏特征空间映射到一个潜在空间。这个潜在空间的维度是2倍的潜在空间维度，这是因为需要从这个空间中分离出均值和对数方差。第二层贝叶斯图卷积的操作可以表示为：

其中，和分别是第二层的权重矩阵和偏置项。通过这两层贝叶斯图卷积，BSTVAE模型将输入特征映射到潜在空间，并为每个节点学习到均值（）和对数方差（）。这些参数可以用于衡量交通网络中的不确定性信息，从而提高行驶时间预测的准确性。

步骤33）重新参数化，重参数化操作是变分自编码器（VAE）的关键组成部分，用于将潜在变量的分布参数化为神经网络的输出。这一操作使得能够在保留梯度传播的同时采样潜在变量，从而实现端到端的模型训练。在编码器部分学到了每个节点的均值（）和对数方差（），然后将这些参数作为重参数化操作的输入。重参数化操作可以表示为：

其中，表示节点的潜在变量，和分别表示节点的均值和标准差（），表示元素间乘法。是一个独立的噪声项，通常采样自标准正态分布：。

步骤34）解码器，解码器部分负责从学到的潜在变量恢复出原始输入特征。解码器采用贝叶斯图卷积（BayesianGCNConv）来实现，将潜在变量映射回原始特征空间。解码器包含两层贝叶斯图卷积层。首先，使用第一层贝叶斯GCN操作将潜在变量映射到一个中间表示，并应用ReLU激活函数：

其中，表示第一层贝叶斯图卷积操作。z是重参数化后的潜在变量，和分别表示图结构和边权重。接下来，使用第二层贝叶斯GCN操作将中间表示映射回原始特征空间，得到重构的节点特征，并应用tanh激活函数：

其中，表示第二层贝叶斯图卷积操作。解码器的目标是最小化重构误差，即原始输入特征与重构特征之间的差异。通常情况下，使用均方误差（MSE）损失来衡量这一差异：

在BSTVAE模型中，解码器的作用是重建交通网络的节点特征，并通过学习潜在空间的结构，捕捉交通网络中的时空依赖关系。这使得模型能够有效地预测交通网络中的行驶时间。边缘行驶时间预测部分是一个重要组成部分，它利用从节点特征中提取的信息来预测交通网络中的行驶时间。具体来说，模型通过计算边缘的两个相邻节点的特征差值来获取边缘特征。然后，这些边缘特征被输入到一个线性层，以预测交通网络中各边缘的行驶时间。首先，从节点特征中计算边缘特征：

其中，和分别表示边缘连接的两个节点的特征。接下来，将计算得到的边缘特征输入到一个线性层中，得到预测的行驶时间：

其中，Linear表示线性层。表示预测的行驶时间。边缘行驶时间预测部分的目标是最小化预测行驶时间与真实行驶时间之间的差异。通常情况下，使用均方误差（MSE）损失来衡量这一差异：

在BSTVAE模型中，边缘行驶时间预测部分的作用是学习交通网络中的时空依赖关系，并利用这些关系预测行驶时间。这使得模型能够有效地应用于交通网络行驶时间预测任务。

所述步骤 4）具体如下：

步骤41）生成器，生成器使用BSTVAE作为基本结构。BSTVAE包括时空图卷积网络（STGCN），贝叶斯图卷积编码器，重参数化操作，贝叶斯图卷积解码器以及边缘行驶时间预测部分。生成器接收交通网络中节点的特征作为输入，输出预测的行驶时间。

步骤42）判别器，判别器采用图注意力网络（GAT）作为基本结构。GAT可以有效地捕获节点之间的复杂依赖关系，提高模型在交通网络行驶时间预测任务中的性能。判别器D包括两个GAT层和一个全连接层。判别器的输入是节点特征和边特征（和），输出是判别器对输入样本的判断结果，即样本为真实数据的概率。

判别器的结构可以表示为：

其中，表示判别器的参数。在交通网络行驶时间预测任务中， 在本研究中将生成器（）和判别器（）进行交替训练，通过最小化生成器生成的数据与真实数据之间的差异来优化生成器。GANs的训练目标函数可以表示为：

其中，表示期望值，表示真实数据，表示随机噪声，表示生成器生成的数据。

在训练过程中，判别器D试图最大化对真实数据和生成数据的分类正确率，而生成器试图最小化判别器对生成数据的分类误差。通过这种相互竞争的训练过程，生成器和判别器将不断提高，最终使生成器生成的数据分布趋近于真实数据分布。

**一****种基于贝叶斯时空图变分自编码器的道路行驶时间预测模型**

**技术领域**

本发明提出一种基于贝叶斯时空图变分自编码器的道路行驶时间预测模型，属于深度学习、智慧交通等交叉技术领域。

**背景技术**

在现代城市中，交通拥堵问题日益严重，这对城市的可持续发展构成了巨大挑战。为了解决这个问题，交通行驶时间预测成为了一个重要的研究课题。准确的行驶时间预测可以为交通管理部门提供决策支持，为驾驶员提供行驶建议，从而提高道路网络的整体效率。然而，由于交通流量、道路条件、天气条件等多种因素的影响，道路行驶时间预测具有很大的不确定性和复杂性。

近年来，深度学习技术在许多领域取得了显著的成果，其中包括交通流量预测。深度学习模型，如图卷积神经网络（CNN）和循环神经网络（RNN），已经被成功应用于交通流量预测。Zhang等人使用大数据和深度学习方法研究了交通预测问题，他们提出了一种基于机器学习的行驶时间预测方法，通过结合交通流特性和大数据技术，实现了高精度的行驶时间预测。研究结果表明，所提出的方法在预测精度和实时性方面均优于现有的基于统计的行驶时间预测方法。Lv等人也采用了深度学习方法来预测交通流量，他们提出了一种基于深度置信网络（DBN）的行驶时间预测方法，通过对大量历史数据进行学习，实现了对城市道路行驶时间的高精度预测。这项研究为交通流预测提供了一种新颖的机器学习方法。Polson和Sokolov在这篇论文中提出了一种基于深度学习的短期交通流量预测方法，他们使用递归神经网络（RNN）对城市道路网络的交通流量进行建模，实现了对行驶时间的高精度预测。这项研究为短期交通流量预测提供了一种有效的机器学习方法。然而，这些模型通常假设交通流量数据是独立同分布的，忽略了交通流量数据的时空依赖性。此外，这些模型通常只能处理规则网格数据，而不能直接处理交通网络数据。

为了解决这些问题，研究人员提出了图神经网络（GNN）模型。GNN模型可以直接处理图结构数据，能够捕捉交通网络中的时空依赖性。然而，现有的GNN模型通常假设图结构是静态的，忽略了交通网络的动态性。此外，现有的GNN模型通常只考虑了交通网络的空间依赖性，忽略了交通网络的时间依赖性。尽管现有的深度学习模型和图神经网络模型在交通流量预测方面取得了一定的成果，但仍存在一些问题需要进一步研究。例如，如何更好地捕捉交通网络的动态性，如何更好地处理交通流量数据的不确定性，如何更好地利用交通网络的结构信息等。这些问题的解决将有助于进一步提高道路行驶时间预测的准确性，推动智能交通系统的发展。

**发明内容**

**技术问题：**发明所要解决的技术问题是设计一种能够准确预测道路行驶时间的预测模型，提高道路行驶时间预测的准确性。

**技术方案：**本发明基于贝叶斯时空图变分自编码器的道路行驶时间预测模型包括以下步骤：

步骤 1）数据预处理，为了使模型更准确地学习和预测，需要对原始数据进行一系列预处理，如OSM文件解析以及图反转，然后提取相关特征并对提取的特征做编码处理。

步骤 2）时空图卷积， 在数据预处理后，将清洗后的数据表达为一个时空图结构，每条道路都被表示为图中的一个节点，相邻的道路之间则通过边连接。通过时空图卷积神经网络，可以在这个图结构上进行信息的传播和融合，以捕捉道路网络的拓扑关系和空间依赖性。

步骤 3）贝叶斯时空图变分自编码器，基于时空图卷积的输出，模型使用贝叶斯图变分自编码器进一步处理数据。通过引入贝叶斯理论，可以考虑行驶时间的不确定性，这使得预测结果不仅可以得到预测值，还可以得到预测值的不确定性，从而更好地反映行驶时间的分布特性。

步骤 4）生成对抗网络优化模型，在生成对抗网络中，基于贝叶斯时空图和变分自编码器的道路行驶时间预测模型作为生成器尝试产生看起来与真实行驶时间数据相似的数据，图注意力机制作为判别器的主要部分，任务是尽可能区分生成的数据和真实数据。通过这种方式，可以进一步优化预测模型，使预测结果更接近真实的行驶时间数据。

其中所述步骤1）具体如下：

步骤11）解析OSM文件，OSM文件记录了选定范围内的所有道路、交叉路口、节点、地理区域以及元数据等信息。由于本文只关注路网结构，因此只关注与道路相关的node标签和way标签。node标签的数据包括node唯一标识、经纬度、修改版本、以及更新时间戳等信息。way标签的数据包括way唯一标识、修改版本、更新时间戳路名以及形状等信息。本文将OSM数据解析成路网图结构 ，其中为节点的集合，表示路口， 为边集合，表示道路，每条道路有name、highway、length、lanes以及maxspeed等特征，节点除经纬度位置信息外无其他特征。

步骤 12）反转图，为更好地适应基于贝叶斯时空图变分自编码器的学习与表示，本文将上一节中得到的路网图 进行反转，即将原有的以道路为边、路口为节点的图结构反转为以道路为节点、道路之间连接关系为边的图结构。反转后的图结构可以使模型更好地学习道路特征并利用道路之间的时空关联进行推理。换言之， 更符合模型的学习机制，有助于模型学习交通网络的时空依赖关系。

步骤 13）特征选取与处理，经过分析，发现能够提取的有效特征包括道路长度、道路类型、限速和平均行驶速度等。使用MinMaxScaler将道路特征即道路长度、道路类型以及限速进行归一化，其中道路类型先使用OneHot进行编码，使用StandardScaler对道路之间连接关系的特征即一条道路行驶到另一条道路的时间特征进行归一化。对这些特征进行归一化处理的目的是以消除量纲的影响。特征选取与处理确保了模型能够有效地处理和学习从数据中提取的特征，从而提高预测的准确性。

所述步骤 2）具体如下：

步骤21）构建时空图，为了充分考虑时空关系对道路行驶时间预测的影响，文将数据集构建为时空图（STG）。时空图能够捕捉到道路网络中不同地点和不同时间的关联关系，通过将历史行驶时间与道路网络的拓扑结构相结合，时空图可以更准确地表示不同地点和不同时段之间的交通状况，更好地捕捉到交通流量的时空相关性。首先将所有道路的历史行驶速度数据按时间戳升序排列，每1800秒为一组，再聚合每一组中的所有数据到对应的路网中，如果一条道路有多条记录，则取平均值作为该条道路对应时间戳下的历史行驶速度，如果一条道路没有数据则使用0填充。即生成了多张时空图，其中节点表示道路，其特征分别为长度和限速；边表示道路之间的连接关系，边特征为一条道路到另一条道路的行驶时间。每个时空图的时间间隔为1800秒，即30分钟。这种构建方式既能够体现道路之间的拓扑关系，又能够包含道路的时空动态变化，为模型提供了丰富的信息。

步骤 22）时空图卷积（STGCN），STGCN用于在图结构中捕捉交通网络中的时空依赖关系，包括时序关系以及局部邻居信息。在交通网络行驶时间预测的背景下，本文采用了一种名为FeaStConv的图卷积操作，用于在图结构数据上捕获局部邻域信息。FeaStConv是一种特殊的图神经网络（GNN）操作，其基本思想是将每个节点的特征与其邻居的特征结合起来，以捕获局部结构信息。在交通网络中，FeaStConv可以提取道路网络中各个路段的时空特征。具体来说，对于节点，表示其邻居集合为，其中表示节点是节点的邻居。然后计算节点的聚合信息，由节点的特征以及其邻居的特征组成：

其中，表示第层节点的特征向量，表示节点的邻接节点集合，是节点和之间的归一化权重，是第层的权重矩阵，是激活函数（如ReLU）。最后，将节点的原始特征与新的节点特征结合起来，得到最终的节点特征表示：

FeaStConv的优点在于其具有良好的局部感知能力，在交通网络图中能够捕捉节点之间的空间相关性以及其随时间的动态变化。

所述步骤 3）具体如下：

步骤31）贝叶斯图卷积层（BayesianGCNConv），图卷积神经网络（GCN）通过在图结构上执行局部卷积操作来处理节点特征。GCN的基本操作定义为：

其中，表示第层的节点特征矩阵，是输入图的邻接矩阵加上自连接，是对应的度矩阵，是第层的权重矩阵，是激活函数。为了捕获节点特征的不确定性，引入贝叶斯神经网络的思想，将权重矩阵替换为随机变量，其中是权重矩阵的后验分布。这使得能够对每个权重参数进行概率建模，从而为节点特征提供更丰富的表示。

步骤32）编码器，编码器部分基于贝叶斯图卷积的神经网络，其目的是将输入的时空特征映射到一个潜在空间，从而捕获交通网络中的不确定性信息。编码器包括两个上一节中提出的贝叶斯图卷积层（BayesianGCNConv），贝叶斯图卷积是一种基于图卷积神经网络（GCN）的扩展，它可以在捕获图结构特征的同时，学习到不确定性信息。对于编码器中的第一层贝叶斯图卷积（BayesianGCNConv），其作用是将输入特征（FeaStConv的输出）映射到一个隐藏特征空间。给定一个图G=(V,E)，其中是节点集合，是边集合。对于每个节点i，其输入特征为，可以表示第一层贝叶斯图卷积的操作为：

其中，表示节点的邻居节点集合，和分别是该层的权重矩阵和偏置项，ReLU为激活函数。第二层贝叶斯图卷积的作用是将隐藏特征空间映射到一个潜在空间。这个潜在空间的维度是2倍的潜在空间维度，这是因为需要从这个空间中分离出均值和对数方差。第二层贝叶斯图卷积的操作可以表示为：

其中，和分别是第二层的权重矩阵和偏置项。通过这两层贝叶斯图卷积，BSTVAE模型将输入特征映射到潜在空间，并为每个节点学习到均值（）和对数方差（）。这些参数可以用于衡量交通网络中的不确定性信息，从而提高行驶时间预测的准确性。

步骤33）重新参数化，重参数化操作是变分自编码器（VAE）的关键组成部分，用于将潜在变量的分布参数化为神经网络的输出。这一操作使得能够在保留梯度传播的同时采样潜在变量，从而实现端到端的模型训练。在编码器部分学到了每个节点的均值（）和对数方差（），然后将这些参数作为重参数化操作的输入。重参数化操作可以表示为：

其中，表示节点的潜在变量，和分别表示节点的均值和标准差（），表示元素间乘法。是一个独立的噪声项，通常采样自标准正态分布：。

步骤34）解码器，解码器部分负责从学到的潜在变量恢复出原始输入特征。解码器采用贝叶斯图卷积（BayesianGCNConv）来实现，将潜在变量映射回原始特征空间。解码器包含两层贝叶斯图卷积层。首先，使用第一层贝叶斯GCN操作将潜在变量映射到一个中间表示，并应用ReLU激活函数：

其中，表示第一层贝叶斯图卷积操作。z是重参数化后的潜在变量，和分别表示图结构和边权重。接下来，使用第二层贝叶斯GCN操作将中间表示映射回原始特征空间，得到重构的节点特征，并应用tanh激活函数：

其中，表示第二层贝叶斯图卷积操作。解码器的目标是最小化重构误差，即原始输入特征与重构特征之间的差异。通常情况下，使用均方误差（MSE）损失来衡量这一差异：

在BSTVAE模型中，解码器的作用是重建交通网络的节点特征，并通过学习潜在空间的结构，捕捉交通网络中的时空依赖关系。这使得模型能够有效地预测交通网络中的行驶时间。边缘行驶时间预测部分是一个重要组成部分，它利用从节点特征中提取的信息来预测交通网络中的行驶时间。具体来说，模型通过计算边缘的两个相邻节点的特征差值来获取边缘特征。然后，这些边缘特征被输入到一个线性层，以预测交通网络中各边缘的行驶时间。首先，从节点特征中计算边缘特征：

其中，和分别表示边缘连接的两个节点的特征。接下来，将计算得到的边缘特征输入到一个线性层中，得到预测的行驶时间：

其中，Linear表示线性层。表示预测的行驶时间。边缘行驶时间预测部分的目标是最小化预测行驶时间与真实行驶时间之间的差异。通常情况下，使用均方误差（MSE）损失来衡量这一差异：

在BSTVAE模型中，边缘行驶时间预测部分的作用是学习交通网络中的时空依赖关系，并利用这些关系预测行驶时间。这使得模型能够有效地应用于交通网络行驶时间预测任务。

所述步骤 4）具体如下：

步骤41）生成器，生成器使用BSTVAE作为基本结构。BSTVAE包括时空图卷积网络（STGCN），贝叶斯图卷积编码器，重参数化操作，贝叶斯图卷积解码器以及边缘行驶时间预测部分。生成器接收交通网络中节点的特征作为输入，输出预测的行驶时间。

步骤42）判别器，判别器采用图注意力网络（GAT）作为基本结构。GAT可以有效地捕获节点之间的复杂依赖关系，提高模型在交通网络行驶时间预测任务中的性能。判别器D包括两个GAT层和一个全连接层。判别器的输入是节点特征和边特征（和），输出是判别器对输入样本的判断结果，即样本为真实数据的概率。

判别器的结构可以表示为：

其中，表示判别器的参数。在交通网络行驶时间预测任务中， 在本研究中将生成器（）和判别器（）进行交替训练，通过最小化生成器生成的数据与真实数据之间的差异来优化生成器。GANs的训练目标函数可以表示为：

其中，表示期望值，表示真实数据，表示随机噪声，表示生成器生成的数据。

在训练过程中，判别器D试图最大化对真实数据和生成数据的分类正确率，而生成器试图最小化判别器对生成数据的分类误差。通过这种相互竞争的训练过程，生成器和判别器将不断提高，最终使生成器生成的数据分布趋近于真实数据分布。

。

**有益效果：**本发明采用的以上技术方案与现有技术相比，具有以下技术效果：

1. 更准确的预测，通过使用贝叶斯时空图，我们的模型能够更好地捕捉交通网络的动态性和时空依赖性，从而提高道路行驶时间预测的准确性。
2. 更强的泛化能力，我们的模型使用变分自编码器来学习交通流量数据的潜在表示，这使得模型具有更强的泛化能力，能够处理不同的交通流量数据。
3. 更高的效率，我们的模型使用深度学习技术，可以利用现代硬件设备（如GPU）进行高效计算，从而提高预测的效率。
4. 更好的解释性，我们的模型不仅可以预测道路行驶时间，还可以生成贝叶斯时空图，这使得我们的模型具有更好的解释性，可以帮助人们理解交通网络的动态性和时空依赖性。
5. 更广泛的应用范围，我们的模型不仅可以用于道路行驶时间预测，还可以用于其他交通网络分析任务，如交通流量预测、交通拥堵检测等，具有更广泛的应用范围。

**附图说明**

图1是本发明的整体流程。

### 具体实施方式

下面结合附图对本发明的技术方案做进一步的详细说明：

在具体实施中，我们选取的GPS轨迹数据坐标系为WGS84坐标系，路网数据为OpenStreetMap提供的OSM格式的路网数据。首先对以上数据进行预处理，GPS轨迹点需删除车辆未移动或移动距离很短的点，因为这种情况下GPS设备采集的数据集中在某一个区域呈现不规则形状，采样点之间的位置信息为概率性波动，这会干扰采样点间的相互影响，导致后续的计算中产生大量噪声且增加计算时长；路网数据需简化结构，合并同一十字路口的所有交叉点视为同一交叉点，降低路网中交点个数。在数据处理完成后，将路网数据与GPS轨迹数据通过转换公式转换为三维笛卡尔坐标，然后将转换后的路网数据构建三维KD树，通过KNN算法在该树中查找每个GPS轨迹点的候选路段，根据每个采样点的若干候选路段，将GPS轨迹点与其对应的若干候选路段通过web墨卡托投影计算出对应的候选点。此时得到了每个GPS轨迹点对应的若干候选点，通过位置上下文分析中时间分析、空间分析以及道路分析，计算出相邻候选点的权重，在计算过程中，舍弃不满足三个约束条件的路径，约束1保证了后续参与运算的数据都符合实际路网结构，不会出现不存在的路径；约束2保证了不会出现下述情况：如果错误的候选路径的累积权重大于前几个正确候选路径的权重，即便在后续候选中权值逐渐减小，但最终累积权重有概率大于正确路径的权重，导致生成错误的结果；约束3由于考虑了路网的三维结构，通过限速与车辆行驶速度的关系提高了匹配候选路径的准确性。以上三点约束能够在参与后续计算之前过滤掉大部分的错误路径，减少运算时间。连续的GPS采样点之间会相互影响，通过对采样点之间进行相互影响建模，得到距离权重矩阵，然后再创建候选图，每条边都拥有累积权重与投票数属性。根据生成的候选图，通过公式计算累计权重，是的权重，受和的距离影响，通过计算累计权重得到了每个候选点的局部最优路径。最后遍历每个候选点的局部最优路径对每条路径进行投票，票数最多的路径为最优匹配路径。



图1 整体流程

文本

描述已自动生成

图2 查找局部最优路径算法

文本, 信件

描述已自动生成

图3 侯选边投票算法