



(12) 发明专利申请

(10) 申请公布号 CN 116934380 A

(43) 申请公布日 2023. 10. 24

(21) 申请号 202310894174.6

(22) 申请日 2023.07.20

(71) 申请人 东南大学

地址 210096 江苏省南京市玄武区四牌楼2号

(72) 发明人 王帅 卢志强 林丽 梅洛瑜

(74) 专利代理机构 南京众联专利代理有限公司  
32206

专利代理师 叶倩

(51) Int.Cl.

G06Q 30/0202 (2023.01)

G06Q 30/0203 (2023.01)

G06Q 30/0204 (2023.01)

G06Q 10/083 (2023.01)

G06F 16/35 (2019.01)

G06F 40/30 (2020.01)

G06F 18/25 (2023.01)

G06F 18/27 (2023.01)

G06N 3/042 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/0455 (2023.01)

G06N 3/0442 (2023.01)

G06N 3/048 (2023.01)

G06N 3/084 (2023.01)

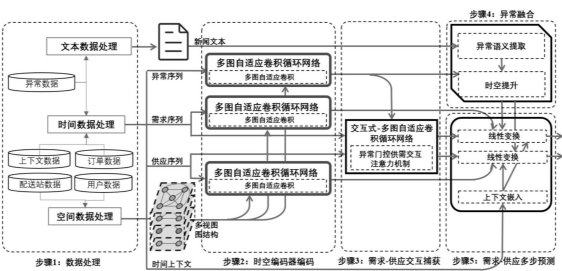
权利要求书4页 说明书14页 附图1页

(54) 发明名称

一种异常事件下的电商物资供需联合预测方法

(57) 摘要

本发明公开了一种异常事件下的电商物资供需联合预测方法,首先,对数据进行预处理,构建多视图的图结构,生成需求、供应、异常时序样本及非序列的辅助特征;其次,设计多图自适应卷积网络模块MGACR并嵌入GRU中作为时空编码器,对序列进行编码,捕获时空关系;再次,使用异常门控注意力机制来捕获需求与供应之间的相互作用;然后,引入预训练Bert增强对文本异常数据的语义理解,并将其经过时空提升进行融合表征;最后,汇集上述步骤表征并做多步解码。本发明解决了需求与供应时空关系的复杂性、多源异构异常数据影响的多样性两点挑战,有利于电商平台在异常事件下针对性地对各区域提供物资物流双重保障,解决各区域物资供需不协调的问题。



1. 一种异常事件下的电商物资供需联合预测方法,其特征在于,所述方法包括如下步骤:

S1,数据预处理:构建多视图的图结构,将图区域作为图结构的节点,将配送站间的距离、配送站区域是否邻接、配送站区域的路区类型相似程度、配送站区域的人群类型相似程度四种关系作为可选的图结构的边,生成时间窗对齐的需求、供应和异常时间序列样本,并使其序列长度保持一致;同时提取了非序列的时间上下文和新闻文本辅助特征;

S2,时空编码器编码:将步骤S1获得的需求序列、供应序列和异常序列输入到各自的多图自适应卷积循环网络MGACR中捕获三者的时空关系,并生成需求嵌入、供应嵌入和异常嵌入;所述多图自适应卷积循环网络的单元具体为:

$$Z_t^1 = \mathbf{MGAC}([X_t || h_{t-1}], G^*)$$

$$u_t = \sigma(Z_t^1 W_u + b_u)$$

$$r_t = \sigma(Z_t^1 W_r + b_r)$$

$$Z_t^2 = \mathbf{MGAC}([X_t || (r_t \odot h_{t-1})], G^*)$$

$$c_t = \tanh(Z_t^2 W_c + b_c)$$

$$h_t = u_t \odot h_{t-1} + (1 - u_t) \odot c_t$$

其中 $X_t \in \mathbb{R}^{N \times F}$ ,  $h_t \in \mathbb{R}^{N \times d_h}$ 表示时间步t时的输入与输出, $u_t$ 与 $r_t$ 分别是时间步t时的更新门和重置门, $G^*$ 表示多视图的图结构, $W_u, W_r, W_c, b_u, b_r, b_c$ 以及多图自适应卷积循环网络MGACR中的参数是可学习的;

S3,需求-供应交互捕获:将需求序列、供应序列及步骤S2捕获的异常嵌入输入到交互式-多图自适应卷积循环网络I-MGACR,学习需求和供应序列的交互关系:

$$h_t^{D_I}, h_t^{S_I} = \mathbf{I-MGACR}(D_t, S_t, h_{t-1}^{D_I}, h_{t-1}^{S_I}, h_{t-1}^A, G_D^*, G_S^*)$$

其中 $h_t^{D_I}$ 表示时间步t时交互的需求嵌入, $h_t^{S_I}$ 表示时间步t时交互的供应嵌入, $D_t$ 表示时间步t时的需求, $S_t$ 表示时间步t时的供应, $h_{t-1}^A$ 表示时间步t-1时异常的时空嵌入, $G_D^*$ 表示需求的多视图图结构, $G_S^*$ 表示供应的多视图图结构;

S4,异常融合:引入预训练的Bert模型,将获取的新闻文本输入到异常融合模块进行异常语义提取,并使用步骤S2捕获的异常嵌入进行时空提升,从而获取异常对需求影响的融合嵌入和异常对供应影响的融合嵌入;

S5,需求-供应多步预测:提取时间上下文信息形成嵌入后,通过连接步骤S2、步骤S3和步骤S4的输出表征,得到历史需求和供应各自的融合表征,通过两个线性变换,生成目标时间窗口长度内未来的需求和供应序列。

2. 根据权利要求1所述的一种异常事件下的电商物资供需联合预测方法,其特征在于:所述步骤S1中构建的四种图结构分别为距离感知图 $G^d = (V, E^{(d)})$ 、邻居感知图 $G^{(n)} = (V, E^{(n)})$ 、类型相似图 $G^{(t)} = (V, E^{(t)})$ 、人群相似图 $G^{(c)} = (V, E^{(c)})$ (;其中V表示区域节点, $E^{(d)}$ 表示节点之间的距离邻接矩阵, $E^{(n)}$ 表示节点之间是否相邻的邻接矩阵, $E^{(t)}$ 表示节点之间路区特

征余弦相似度的邻接矩阵,  $E^{(c)}$  表示节点之间人群特征余弦相似度的邻接矩阵。

3. 根据权利要求2所述的一种异常事件下的电商物资供需联合预测方法, 其特征在于: 所述步骤S2中, 对于多视图的图结构  $G^*$ , 集成了  $m$  个视图, 相应地设计  $m$  个独立的GCN层, 用以生成预定义图结构的节点嵌入表征:

$$Z^{(m)} = \left( I_N + D^{-\frac{1}{2}} A^{(m)} D^{-\frac{1}{2}} \right) X \Theta^{(m)} + b^{(m)}$$

残差自适应图结构学习, 与残差自适应图结构的节点嵌入表征生成:

$$\widetilde{E}^r = \text{Softmax}(\text{Relu}(M^r M^{rT}))$$

$$Z^r = \widetilde{E}^r (X^i - H^b) W^r + b^r$$

其中, Relu被用于缓解弱连接, Softmax被用于归一化自适应图结构,  $Z^r \in R^{N \times d_o}$  表示残差图节点嵌入,  $N$  是节点数量,  $d_o$  是残差图节点嵌入的维度;

全量自适应图结构学习, 与全量自适应图结构的节点嵌入表征生成:

$$\widetilde{E}^f = \text{Softmax}(\text{Relu}(M^f M^{fT}))$$

$$Z^f = \widetilde{E}^f X^i W^f + b^f$$

其中,  $Z^f \in R^{N \times d_o}$  表示全量图节点嵌入,  $N$  是节点数量,  $d_o$  是全量图节点嵌入的维度;

为实现多视图之间的信息共享, 使用自注意力层将预定义图结构与残差自适应图结构的节点嵌入进行融合表征, 公式如下:

$$Z^s = \text{Softmax} \left( \frac{[Z^p || Z^r] W^Q ([Z^p || Z^r] W^K)^T}{\sqrt{d_k}} \cdot [Z^p || Z^r] W^V \right)$$

其中,  $Z^s \in R^{N \times d_o}$  表示自注意力层的输出,  $W^Q, W^K, W^V \in R^{(m+1)d_o \times d_o}$  是可学习的Query、Key、Value矩阵,  $d_k = d_o$  是  $W^K$  的维度;

最终得到多图自适应卷积的输出:

$$Z = \sigma[Z^s || Z^f] W + b.$$

4. 根据权利要求3所述的一种异常事件下的电商物资供需联合预测方法, 其特征在于: 所述步骤S3中, 交互式-多图自适应卷积循环网络I-MGACR单元包括两个MGACR单元和一个InterModule模块, 两个MGACR单元用以捕获需求-供应交互序列的时空信息, 并受InterModule的影响进行信息交互:

$$h_t^{D_I'} = \text{MGACR}(D_t, h_{t-1}^{D_I}, G_D^*)$$

$$h_t^{S_I'} = \text{MGACR}(S_t, h_{t-1}^{S_I}, G_S^*)$$

$$h_t^{D_I}, h_t^{S_I} = \text{InterModule} \left( h_t^{D_I'}, h_t^{S_I'}, h_{t-1}^A \right)$$

其中  $h_t^{D_I'}$  表示时间步  $t$  时InterModule模块的交互的需求嵌入输入,  $h_t^{S_I'}$  表示时间步  $t$  时InterModule模块的交互的供应嵌入输入。

5. 根据权利要求4所述的一种异常事件下的电商物资供需联合预测方法,其特征在于:所述步骤S4中,InterModule模块包括异常门控的供需交互注意力机制,首先计算并生成需求-供应的自注意力嵌入:

$$\hat{h}_t^D || \hat{h}_t^S = \text{Softmax} \left( \frac{[h_t^{D_I'} || h_t^{S_I'}] W_1^Q \left( [h_t^{D_I'} || h_t^{S_I'}] W_1^K \right)^T}{\sqrt{d_k}} \cdot [h_t^{D_I'} || h_t^{S_I'}] W_1^V \right)$$

其中 $W_1^Q, W_1^K, W_1^V$ 是自注意力机制的参数权重;

其次,使用交叉注意力机制来计算需求-供应受异常影响的程度,并生成受异常影响的需求-供应的交叉注意力嵌入:

$$\hat{h}_t^A = \text{Softmax} \left( \frac{h_t^A W_2^Q ([h_t^{D_I'} || h_t^{S_I'}] W_2^K)^T}{\sqrt{d_k}} \cdot [h_t^{D_I'} || h_t^{S_I'}] W_2^V \right)$$

其中 $W_2^Q, W_2^K, W_2^V$ 是交叉注意力机制的参数权重;

最后,使用异常门控来控制上述自注意力机制、交叉注意力机制产生的嵌入比例,以形成最终的交互需求嵌入、交互供应嵌入:

$$h_t^{D_I} = \sigma(h_t^A) \odot \hat{h}_t^A + (1 - \sigma(h_t^A)) \odot \hat{h}_t^D$$

$$h_t^{S_I} = \sigma(h_t^A) \odot \hat{h}_t^A + (1 - \sigma(h_t^A)) \odot \hat{h}_t^S$$

其中 $\sigma$ 表示sigmoid函数, $\odot$ 表示矩阵的元素乘积。

6. 根据权利要求1所述的一种异常事件下的电商物资供需联合预测方法,其特征在于:所述步骤S4中,基于语义Bert的异常融合模块分为两个子模块:异常语义提取和时空增强; Bert首先由一个简单的MLP层将生成的嵌入映射为是否异常的二分类问题:

$$E_{FN} \xleftarrow{C} \text{BERT}(\text{Tokens}_{FN})$$

$$\hat{y} = \text{MLP}(E_{FN})$$

其中 $E_{FN}$ 表示微调的嵌入,由Bert生成的C是分类器token[CLS]的表征, $\text{Tokens}_{FN}$ 表示微调新闻FN转化得到的Tokens;

在上述模型收敛之后,得到了一个具备识别异常新闻能力的微调好的Bert,然后将该模型的参数部分冻结,并传递给需求-供应预测任务,从Bert的输出中,使用第一个Token([CLS])的最后一个隐藏层状态C来表示文本输入 $\text{PN}_T$ 中的每一条新闻标题的聚合语义信息,如下所示:

$$E_p \xleftarrow{C} \text{BERT}_f(\text{Tokens}_{PN_T})$$

其中 $E_p$ 表示条级新闻语义嵌入, $BERT_f$ 表示初始化后固定参数的Bert模型, $Tokens_{PN_T}$ 表示历史最近的预测任务新闻 $PN_T$ 转化得到的Tokens;

然后,使用微调好的MLP中的Softmax来计算得到天级的语义表征:

$$\hat{y}' = MLP(E_p)|_{label=1}$$

$$E_d = \sum_{i=1}^{N_{news}} \hat{y}_i' \odot E_{p_i}$$

其中 $\hat{y}'$ 表示 $E_p$ 的权重, $E_d$ 表示融合的天级新闻语义嵌入, $N_{news}$ 表示一天中最大的新闻条数。

7.根据权利要求6所述的一种异常事件下的电商物资供需联合预测方法,其特征在于:所述步骤S4中设计了三个FC全连接层来建模异常的不同影响:

$$h_T^{AD} = FC_1([h_T^A \parallel N * FC_3(E_d)])$$

$$h_T^{AS} = FC_2([h_T^A \parallel N * FC_3(E_d)])$$

其中 $h_T^{AD}$ 表示异常对需求影响的融合嵌入, $h_T^{AS}$ 表示异常对供应影响的融合嵌入。

8.根据权利要求5或7所述的一种异常事件下的电商物资供需联合预测方法,其特征在于:所述步骤S5具体为:

首先生成了上下文信息的嵌入来捕获需求和供应序列的周期性:

$$h_T^C = N * MLP([FC(DoW_T) \parallel FC(FT_T)])$$

其中 $DoW_T$ 表示对历史最近一天是一周中第几天的编码, $FT_T$ 表示对历史最近一天的节假日类型的编码;

然后,将来自步骤S2、S3、S4的嵌入进行拼接形成需求和供应的最终表征:

$$H_T^D = [h_T^D \parallel h_t^{DI} \parallel h_T^{AD} \parallel h_T^C]$$

$$H_T^S = [h_T^S \parallel h_t^{SI} \parallel h_T^{AS} \parallel h_T^C]$$

最后,使用两个线性变换来得到未来1天的需求和供应序列,并使用L1损失函数作为训练目标:

$$\mathcal{L}_D = \sum_{t=T+1}^{T+l} |D_t - D_t'|$$

$$\mathcal{L}_S = \sum_{t=T+1}^{T+l} |S_t - S_t'|$$

$$\mathcal{L}_\theta = \mathcal{L}_D + \mathcal{L}_S$$

经过反向传播、Adam优化器更新模型参数,直到损失函数收敛,得到模型的预测结果:未来的各配送站区域需求序列 $D_f'$ 与未来的各配送站区域供应序列 $S_f'$ 。

## 一种异常事件下的电商物资供需联合预测方法

### 技术领域

[0001] 本发明属于计算机应用数据科学领域,主要涉及了一种异常事件下的电商物资供需联合预测方法。

### 背景技术

[0002] 仓配一体化(WDI)是随着O2O电子商务发展应运而生的一种新型物流模式。在传统的仓配分离(WDS)模式下,电商商家将物资放在自己的仓库,根据收到的订单进行打包,并通知第三方快递派送包裹。然而,不断增加的物资订单量和销售品类将会导致WDS模式的低效率和高成本。在WDI模式下,电商企业凭借强大的库存管理和批量订单处理的能力,可以全面收集各地区物资的购买需求和配送供应数据。WDI模式的成功高度依赖于供需预测方法,然而在现实世界场景中,供需关系容易受到异常事件的影响。在异常事件(如节假日、促销活动等)的影响下,需求和供应往往会呈现不规律的变化。为了促进异常事件下WDI的应用,不仅需要对供需之间规律性的相互作用进行建模,还需要将演化的异常事件对供需关系和供需本身的影响进行建模。

[0003] 虽然需求和供应的联合预测引起了研究人员的关注,然而很少有工作研究电子商务场景中购买需求和配送供应的联合预测。现有的工作集中在打车系统、劳动力市场等场景。在打车场景中,研究将需求和供应分开预测或将问题简化为需求供应差值的预测。在劳动力市场场景中,研究设计需求-供应交互模块来同时预测需求和供应。然而,上述方法无法应对由异常事件引起的供需急剧失衡。虽然已经有一些研究认识到异常事件的重大影响,并在预测过程中考虑它们,这些工作解决了异常场景下需求序列建模的稀疏性问题或使用外部数据进行异常表示学习,但未能理解配送供应对购买需求的抑制作用。

[0004] 对于电商物资购买需求和配送供应的实操来说,有以下两个机会点:购买需求和配送供应在时间和空间上相互关联,协同预测可以提高两者的预测精度;异常事件通过分别作用于各区域的供需交互和供需自身,严重破坏了供需的一致性。

[0005] 考虑上述机遇,解决异常事件下的供需预测问题面临以下挑战:首先,建立购买需求和配送供应的时空关系是复杂的,两者在不同时期表现出不同的时间依赖模式,且在不同区域之间具有多种多样的空间相似性;其次,异常事件的影响是多种多样的,不能用完全相同的方式来描述。同时,异常数据在现实世界中往往是多源异构的,针对特定需求-供应预测目标的多源异构数据的采集、挖掘和融合是非常具有挑战性的。

### 发明内容

[0006] 本发明正是针对上述现有技术中存在的问题,设计一种异常事件下的电商物资供需联合预测方法,首先,对数据进行预处理,构建多视图的图结构,生成需求、供应、异常时序样本及非序列的辅助特征;其次,设计多图自适应卷积网络模块MGACR并嵌入GRU中作为时空编码器,对序列进行编码,捕获时空关系;再次,使用异常门控注意力机制来捕获需求与供应之间的相互作用;然后,引入预训练Bert增强对文本异常数据的语义理解,并将其经

过时空提升进行融合表征；最后，汇集上述步骤表征并做多步解码。本发明解决了需求与供应时空关系的复杂性、多源异构异常数据影响的多样性两点挑战，有利于电商平台在异常事件下针对性地对各区域提供物资物流双重保障，解决各区域物资供需不协调的问题。

[0007] 为了实现上述目的，本发明采取的技术方案是：一种异常事件下的电商物资供需联合预测方法，所述方法包括如下步骤：

[0008] S1,数据预处理：构建多视图的图结构，将配送站区域作为图结构的节点，将配送站间的距离、配送站区域是否邻接、配送站区域的路区类型相似程度、配送站区域的人群类型相似程度四种关系作为可选的图结构的边，生成时间窗对齐的需求、供应和异常时间序列样本，并使其序列长度保持一致；同时提取了非序列的时间上下文和新闻文本辅助特征；

[0009] S2,时空编码器编码：将步骤S1获得的需求序列、供应序列和异常序列输入到各自的多图自适应卷积循环网络MGACR中捕获三者的时空关系，并生成需求嵌入、供应嵌入和异常嵌入；所述多图自适应卷积循环网络的单元具体为：

$$[0010] \quad Z_t^1 = \mathbf{MGAC}([X_t || h_{t-1}], G^*)$$

$$[0011] \quad u_t = \sigma(Z_t^1 W_u + b_u)$$

$$[0012] \quad r_t = \sigma(Z_t^1 W_r + b_r)$$

$$[0013] \quad Z_t^2 = \mathbf{MGAC}([X_t || (r_t \odot h_{t-1})], G^*)$$

$$[0014] \quad c_t = \tanh(Z_t^2 W_c + b_c)$$

$$[0015] \quad h_t = u_t \odot h_{t-1} + (1 - u_t) \odot c_t$$

[0016] 其中  $X_t \in \mathbb{R}^{N \times F}$ ,  $h_t \in \mathbb{R}^{N \times d_h}$  表示时间步  $t$  时的输入与输出,  $u_t$  与  $r_t$  分别是时间步  $t$  时的更新门和重置门,  $G^*$  表示多视图的图结构,  $W_u, W_r, W_c, b_u, b_r, b_c$  以及多图自适应卷积循环网络MGACR中的参数是可学习的；

[0017] S3,需求-供应交互捕获：将需求序列、供应序列及步骤S2捕获的异常嵌入输入到交互式-多图自适应卷积循环网络I-MGACR,学习需求和供应序列的交互关系：

$$[0018] \quad h_t^{D_I}, h_t^{S_I} = \mathbf{I} - \mathbf{MGACR}(D_t, S_t, h_{t-1}^{D_I}, h_{t-1}^{S_I}, h_{t-1}^A, G_D^*, G_S^*)$$

[0019] 其中  $h_t^{D_I}$  表示时间步  $t$  时交互的需求嵌入,  $h_t^{S_I}$  表示时间步  $t$  时交互的供应嵌入,  $D_t$  表示时间步  $t$  时的需求,  $S_t$  表示时间步  $t$  时的供应,  $h_{t-1}^A$  表示时间步  $t-1$  时异常的时空嵌入,  $G_D^*$  表示需求的多视图图结构,  $G_S^*$  表示供应的多视图图结构；

[0020] S4,异常融合：引入预训练的Bert模型，将获取的新闻文本输入到异常融合模块进行异常语义提取，并使用步骤S2捕获的异常嵌入进行时空提升，从而获取异常对需求影响的融合嵌入和异常对供应影响的融合嵌入；

[0021] S5,需求-供应多步预测：提取时间上下文信息形成嵌入后，通过连接步骤S2、步骤S3和步骤S4的输出表征，得到历史需求和供应各自的融合表征，通过两个线性变换，生成目标时间窗口长度内未来的需求和供应序列。

[0022] 作为本发明的一种改进，所述步骤S1中构建的四种图结构分别为距离感知图  $G^d = (V, E^{(d)})$ 、邻居感知图  $G^{(n)} = (V, E^{(n)})$ 、类型相似图  $G^{(t)} = (V, E^{(t)})$ 、人群相似图  $G^{(c)} = (V, E^{(c)})$ ；

其中V表示区域节点,  $E^{(d)}$ 表示节点之间的距离邻接矩阵,  $E^{(n)}$ 表示节点之间是否相邻的邻接矩阵,  $E^{(t)}$ 表示节点之间路区特征余弦相似度的邻接矩阵,  $E^{(c)}$ 表示节点之间人群特征余弦相似度的邻接矩阵。

[0023] 作为本发明的一种改进,所述步骤S2中,对于多视图的图结构 $G^*$ ,集成了m个视图,相应地设计m个独立的GCN层,用以生成预定义图结构的节点嵌入表征:

$$[0024] \quad Z^{(m)} = \left( I_N + D^{-\frac{1}{2}} A^{(m)} D^{-\frac{1}{2}} \right) X \Theta^{(m)} + b^{(m)}$$

[0025] 残差自适应图结构学习,与残差自适应图结构的节点嵌入表征生成:

$$[0026] \quad \widetilde{E}^r = \text{Softmax}(\text{Relu}(M^r M^{rT}))$$

$$[0027] \quad Z^r = \widetilde{E}^r (X^i - H^b) W^r + b^r$$

[0028] 其中, Relu被用于缓解弱连接, Softmax被用于归一化自适应图结构,  $Z^r \in R^{N \times d_o}$ 表示残差图节点嵌入, N是节点数量,  $d_o$ 是残差图节点嵌入的维度;

[0029] 全量自适应图结构学习,与全量自适应图结构的节点嵌入表征生成:

$$[0030] \quad \widetilde{E}^f = \text{Softmax}(\text{Relu}(M^f M^{fT}))$$

$$[0031] \quad Z^f = \widetilde{E}^f X^i W^f + b^f$$

[0032] 其中,  $Z^f \in R^{N \times d_o}$ 表示全量图节点嵌入, N是节点数量,  $d_o$ 是全量图节点嵌入的维度;

[0033] 为实现多视图之间的信息共享,使用自注意力层将预定义图结构与残差自适应图结构的节点嵌入进行融合表征,公式如下:

$$[0034] \quad Z^s = \text{Softmax} \left( \frac{[Z^p || Z^r] W^Q ([Z^p || Z^r] W^K)^T}{\sqrt{d_k}} \cdot [Z^p || Z^r] W^V \right)$$

[0035] 其中,  $Z^s \in R^{N \times d_o}$ 表示自注意力层的输出,  $W^Q, W^K, W^V \in R^{(m+1)d_o \times d_o}$ 是可学习的Query、Key、Value矩阵,  $d_k = d_o$ 是 $W^K$ 的维度;

[0036] 最终得到多图自适应卷积的输出:

$$[0037] \quad Z = \sigma[Z^s || Z^f] W + b。$$

[0038] 作为本发明的另一种改进,所述步骤S3中,交互式-多图自适应卷积循环网络I-MGACR单元包括两个MGACR单元和一个InterModule模块,两个MGACR单元用以捕获需求-供应交互序列的时空信息,并受InterModule的影响进行信息交互:

$$[0039] \quad h_t^{D_I'} = \text{MGACR}(D_t, h_{t-1}^{D_I}, G_D^*)$$

$$[0040] \quad h_t^{S_I'} = \text{MGACR}(S_t, h_{t-1}^{S_I}, G_S^*)$$

$$[0041] \quad h_t^{D_I}, h_t^{S_I} = \text{InterModule}(h_t^{D_I'}, h_t^{S_I'}, h_{t-1}^A)$$

[0042] 其中  $h_t^{D_I'}$ 表示时间步t时InterModule模块的交互的需求嵌入输入,  $h_t^{S_I'}$ 表示时间



步t时InterModule模块的交互的供应嵌入输入。

[0043] 作为本发明的另一种改进,所述步骤S4中,InterModule模块包括异常门控的供需交互注意力机制,首先计算并生成需求-供应的自注意力嵌入:

$$[0044] \quad \hat{h}_t^D || \hat{h}_t^S = \text{Softmax} \left( \frac{[h_t^{D_I'} || h_t^{S_I'}] W_1^Q ([h_t^{D_I'} || h_t^{S_I'}] W_1^K)^T}{\sqrt{d_k}} \cdot [h_t^{D_I'} || h_t^{S_I'}] W_1^V \right)$$

[0045] 其中 $W_1^Q, W_1^K, W_1^V$ 是自注意力机制的参数权重;

[0046] 其次,使用交叉注意力机制来计算需求-供应受异常影响的程度,并生成受异常影响的需求-供应的交叉注意力嵌入:

$$[0047] \quad \hat{h}_t^A = \text{Softmax} \left( \frac{h_t^A W_2^Q ([h_t^{D_I'} || h_t^{S_I'}] W_2^K)^T}{\sqrt{d_k}} \cdot [h_t^{D_I'} || h_t^{S_I'}] W_2^V \right)$$

[0048] 其中 $W_2^Q, W_2^K, W_2^V$ 是交叉注意力机制的参数权重;

[0049] 最后,使用异常门控来控制上述自注意力机制、交叉注意力机制产生的嵌入比例,以形成最终的交互需求嵌入、交互供应嵌入:

$$[0050] \quad h_t^{D_I} = \sigma(h_t^A) \odot \hat{h}_t^A + (1 - \sigma(h_t^A)) \odot \hat{h}_t^D$$

$$[0051] \quad h_t^{S_I} = \sigma(h_t^A) \odot \hat{h}_t^A + (1 - \sigma(h_t^A)) \odot \hat{h}_t^S$$

[0052] 其中 $\sigma$ 表示sigmoid函数, $\odot$ 表示矩阵的元素乘积。

[0053] 作为本发明的又一种改进,所述步骤S4中,基于语义Bert的异常融合模块分为两个子模块:异常语义提取和时空增强;Bert首先由一个简单的MLP层将生成的嵌入映射为是否异常的二分类问题:

$$[0054] \quad E_{FN} \xleftarrow{C} \text{BERT}(\text{Tokens}_{FN})$$

$$[0055] \quad \hat{y} = \text{MLP}(E_{FN})$$

[0056] 其中 $E_{FN}$ 表示微调的嵌入,由Bert生成的C是分类器token[CLS]的表征, $\text{Tokens}_{FN}$ 表示微调新闻FN转化得到的Tokens;

[0057] 在上述模型收敛之后,得到了一个具备识别异常新闻能力的微调好的Bert,然后将该模型的参数部分冻结,并传递给需求-供应预测任务,从Bert的输出中,使用第一个Token([CLS])的最后一个隐藏层状态C来表示文本输入 $\text{PN}_T$ 中的每一条新闻标题的聚合语义信息,如下所示:

$$[0058] \quad E_p \xleftarrow{C} \text{BERT}_f(\text{Tokens}_{PN_T})$$

[0059] 其中 $E_p$ 表示条级新闻语义嵌入, $\text{BERT}_f$ 表示初始化后固定参数的Bert模型, $\text{Tokens}_{PN_T}$ 表示历史最近的预测任务新闻 $\text{PN}_T$ 转化得到的Tokens;

[0060] 然后,使用微调好的MLP中的Softmax来计算得到天级的语义表征:

$$[0061] \quad \hat{y}' = \text{MLP}(E_p)|_{\text{label}=1}$$

$$[0062] \quad E_d = \sum_{i=1}^{N_{news}} \hat{y}_i' \odot E_{p_i}$$

[0063] 其中 $\hat{y}'$ 表示 $E_p$ 的权重, $E_d$ 表示融合的天级新闻语义嵌入, $N_{news}$ 表示一天中最大的新闻条数。

[0064] 作为本发明的又一种改进,所述步骤S4中设计了三个FC全连接层来建模异常的不同影响:

$$[0065] \quad h_T^{Ad} = FC_1([h_T^A \parallel N * FC_3(E_d)])$$

$$[0066] \quad h_T^{As} = FC_2([h_T^A \parallel N * FC_3(E_d)])$$

[0067] 其中 $h_T^{Ad}$ 表示异常对需求影响的融合嵌入, $h_T^{As}$ 表示异常对供应影响的融合嵌入。

[0068] 作为本发明的更进一步改进,所述步骤S5具体为:

[0069] 首先生成了上下文信息的嵌入来捕获需求和供应序列的周期性:

$$[0070] \quad h_T^C = N * MLP([FC(DoW_T) \parallel FC(FT_T)])$$

[0071] 其中 $DoW_T$ 表示对历史最近一天是一周中第几天的编码, $FT_T$ 表示对历史最近一天的节假日类型的编码;

[0072] 然后,将来自步骤S2、S3、S4的嵌入进行拼接形成需求和供应的最终表征:

$$[0073] \quad H_T^D = [h_T^D \parallel h_t^{D_I} \parallel h_T^{Ad} \parallel h_T^C]$$

$$[0074] \quad H_T^S = [h_T^S \parallel h_t^{S_I} \parallel h_T^{As} \parallel h_T^C]$$

[0075] 最后,使用两个线性变换来得到未来1天的需求和供应序列,并使用L1损失函数作为训练目标:

$$[0076] \quad \mathcal{L}_D = \sum_{t=T+1}^{T+l} |D_t - D_t'|$$

$$[0077] \quad \mathcal{L}_S = \sum_{t=T+1}^{T+l} |S_t - S_t'|$$

$$[0078] \quad \mathcal{L}_\theta = \mathcal{L}_D + \mathcal{L}_S$$

[0079] 经过反向传播、Adam优化器更新模型参数,直到损失函数收敛,得到模型的预测结果:未来的各配送站区域需求序列 $D_f'$ 与未来的各配送站区域供应序列 $S_f'$

[0080] 与现有技术相比,本发明具有的有益效果:

[0081] (1) 本发明首次联合预测异常事件影响下的需求和供应,考虑了供需的时空相互作用和异常的多种影响,弥补了传统供需预测方法中,未考虑异常事件对供应和需求的影响以及未考虑供应对需求抑制作用的不足。

[0082] (2) 为了解决复杂时空关系的挑战,本发明设计了一种新的多图自适应卷积循环网络(MGACR)来捕获需求、供应和异常的时空信息。为应对异常影响多元化、异常数据多源异构性的挑战,本发明设计了一种改进的I-MGACR网络,利用异常门控注意力机制学习异常对供需交互的影响,并通过生成异常融合表征,分别表示异常对需求和供应的影响。

[0083] (3) 本发明在真实世界的数据集上进行了评估,其中包括从国内某大型电商平台收集的上海282个配送站、1985万用户和2.7302亿订单数据,以及从政府官方网站获得的异常事件数据集。实验结果表明,本发明提出的方法比现有最先进方法在MAE、RMSE指标上分别提高了10%、13%以上。消融实验证实,设计的多图自适应卷积、异常门控注意力机制和异常融合模块明显提升了需求和供应预测的准确程度。

## 附图说明

[0084] 图1为本发明一种异常事件下的电商物资供需联合预测方法的结构架构图;

[0085] 图2为本发明一种异常事件下的电商物资供需联合预测方法的设计细节图。

## 具体实施方式

[0086] 以下结合附图和具体实施例对本发明作进一步详细说明。

[0087] 实施例1

[0088] 一种异常事件下的电商物资供需联合预测方法,包括如下步骤:

[0089] 步骤S1,数据预处理:首先构建了一个多视图的图结构,将配送站区域作为图结构的节点,将配送站间的距离、配送站区域是否邻接、配送站区域的路区类型相似程度、配送站区域的人群类型相似程度四种关系作为可选的图结构的边。接着,生成了时间窗对齐的需求、供应和异常时间序列样本,使它们的序列长度保持一致。同时,还提取了非序列的时间上下文和新闻文本辅助特征。

[0090] 构建四种预定义的图结构,分别为距离感知图 $G^d = (V, E^{(d)})$ 、邻居感知图 $G^{(n)} = (V, E^{(n)})$ 、类型相似图 $G^{(t)} = (V, E^{(t)})$ 、人群相似图 $G^{(c)} = (V, E^{(c)})$ ,其中 $V$ 表示区域节点, $E^{(d)}$ 表示节点之间的距离邻接矩阵, $E^{(n)}$ 表示节点之间是否相邻的邻接矩阵, $E^{(t)}$ 表示节点之间路区特征余弦相似度的邻接矩阵, $E^{(c)}$ 表示节点之间人群特征余弦相似度的邻接矩阵,并生成窗口长度为 $T+1$ 的需求 $D$ 、供应 $S$ 、异常 $A$ 序列,其中 $T$ 表示历史时间步,1表示未来时间步。获取时间上下文,一周内第几天DoW和节假日类型FT。收集政府新闻作为文本异常数据,并将其划分为两部分:微调新闻FN和预测任务新闻PT,分别对预训练的Bert模型和下游预测任务进行微调。

[0091] 步骤S2,时空编码器编码:在构建好预定义多视图的图结构的基础上,针对每种数据(需求、供应、异常)设计包含两个自适应图结构的多图自适应卷积模块,用于自适应捕获预定义多视图的图结构不能包含的剩余信息,并自适应学习与预测任务直接相关的完整图结构;在每个时间步将多图自适应卷积嵌入到GRU中,进而得到时空编码器——多图自适应卷积循环网络(MGACR),用于捕获序列的时空关系。

[0092] 多图自适应卷积的核心可以表述如下,对于预先定义的多视图的图结构 $G^*$ ,它集成了 $m$ 个视图,相应地设计 $m$ 个独立的GCN层,用以生成预定义图结构的节点嵌入表征:

$$[0093] \quad Z^{(m)} = \left( I_N + D^{-\frac{1}{2}} A^{(m)} D^{-\frac{1}{2}} \right) X \Theta^{(m)} + b^{(m)}$$

[0094] 残差自适应图结构学习,与残差自适应图结构的节点嵌入表征生成:

$$[0095] \quad \widetilde{E}^r = \text{Softmax}(\text{Relu}(M^r M^{rT}))$$

$$[0096] \quad Z^r = \widetilde{E}^r (X^i - H^b) W^r + b^r$$

[0097] 其中, Relu被用于缓解弱连接, Softmax被用于归一化自适应图结构,  $Z^r \in R^{N \times d_o}$  表示残差图节点嵌入, N是节点数量,  $d_o$ 是残差图节点嵌入的维度。

[0098] 全量自适应图结构学习, 与全量自适应图结构的节点嵌入表征生成:

$$[0099] \quad \widetilde{E}^f = \text{Softmax}(\text{Relu}(M^f M^{fT}))$$

$$[0100] \quad Z^f = \widetilde{E}^f X^i W^f + b^f$$

[0101] 其中,  $Z^f \in R^{N \times d_o}$  表示全量图节点嵌入, N是节点数量,  $d_o$ 是全量图节点嵌入的维度。

[0102] 为实现多视图之间的信息共享, 使用自注意力层将预定义图结构与残差自适应图结构的节点嵌入进行融合表征, 公式如下:

$$[0103] \quad Z^s = \text{Softmax} \left( \frac{[Z^p || Z^r] W^Q ([Z^p || Z^r] W^K)^T}{\sqrt{d_k}} \cdot [Z^p || Z^r] W^V \right)$$

[0104] 其中,  $Z^s \in R^{N \times d_o}$  表示自注意力层的输出,  $W^Q, W^K, W^V \in R^{(m+1)d_o \times d_o}$  是可学习的Query、Key、Value矩阵,  $d_k = d_o$  是  $W^K$  的维度。

[0105] 最终得到多图自适应卷积的输出:

$$[0106] \quad Z = \sigma[Z^s || Z^f] W + b$$

[0107] 进而可以得到多图自适应卷积循环网络的单元:

$$[0108] \quad Z_t^1 = \text{MGAC}([X_t || h_{t-1}], G^*)$$

$$[0109] \quad u_t = \sigma(Z_t^1 W_u + b_u)$$

$$[0110] \quad r_t = \sigma(Z_t^1 W_r + b_r)$$

$$[0111] \quad Z_t^2 = \text{MGAC}([X_t || (r_t \odot h_{t-1})], G^*)$$

$$[0112] \quad c_t = \tanh(Z_t^2 W_c + b_c)$$

$$[0113] \quad h_t = u_t \odot h_{t-1} + (1 - u_t) \odot c_t$$

[0114] 其中  $X_t \in R^{N \times F}$ ,  $h_t \in R^{N \times d_h}$  表示时间步t时的输入与输出,  $u_t$  与  $r_t$  分别是时间步t时的更新门和重置门,  $G^*$  表示多视图的图结构,  $W_u, W_r, W_c, b_u, b_r, b_c$  以及MGAC中的参数是可学习的。

[0115] 用于捕获需求、供应、异常三者的时空关系, 并生成需求嵌入、供应嵌入和异常嵌入。

[0116] 步骤S3, 需求-供应交互捕获: 使用异常门控注意力机制来捕获需求与供应之间的相互作用, 在此过程中, 异常的严重程度被设计来控制每个时间步中需求与供应之间的自相互作用和异常与供需之间的交叉相互作用。遵循上述交互方式, 一个需求MGACR单元和一个供应MGACR单元将形成一个交互式-多图自适应卷积循环网络(I-MGACR)单元, 用于捕获需求序列与供应序列的交互关系。

[0117] 为了建模供需之间的交互关系, 设计了一个基于MGACR的供需交互捕获模块, 命名

为交互式MGACR(I-MGACR),目标是学习需求和供应序列的交互状态:

$$[0118] \quad h_t^{D_I}, h_t^{S_I} = \mathbf{I} - \mathbf{MGACR}(D_t, S_t, h_{t-1}^{D_I}, h_{t-1}^{S_I}, h_{t-1}^A, G_D^*, G_S^*)$$

[0119] 其中  $h_t^{D_I}$  表示时间步  $t$  时交互的需求嵌入,  $h_t^{S_I}$  表示时间步  $t$  时交互的供应嵌入,  $D_t$  表示时间步  $t$  时的需求,  $S_t$  表示时间步  $t$  时的供应,  $h_{t-1}^A$  表示时间步  $t-1$  时异常的时空嵌入,  $G_D^*$  表示需求的多视图图结构,  $G_S^*$  表示供应的多视图图结构。

[0120] I-MGACR单元由如下三个组件组成,两个MGACR单元用以捕获需求-供应交互序列的时空信息,并受InterModule的影响进行信息交互:

$$[0121] \quad h_t^{D_I'} = \mathbf{MGACR}(D_t, h_{t-1}^{D_I}, G_D^*)$$

$$[0122] \quad h_t^{S_I'} = \mathbf{MGACR}(S_t, h_{t-1}^{S_I}, G_S^*)$$

$$[0123] \quad h_t^{D_I}, h_t^{S_I} = \mathbf{InterModule}(h_t^{D_I'}, h_t^{S_I'}, h_{t-1}^A)$$

[0124] 其中  $h_t^{D_I'}$  表示时间步  $t$  时InterModule模块的交互的需求嵌入输入,  $h_t^{S_I'}$  表示时间步  $t$  时InterModule模块的交互的供应嵌入输入。

[0125] 设计异常门控的供需交互注意力机制作为InterModule,首先计算并生成需求-供应的自注意力嵌入:

$$[0126] \quad \hat{h}_t^D || \hat{h}_t^S = \text{Softmax} \left( \frac{[h_t^{D_I'} || h_t^{S_I'}] W_1^Q \left( [h_t^{D_I'} || h_t^{S_I'}] W_1^K \right)^T}{\sqrt{d_k}} \cdot [h_t^{D_I'} || h_t^{S_I'}] W_1^V \right)$$

[0127] 其中  $W_1^Q, W_1^K, W_1^V$  是自注意力机制的参数权重。

[0128] 其次,使用交叉注意力机制来计算需求-供应受异常影响的程度,并生成受异常影响的需求-供应的交叉注意力嵌入:

$$[0129] \quad \hat{h}_t^A = \text{Softmax} \left( \frac{h_t^A W_2^Q ([h_t^{D_I'} || h_t^{S_I'}] W_2^K)^T}{\sqrt{d_k}} \cdot [h_t^{D_I'} || h_t^{S_I'}] W_2^V \right)$$

[0130] 其中  $W_2^Q, W_2^K, W_2^V$  是交叉注意力机制的参数权重。

[0131] 最后,使用异常门控来控制上述自注意力机制、交叉注意力机制产生的嵌入比例,以形成最终的交互需求嵌入、交互供应嵌入:

$$[0132] \quad h_t^{D_I} = \sigma(h_t^A) \odot \hat{h}_t^A + (1 - \sigma(h_t^A)) \odot \hat{h}_t^D$$

$$[0133] \quad h_t^{S_I} = \sigma(h_t^A) \odot \hat{h}_t^A + (1 - \sigma(h_t^A)) \odot \hat{h}_t^S$$

[0134] 其中 $\sigma$ 表示sigmoid函数,  $\odot$ 表示矩阵的元素乘积。

[0135] 步骤S4, 异常融合: 以文本形式描述的异常数据通常包含丰富的语义信息, 引起用户行为的变化, 进而导致需求和供应的波动。引入预训练的Bert模型来增强对文本异常数据的理解, 并应用经过时空提升的语义信息来建模异常对需求和供应模式的影响。

[0136] 基于语义Bert的异常融合模块可分为两个子模块: 异常语义提取和时空增强。Bert首先由一个简单的MLP层将生成的嵌入映射为是否异常的二分类问题:

$$[0137] \quad E_{FN} \xleftarrow{C} BERT(Tokens_{FN})$$

$$[0138] \quad \hat{y} = MLP(E_{FN})$$

[0139] 其中 $E_{FN}$ 表示微调的嵌入, 由Bert生成的C是分类器token[CLS]的表征,  $Tokens_{FN}$ 表示微调新闻FN转化得到的Tokens。

[0140] 在上述模型收敛之后, 得到了一个具备识别异常新闻能力的微调好的Bert, 然后将该模型的参数部分冻结, 并传递给需求-供应预测任务。从Bert的输出中, 使用第一个Token([CLS])的最后一个隐藏层状态C来表示文本输入 $PN_T$ 中的每一条新闻标题的聚合语义信息, 如下所示:

$$[0141] \quad E_p \xleftarrow{C} BERT_f(Tokens_{PN_T})$$

[0142] 其中 $E_p$ 表示条级新闻语义嵌入,  $BERT_f$ 表示初始化后固定参数的Bert模型,  $Tokens_{PN_T}$ 表示历史最近的预测任务新闻 $PN_T$ 转化得到的Tokens。

[0143] 然后, 使用微调好的MLP中的Softmax来计算得到天级的语义表征:

$$[0144] \quad \hat{y}' = MLP(E_p)|_{label=1}$$

$$[0145] \quad E_d = \sum_{i=1}^{N_{news}} \hat{y}'_i \odot E_{p_i}$$

[0146] 其中 $\hat{y}'$ 表示 $E_p$ 的权重,  $E_d$ 表示融合的天级新闻语义嵌入,  $N_{news}$ 表示一天中最大的新闻条数。

[0147] 由于异常以时空的方式分别作用于需求和供应序列, 设计了三个FC全连接层来建模异常的不同影响:

$$[0148] \quad h_T^{A_D} = FC_1([h_T^A \parallel N * FC_3(E_d)])$$

$$[0149] \quad h_T^{A_S} = FC_2([h_T^A \parallel N * FC_3(E_d)])$$

[0150] 其中 $h_T^{A_D}$ 表示异常对需求影响的融合嵌入,  $h_T^{A_S}$ 表示异常对供应影响的融合嵌入。

[0151] 步骤S5, 需求-供应多步预测: 通过连接步骤S2、步骤S3和步骤S4的输出表征, 得到历史需求和供应各自的融合表征。随后, 使用两个简单的线性层来生成目标时间窗口长度内未来的需求和供应序列。

[0152] 首先生成了上下文信息的嵌入来捕获需求和供应序列的周期性:

$$[0153] \quad h_T^C = N * MLP([FC(DoW_T) \parallel FC(FT_T)])$$

[0154] 其中 $DoW_T$ 表示对历史最近一天是一周中第几天的编码, $FT_T$ 表示对历史最近一天的节假日类型的编码。

[0155] 然后,将来自步骤S2、S3、S4的嵌入进行拼接形成需求和供应的最终表征:

$$[0156] \quad H_T^D = [h_T^D \parallel h_t^{D_I} \parallel h_T^{A_D} \parallel h_T^C]$$

$$[0157] \quad H_T^S = [h_T^S \parallel h_T^{S_I} \parallel h_T^{A_S} \parallel h_T^C]$$

[0158] 最后,使用两个线性变换来得到未来1天的需求和供应序列,并使用L1损失函数作为训练目标:

$$[0159] \quad \mathcal{L}_D = \sum_{t=T+1}^{T+l} |D_t - D'_t|$$

$$[0160] \quad \mathcal{L}_S = \sum_{t=T+1}^{T+l} |S_t - S'_t|$$

$$[0161] \quad \mathcal{L}_\theta = \mathcal{L}_D + \mathcal{L}_S$$

[0162] 经过反向传播、Adam优化器更新模型参数,直到损失函数收敛,得到模型的预测结果:未来的各配送站区域需求序列 $D_f'$ 与未来的各配送站区域供应序列 $S_f'$ 。

[0163] 实施例2

[0164] 定义需求为区域订单创建量,供应为区域订单配送量,异常为区域爆发地数量,分别预测需求和供应。具体来说,以1天为时间步长,使用历史7天各营业部区域的需求、供应、异常序列,历史最近1天的非序列辅助特征,以及预定义的多视图结构,分别预测未来各营业部区域3天、5天、7天的需求和供应。

[0165] 如图1所示,分为五个步骤:(1)数据处理:对上海市人民政府官方网站获取到的异常数据,以及国内某大型电商平台获取的订单数据、配送站数据、用户数据、上下文数据进行处理后,得到需求序列、供应序列、异常序列、新闻文本、时间上下文、多视图图结构;(2)时空编码器编码:将获得到的需求序列、供应序列、异常序列输入到各自的多图自适应卷积循环网络中捕获三者的时空关系,并生成需求嵌入、供应嵌入和异常嵌入;(3)需求-供应交互捕获:将需求序列、供应序列、以及步骤(2)捕获的异常时空关系嵌入输入到交互式-多图自适应卷积循环网络,以获得交互的需求嵌入和交互的供应嵌入;(4)异常融合:将获取的新闻文本输入到异常融合模块,进行异常语义提取,并且使用步骤(2)捕获的异常时空关系嵌入进行时空提升,从而获取异常对需求影响的融合嵌入和异常对供应影响的融合嵌入;(5)需求-供应多步预测:提取时间上下文信息形成嵌入后,与步骤(2)(3)(4)生成的嵌入进行拼接,并通过两个线性变换,分别得到最终的需求表征和供应表征。

[0166] 参阅图2所示,对需求、供应、异常序列进行时空编码器编码的过程为:

$$[0167] \quad h_t^D = \mathbf{MGACR}(D_t, h_{t-1}^D, G_D^*)$$

$$[0168] \quad h_t^S = \mathbf{MGACR}(S_t, h_{t-1}^S, G_S^*)$$

$$[0169] \quad h_t^A = \mathbf{MGACR}(A_t, h_{t-1}^A, G_A^*)$$

[0170] 其中  $h_t^D, h_t^S, h_t^A$  分别表示t时间步时的需求、供应、异常的时空关系嵌入, MGACR表示多图自适应卷积循环网络的单元,  $D_t, S_t, A_t$  分别表示t时间步时的需求值、供应值、异常值,  $G_D^*, G_S^*, G_A^*$  分别表示需求、供应、异常的多视图图结构。

[0171] 对需求-供应交互捕获的过程为:

$$[0172] \quad h_t^{D_I}, h_t^{S_I} = \mathbf{I} - \mathbf{MGACR}(D_t, S_t, h_{t-1}^{D_I}, h_{t-1}^{S_I}, h_{t-1}^A, G_D^*, G_S^*)$$

[0173] 其中  $h_t^{D_I}, h_t^{S_I}$  分别表示t时间步时交互的需求嵌入和交互的供应嵌入, I-MGACR表示交互式-多图自适应卷积循环网络的单元,  $D_t, S_t$  分别表示t时间步时的需求值、供应值,  $G_D^*, G_S^*$  分别表示需求、供应的多视图图结构。

[0174] 对异常进行融合的过程为:

$$[0175] \quad h_T^{A_D}, h_T^{A_S} = \text{AnomalyFusion}(h_T^A, \text{Aux}_1)$$

[0176] 其中  $h_T^{A_D}, h_T^{A_S}$  分别表示历史最近的异常对需求影响的融合嵌入、异常对供应影响的融合嵌入, AnomalyFusion表示异常融合模块,  $h_T^A$  表示历史最近的异常时空关系嵌入,  $\text{Aux}_1$  表示新闻文本辅助特征。

[0177] 对需求-供应多步预测的过程为:

$$[0178] \quad H_T^D = [h_T^D \parallel h_t^{D_I} \parallel h_T^{A_D} \parallel h_T^C]$$

$$[0179] \quad H_T^S = [h_T^S \parallel h_t^{S_I} \parallel h_T^{A_S} \parallel h_T^C]$$

$$[0180] \quad D_f' = \text{Linear1}(H_T^D)$$

$$[0181] \quad S_f' = \text{Linear2}(H_T^S)$$

[0182] 其中  $H_T^D, H_T^S$  分别表示最终的需求表征和最终的供应表征,  $D_f', S_f'$  分别表示模型预测出的未来1天的需求序列和未来1天的供应序列。

[0183] 最后使用反向传播、Adam优化器更新模型参数,直到损失函数收敛。

[0184] 实施例3

[0185] 定义需求为区域订单创建量,供应为同一批订单配送量,异常为区域爆发地数量,预测供需缺口(差距)。具体来说,以1天为时间步长,使用历史7天各营业部区域的需求、供应、异常序列,历史最近1天的非序列辅助特征,以及预定义的多视图结构,预测未来各营业部区域3天、5天、7天的供需缺口(差距)。

[0186] 参阅图1所示,分为五个步骤:(1)数据处理:对上海市人民政府官方网站获取到的异常数据,以及国内某大型电商平台获取的订单数据、配送站数据、用户数据、上下文数据进行处理后,得到需求序列、供应序列、异常序列、新闻文本、时间上下文、多视图图结构;(2)时空编码器编码:将获得到的需求序列、供应序列、异常序列输入到各自的多图自适应卷积循环网络中捕获三者的时空关系,并生成需求嵌入、供应嵌入和异常嵌入;(3)需求-供应交互捕获:将需求序列、供应序列、以及步骤(2)捕获的异常时空关系嵌入输入到交互式-



多图自适应卷积循环网络,以获得交互的需求嵌入和交互的供应嵌入;(4)异常融合:将获取的新闻文本输入到异常融合模块,进行异常语义提取,并且使用步骤(2)捕获的异常时空关系嵌入进行时空提升,从而获取异常对需求影响的融合嵌入和异常对供应影响的融合嵌入;(5)需求-供应多步预测:提取时间上下文信息形成嵌入后,与步骤(2)(3)(4)生成的嵌入进行拼接,并通过两个线性变换,分别得到最终的需求表征和供应表征。

[0187] 参阅图2所示,对需求、供应、异常序列进行时空编码器编码的过程为:

$$[0188] \quad h_t^D = \mathbf{MGACR}(D_t, h_{t-1}^D, G_D^*)$$

$$[0189] \quad h_t^S = \mathbf{MGACR}(S_t, h_{t-1}^S, G_S^*)$$

$$[0190] \quad h_t^A = \mathbf{MGACR}(A_t, h_{t-1}^A, G_A^*)$$

[0191] 其中 $h_t^D, h_t^S, h_t^A$ 分别表示t时间步时的需求、供应、异常的时空关系嵌入, MGACR表示多图自适应卷积循环网络的单元,  $D_t, S_t, A_t$ 分别表示t时间步时的需求值、供应值、异常值,  $G_D^*, G_S^*, G_A^*$ 分别表示需求、供应、异常的多视图图结构。

[0192] 对需求-供应交互捕获的过程为:

$$[0193] \quad h_t^{D_I}, h_t^{S_I} = \mathbf{I} - \mathbf{MGACR}(D_t, S_t, h_{t-1}^{D_I}, h_{t-1}^{S_I}, h_{t-1}^A, G_D^*, G_S^*)$$

[0194] 其中 $h_t^{D_I}, h_t^{S_I}$ 分别表示t时间步时交互的需求嵌入和交互的供应嵌入, I-MGACR表示交互式-多图自适应卷积循环网络的单元,  $D_t, S_t$ 分别表示t时间步时的需求值、供应值,  $G_D^*, G_S^*$ 分别表示需求、供应的多视图图结构。

[0195] 对异常进行融合的过程为:

$$[0196] \quad h_T^{A_D}, h_T^{A_S} = \mathbf{AnomalyFusion}(h_T^A, \text{Aux}_1)$$

[0197] 其中 $h_T^{A_D}, h_T^{A_S}$ 分别表示历史最近的异常对需求影响的融合嵌入、异常对供应影响的融合嵌入, AnomalyFusion表示异常融合模块,  $h_T^A$ 表示历史最近的异常时空关系嵌入,  $\text{Aux}_1$ 表示新闻文本辅助特征。

[0198] 对需求-供应多步预测的过程为:

$$[0199] \quad H_T^D = [h_T^D \parallel h_t^{D_I} \parallel h_T^{A_D} \parallel h_T^C]$$

$$[0200] \quad H_T^S = [h_T^S \parallel h_t^{S_I} \parallel h_T^{A_S} \parallel h_T^C]$$

$$[0201] \quad D_f' = \mathbf{Linear1}(H_T^D)$$

$$[0202] \quad S_f' = \mathbf{Linear2}(H_T^S)$$

[0203] 其中 $H_T^D, H_T^S$ 分别表示最终的需求表征和最终的供应表征,  $D_f'$ 与 $S_f'$ 的差值即为模型预测出的未来1天的供需缺口(差距)。

[0204] 最后使用反向传播、Adam优化器更新模型参数,直到损失函数收敛。

[0205] 测试例

[0206] 本发明在真实世界的数据集上进行了评估,其中包括从国内某大型电商平台收集的上海282个配送站、1985万用户和2.7302亿订单数据,以及从政府官方网站获得的异常事件数据集(含3.635万地址和8070条新闻)。每个时间步长设为1天,从2021年1月1日到2022年8月31日,共有608个时间间隔。并按照7:2:1的分割比例将数据集随机划分训练集、验证集和测试集。

[0207] 对比模型:

[0208] (1)经典时间序列预测模型:HI、ARIMA和SVR

[0209] (2)基于图的时空预测模型:Graph WaveNet、MTGNN、DCRNN和AGCRN

[0210] (3)需求-供应联合预测模型:DH-GEM(-s)、DH-GEM(-r)

[0211] 评价指标:

[0212] 本发明选取回归预测任务中典型的三个指标作为评价指标,如下所示。指标值越小,说明模型的供需预测精度越高、预测误差越低。

[0213] (1)平均绝对误差(MAE)

[0214] (2)均方根误差(RMSE)

[0215] (3)对称平均绝对百分比误差(sMAPE)

[0216] 实验结果:

[0217] 本发明与三类对比模型进行对比实验,相应的实验结果如表1所示。

[0218] 表1对比实验结果

模型	需求			供应			
	MAE	RMSE	sMAPE	MAE	RMSE	sMAPE	
[0219]	HI	433.57	758.44	25.47%	373.26	585.67	27.30%
	ARIMA	394.36	921.16	22.34%	316.99	552.00	24.15%
	SVR	352.89	800.22	20.61%	284.07	482.35	21.05%
	GWN	302.46	523.04	18.61%	<u>264.56</u>	<u>428.87</u>	20.64%
	MTGNN	<u>289.20</u>	501.64	<u>18.04%</u>	268.99	433.94	<u>20.61%</u>
	DCRNN	327.80	538.22	20.98%	320.24	487.11	22.36%
	AGCRN	293.27	<u>499.80</u>	18.51%	274.70	429.68	20.79%
	Transformer	381.97	588.77	24.48%	371.72	545.70	24.37%
	DH-GEM (-s)	371.12	568.50	23.88%	361.69	533.11	23.78%
[0220]	DH-GEM (-r)	369.76	569.25	23.86%	366.16	539.77	24.02%
	本发明	<b>274.19</b>	<b>459.26</b>	<b>17.76%</b>	<b>225.37</b>	<b>355.19</b>	<b>19.32%</b>

[0221] 结果分析:

[0222] 上表中可看出,与经典时间序列预测对比模型(即HI、ARIMA和SVR)的比较:本发明远比统计数据或传统的机器学习模型更有效,需求的MAE和RMSE的表现平均提升24.80%,供应的MAE和RMSE的表现平均提升39.46%。由于非线性依赖性和缺乏手工特征,这些方法比深度模型基本都更差。

[0223] 与基于图的时空预测对比模型(即Graph WaveNet、MTGNN、DCRNN和AGCRN)的比较:这些模型经过精心设计,可以同时挖掘需求或供应序列的时间和空间数据,但未考虑需求和供应之间的关系。初步发现AGCRN、MTGNN和Graph WaveNet的性能优于DCRNN,这证明了自适应图学习的有效性。在所有对比模型中,MTGNN的性能最好,仅次于本发明提出的方法。与MTGNN相比,本发明的MAE的表现平均提升了10.71%,RMSE的表现平均提升了13.30%。

[0224] 与需求-供应联合预测对比模型(即DH-GEM(-s)和DH-GEM(-r))的比较:表中结果证明了异常信息对需求和供应之间相互作用建模的贡献。此外,在供需序列联合建模的帮助下,即使在DH-GEM(-r)中去除了图结构,DH-GEM也比单独的Transformer表现更好,这确认了供需交互设计的必要性。然而,由于DH-GEM忽略了需求和供应的时空特征,即使用自适应的时空同构图代替原始的异构图,DH-GEM(-s)的效果仍然比本发明差。这进一步反映了本发明设计的交互模块可以提取更精确的需求和供应表示,并且解码器仅通过使用两个简单的线性变换来实现,更为轻量。

[0225] 以上所述仅是本发明的优选实施方式,应当指出:对本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

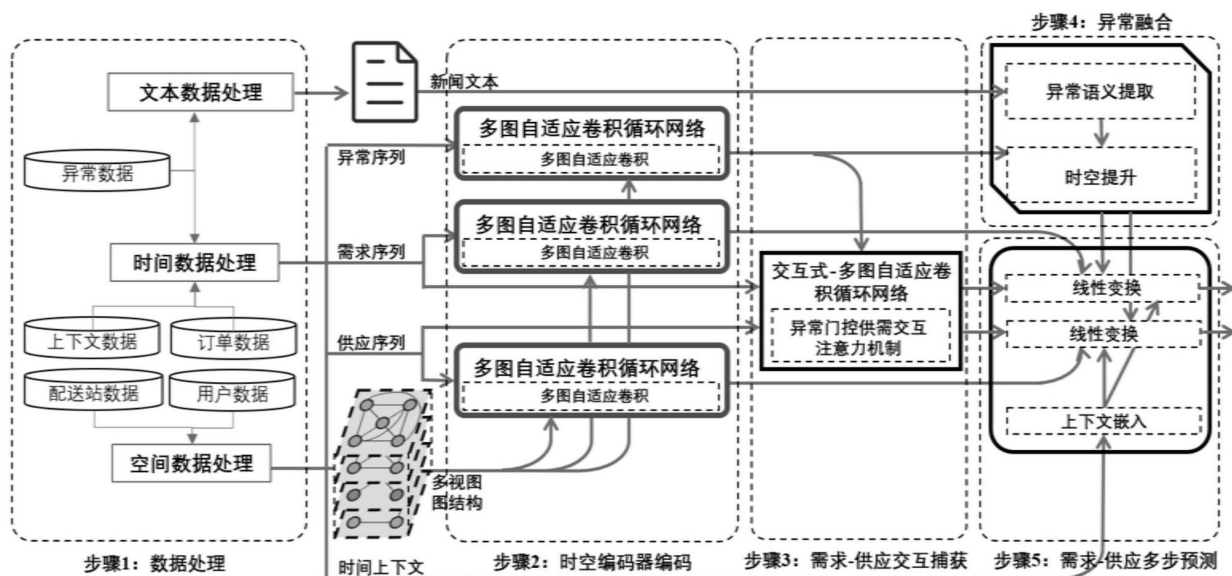


图1

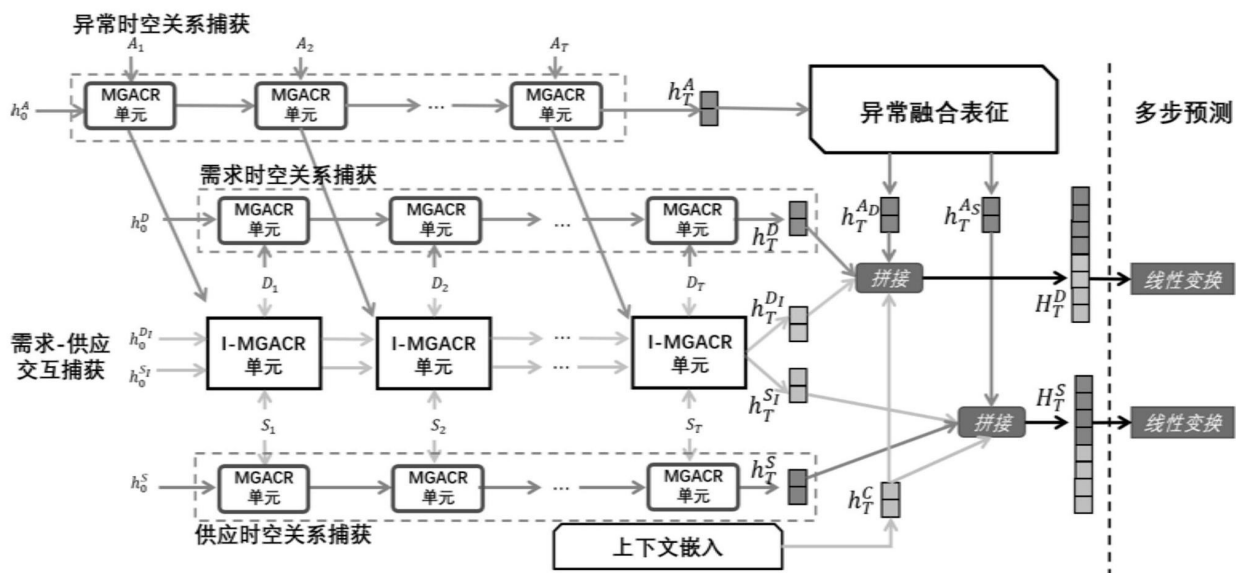


图2