



(12) 发明专利申请

(10) 申请公布号 CN 116525132 A

(43) 申请公布日 2023. 08. 01

(21) 申请号 202310318142.1

G06N 7/01 (2023.01)

(22) 申请日 2023.03.29

(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 胡世杰 林丽 夏凯文  
梅洛瑜

(74) 专利代理机构 南京众联专利代理有限公司  
32206

专利代理人 叶倩

(51) Int. Cl.

G16H 50/80 (2018.01)

G06N 3/0442 (2023.01)

G06N 3/047 (2023.01)

G06N 3/048 (2023.01)

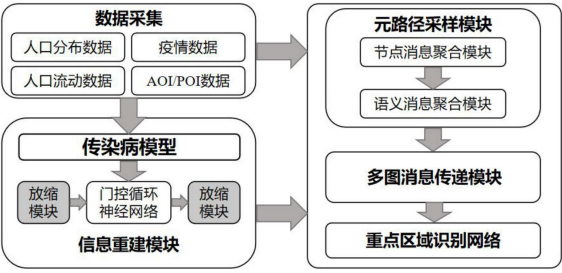
权利要求书4页 说明书10页 附图1页

(54) 发明名称

面向POI的突发公共卫生事件重点区域预测方法及系统

(57) 摘要

本发明公开了一种面向POI的突发公共卫生事件重点区域预测方法及系统,首先对待测城市的细粒度区域进行预测信息的收集与统计,并利用预测信息进行传染病模型构建,进行公共卫生事件演变规律的挖掘;针对演变规律,搭建公共卫生事件不可观测信息重建网络;以24小时为单位,构建以细粒度区域为节点的人口流动异质有向多图,基于异质有向图的特点,在空间上捕捉每个人口流动图内各区域间与关联区域的信息,得到每个人口流动图的空间嵌入表示;再根据24张人口流动图的时序关系,在时间维度进行各节点的信息融合,得到最终的城市细粒度区域嵌入表示,将其输入城市重点区域识别网络,得到各个细粒度区域潜在的风险等级,实现风险预测。



1. 面向POI的突发公共卫生事件重点区域预测方法,其特征在于,包括如下步骤:

S1:确定待预测城市的范围,将城市范围划分为多个细粒度区域,并基于细粒度区域进行预测信息的收集与统计;所述预测信息至少包括城市粒度公共卫生事件可观测信息、区域粒度公共卫生事件可观测信息和区域间人口流动数据;

S2:基于城市粒度公共卫生事件可观测信息,进行传染病模型构建;所述传染病模型包含不同状态的人群,并由微分方程组成,每一条微分方程表示人群在每个时间步下的状态变化概率,传染病模型用于不同状态人群之间状态转移概率的构建,根据城市粒度公共卫生事件可观测信息拟合的传染病模型挖掘公共卫生事件演变规律;

S3:针对步骤S2获得的公共卫生事件演变规律,搭建公共卫生事件不可观测信息重建网络,所述信息重建网络由门控循环神经网络构成,基于区域粒度公共卫生事件的可观测信息,通过门控循环神经网络对可观测信息时序上的特征进行捕捉,输出区域粒度公共卫生事件下的不可观测信息,所述不可观测信息包括区域粒度未隔离的潜伏期病例数和确诊病例数;

S4:基于区域间人口流动数据、区域人口信息和区域类别信息,结合步骤S1和步骤S3中的区域粒度公共卫生事件信息,以小时为单位将一天时间进行划分,构建以区域作为节点,以人口流动作为边的24小时人口流动异质有向多图;

S5:根据步骤S4获得的人口流动异质有向多图,基于所构建的异质有向图的特点,选取合理的元路径,一种元路径代表图中一种人口流动模式,在空间上捕捉各小时人口流动图内每个区域与其关联区域的信息,得到各小时人口流动图区域节点的空间嵌入表示;

S6:根据24张人口流动图的时序关系,在时间维度进行各节点的信息融合,得到最终的城市细粒度区域嵌入表示;

S7:构建城市重点区域识别网络,通过该网络将区域节点嵌入表示映射至各区域风险程度;城市重点区域识别网络为一个多层神经网络,包含输入层,隐藏层和输出层;输入层将步骤S6得到的区域嵌入表示输入该网络;隐藏层有多层全连接层组成,逐层提取特征;输出层最终输出各个区域潜在的风险等级,实现风险预测。

2. 如权利要求1所述的面向POI的突发公共卫生事件重点区域预测方法,其特征在于:所述步骤S1中,预测信息还至少包括城市粒度公共卫生事件可观测信息、区域粒度公共卫生事件可观测信息、区域人口信息、区域间人口流动数据和区域类别信息,其中,

所述城市粒度公共卫生事件可观测信息:至少包括一个城市范围内总的每天的可观测的确诊病例数据、潜伏期病例数据、死亡病例数据和康复病例数据;

所述区域粒度公共卫生事件可观测信息:至少包括城市范围内各细粒度区域每天的可观测的确诊病例数据、潜伏期病例数据;

所述区域人口信息:包括城市范围内各细粒度区域每天的人流量数据;

所述区域间人口流动数据:包括城市范围内各细粒度区域之间一天各小时的人口流动数据,且为一个区域至另一个区域的有向人口流动;

所述区域类别信息:包括城市范围内各细粒度区域属于AOI区域或POI区域的类别划分。

3. 如权利要求2所述的面向POI的突发公共卫生事件重点区域预测方法,其特征在于:所述步骤S2中的传染病模型包含多种不同状态的人群,具体包括易感者、潜伏者、患病者、

隔离潜伏者、隔离患病者、康复者和死亡者共七类人群,传染病模型通过多条微分方程构建了七类人群转化的规律,每条微分方程代表了每类人群由其他人群转化的概率;微分方程中包括有潜伏者和患病者传染易感者的概率 $\beta_E$ 和 $\beta_I$ 、潜伏者转化为发病者的概率 $\delta$ 、潜伏者和患病者被检测隔离的概率 $\gamma_E$ 和 $\gamma_I$ 、及患病者康复概率 $\mu$ 等参数,通过步骤S1中统计收集的城市粒度公共卫生事件可观测信息对以上参数进行拟合,得到公共卫生事件演变规律。

4.如权利要求2或3所述的面向POI的突发公共卫生事件重点区域预测方法,其特征在于:所述步骤S3中,信息重建网络接收区域粒度公共卫生事件可观测信息,可观测的确诊病例数据、潜伏期病例数据在传染病模型中分别对应隔离患病者和隔离潜伏者,经过放缩模块进行数据放缩;将数据结合步骤S2中获得的传染病模型演变规律,即潜伏者和患病者转化为隔离潜伏者和隔离患病者,分别输入至两个门控循环神经网络框架的分支中,分别对各个区域每日的潜伏者和患病者的数量信息进行重建;所述门控循环神经网络框架的每一条分支,是研究区域数个门控循环单元的串行拼接,并在末尾处,拼接一个全连接层,将循环神经网络隐藏状态映射至需要重建的两类人群的数据,即区域粒度未隔离的潜伏期病例数和确诊病例数,在传染病模型中分别对应潜伏者和患病者,并对输出进行放缩,使得数据数量级符合现实数据。

5.如权利要求4所述的面向POI的突发公共卫生事件重点区域预测方法,其特征在于:所述步骤S4中,以每个区域为节点,节点信息包含区域人口信息、区域类别信息以及区域公共卫生事件信息;以区域间的24小时人口流动作为节点的连接关系,根据人口流动的有向性,构建出24小时人口流动有向图;结合区域类别信息,对每个节点进行AOI和POI类别划分,将24小时人口流动有向图转变为包含两类节点的24小时人口流动异质有向图。

6.如权利要求5所述的面向POI的突发公共卫生事件重点区域预测方法,其特征在于:所述步骤S5具体包括如下步骤:

S51:使用元路径进行节点级别的聚合;

基于选取的每一条元路径分别进行节点信息聚合,根据元路径进行每个节点的邻居节点的采样;节点聚合在同一种元路径下对节点的信息进行传递,根据当前节点的采样进行节点间逐层的消息传递,得到当前节点的信息聚合表示;节点级别的消息传递采用自注意力机制对不同节点的消息赋予不同的权重,基于元路径P连接的一个节点对 $(i, j)$ ,节点的注意力 $An_{ij}^P$ 可以学习节点j对于节点i的重要程度:

$$An_{ij}^P = node\_att(n_i', n_j'; P)$$

其中,node\_att是节点级别信息聚合表示注意力的深度神经网络;

对每个节点根据其相应的元路径P可得到相应的元路径邻居节点集合 $N_i^P$ ,使用node\_att得到每个节点 $j \in N_i^P$ 对于节点i的重要性指数,并对同一元路径P下得到的节点重要性进行归一化:

$$Wn_{ij}^P = softmax_j(An_{ij}^P) = \frac{\exp(\sigma(An_{ij}^P))}{\sum_{k \in N_i^P} \exp(\sigma(An_{ik}^P))}$$

其中, $\sigma$ 是激活函数,得到每个从元路径P采样得到的邻居节点的权重 $Wn_{ij}^P$ ,根据权重,

对每个邻居节点进行特征的聚合,得到:

$$info_i^P = \sigma(\sum_{j \in N_i^P} W n_{ij}^P \cdot n_j')$$

$info_i^P$  是在元路径P下学习到的关于节点i的嵌入表示;

重复N次节点级别注意力,并将学习到的N次节点的嵌入表示进行拼接:

$$info_i^P = ||_{n=1}^N (\sum_{j \in N_i^P} W_{ij}^P \cdot n_j')$$

对每条元路径  $\{P_0, P_1, \dots, P_q\}$ , 采用上述的节点多头注意力,求得各个节点在每条元路径下的嵌入表示,得到  $\{Info_{P_0}, Info_{P_1}, \dots, Info_{P_q}\}$ ;

S52: 对不同元路径聚合得到的表示进行语义级别的聚合;

考量不同的元路径P对于节点嵌入表示的重要性,使用语义级别注意力机制获取不同元路径重要程度:

$$(As_{P_0}, As_{P_1}, \dots, As_{P_q}) = sem\_att(Info_{P_0}, Info_{P_1}, \dots, Info_{P_q})$$

其中,  $sem\_att$  是语义级别信息聚合表示注意力的深度神经网络;

对每条路径获得的重要程度取归一化,获得元路径权重:

$$Ws_{P_i} = softmax(As_{P_i}) = \frac{\exp(As_{P_i})}{\sum_{i=1}^q \exp(As_{P_i})}$$

根据所得的元路径权重,对每条元路径下节点的嵌入表示进行聚合,得到单个图的总的嵌入表示Info如下:

$$Info = \sum_{i=1}^P Ws_{P_i} \cdot Info_{P_i}。$$

7. 如权利要求6所述的面向POI的突发公共卫生事件重点区域预测方法,其特征在于:所述步骤S6中,根据24张人口流动图的时序关系,在时间维度上构建了不同节点的人口流动关系;

考虑不同图  $G_i$  和  $G_j$  下的节点  $n_k^i$  和  $n_k^j$ , 使用注意力机制获得  $n_k^j$  对  $n_k^i$  的注意力权重,同时考虑到图  $G_i$  和  $G_j$  间时序关系:

$$\alpha_{ij}^{n_k} = \frac{\exp(graph\_att(n_k^i, n_k^j, t_{ij}))}{\sum_{l=1}^{24} \exp(graph\_att(n_k^i, n_k^l, t_{il}))}$$

应用多头注意力得到每个图中的节点嵌入表示,最后对每个节点使用加权平均求得所有节点最后的嵌入表示Z。

8. 实施如权利要求1所述方法的面向POI的突发公共卫生事件重点区域预测系统,其特征在于:至少包括数据采集模块、信息重建模块及图信息融合模块,

所述数据采集模块:用于采集及统计城市范围内多个细粒度区域的预测信息,所述预测信息包括但不限于城市粒度公共卫生事件可观测信息、区域粒度公共卫生事件可观测信

息、区域人口信息、区域间人口流动数据和区域类别信息；

所述信息重建模块：根据突发公共卫生事件数据拟合城市范围内的传染病模型，挖掘公共卫生事件演变规律特点；

所述图信息融合模块：将信息重建模块重建的信息与现有信息进行聚合，构建城市细粒度区域的嵌入表示；所述图信息融合模块至少包括元路径采样模块、多图消息传递模块和重点区域识别网络，其中，

元路径采样模块用于对人口流动图进行图内的信息聚合得到各个图内的各个节点的嵌入表示；多图消息传递模块用于根据每个图内的节点聚合进行多图间信息聚合；重点区域识别网络用于进行各个细粒度区域潜在的风险等级预测。

9. 如权利要求8所述方法的面向POI的突发公共卫生事件重点区域预测系统，其特征在于：所述元路径采样模块包括节点消息聚合模块和语义消息聚合模块，

所述节点消息聚合模块：选取合适元路径，聚合邻居节点信息，通过元路径提取人口流动图的人口流动特征，挖掘节点间人口潜在关联，输出各节点嵌入表示。

所述语义消息聚合模块：聚合不同基于元路径聚合得到的节点嵌入表示，综合不同元路径表示的人口流动模式得到各小时的人口流动图的节点嵌入表示。

## 面向POI的突发公共卫生事件重点区域预测方法及系统

### 技术领域

[0001] 本发明属于计算机智能计算与运用的技术领域，主要涉及了一种面向POI的突发公共卫生事件重点区域预测方法及系统。

### 背景技术

[0002] 突发公共卫生事件是指突然发生，造成或者可能造成社会公众健康严重损害的重大传染病疫情、群体性不明原因疾病、重大食物和职业中毒以及其他严重影响公众健康的事件，对突发公共卫生事件的演变情况的掌握是后续进行有效决策的重要基础，根据公共卫生事件的传播性特点，使用传染病模型对事件进行建模并预测是一种主流的研究方法。现有的关于公共卫生事件的事件演变的预测研究，通常使用大量的数据，构建优良的神经网络去拟合传染病模型，以此挖掘事件数据中隐含的关于事件演化的机理与规律，并预测后续时间的事件演变情形，并以此为基础进行后续的决策。但使用数据直接拟合的模型，通常受真实数据滞后性的影响，且没有考虑公共卫生事件中不可观测信息在每个时间步下的直接影响，对于精确的城市内细粒度重点区域(POI/AOI)识别存在着不足。

[0003] 另一方面，在突发公共卫生事件中，人群的移动性是影响事件演变的最主要因素，公共卫生事件的扩散与事件源处的人口流动呈现正相关关系，因此，在人群移动性对于公共卫生事件演变的影响方面也有许多研究。现有关于移动性的研究主要基于大范围的城市间或区级间的人口的流动，没有对更细粒度的区域之间的人群移动性对于公共卫生事件的影响进行深入研究。细粒度区域的人群移动性存在着更丰富的区域级别特征，现有研究主要聚焦粗粒度区域的模式和统计方法，无法直接适用于细粒度区域级别。

[0004] 此外，公共卫生事件的不可观测信息，会随着时间演变，在未来时间周期中可以显性捕捉，存在着一定的时间规律，可从可观测数据对不可观测数据进行直接的推测。城市细网格人口流动数据能够动态反映城市的人流、细粒度区域的访问情况，为细粒度区域下公共卫生事件演变的研究也提供了足够多的机会。

### 发明内容

[0005] 本发明正是针对现有技术中存在的问题，提供一种面向POI的突发公共卫生事件重点区域预测方法及系统，首先将待预测城市的范围划分为多个细粒度区域，并基于细粒度区域进行预测信息的收集与统计；然后基于城市粒度公共卫生事件可观测信息进行传染病模型构建，进行公共卫生事件演变规律的挖掘；针对演变规律，搭建公共卫生事件不可观测信息重建网络，通过门控循环神经网络对可观测信息时序上的特征进行捕捉，输出公共卫生事件下的不可观测信息；同时，根据区域间人口流动数据和区域人口信息，结合S3得到的区域粒度公共卫生事件下的不可观测信息，以一天24小时为时间单位划分，构建以细粒度区域为节点的人口流动异质有向多图，基于所构建的异质有向图的特点，选取合理的元路径，在空间上捕捉每个人口流动图内各区域间与关联区域的信息，得到每个人口流动图的空间嵌入表示；再根据24小时人口流动有向图的时序关系，在时间维度进行各节点的信

息融合,得到最终的城市细粒度区域嵌入表示,并将其输入城市重点区域识别网络,得到各个细粒度区域潜在的风险等级,实现风险预测。

[0006] 为了实现上述目的,本发明采取的技术方案是:面向POI的突发公共卫生事件重点区域预测方法,包括如下步骤:

[0007] S1:确定待预测城市的范围,将城市范围划分为多个细粒度区域,并基于细粒度区域进行预测信息的收集与统计;所述预测信息至少包括城市粒度公共卫生事件可观测信息、区域粒度公共卫生事件的可观测信息和城市内细粒度区域间人口流动数据;

[0008] S2:基于城市粒度公共卫生事件可观测信息,进行传染病模型构建;所述传染病模型包含不同状态的人群,并由微分方程组成,每一条微分方程表示人群在每个时间步下的状态变化的概率,传染病模型用于不同状态人群之间状态转移概率的构建,根据城市粒度公共卫生事件可观测信息拟合的传染病模型可以挖掘公共卫生事件演变规律;

[0009] S3:针对步骤S2获得的公共卫生事件演变规律,搭建公共卫生事件不可观测信息重建网络,所述信息重建网络由门控循环神经网络构成,基于区域粒度公共卫生事件的可观测信息,通过门控循环神经网络对可观测信息时序上的特征进行捕捉,输出区域粒度公共卫生事件下的不可观测信息,所述不可观测信息包括区域粒度潜伏期人群数和康复人群数;

[0010] S4:基于区域间人口流动数据、区域人口信息和区域类别信息,结合步骤S3得到的区域粒度公共卫生事件下的不可观测信息,以小时为单位将一天时间进行划分,构建以区域作为节点,以人口流动作为边的24小时人口流动异质有向多图;

[0011] S5:根据步骤S4获得的人口流动异质有向多图,基于所构建的异质有向图的特点,选取合理的元路径,在空间上捕捉各小时人口流动图内每个区域与其关联区域的信息,得到各小时人口流动图区域节点的空间嵌入表示;

[0012] S6:根据24小时人口流动图的时序关系,在时间维度进行各区域节点的信息融合,得到最终的区域节点嵌入表示;

[0013] S7:构建城市重点区域识别网络,通过该网络将区域节点嵌入表示映射至各区域风险程度;城市重点区域识别网络为一个多层神经网络,包含输入层,隐藏层和输出层;输入层将步骤S6得到的区域嵌入表示输入该网络;隐藏层有多层全连接层组成,逐层提取特征;输出层最终输出各个区域潜在的风险等级,实现风险预测。

[0014] 作为本发明的一种改进,所述步骤S1中,预测信息还至少包括城市粒度公共卫生事件可观测信息、区域粒度公共卫生事件可观测信息、区域人口信息、区域间人口流动数据和区域类别信息,其中,

[0015] 所述城市粒度公共卫生事件可观测信息:至少包括一个城市范围内总的每天的可观测的确诊病例数据、潜伏期病例数据、死亡病例数据和康复病例数据;

[0016] 所述区域粒度公共卫生事件可观测信息:至少包括城市范围内各细粒度区域每天的可观测的确诊病例数据、潜伏期病例数据;

[0017] 所述区域人口信息:包括城市范围内各细粒度区域每天的人流量数据;

[0018] 所述区域间人口流动数据:包括城市范围内各细粒度区域之间一天各小时的人口流动数据,且为一个区域至另一个区域的有向人口流动;

[0019] 所述区域类别信息:包括城市范围内各细粒度区域属于AOI区域或POI区域的类别

划分。

[0020] 作为本发明的一种改进,所述步骤S2中的传染病模型包含多种不同状态的人群,具体包括易感者、潜伏者、患病者、隔离潜伏者、隔离患病者、康复者和死亡者共七类人群,传染病模型通过多条微分方程构建了七类人群转化的规律,每条微分方程代表了每类人群由其他人群转化的概率;微分方程中包括有潜伏者和患病者传染易感者的概率 $\beta_E$ 和 $\beta_I$ 、潜伏者转化为发病者的概率 $\delta$ 、潜伏者和患病者被检测隔离的概率 $\gamma_E$ 和 $\gamma_I$ 、及患病者康复概率 $\mu$ 等参数,通过步骤S1中统计收集的城市粒度公共卫生事件可观测信息对以上参数进行拟合,得到公共卫生事件演变规律。

[0021] 作为本发明的另一种改进,所述步骤S3中,信息重建网络接收区域粒度公共卫生事件可观测信息,可观测的确诊病例数据、潜伏期病例数据在传染病模型中分别对应隔离患病者和隔离潜伏者,经过放缩模块进行数据放缩;将数据结合步骤S2中获得的传染病模型演变规律,即潜伏者和患病者转化为隔离潜伏者和隔离患病者,分别输入至两个门控循环神经网络框架的分支中,分别对各个区域每日的潜伏者和患病者的数量信息进行重建;所述门控循环神经网络框架的每一条分支,是研究区域数个门控循环单元的串行拼接,并在末尾处,拼接一个全连接层,将循环神经网络隐藏状态映射至需要重建的两类人群的数据,即区域粒度未隔离的潜伏期病例数和确诊病例数,在传染病模型中分别对应潜伏者和患病者,并对输出进行放缩,使得数据数量级符合现实数据。

[0022] 作为本发明的另一种改进,所述步骤S4中,以每个区域为节点,节点信息包含区域人口信息、区域类别信息以及区域公共卫生事件信息;以区域间的24小时人口流动作为节点的连接关系,根据人口流动的有向性,构建出24小时人口流动有向图;结合区域类别信息,对每个节点进行A0I和POI类别划分,将24小时人口流动有向图转变为包含两类节点的24小时人口流动异质有向图。

[0023] 作为本发明的又一种改进,所述步骤S5具体包括如下步骤:

[0024] S51:使用元路径进行节点级别的聚合;

[0025] 基于选取的每一条元路径分别进行节点信息聚合,根据元路径进行每个节点的邻居节点的采样;节点聚合在同一种元路径下对节点的信息进行传递,根据当前节点的采样进行节点间逐层的消息传递,得到当前节点的信息聚合表示;节点级别的消息传递采用自注意力机制对不同节点的消息赋予不同的权重,基于元路径P连接的一个节点对 $(i, j)$ ,节点的注意力 $An_{ij}^P$ 可以学习节点j对于节点i的重要程度:

$$[0026] \quad An_{ij}^P = node\_att(n_i', n_j'; P)$$

[0027] 其中,node\_att是节点级别信息聚合表示注意力的深度神经网络;

[0028] 对每个节点根据其相应的元路径P可得到相应的元路径邻居节点集合 $N_i^P$ ,使用node\_att得到每个节点 $j \in N_i^P$ 对于节点i的重要性指数,并对同一元路径P下得到的节点重要性进行归一化:

$$[0029] \quad Wn_{ij}^P = softmax_j(An_{ij}^P) = \frac{\exp(\sigma(An_{ij}^P))}{\sum_{k \in N_i^P} \exp(\sigma(An_{ik}^P))}$$



[0030] 其中,  $\sigma$  是激活函数, 得到每个从元路径  $P$  采样得到的邻居节点的权重  $Wn_{ij}^P$ , 根据权重, 对每个邻居节点进行特征的聚合, 得到:

$$[0031] \quad info_i^P = \sigma \left( \sum_{j \in N_i^P} Wn_{ij}^P \cdot n_j' \right)$$

[0032]  $info_i^P$  是在元路径  $P$  下学习到的关于节点  $i$  的嵌入表示;

[0033] 重复  $N$  次节点级别注意力, 并将学习到的  $N$  次节点的嵌入表示进行拼接:

$$[0034] \quad info_i^P = ||_{n=1}^N \left( \sum_{j \in N_i^P} W_{ij}^P \cdot n_j' \right)$$

[0035] 对每条元路径  $\{P_0, P_1, \dots, P_q\}$ , 采用上述的节点多头注意力, 求得各个节点在每条元路径下的嵌入表示, 得到  $\{Info_{P_0}, Info_{P_1}, \dots, Info_{P_q}\}$ ;

[0036] S52: 对不同元路径聚合得到的表示进行语义级别的聚合;

[0037] 考量不同的元路径  $P$  对于节点嵌入表示的重要性, 使用语义级别注意力机制获取不同元路径重要程度:

$$[0038] \quad (As_{P_0}, As_{P_1}, \dots, As_{P_q}) = sem\_att(Info_{P_0}, Info_{P_1}, \dots, Info_{P_q})$$

[0039] 其中,  $sem\_att$  是语义级别信息聚合表示注意力的深度神经网络;

[0040] 对每条路径获得的重要程度取归一化, 获得元路径权重:

$$[0041] \quad Ws_{P_i} = softmax(As_{P_i}) = \frac{\exp(As_{P_0})}{\sum_{i=1}^q \exp(As_{P_i})}$$

[0042] 根据所得的元路径权重, 对每条元路径下节点的嵌入表示进行聚合, 得到单个图的总的嵌入表示  $Info$  如下:

$$[0043] \quad Info = \sum_{i=1}^P Ws_{P_i} \cdot Info_{P_i}.$$

[0044] 作为本发明的又一种改进, 所述步骤 S6 中, 根据 24 张人口流动图的时序关系, 在时间维度上构建了不同节点的人口流动关系;

[0045] 考虑不同图  $G_i$  和  $G_j$  下的节点  $n_k^i$  和  $n_k^j$ , 使用注意力机制获得  $n_k^j$  对  $n_k^i$  的注意力权重, 同时考虑到图  $G_i$  和  $G_j$  间时序关系:

$$[0046] \quad \alpha_{ij}^{n_k} = \frac{\exp(graph\_att(n_k^i, n_k^j, t_{ij}))}{\sum_{l=1}^{24} \exp(graph\_att(n_k^i, n_k^l, t_{il}))}$$

[0047] 应用多头注意力得到每个图中的节点嵌入表示, 最后对每个节点使用加权平均求得所有节点最后的嵌入表示  $Z$ 。

[0048] 为了实现上述目的, 本发明还采取的技术方案是: 面向 POI 的突发公共卫生事件重点区域预测系统, 至少包括数据采集模块、信息重建模块及图信息融合模块,

[0049] 所述数据采集模块: 用于采集及统计城市范围内多个细粒度区域的预测信息, 所

述预测信息包括但不限于城市粒度公共卫生事件可观测信息、区域粒度公共卫生事件可观测信息、区域人口信息、区域间人口流动数据和区域类别信息；

[0050] 所述信息重建模块：根据突发公共卫生事件数据拟合城市范围内的传染病模型，挖掘公共卫生事件演变规律特点；

[0051] 所述图信息融合模块：将信息重建模块重建的信息与现有信息进行聚合，构建城市细粒度区域的嵌入表示；所述图信息融合模块至少包括元路径采样模块、多图消息传递模块和重点区域识别网络，其中，

[0052] 元路径采样模块用于对人口流动图进行图内的信息聚合得到各个图内的各个节点的嵌入表示；多图消息传递模块用于根据每个图内的节点聚合进行多图间信息聚合；重点区域识别网络用于进行各个细粒度区域潜在的风险等级预测。

[0053] 作为本发明的一种改进，所述元路径采样模块包括节点消息聚合模块和语义消息聚合模块，

[0054] 所述节点消息聚合模块：选取合适元路径，聚合邻居节点信息，通过元路径提取人口流动图的人口流动特征，挖掘节点间人口潜在关联，输出各节点嵌入表示。

[0055] 所述语义消息聚合模块：聚合不同基于元路径聚合得到的节点嵌入表示，综合不同元路径表示的人口流动模式得到各小时的人口流动图的节点嵌入表示。

[0056] 与现有技术相比，本发明具有的有益效果：提供了一种面向POI的突发公共卫生事件重点区域预测方法及系统，首先，在传染病模型模块，考虑了更符合现实场景的模型进行构建；其次，对于公共卫生事件的信息，使用信息重建模块着重考虑了不可观测信息对于未来事件演变的影响，结合针对性设计的传染病模型，能更好地重建不可观测信息；除此之外，使用细粒度区域相关信息，使得图融合模块根据我们图内和图间的信息融合得到节点嵌入表示能很好地获取细粒度特征，有效精准地对细粒度区域风险等级进行识别。本发明的方法和系统可充分用于公共卫生事件场景，对于图信息融合模块，可更广泛的适用于城市范围内的事件推演。

## 附图说明

[0057] 图1为本发明面向POI的突发公共卫生事件重点区域预测系统的结构框图。

## 具体实施方式

[0058] 下面结合附图和具体实施方式，进一步阐明本发明，应理解下述具体实施方式仅用于说明本发明而并不用于限制本发明的范围。

[0059] 实施例1

[0060] 面向POI的突发公共卫生事件重点区域预测方法，主要用于存在突发公共卫生事件的城市风险判断，通过收集细粒度区域的相关信息和数据，逐步进行模型构建、数据重建，得到图信息融合模块必要输入，再经过数据信息的传递最后得到嵌入表达，通过最后的重点区域识别网络得到各个区域的风险识别等级，辅助后续的上层决策进行。该方法具体包括如下步骤：

[0061] 步骤S1：明确方法作用的城市范围，划分为多个细粒度区域，并基于细粒度区域进行相关信息的收集与统计；

[0062] 城市细粒度区域以城市POI与AOI范围进行划分,并以细粒度区域进行方法所需信息的收集与统计,至少包括:城市粒度公共卫生事件可观测信息、区域粒度公共卫生事件可观测信息、区域人口信息、区域间人口流动数据和区域类别信息具体信息内容如下:

[0063] 所述城市粒度公共卫生事件可观测信息:至少包括一个城市范围内总的每天的可观测的确诊病例数据、潜伏期病例数据、死亡病例数据和康复病例数据;

[0064] 所述区域粒度公共卫生事件可观测信息:至少包括城市范围内各细粒度区域每天的可观测的确诊病例数据、潜伏期病例数据;

[0065] 所述区域人口信息:包括城市范围内各细粒度区域每天的人流量数据;

[0066] 所述区域间人口流动数据:包括城市范围内各细粒度区域之间一天各小时的人口流动数据,且为一个区域至另一个区域的有向人口流动;

[0067] 所述区域类别信息:包括城市范围内各细粒度区域属于AOI区域或POI区域的类别划分。

[0068] S2:基于城市粒度公共卫生事件可观测信息,进行传染病模型构建,进行公共卫生事件演变规律的挖掘。

[0069] 传染病模型可以用于捕捉真实世界突发公共卫生事件的演变规律,传染病模型将人群划分为了七种人群,包括易感者、潜伏者、患病者、隔离潜伏者、隔离患病者、康复者和死亡者,人群具体含义如下所述:

[0070] (1)易感者(S):在公共卫生事件中,能被传染病感染的一类人群,人群基数最大的群体;

[0071] (2)潜伏者(E):携带着传染病病毒的一类人群,但还未出现具体的发病症状,由易感者被感染后转化为潜伏者,根据传染病病毒种类的不同,病毒潜伏期有着不同的时间,同时病毒种类也决定了潜伏者是否具有传染性,潜伏者具有传染性,有概率将易感者转化为潜伏者。

[0072] (3)患病者(I):携带传染病病毒并已出现发病症状的一类人群,由潜伏者在数天的潜伏期后发病转化为患病者,患病者有病毒传染性,有概率将易感者转化为潜伏者。

[0073] (4)隔离潜伏者(QE):被进行住院隔离的潜伏者,公共卫生事件中,通过患病者的密接排查或全员检测,会采取将一部分潜在人群进行隔离操作。

[0074] (5)隔离患病者(QI):由隔离潜伏者转化而来,或由对确诊的患病者进行隔离操作得到。

[0075] (6)康复者(R):由患病者或隔离患病者经过一段时间的治疗或身体免疫系统自愈后而康复的一类人群,根据传染病病毒种类的不同,康复者体内的病毒免疫时间也不同,免疫时间后,康复者重新转化为易感者,能再次被病毒传染。

[0076] (7)死亡者(D):死于传染病的一类人群。

[0077] 传染病模型通过多条微分方程构建了七类人群转化的规律,每条微分方程代表了每类人群由其他人群转化的概率。微分方程中包括有潜伏者和患病者传染易感者的概率 $\beta_E$ 和 $\beta_I$ ,潜伏者转化为发病者的概率 $\delta$ ,潜伏者和患病者被检测隔离的概率 $\gamma_E$ 和 $\gamma_I$ ,患病者康复概率 $\mu$ 等参数,通过步骤S1中统计收集的城市粒度公共卫生事件可观测信息对以上参数进行拟合,得到公共卫生事件演变规律拟合后得到的传染病模型,基本包含了该城市这段时间的传染病传播规律,为后续信息重建部分提供机理辅助。

[0078] S3:针对步骤S2获得的公共卫生事件演变规律,搭建公共卫生事件不可观测信息重建网络,所述信息重建网络由门控循环神经网络构成,基于区域粒度公共卫生事件的可观测信息,通过门控循环神经网络对可观测信息时序上的特征进行捕捉,输出区域粒度公共卫生事件下的不可观测信息,所述不可观测信息包括区域粒度未隔离的潜伏期病例数和确诊病例数,具体为:

[0079] 信息重建网络接收区域粒度公共卫生事件可观测信息,可观测的确诊病例数据、潜伏期病例数据在传染病模型中分别对应隔离患病者和隔离潜伏者,为了消除数据经过门控循环神经网络的激活函数Tanh后导致的一定的缩放比例,因此数据的输入首先经过一个放缩模块进行数据的放缩;将数据结合结合步骤S2中获得的传染病模型的演变规律,即潜伏者和患病者转化为隔离潜伏者和隔离患病者,分别输入至两个门控循环神经网络框架的分支中,分别对各个区域每日的潜伏者和患病者的数量信息进行重建;所述门控循环神经网络框架的每一条分支,是研究区域数个门控循环单元的串行拼接,并在末尾处,拼接一个全连接层,将循环神经网络隐藏状态映射至需要重建的两类人群的数据,即区域粒度未隔离的潜伏期病例数和确诊病例数,在传染病模型中分别对应潜伏者和患病者,并对输出进行放缩,使得数据数量级符合现实数据。

[0080] S4:基于区域间人口流动数据、区域人口信息和区域类别信息,结合步骤S1和步骤S3中的区域粒度公共卫生事件信息,以小时为单位将一天时间进行划分,构建以区域作为节点,以人口流动作为边的24小时人口流动异质有向多图;

[0081] 以每个区域为节点,节点信息包含区域人口信息、区域类别信息以及区域公共卫生事件信息;以区域间的24小时人口流动作为节点的连接关系,根据人口流动的有向性,构建出24小时人口流动有向图;结合区域类别信息,对每个节点进行A0I和POI类别划分,将24小时人口流动有向图转变为包含两类节点的24小时人口流动异质有向图。

[0082] S5:根据步骤S4获得的人口流动异质有向多图,基于所构建的异质有向图的特点,选取合理的元路径,一种元路径代表图中一种人口流动模式,在空间上捕捉各小时人口流动图内每个区域与其关联区域的信息,得到各小时人口流动图区域节点的空间嵌入表示;

[0083] 对人口流动图进行图内的信息聚合得到各个图内的各个节点的嵌入表示,图内信息融合具体设计如下:

[0084] 所构建的人口流动图是异质图,因此对于异质图的信息聚合采取两阶段方法,首先使用元路径进行节点级别的聚合,而后对不同元路径聚合得到的表示再进行语义级别的聚合。

[0085] 由于不同类型的节点对应的特征空间不同,因此需要将A0I或POI节点*i*根据类型 $type_i$ 构造不同的变换矩阵 $M_{type_i}$ ,映射至相同的特征空间:

$$[0086] \quad n'_i = M_{type_i} n_i$$

[0087] 其中 $n_i$ 和 $n'_i$ 分别为映射前后的节点的特征表示。

[0088] 节点聚合:在节点进行特征空间映射后,选取多条元路径,如A-P-A(表示按A0I-POI-A0I的节点类型选取),并根据元路径进行每个节点的邻居节点的采样。节点聚合在同一种元路径下对节点的信息进行传递,根据当前节点的采样进行节点间逐层的消息传递,得到当前节点的信息聚合表示。节点级别的消息传递采用自注意力机制对不同节点的消息

赋予不同的权重,基于元路径P连接的一个节点对(i,j),节点的注意力 $An_{ij}^P$ 可以学习节点j对于节点i的重要程度:

$$[0089] \quad An_{ij}^P = node\_att(n_i', n_j'; P)$$

[0090] node\_att是节点级别信息聚合表示注意力的深度神经网络。

[0091] 对每个节点根据其相应的元路径P可得到相应的元路径邻居节点集合 $N_i^P$ ,使用node\_att得到每个节点 $j \in N_i^P$ 对于节点i的重要性指数,并对同一元路径P下得到的节点重要性进行归一化:

$$[0092] \quad Wn_{ij}^P = softmax_j(An_{ij}^P) = \frac{\exp(\sigma(An_{ij}^P))}{\sum_{k \in N_i^P} \exp(\sigma(An_{ik}^P))}$$

[0093]  $\sigma$ 是激活函数,得到每个从元路径P采样得到的邻居节点的权重 $Wn_{ij}^P$ ,根据权重,对每个邻居节点进行特征的聚合,得到:

$$[0094] \quad info_i^P = \sigma\left(\sum_{j \in N_i^P} Wn_{ij}^P \cdot n_j'\right)$$

[0095]  $info_i^P$ 是在元路径P下学习到的关于节点i的嵌入表示,为解决异构图存在的图数据方差高的问题,将节点级别注意力node\_att拓展为多头注意力,使训练过程更稳定,即重复N次节点级别注意力,并将学习到的N次节点的嵌入表示进行拼接:

$$[0096] \quad info_i^P = ||_{n=1}^N \left( \sum_{j \in N_i^P} W_{ij}^P \cdot n_j' \right)$$

[0097] 对每条元路径 $\{P_0, P_1, \dots, P_q\}$ ,采用上述的节点多头注意力,求得各个节点在每条元路径下的嵌入表示,得到 $\{Info_{P_0}, Info_{P_1}, \dots, Info_{P_q}\}$ 。

[0098] 语义级别的聚合:在节点级别进行信息聚合得到各元路径下的节点嵌入表示后,针对不同元路径进行了语义级别的消息传递,得到最终的节点嵌入表示。

[0099] 对于不同的元路径P,需要考量分别对于节点嵌入表示的重要性,因此使用语义级别注意力机制获取不同元路径重要程度:

$$[0100] \quad (As_{P_0}, As_{P_1}, \dots, As_{P_q}) = sem\_att(Info_{P_0}, Info_{P_1}, \dots, Info_{P_q})$$

[0101] sem\_att是语义级别信息聚合表示注意力的深度神经网络。

[0102] 对每条路径获得的重要程度取归一化,获得元路径权重:

$$[0103] \quad Ws_{P_i} = softmax(As_{P_i}) = \frac{\exp(As_{P_0})}{\sum_{i=1}^q \exp(As_{P_i})}$$

[0104] 元路径的权重越高,表明元路径对相应任务的贡献度越高,越重要。根据所得的元路径权重,对每条元路径下节点的嵌入表示进行聚合,得到单个图的总的嵌入表示Info如

下:

$$[0105] \quad Info = \sum_{i=1}^P W_{s_{p_i}} \cdot Info_{p_i} \quad .$$

[0106] S6:根据24张人口流动图的时序关系,在时间维度进行各节点的信息融合,得到最终的城市细粒度区域嵌入表示;

[0107] 得到每个图的图内节点聚合表示后,进行多图间信息聚合,得到最终各个节点的嵌入表示。

[0108] 考虑不同图 $G_i$ 和 $G_j$ 下的节点 $n_k^i$ 和 $n_k^j$ ,同样使用注意力机制获得 $n_k^j$ 对 $n_k^i$ 的注意力权重,同时考虑到图 $G_i$ 和 $G_j$ 间时序关系:

$$[0109] \quad \alpha_{ij}^{n_k} = \frac{\exp(\text{graph\_att}(n_k^i, n_k^j, t_{ij}))}{\sum_{l=1}^{24} \exp(\text{graph\_att}(n_k^i, n_k^l, t_{il}))}$$

[0110] 再次应用多头注意力得到每个图中的节点嵌入表示,最后对每个节点使用加权平均求得所有节点最后的嵌入表示Z。

[0111] S7:构建城市重点区域识别网络,通过该网络将区域节点嵌入表示映射至各区域风险程度;城市重点区域识别网络为一个多层神经网络,包含输入层,隐藏层和输出层;输入层将步骤S6得到的区域嵌入表示输入该网络;隐藏层有多层全连接层组成,逐层提取特征;输出层最终输出各个区域潜在的风险等级,实现风险预测。

[0112] 实施例2

[0113] 实施上述实施例1的一种面向POI的突发公共卫生事件重点区域预测系统,如图1所示,至少包括数据采集模块、信息重建模块及图信息融合模块,

[0114] 所述数据采集模块用于采集及统计城市范围内多个细粒度区域的预测信息,所述预测信息包括但不限于城市范围公共卫生事件可观测信息,城市内细粒度区域公共卫生事件可观测信息、城市内细粒度区域人口信息、城市内细粒度区域间人口流动数据和区域类别信息;

[0115] 所述信息重建模块:根据突发公共卫生事件数据拟合城市范围内的传染病模型,挖掘公共卫生事件演变规律特点;传染病模型模块至少包括城市范围公共卫生事件可观测信息,通过信息构建城市范围传染病模型,挖掘传染病演变特点,结合传染病模型辅助门控循环神经网络,重建突发公共卫生事件中的不可观测信息,至少包括城市内细粒度区域公共卫生事件可观测信息和人口信息,通过可观测信息,结合传染病模型,重建细粒度区域公共卫生事件不可观测信息;

[0116] 所述图信息融合模块:将信息重建模块重建的信息与现有信息进行聚合,构建城市细粒度区域的嵌入表示;以此挖掘突发公共卫生事件的发展可能,识别具有潜在风险的城市细粒度重点区域,为后续上层决策提供更精确地决策前提;

[0117] 所述图信息融合模块至少包括元路径采样模块、多图消息传递模块和重点区域识别网络,其中,元路径采样模块用于对人口流动图进行图内的信息聚合得到各个图内的各个节点的嵌入表示;多图消息传递模块用于根据每个图内的节点聚合进行多图间信息聚合;重点区域识别网络用于进行各个细粒度区域潜在的风险等级预测。

[0118] 实施例3

[0119] 数据获取:以20年初的某城市为例,根据居民区、写字楼、商场、医院等真实地理实体,获取该城市AOI、POI级别的细粒度区域划分,得到区域类别信息;基于区域粒度统计获取一个月各个区域的人口密度数据,得到区域人口信息;统计24小时区域间人口流动量,得到区域间人口流动数据;获取突发公共卫生事件相关信息,获取该城市一个月城市粒度和区域粒度的病例数据。

[0120] 传染病模型拟合:输入该城市一个月城市粒度的病例数据,拟合传染病模型中各条微分方程的参数,挖掘事件演变规律,在本例中即为挖掘病毒传染性,在出现N个病例时,后续每天感染人数预测。

[0121] 信息重建:基于传染病模型拟合得到的传染病规律,结合每日得到的可观测的病例数据,重建不可直接观测的潜伏期病例以及未被检测的病例数量信息,以辅助更好地进行区域重点区域识别。

[0122] 元路径采样:

[0123] 以获取的居民区、商场等细粒度区域为节点,以每个小时不同区域间人口流动量为边权重构建24小时人口流动图,通过人口流动获取和挖掘不同区域间显式和隐式的关联,显式关联为两个区域间存在直接人口流动,隐式关联为两个不存在直接人口流动的区域通过其他区域而产生的的间接人口流动。

[0124] 节点消息聚合:选取元路径,如AOI-POI-AOI,这条元路径表示人口流动从一个AOI区域到一个POI区域再到一个AOI区域,例如从一个居民区前往一个商场,并随后再前往另一个居民区,两个不直接关联的AOI区域通过POI区域建立了隐式关联,当前者AOI出现病例,则后者AOI通过该连接也有潜在病例可能。通过不同元路径挖掘不同的区域连接序列的关联,并将其他区域信息通过元路径聚合至当前节点构建当前节点的嵌入表示。

[0125] 语义消息聚合:不同元路径存在着不同的权重,元路径1(AOI-POI-AOI)和元路径2(AOI-AOI-AOI)对于区域的影响不同,若元路径1出现频次更多,权重更大,则元路径1对于区域的潜在影响大于元路径2,语义消息聚合通过不同的权重对节点消息聚合中每条元路径获得的节点嵌入表示进行加权聚合,获得每张图下节点的嵌入表示,代表了节点空间上的信息表示。

[0126] 多图消息传递:构建的24小时人口流动图,意味着不同时间下区域间的关联也不同,如早上和下午,分别对应与上班和下班时间,则区域间人口流动方向也会存在不同,不仅有从居民区至办公区的人口流动,同样存在从办公区至居民区的人口流动。多图消息传递聚合了区域节点时间上的信息,得到每个细粒度区域时空嵌入表示,包含了区域自身信息和时间和空间上的区域间的关联。

[0127] 重点区域识别:输入区域时空嵌入表示,当一个或多个区域出现公共卫生事件时,基于包含区域关联的区域时空嵌入表示,能发掘与出现公共卫生事件区域关联强的其他区域,并以此识别潜在的重点区域。

[0128] 需要说明的是,以上内容仅仅说明了本发明的技术思想,不能以此限定本发明的保护范围,对于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,这些改进和润饰均落入本发明权利要求书的保护范围之内。

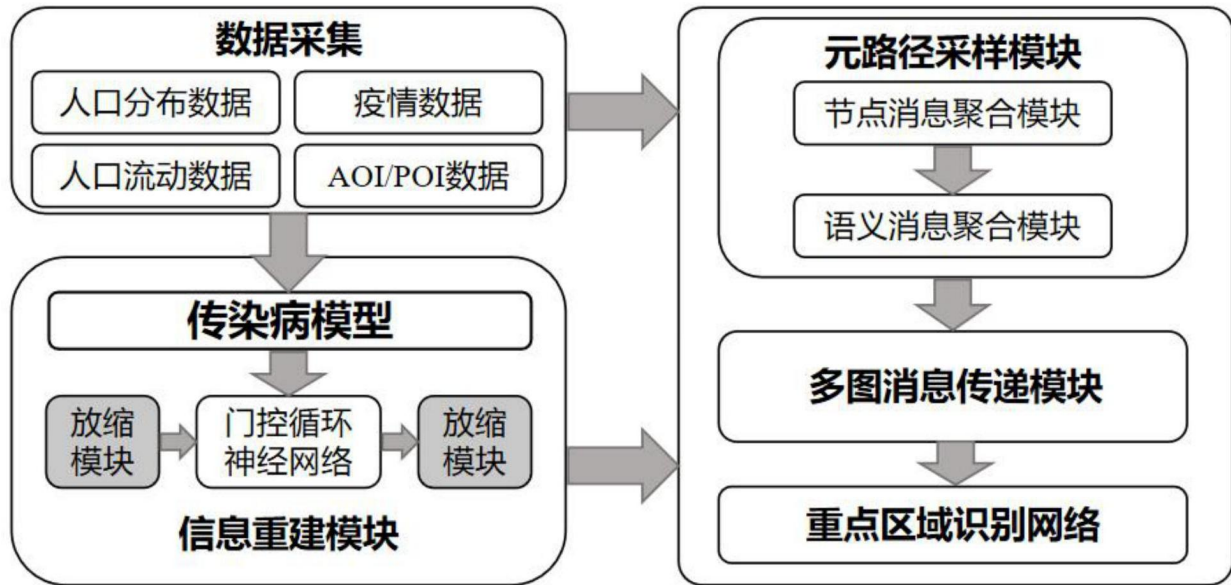


图1