



(21) 申请号 202111365159.X

(22) 申请日 2021.11.17

(65) 同一申请的已公布的文献号

申请公布号 CN 114048783 A

(43) 申请公布日 2022.02.15

(73) 专利权人 东南大学

地址 210000 江苏省南京市麒麟科创园智  
识路26号启迪城立业园04幢

(72) 发明人 王帅 王海 梅洛瑜 徐鑫

(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 薛雨妍

(51) Int. Cl.

G06F 18/25 (2023.01)

G06F 18/213 (2023.01)

G06F 18/10 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/045 (2023.01)

G06N 3/0499 (2023.01)

G06N 3/084 (2023.01)

(56) 对比文件

CN 103745489 A, 2014.04.23

CN 110223509 A, 2019.09.10

CN 111757258 A, 2020.10.09

CN 112532439 A, 2021.03.19

US 2005246334 A1, 2005.11.03

谭朝文; 宗容; 席先强; 武浩. 基于贝叶斯张  
量分解的改进QoS预测. 计算机工程与应用  
.2018, (20), 全文.陈文娟; 周小平; 王家南; 李莉; 杨哲. 基于张  
量分解的大规模多输入多输出天线预编码. 上海  
师范大学学报(自然科学版). 2018, (02), 全文.

审查员 姚子琪

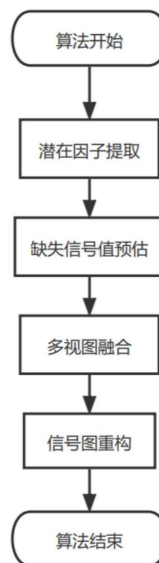
权利要求书3页 说明书9页 附图3页

(54) 发明名称

一种基于移动群体感知的蜂窝信号地图构  
建方法

(57) 摘要

本发明公开了一种基于移动群体感知的关于蜂窝信号地图的构建方法,以三维张量存储信号数据的时空特征,其张量的三个维度分别表示长、宽及时间维度。在构建信号地图时,首先进行MCS信号数据的潜在因子特征采集,然后提取其内部特征。接着进行缺失信号值预估,在潜在因子特征采集的基础上,利用贝叶斯概率张量分解来实现信号恢复,得到MCS数据的内部特征。最后,引入了一种新的基于学习的多视图特征融合模块,将外部特征嵌入到一个向量中,并将其与内部特征相结合,然后将其输入全连接层,获取信号特征并构建细粒度信号图。



1.一种基于移动群体感知的蜂窝信号地图构建方法,其特征在于,该方法包括以下步骤:

步骤1.将信号数据的时空特征表示为三维张量 $S \in \mathbb{R}^{I \times J \times K}$ ,其中,I和J分别表示矩形区域的长和宽的网格数,K表示网格内信号采集时间跨度, $S_{ijk} \in S$ 表示移动设备从基站接收到的信号强度值;

步骤2.MCS信号数据的潜在因子特征提取,在基站覆盖交叉区域的部分标定信号值;

步骤2中,进行潜在因子特征提取方法如下:

步骤201.利用欧几里得距离来描述空间相关性,然后把地图上的地理位置联系起来,利用公式(1)表示基站之间的距离相关性:

$$\Phi^{bs}(i, j) = \exp - \frac{\|l_i^{bs} - l_j^{bs}\|_2}{\sigma^2} \quad (1)$$

其中, $\sigma^2$ 代表方差, $\Phi^{bs}(i, j)$ 代表网格i和j的关系, $\|l_i^{bs} - l_j^{bs}\|_2$ 表示欧几里得距离,即

$$\|l_i^{bs} - l_j^{bs}\|_2 = \sqrt{(l_i^x - l_j^x)^2 + (l_i^y - l_j^y)^2},$$

步骤202.利用公式(2)表示两个高维信号向量之间的余弦相似度,进行信号维度特征提取:

$$\Phi^s(i, j) = \frac{\overrightarrow{slg_i} \cdot \overrightarrow{slg_j}}{\|\overrightarrow{slg_i}\| \|\overrightarrow{slg_j}\|} \quad (2)$$

其中, $\overrightarrow{slg} \in \mathbb{R}^H$ ,H表示维度, $\overrightarrow{slg_i}$ 表示在网格i中从H个基站中接收到的信号向量;

步骤3.采用贝叶斯概率张量分解恢复未观测区域的信号值,即对缺失信号的预估;

步骤3中,采用贝叶斯概率张量分解对缺失信号值的预估,具体方法如下:

步骤301.假设每个观测项 $S_{ijk} \in S$ 近似服从独立高斯分布,即 $S_{ijk} \sim N(\hat{S}_{ijk}, \tau^{-1})$ ;进一步对因子矩阵U,V,W的行向量 $u_i, v_j, w_k$ 进行先验分布,并且均服从多元高斯分布;因子矩阵U的第i行向量 $u_i \sim N(\mu_u, \Lambda_u^{-1})$ ;其中, $\mu_u \in \mathbb{R}^R, \Lambda_u \in \mathbb{R}^{R \times R}$ ,并且满足公式(3), $(\mu_u, \Lambda_u)$ 的概率分布函数为公式(4):

$$(\mu_u, \Lambda_u) \sim \text{Gaussian-Wishart}(\mu_0, \beta_0, W_0, v_0) \quad (3)$$

$$p(\mu_u, \Lambda_u | \Theta) = N(\mu_u | \mu_0, (\beta_0 \Lambda_u)^{-1}) \times \text{Wishart}(\Lambda_u | W_0, v_0) \quad (4)$$

步骤302.在精度参数 $\tau$ 上引入了柔性共轭Gamma先验,利用公式(5)表示精度参数 $\tau$ 先验函数分布:

$$p(\tau | a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \tau^{a_0-1} \exp(-b_0 \tau) \quad (5)$$

其中,参数U,V,W, $\tau$ 和超参数 $\mu_u, \Lambda_u, \mu_v, \Lambda_v, \mu_w, \Lambda_w, a_0, b_0$ ,利用吉布斯抽样求解;

步骤303.因子矩阵U,V,W通过 $u_i, v_j, w_k$ 进行求解,计算方法如下:

步骤30301.由于每个观测项 $S_{ijk}$ 近似服从独立高斯分布,可得到似然函数,表达式如公式(6)所示:

$$L(S|u_i, V, W, \tau) = \exp \left\{ -\frac{\tau}{2} \sum_j^J \sum_k^K b_i \left( S_{ijk} - (u_i)^T (v_j \odot w_k) \right)^2 \right\} \quad (6)$$

公式中  $\odot$  表示Hadamard积;

步骤30302. 由于因子矩阵U的行向量 $u_i$ 服从多元高斯分布,联合步骤30301中的似然函数,通过贝叶斯公式计算得到行向量 $u_i$ 的后验分布,表达式如公式(7)所示:

$$p(u_i | S, V, W, \tau, \mu_u, \Lambda_u^{-1}) \propto L(S|u_i, V, W, \tau) \times N(u_i | \mu_u, \Lambda_u^{-1}) \propto N(u_i | \mu_u, \Lambda_u^{-1}) \quad (7)$$

步骤30303. 根据步骤30302得到的后验分布,按公式(8)、公式(9)求得 $\widehat{\Lambda}_u, \widehat{u}_i$ :

$$\widehat{\Lambda}_u = \tau \sum_j^J \sum_k^K b_i (v_j \odot w_k) (v_j \odot w_k)^T + \Lambda_u \quad (8)$$

$$\widehat{u}_i = \widehat{\Lambda}_u^{-1} \left( \tau \sum_j^J \sum_k^K S_{ijk} (v_j \odot w_k) + \Lambda_u \mu_u \right) \quad (9)$$

步骤304. 通过公式(3) - (9) 求解得到因子矩阵V, W的行向量 $v_j, w_k$ ,最后,在吉布斯采样算法达到平稳状态后,通过蒙特卡罗近似可以估计出所有缺失值,由此实现对缺失信号数据的估计;

步骤4. 首先进行内部和外部视图的构造;然后利用两个嵌入网络分别对每个视图提取高维特征并进行向量表示;最后利用两个全连接层对不同视图向量进行拼接和融合,最终生成精细化信号地图;

步骤401. 分别进行内部信号视图和外部环境视图的构建,具体方法如下:

步骤40101. 将步骤(3)中得到的恢复后的信号张量分布及其特征作为内部视图;

步骤40102. 将收集到的信息数据,包括人口、道路网和兴趣点的外部特征信号映射成张量作为外部视图;

步骤402. 利用两个嵌入网络分别对内外部视图提取高维特征并进行向量表示,具体方法如下:

步骤40201. 对于内部视图,根据信号的更新周期,按照时间维度将步骤3中得到的恢复后的张量信号分布划分为矩阵,形成粗粒度信号地图 $S_{::K}^{(c)}$ ;引入平均池化层,整合从历史信号图中提取的特征,得到细粒度信号图 $S_{::K}^{(f)}$ ;

步骤40202. 对于外部视图,提出了一种基于上下文感知的外部特征嵌入表示;人口、道路网络和兴趣点的分布特征被映射成张量,并用嵌入向量表示;具体来说,首先收集外部特征,并将外部视图按照与张量S相同的地理空间进行划分,即长宽分别为I和J个网格的地理空间;然后,为了整合这些不同类型的环境特征,嵌入每个小网格区域的外部特征集,得到外部特征的嵌入向量;

步骤40203. 将步骤40201得到的细粒度信号地图与步骤40202得到的外部环境特征向量映射起来,利用卷积层进行处理,增强对外部信息的重构;

步骤403. 利用两个全连接层对不同视图进行拼接和融合,生成精细、准确的信号地图,

具体方法如下：

步骤40301.将内部信号特征和外部环境特征分别输入到两个多层感知层中,然后将输出结果拼接融合并经过全连接层得到精细精确的信号地图；

步骤40302.使用Adam优化器通过提供训练对 $(S_{::K}^{(c)}, S_{::K}^{(f)})$ 来学习所提出的多视图融合模型,并且计算反向传播的像素方向均方误差损失；

步骤40303.引入掩模矩阵 $M \in \mathbb{R}^{(NI \times NJ)}$ 在细粒度信号图中表示有可观测数据的区域,定义了损失函数 $L = \frac{1}{\|M\|_1} \left\| \left( S_{::K}^{(f)} - \hat{S}_{::K}^{(f)} \right) \times M \right\|^2$ ,对多视图融合模型进行训练,由步骤40201输出的内部时变信号特征与步骤40202输出的外部环境特征相融合,在张量恢复的粗粒度信号地图 $S_{::K}^{(c)}$ 的基础上构建细粒度精准信号地图。

## 一种基于移动群体感知的蜂窝信号地图构建方法

### 技术领域

[0001] 本发明涉及信号地图构建领域,具体涉及一种基于移动群体感知的蜂窝信号地图构建方法。

### 背景技术

[0002] 随着物联网在智慧城市和自动驾驶服务领域的快速发展,对无线频谱资源的需求日益增加,无线覆盖环境更加复杂。复杂而拥挤的无线环境使得频谱资源显得尤为宝贵。为了迎合这一趋势,信号图是衡量无线环境的有效指标,它可以应用于各种现实世界的应用,包括网络运维、频谱监控和基于位置的服务。信号图是标记了特定区域的位置信息的指纹数据库。我们通过与信号图上的指纹进行比较来估计在线指纹测量的位置。此外,构建和分析信号图是评估不同地区移动宽带技术(如4G/LTE、5G)可用性和比较不同运营商的典型方法。

[0003] 由于无线信号容易受到各种因素的影响,如传输功率的自动调节、多径衰落以及非视距传播等,在城市尺度上很难及时有效地维护最新的信号图。许多关于信号图建设的研究都是在假定参考点是预先定义的前提下,依靠劳动密集型和耗时的现场调查,必须定期进行以保持最新的信号图,导致成本极高。为了减少密集的现场调查造成的费用,也有一些研究试图利用基于模型的方法和数据驱动的方法来构建信号图。

[0004] 现有的信号图构建方法存在两个局限性:基于模型的方法通常对信号图的分布(如理想传播空间)有很强的假设,这极大地限制了其对复杂的现实无线环境建模的能力。数据驱动方法需要大量的测量数据,这些数据通常由战争驱动和群体感知收集。但是这些数据没有覆盖网络的所有区域,不能密切反映用户的实际体验。群众感知方法不处理多层次的同时相互依存关系,在有限的局部区域内提取空间特征。此外,信号内部特征与外部环境特征的融合缺乏统一表示。

### 发明内容

[0005] 为解决上述问题,本发明公开了一种基于移动群体感知的蜂窝信号地图构建方法,该方法包括以下步骤:

[0006] 步骤1.将信号数据的时空特征表示为三维张量 $S \in \mathbb{R}^{I \times J \times K}$ ,其中,I、J表示矩形区域的长、宽网格数;K表示网格内信号采集时间跨度; $S_{ijk}$ 表示移动设备从基站接收到的信号值;

[0007] 步骤2.MCS信号数据的潜在因子特征提取,在BS覆盖交叉区域的部分标定信号值;

[0008] 步骤3.采用贝叶斯概率张量分解恢复未观测区域的信号值,即对缺失信号的预估;

[0009] 步骤4.先进行内部和外部视图构造;然后引入一种新的基于学习的多视图特征融合模块,利用两个网络分别对每个视图提取并嵌入高维特征,然后利用两个全连接层对不同视图进行拼接和融合,生成精细、准确的信号图。

[0010] 本发明进一步改进在于：

[0011] 与其他数据源(如城市交通数据、噪声数据)相比,信号数据更容易受到环境影响,表现出波动性特征。我们通常通过对给定面积取平均值来求得近似。但是,由于每个BS的特性(如覆盖范围、天线发射功率等)不同,在某些重叠的BS覆盖区域直接取平均信号是不合适的,会导致信号漂移和误差。更重要的是,由于各种未知因素,我们无法有效地利用同一网格内多个BS接收到的信号之间的关系。因此,在考虑基于张量的方法时,我们需要提取满足张量低秩特征的数据,从而构造适合信号数据补全的张量分解方法。也就是说,在构造张量和分解之前,我们需要在交叉区域的部分标定信号值。

[0012] 步骤(2)中,进行潜在因子特征提取方法如下：

[0013] 步骤201.利用欧几里得距离来描述空间相关性,然后把地图上的地理位置联系起来,利用公式(1)表示BSs之间的距离相关性：

$$[0014] \quad \Phi^{bs}(i, j) = \exp\left(-\frac{\|l_i^{bs} - l_j^{bs}\|_2^2}{\sigma^2}\right) \quad (1)$$

[0015] 其中, $\sigma^2$ 代表方差, $\Phi(i, j)$ 代表位置*i*、*j*的关系, $\|l_i - l_j\|_2$ 表示表示欧几里得距离,即 $\|l_i - l_j\|_2 = \sqrt{(l_i^x - l_j^x)^2 + (l_i^y - l_j^y)^2}$ ;

[0016] 步骤202.利用公式(2)表示两个高维信号向量之间的余弦相似度,进行另一种特征提取：

$$[0017] \quad \Phi^s(i, j) = \frac{\overrightarrow{sig_i} \cdot \overrightarrow{sig_j}}{\|\overrightarrow{sig_i}\| \|\overrightarrow{sig_j}\|} \quad (2)$$

[0018] 其中, $\overrightarrow{sig} \in \mathbb{R}^H$ ,*H*表示维度, $\overrightarrow{sig_i}$ 表示在网格*i*中从*H*个BSs中接受到的信号向量。

[0019] 本发明进一步改进在于:通过数据分析,我们知道信号数据近似服从正态分布;

[0020] 步骤(3)中,采用贝叶斯概率张量分解对缺失信号值的预估,具体方法如下：

[0021] 步骤301.假设每个观测项的噪声项近似服从独立高斯分布,即

$S_{ijk} \sim N(\hat{S}_{ijk}, \tau^{-1})$ ,进一步对所有因子矩阵的行向量进行先验分布,并假定它们是多元

高斯函数,即 $u_i \sim N(\mu_u, \Lambda_u^{-1})$ ;其中, $\mu_u \in \mathbb{R}^R$ , $\Lambda_u \in \mathbb{R}^{R \times R}$ ,并且满足公式(3),其概率分布函数为公式(4)：

$$[0022] \quad (\mu_u, \Lambda_u) \sim \text{Gaussian-Wishart}(\mu_0, \beta_0, W_0, v_0) \quad (3)$$

$$[0023] \quad p(\mu_u, \Lambda_u | \Theta) = N(\mu_u | \mu_0, (\beta_0 \Lambda_u)^{-1}) \times \text{Wishart}(\Lambda_u | W_0, v_0) \quad (4)$$

[0024] 步骤302.由于信号数据的精度是未知的,并且它不能完全通过所有观测值方差的倒数来获取,为了提高模型的鲁棒性,我们在精度参数 $\tau$ 上引入了柔性共轭Gamma先验,利用公式(5)表示其先验函数分布：

$$[0025] \quad p(\tau | a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \tau^{a_0-1} \exp(-b_0 \tau) \quad (5)$$

[0026] 其中,参数*U*,*V*,*W*, $\tau$ 和超参数 $\mu_u, \Lambda_u, \mu_v, \Lambda_v, \mu_w, \Lambda_w, a_0, b_0$ ,利用吉布斯抽样求解;

[0027] 步骤303.因子矩阵 $U, V, W$ 可通过 $u_i, v_i, w_i$ 进行求解,其计算方法如下:

[0028] 步骤30301.由于每个观测项的噪声项近似服从独立高斯分布,得到其似然函数,其表达式如公式(6)所示:

$$[0029] \quad L(S|u_i, V, W, \tau) = \exp\left\{-\frac{\tau}{2} \sum_j^J \sum_k^K b_i (S_{ijk} - (u_i)^T (v_i * w_k))^2\right\} \quad (6)$$

[0030] 公式中 $*$ 表示Hadamard积。

[0031] 步骤30302.由于其因子矩阵的行向量服从多元高斯分布,联合其步骤30401中的似然函数,得到其后验分布,其表达式如公式(7)所示:

$$[0032] \quad p(u_i | S, V, W, \tau, \mu_u, \Lambda_u^{-1}) \propto L(S|u_i, V, W, \tau) \times N(u_i | \mu_u, \Lambda_u^{-1}) \propto N(u_i | \mu_u, \Lambda_u^{-1}) \quad (7)$$

[0033] 步骤30303.根据步骤30402得到的后验分布,按公式(8)、公式(9)求得 $\widehat{\Lambda}_u, \widehat{u}_i$ :

$$[0034] \quad \widehat{\Lambda}_u = \tau \sum_j^J \sum_k^K b_i (v_j * w_k)(v_j * w_k)^T + \Lambda_u \quad (8)$$

$$[0035] \quad \widehat{u}_i = \widehat{\Lambda}_u^{-1} \left( \tau \sum_j^J \sum_k^K S_{ijk} (v_j * w_k) + \Lambda_u \mu_u \right) \quad (9)$$

[0036] 步骤304.类似地,通过上述方法求得 $u_i, v_i, w_i$ ,最后,在吉布斯采样算法达到平稳状态后,通过蒙特卡罗近似可以估计出所有缺失值,然后实现对缺失信号数据的估计。

[0037] 本发明进一步改进在于:步骤(4)中,先进行内部试图和外部试图地构建,具体方法如下:

[0038] 步骤40101.将步骤(3)中得到的恢复后地信号张量分布及其特征作为内部视图;

[0039] 步骤40102.将收集到的信息数据,包括人口、道路网和POI地外部特征信号映射成张量作为外部试图。

[0040] 5.根据权利要求4所述的一种基于移动群体感知的蜂窝信号地图构建方法,其特征在于,步骤(4)中,利用两个网络分别对每个视图提取并嵌入高维特征,具体方法如下:

[0041] 步骤40201.对于内部视图,根据信号的更新周期,按照时间维度将张量划分为矩阵,形成粗粒度信号图 $S_{::K}^{(c)}$ ;引入平均池化层,整合从历史信号图中提取的特征,得到细粒度信号图 $S_{::K}^{(f)}$ ;

[0042] 步骤40202.对于外部视图,提出了一种新的上下文感知神经网络,嵌入外部特征;人口、道路网络和POI的分布特征,并用嵌入向量表示它们;具体来说,我们首先收集外部特征,并将其按照与张量相同的地理空间进行划分。然后,为了整合这些不同类型的背景特征,增强特征的表达能力,我们嵌入每个小网格区域的这些外部特征集,得到背景外部特征的嵌入向量;

[0043] 步骤40203.将步骤40201得到的细粒度信号图与步骤40202得到的外部特征向量映射起来,利用卷积层进行处理,增强外部信息的重构;

[0044] 本发明进一步改进在于:步骤(4)中,利用两个全连接层对不同视图进行拼接和融合,生成精细、准确的信号图,具体方法如下:

[0045] 步骤40301.将内部特征和外部特征分别输入到两个多感知层中,然后将它们与连接层进行融合,得到精细精确的信号映射;

[0046] 步骤40302.使用Adam优化器通过提供训练对 $(S_{::K}^{(c)}, S_{::K}^{(f)})$ 来学习所提出的模型,并且计算反向传播的像素方向均方误差损失;

[0047] 步骤40303.引入掩模矩阵 $M \in \mathbb{R}^+(NI \times NJ)$ 在细粒度信号图中表示有可用数据的区域,定义了损失函数 $L = \frac{1}{\|M\|_1} \left\| \left( S_{::K}^{(f)} - \hat{S}_{::K}^{(f)} \right) \times M \right\|^2$ ,对融合网络进行训练,使信号特征

与外部特征相融合,在张量恢复的粗粒度信号图的基础上构建细粒度信号图

[0048] 本发明的有益效果:

[0049] 解决了两个挑战,包括:(1)缺失和不可靠的MCS数据问题;(2)信号传播的时空不确定性。蜂窝信号图重建方法同时捕捉蜂窝基站间和蜂窝基站内信号的时空特征,利用贝叶斯张量分解进行缺失信号恢复,构建大面积的细粒度信号图。此外,该方法开发了一种上下文感知的多视图融合网络,以充分利用外部信息,提高信号图构建精度。

## 附图说明

[0050] 图1:基于移动群体感知的蜂窝信号地图构建方法流程图;

[0051] 图2:蜂窝信号地图构建系统框架图;

[0052] 图3:贝叶斯张量分解生成过程的总体图结构;

[0053] 图4:多视图融合网络结构;

[0054] 图5:基于移动群体感知的蜂窝信号地图构建方法流程图。

## 具体实施方式

[0055] 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不适用于限制本发明的范围。需要说明的是,下面描述中使用的词语“前”、“后”、“左”、“右”、“上”和“下”指的是附图中的方向,词语“内”和“外”分别指的是朝向或远离特定部件几何中心的方向。

[0056] 如图1-5所示,本实施例提供了一种基于移动群体感知的蜂窝信号地图构建方法,包括以下步骤:

[0057] 步骤1.将信号数据的时空特征表示为三维张量 $S \in \mathbb{R}^{I \times J \times K}$ ,其中,I、J表示矩形区域的长、宽网格数;K表示网格内信号采集时间跨度; $S_{ijk}$ 表示移动设备从基站接收到的信号值;

[0058] 步骤2.MCS信号数据的潜在因子特征提取,在BS覆盖交叉区域的部分标定信号值;与其他数据源(如城市交通数据、噪声数据)相比,信号数据更容易受到环境影响,表现出波动性特征。我们通常通过对给定面积取平均值来求得近似。但是,由于每个BS的特性(如覆盖范围、天线发射功率等)不同,在某些重叠的BS覆盖区域直接取平均信号是不合适的,会导致信号漂移和误差。更重要的是,由于各种未知因素,我们无法有效地利用同一网格内多



个BS接收到的信号之间的关系。因此,在考虑基于张量的方法时,我们需要提取满足张量低秩特征的数据,从而构造适合信号数据补全的张量分解方法。也就是说,在构造张量和分解之前,我们需要在交叉区域的部分标定信号值。受BS知识提取的启发,我们设计了一种合适的潜在因素特征提取方法:一般来说,两个位置越近,两个样本的相关性越高。我们利用欧几里得距离来描述空间相关性,然后把地图上的地理位置联系起来。BS之间的距离相关性定义为:

$$[0059] \quad \Phi^{bs}(i, j) = \exp\left(-\frac{\|l_i^{bs} - l_j^{bs}\|^2}{\sigma^2}\right)$$

[0060] 此外,我们还利用两个高维信号向量之间的余弦相似度进行另一种特征提取,定义为:

$$[0061] \quad \Phi^s(i, j) = \frac{\vec{sig}_i \cdot \vec{sig}_j}{\|\vec{sig}_i\| \|\vec{sig}_j\|}$$

[0062] 步骤3.采用贝叶斯概率张量分解恢复未观测区域的信号值,即对缺失信号的预估;

[0063] 通过数据分析,我们知道信号数据近似服从正态分布。因此,我们假设每个观测项的噪声项近似服从独立高斯分布,并定义为  $S_{ijk} \sim N(\hat{S}_{ijk}, \tau^{-1})$ 。我们进一步对所有因子

矩阵的行向量进行先验分布,并假定它们服从多元高斯分布  $u_i \sim N(\mu_u, \Lambda_u^{-1})$ ,超参数  $\mu_u \in \mathbb{R}^R, \Lambda_u \in \mathbb{R}^{R \times R}$ 。这将增强模型的鲁棒性,并加快收敛速度时,使用抽样算法进行模型推理。其中超参数  $\mu_u, \Lambda_u$  定义为

$$[0064] \quad (\mu_u, \Lambda_u) \sim \text{Gaussian-Wishart}(\mu_0, \beta_0, W_0, v_0)$$

$$[0065] \quad p(\mu_u, \Lambda_u | \Theta) = N(\mu_u | \mu_0, (\beta_0 \Lambda_u)^{-1}) \times \text{Wishart}(\Lambda_u | W_0, v_0)$$

[0066] 由于信号数据的精度未知,并且它不能完全通过所有观测值方差的倒数来捕捉。为了提高模型的鲁棒性,我们在模型上引入了柔性共轭Gamma先验:

$$[0067] \quad p(\tau | a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \tau^{a_0-1} \exp(-b_0 \tau)$$

[0068] 图2是描述上述贝叶斯张量分解生成过程的总体图结构。其中的参数  $U, V, W, \tau$  和超参数  $\mu_u, \Lambda_u, \mu_v, \Lambda_v, \mu_w, \Lambda_w, a_0, b_0$  都可以通过吉布斯抽样求解。由于每个观测项服从独立高斯分布,可以写出其可能性:

$$\begin{aligned}
& L(S|u_i, V, W, \tau) \\
& \propto \prod_j^J \prod_k^K \exp \left\{ -\frac{\tau b_i}{2} (S_{ijk} - \hat{S}_{ijk})^2 \right\} \\
& = \exp \left\{ -\frac{\tau}{2} \sum_j^J \sum_k^K b_i (S_{ijk} - (u_i)^T (v_i * w_k))^2 \right\}
\end{aligned}$$

[0070] 得到后验分布：

$$\begin{aligned}
p(u_i | S, V, W, \tau, \mu_u, \Lambda_u^{-1}) & \propto L(S|u_i, V, W, \tau) \times N(u_i | \mu_u, \Lambda_u^{-1}) \\
& \propto N(u_i | \mu_u, \Lambda_u^{-1})
\end{aligned}$$

[0072] 然后求得：

$$\hat{\Lambda}_u = \tau \sum_j^J \sum_k^K b_i (v_j * w_k)(v_j * w_k)^T + \Lambda_u$$

$$\hat{u}_i = \hat{\Lambda}_u^{-1} \left( \tau \sum_j^J \sum_k^K S_{ijk} (v_j * w_k) + \Lambda_u \mu_u \right)$$

[0075] 类似的,  $v_j, w_k$  也可以按照上述方法求得。最后, 在吉布斯采样算法达到平稳状态后, 通过蒙特卡罗近似可以估计出所有缺失值, 然后实现对缺失信号数据的估计。

[0076] 步骤4. 先进行内部和外部视图构造; 然后引入一种新的基于学习的多视图特征融合模块, 利用两个网络分别对每个视图提取并嵌入高维特征, 然后利用两个全连接层对不同视图进行拼接和融合, 生成精细、准确的信号图。

[0077] 内部和外部视图构造

[0078] 张量恢复缺失信号后, 得到指定区域的粗粒度信号覆盖图。尽管张量具有良好的空间结构来挖掘时空信息, 但由于动态环境的影响, 同一网格中恢复的信号可能在不同时间存在显著差异。信号图被认为在一定的更新周期内是不变的。但是, 仅对不同时期的历史信号图进行平均, 得到最终的信号图是粗粒度的, 是不准确的。由于相邻网格中的外部特征(人口、POI、路网)相似, 因此网格中接收到的信号也具有一定的相似性, 有助于构建细粒度的信号图。因此, 为了提高信号图的精度, 我们进一步实现了多视图内部和外部特征的融合, 我们把恢复后的信号分布张量及其特征视图作为A视图。把收集的信息, 包括人口、道路网和POI的外部特征信号映射成张量作为B视图。

[0079] 多视点神经融合

[0080] 融合不同视图构建最终信号图的网络如图3所示。我们首先利用两个网络分别对每个视图提取并嵌入高维特征, 然后利用两个全连接层对不同视图进行拼接和融合, 生成

精细、准确的信号图。

[0081] 图4中,对于A视图,一个完整的信号空时特征用张量形式表示。根据信号图的更新周期,将张量划分为矩阵,即粗粒度信号图 $S_{::K}^{(c)}$ 。引入平均池化层,整合从历史信号图中提取的特征,得到细粒度信号图 $S_{::K}^{(f)}$ 。对于B视图,我们提出了一种新的上下文感知神经网络,嵌入外部特征(人口、道路网络和POI)的分布特征,并用嵌入向量表示它们。具体来说,我们首先收集外部特征,并将其按照与张量相同的地理空间进行划分。然后,为了整合这些不同类型的背景特征,增强特征的表达能力,我们嵌入每个小网格区域的这些外部特征集,得到背景外部特征的嵌入向量。最后,我们将其与细粒度信号映射连接起来,然后用卷积层进行处理,增强了外部信息的重构过程。

[0082] 为了融合两部分特征嵌入,我们首先将内部特征和外部特征分别输入到两个多感知层中,然后将它们与连接层进行融合,得到精细精确的信号映射。我们使用Adam优化器通过提供训练对 $(S_{::K}^{(c)}, S_{::K}^{(f)})$ 来学习所提出的模型,并且计算反向传播的像素方向均方误差损失。我们引入一个掩模矩阵 $M \in R^{N_I \times N_J}$ 在细粒度信号图中表示有可用数据的区域。我们定义了损失函数:

$$[0083] \quad L = \frac{1}{\|M\|_1} \left\| \left( S_{::K}^{(f)} - \hat{S}_{::K}^{(f)} \right) \times M \right\|^2$$

[0084] 融合网络解决了空间粒度划分问题,使信号特征与外部特征融合成为可能,在张量恢复的粗粒度信号图的基础上构建细粒度信号图。

[0085] 本实施例的技术方案可以应用到以下的实施例中:

[0086] 实施例1:

[0087] 用于构建高质量、细粒度的信号地图。

[0088] 根据MCS时空数据信息,结合外部环境信息(如人口、道路、POI和BS位置),实现信号数据的恢复,从而构建细粒度的信号地图。

[0089] 实施例2:

[0090] 蜂窝信号网络的实时更新与维护。

[0091] 由于环境影响与人口流动性,导致信号数据有较强的时空依赖性,从而需要对信号图的实时更新。利用蜂窝信号图构建系统,根据当前的数据信息进行构建信号图,并向用户提供更新。

[0092] 实施例3:

[0093] 用于空间定位,提供地理信息服务。

[0094] 利用蜂窝信号图构建系统,获取信号数据,分析其信号强度与基站的距离相关性,从而实现信号源空间定位,向用户提供相应的地理信息服务。

[0095] 本实施例为了评估CSMC的性能,我们在上海采集的超过200GB MCS信号记录的大规模数据集上进行了广泛的实验和消融研究。实验结果如表1、表2所示,与现有方法相比,该方法的信号估计误差性能提高了29%。

[0096] 表1采样率30%、60%、90%下不同方法MAPE和RMSE结果(城市)

Models	30%Sampling Rate		60%Sampling Rate		90%Sampling Rate	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
GP-WTLS	0.1030	10.3288	0.0996	9.9910	0.0956	9.5872
RFs	0.1031	10.3363	0.0985	9.8759	0.0950	9.5222
SSOA	0.0939	9.4145	0.0915	9.1773	0.0881	8.8369
BCCS	0.0881	8.8313	0.0841	8.4298	0.0806	8.0815
AER	0.0801	8.0346	0.0776	7.7876	0.0754	7.5604
CSMC(30)	0.0714	7.1592	0.0674	6.7574	0.0645	6.4642
CSMC(60)	0.0662	6.6404	0.0627	6.2913	0.0610	6.1167
CSMC(90)	<b>0.0622</b>	<b>6.2381</b>	<b>0.0603</b>	<b>6.0506</b>	<b>0.0593</b>	<b>5.9479</b>

[0098] 表2采样率30%、60%、90%下不同方法MAPE和RMSE结果(郊区)

Models	30%Sampling Rate		60%Sampling Rate		90%Sampling Rate	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
GP-WTLS	0.1026	10.2826	0.0986	9.8849	0.0949	9.5203
RFs	0.1009	10.1080	0.0973	9.7618	0.0944	9.4667
SSOA	0.0910	9.1261	0.0875	8.7779	0.0847	8.4934
BCCS	0.0863	8.6622	0.0835	8.3720	0.0801	8.0289

	AER	0.0771	7.7305	0.0759	7.6129	0.0737	7.3990
	CSMC(30)	0.0657	6.5891	0.0621	6.2323	0.0592	5.9408
[0100]	CSMC(60)	0.0622	6.2395	0.0587	5.8880	0.0553	5.5438
	CSMC(90)	<b>0.0581</b>	<b>5.8293</b>	<b>0.0547</b>	<b>5.4876</b>	<b>0.0524</b>	<b>5.2536</b>

[0101] 本发明方案所公开的技术手段不仅限于上述实施方式所公开的技术手段,还包括由以上技术特征任意组合所组成的技术方案。

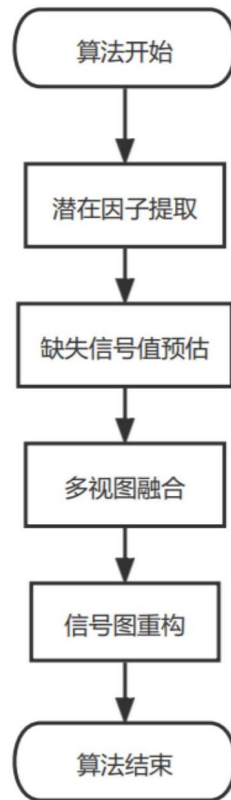


图1

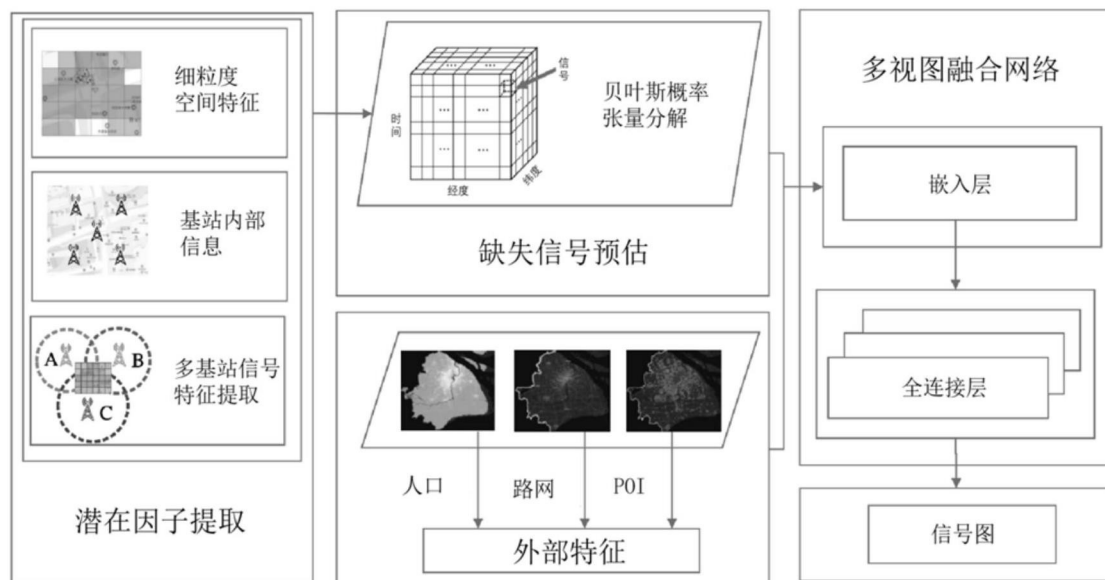


图2

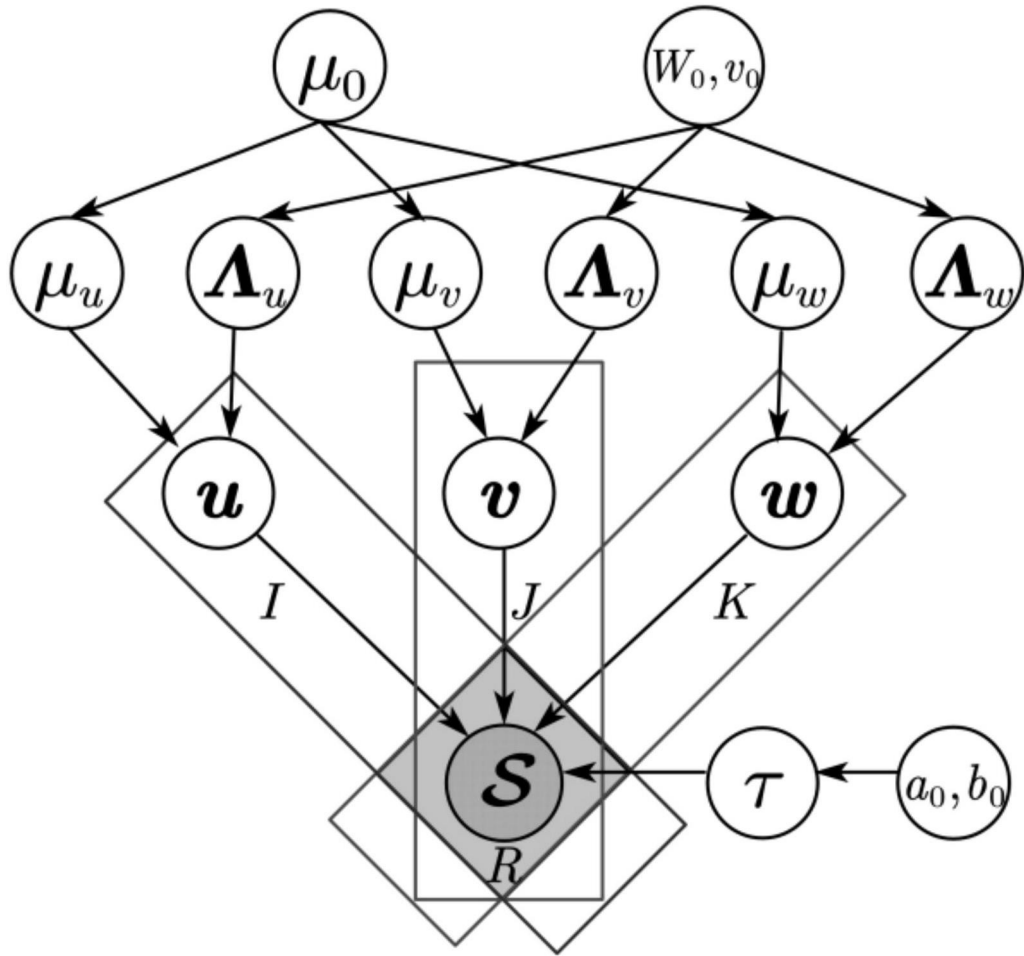


图3

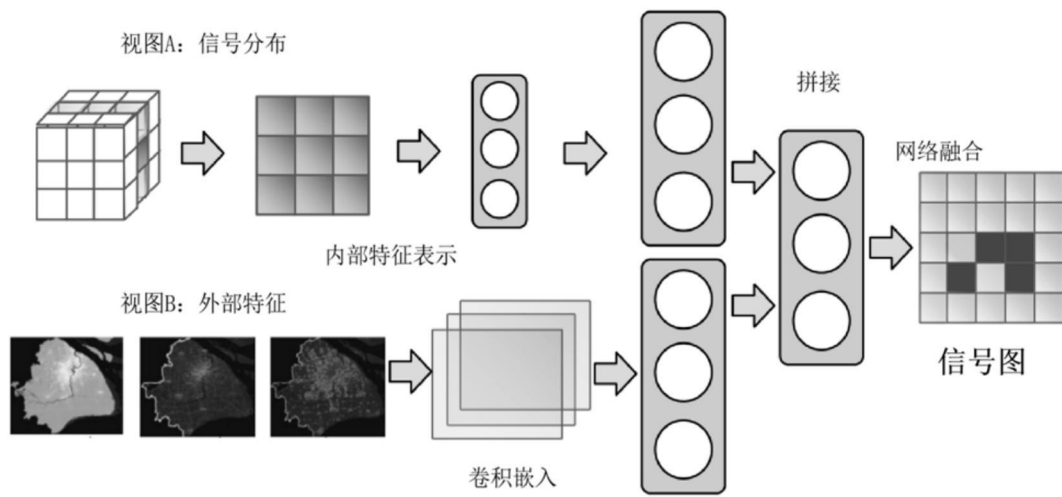


图4

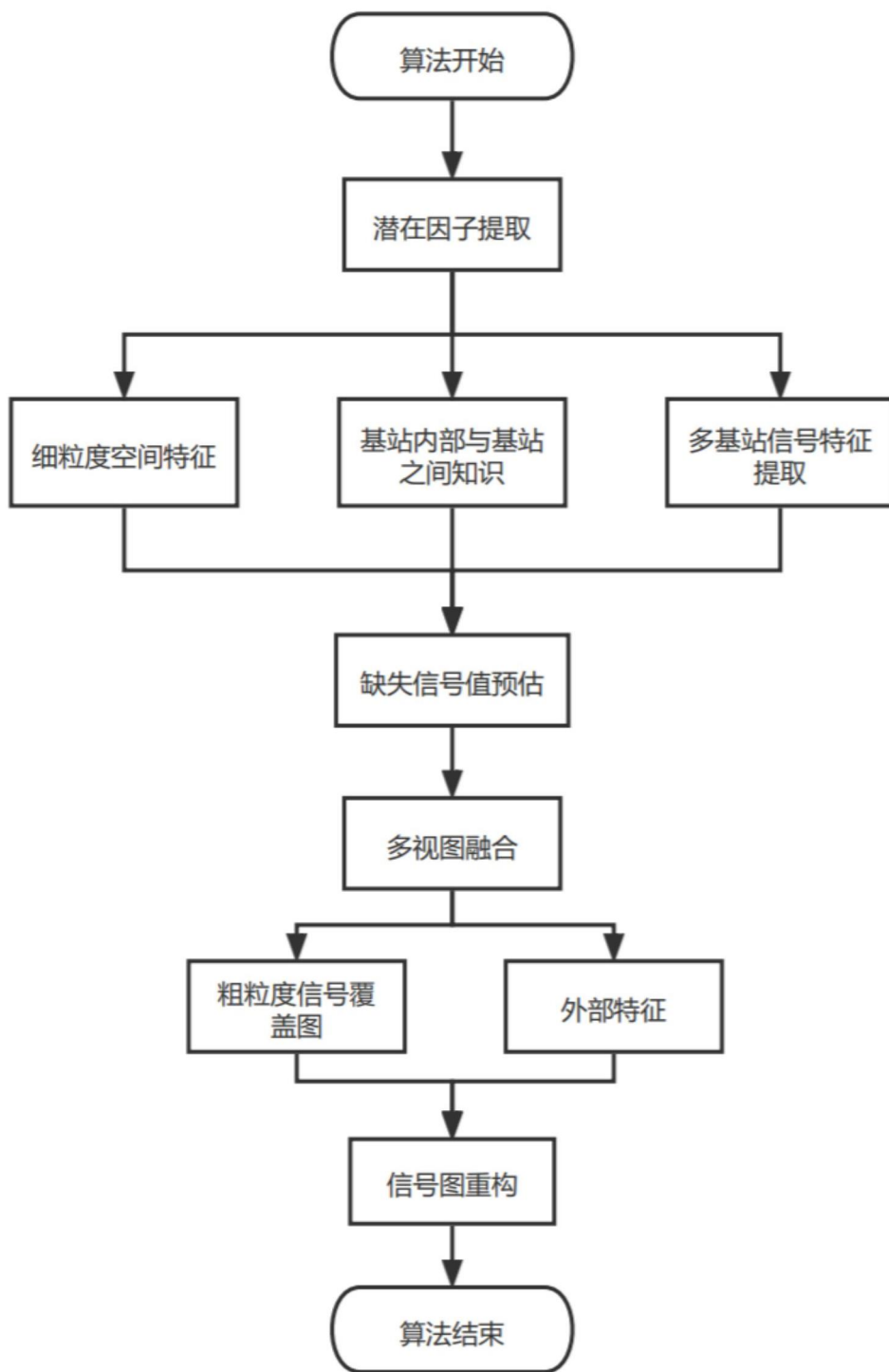


图5