



(21) 申请号 202410022968.8

(22) 申请日 2024.01.08

(65) 同一申请的已公布的文献号

申请公布号 CN 117528552 A

(43) 申请公布日 2024.02.06

(73) 专利权人 深圳市英迈通信技术有限公司

地址 518000 广东省深圳市福田区沙头街

道天安社区泰然四路25号天安创新科

技广场一期B座1108

(72) 发明人 刘春明 沈娟

(74) 专利代理机构 天津创信方达专利代理事务

所(普通合伙) 12247

专利代理师 李京京

(51) Int. Cl.

H04W 16/22 (2009.01)

H04L 41/16 (2022.01)

H04W 24/02 (2009.01)

H04W 24/06 (2009.01)

H04W 36/08 (2009.01)

G06N 3/092 (2023.01)

(56) 对比文件

CN 114707664 A, 2022.07.05

CN 116846592 A, 2023.10.03

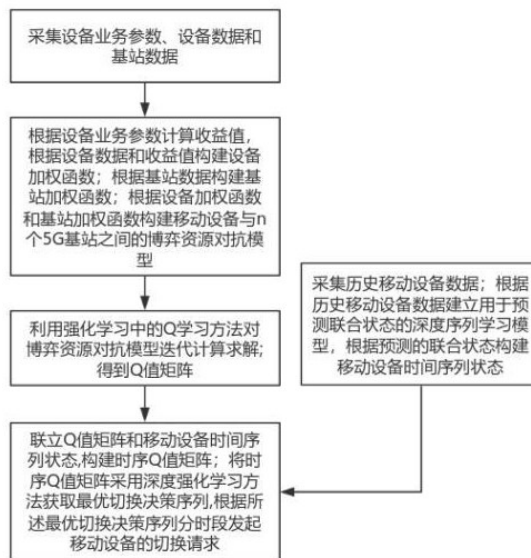
审查员 陈莹

(54) 发明名称

一种基于5G的手机数据传输方法和系统

(57) 摘要

本发明属于数据传输技术领域,本发明公开了一种基于5G的手机数据传输方法和系统;包括接收规划数据;构建移动设备与n个5G基站之间的博弈资源对抗模型;利用Q学习方法对博弈资源对抗模型迭代计算求解;得到Q值矩阵;采集历史移动设备数据;根据历史移动设备数据建立用于预测联合状态的深度序列学习模型,根据预测的联合状态构建移动设备时间序列状态;联立Q值矩阵和移动设备时间序列状态,构建时序Q值矩阵;将时序Q值矩阵采用深度强化学习方法获取最优切换决策序列,根据所述最优切换决策序列分时段发起移动设备的切换请求,融合个体建模、业务感知和全局优化为一体实现了稳定的5G网络运维。



1.一种基于5G的手机数据传输方法,其特征在于,包括:

S1、采集设备业务参数、设备数据和基站数据;

S2、根据设备业务参数计算收益值,根据设备数据和收益值构建设备加权函数;根据基站数据构建基站加权函数;根据设备加权函数和基站加权函数构建移动设备与n个5G基站之间的博弈资源对抗模型;

S3、利用强化学习中的Q学习方法对博弈资源对抗模型迭代计算求解;得到Q值矩阵;

S4、采集历史移动设备数据;根据历史移动设备数据建立用于预测联合状态的深度序列学习模型,根据预测的联合状态构建移动设备时间序列状态;

S5、联立Q值矩阵和移动设备时间序列状态,构建时序Q值矩阵;将时序Q值矩阵采用深度强化学习方法获取最优切换决策序列,根据所述最优切换决策序列分时段发起移动设备的切换请求;

所述设备业务参数包括数据业务平均吞吐量、单位数据吞吐量收入、语音业务平均持续时间、单位时间语音收入、视频业务平均码率和单位视频码率收入;

所述设备数据包括平均时延和移动性开销;所述基站数据包括基站负载提升值和平均传输功率;

所述收益值的计算方式包括:

收益值 $R = T_data \times R_T + T_voice \times R_V + B_video \times R_B$;

其中, T_data 为数据业务平均吞吐量; R_T 为单位数据吞吐量收入; T_voice 为语音业务平均持续时间; R_V 为单位时间语音收入; B_video 为视频业务平均码率; R_B 为单位视频码率收入;

所述平均时延D是一个预定义的业务质量统计参数,由网络侧的质量监控系统实时统计汇总,直接调用;

测算移动设备切换前后的位置变化,即为移动性开销C;

获取移动设备在切换前后的位置坐标;通过移动设备自身的GPS或定位模块获取切换前后的位置坐标,并上报网络侧;网络侧收到位置坐标后,按照坐标系的计算方法获得两位置之间的距离 d_s ;查询移动设备的类型参数,获取其平均移动速度V;移动时间开销=距离 d_s /平均速度V;移动时间开销即为移动性开销C;

构建设备加权函数 $U = \omega_1 \times R + \omega_2 \times D + \omega_3 \times C$;

其中, ω_1 、 ω_2 和 ω_3 为设备权重参数;

所述博弈资源对抗模型的构建方式包括:

获取参与博弈决策的每个移动设备,确定移动设备的唯一标识ID;获取每个移动设备当前所运行的业务类型;将每个移动设备定义设备动作空间,设备动作空间包括移动设备可申请切换的L个相邻基站;

获取移动设备周边参与博弈决策的每个基站,确定基站的唯一标识ID;获取每个基站的网络频谱资源信息和基站参数;将每个基站定义基站动作空间,基站动作空间包括接受切换请求和拒绝切换请求;

移动设备接入后,计算基站负载提升值 ΔL ;

基站负载提升值 ΔL 的计算过程包括:

获取基站的当前业务数N和移动设备的业务类型;根据业务类型查询对应类业务的平

均占用带宽BW_语音、BW_数据和BW_视频；

计算基站当前总占用带宽BW_总= $N \times BW_语音 + N \times BW_数据 + N \times BW_视频$ ；

移动设备接入后，总业务数变为N+1，则总占用带宽为BW_新总= $(N+1) \times BW_语音 + (N+1) \times BW_数据 + (N+1) \times BW_视频$ ；带宽提升量BW_提升=BW_新总-BW_总；

带宽提升量与基站总带宽计算负载提升 $\Delta L = BW_提升 / 基站总带宽$ ；

查表获取移动设备的平均传输功率P；

构建基站加权函数 $V = \alpha_1 \times \Delta L + \alpha_2 \times P$ ；

其中， α_1 和 α_2 为基站权重参数；

结合设备加权函数和基站加权函数获取博弈资源对抗模型；

博弈资源对抗模型 $J = \max \sum U - \sum V$ ；

所述Q值矩阵的获取过程包括：

S301、定义博弈资源对抗模型中的状态空间S和动作空间A；其中，状态空间S表示场景数量；动作空间A表示每种状态下可供移动设备或基站选择执行的不同行为数量；初始化一个存储状态—动作值函数的Q(s,a)矩阵，其中Q(s,a)矩阵的大小为 $|S| \times |A|$ ，Q(s,a)矩阵的行表示状态s，列表示动作a；矩阵中的每个元素Q(s,a)表示在状态s下选择执行动作a对应的价值，初始化全部矩阵元素为0；

S302、设定迭代终止条件，迭代终止条件为最大迭代数M或Q矩阵元素值变化小于设置阈值p；设定Q学习算法的参数，参数包括学习率 α ，折扣因子 γ ，在 ϵ 贪婪策略中探索概率 ϵ ；

S303、基于 ϵ 贪婪策略以固定概率 ϵ 选择当前Q值最大的动作a，同时以 $1-\epsilon$ 概率随机选择一个动作a；执行所选择的动作a后，环境状态从s转移到s'，并产生一个即时收益r；收集执行动作a后环境的反馈，反馈包括新的状态s'和即时收益r；

S304、更新Q(s,a)；

更新公式为 $Q(s,a) \leftarrow Q(s,a) + \alpha \times (r + \gamma \times \max_{a'} Q(s') - Q(s,a))$ ；

式中，s为环境状态；a为在状态s下可选择执行的动作；r为即时收益；s'为执行完动作后的新的环境状态； α 为学习率； γ 为折扣因子； $\max(Qs')$ 为下一状态s'下所有可能的动作对应的Q值最大值； $\leftarrow$ 为更新符号；

S305、判断是否满足终止条件，若不满足则继续更新Q(s,a)；若满足则得到状态空间下的Q值矩阵；

所述历史移动设备数据包括历史定位轨迹数据和历史业务数据；

历史定位轨迹数据包括设备经纬度坐标及其对应的设备定位时间戳；

历史业务数据包括业务类型、业务流量和网络连接数及其对应的时间戳；

联合状态为未来f天的定位轨迹数据和业务数据；

深度序列学习模型为联立定位预测模型和参数预测模型获取；

定位预测模型获取方式包括：

将历史定位轨迹数据作为测试数据集，并进行曲线拟合；

曲线拟合过程包括：

用线性多项式拟合曲线、二次多项式拟合曲线和三次多项式分别进行拟合曲线；线性多项式拟合曲线为 $y = ax + b$ ；二次多项式拟合曲线 $y = ax^2 + bx + c$ ；三次多项式拟合曲线 $y = ax^3 + bx^2 + cx + d$ ；式中，x为设备定位时间戳；y为设备经纬度坐标；a、b、c和d为待求的参数，均通

过曲线拟合得出；

计算线性多项式、二次多项式和三次多项式拟合曲线的预测点位误差；预测点位误差

$$E = \sum \frac{|\text{实际值} - \text{预测值}|}{h};$$

其中，实际值为测试数据集中的真实值；预测值为拟合曲线计算的预测值；h为测试数据集大小；

选取多项式拟合曲线、二次多项式拟合曲线和三次多项式拟合曲线中预测点位误差最小的拟合曲线，即最佳拟合的曲线；最佳拟合的曲线即为定位预测模型；

参数预测模型为ARIMA时间序列预测模型，其获取方式包括：

采集最近r天内的历史业务数据作为历史样本，历史样本中包括4个字段；4个字段分别为时间戳、业务类型、业务流量和网络连接数；将历史样本按照时间顺序构建输入样本矩阵，每一行代表一个采样点位；

将输入样本矩阵的数据进行预处理；将24小时的全天数据依照分时段划分为g个时段区间，统计计算每个时段区间内各类业务数据所占比重，得到不同业务类别在各个时间段内的分布统计量；用于相应调整预测结果；

将网络连接数的时间序列数据建立ARIMA自回归滑动平均模型；进行平稳性检验，对不满足稳定条件的序列进行差分处理，直至达到稳定条件；

确定时间序列的AR和MA阶数，配置ARIMA(p, d, q)模型的参数；

利用配置好的参数将ARIMA自回归滑动平均模型训练和优化，输入训练数据为预处理的输入样本矩阵，训练目标为预测未来f天内每小时的网络连接数；获得未来f天内网络连接数的时间序列；训练完成的ARIMA自回归滑动平均模型即为参数预测模型；

所述联立定位预测模型和参数预测模型获取深度序列学习模型的方式包括：

将过去r天的历史定位轨迹数据和历史业务数据作为历史训练集；将预测得到的未来f天内设备经纬度坐标及其对应的定位时间戳、业务类型、业务流量和网络连接数及其对应的时间戳；利用时间戳进行时间对齐，并作为未来数据集；

构建卷积神经网络LSTM模型；输入层为历史训练集和未来数据集拼接而成的三维矩阵；

卷积神经网络LSTM模型包括卷积层、池化层、LSTM层和全连接层；编译卷积神经网络LSTM模型，配置优化器和损失函数；

优化器选用Adam算法；Adam算法基于优化随机梯度下降，自适应地对学习率进行调整；

$$\text{损失函数选用均方误差 } MSE = \frac{1}{z} \sum_{i=1}^z (Y_{pred,i} - Y_{true,i})^2;$$

其中，z为样本数量； $Y_{pred,i}$ 为模型预测得到的第i个样本输出； $Y_{true,i}$ 为第i个样本真实的标签输出；

MSE函数计算预测值和真实值之间的均方差，惩罚得到的损失数值越小模型效果越好；以历史训练集为卷积神经网络LSTM模型输入，训练目标是预测出未来数据集；

将历史训练集按照批大小分批输入卷积神经网络LSTM模型，卷积神经网络LSTM模型输出一批预测结果；计算这一批预测结果和未来数据集批之间的均方误差MSE函数值；利用反

向传播算法,将MSE损失误差沿着模型计算图反方向回传,更新每层的卷积神经网络LSTM模型的参数;

当训练达到预定的迭代轮数后,获得最终训练好的深度序列学习模型;使用训练好的深度序列学习模型进行预测;输入给定的过去r天历史数据,深度序列学习模型预测输出未来f天的联合状态;

所述移动设备时间序列状态的构建方式包括:

根据预测的联合状态,提取未来f天内的时间戳、经纬度坐标、业务类型、业务流量和网络连接数;将提取的数据重新组织成v个时间序列,时间序列状态包括时间戳序列、经纬度坐标时间序列、业务类型时间序列、业务流量时间序列和网络连接数时间序列;即构建出移动设备在未来f天内的移动设备时间序列状态;

所述时序Q值矩阵的构建方式包括:

从移动设备时间序列状态中,提取所有时刻的状态参数,构成状态空间S,状态参数包括移动设备当前所运行的业务类型、移动设备的位置坐标和运动速度;业务类型包括语音业务、视频业务和数据业务;

从Q值矩阵中,提取移动设备所有可选择的基站切换动作,构成动作空间A;初始化存储状态—动作值的时序Q(St,At)矩阵;时序Q(St,At)矩阵的行表示不同时刻t的设备状态St,时序Q(St,At)矩阵表示L个相邻基站作为基站切换动作At;矩阵元素Q(St,At)表示在状态St下选择动作At的价值,初始化为0;

所述时序Q值矩阵采用深度强化学习方法获取最优切换决策序列的过程包括:

构建深度强化学习模型为端到端的深度神经网络,输入为设备的状态St,输出为当前状态下各个基站切换动作的Q值;

深度强化学习模型内部结构包括卷积层、LSTM循环层和全连接层,卷积层负责提取输入状态的时空特征表示,LSTM层负责挖掘状态序列的时序关联特征,全连接层完成状态动作值函数Q(St,At)的估计;

令马尔可夫决策过程(MDP)的五元组中即时奖励为rt;通过采取时序上的折扣累积方法定义总奖励TY,公式为: $TY = \sum_{t=0}^{tr} \gamma^t r_t$;

其中, γ 是数值为0-1的折扣因子;tr为决策序列终止时刻;t表示的是当前的时间点,从该时刻t开始计算一个序列中的累积折扣收益;模型的目标是配置参数使得总奖励TY最大化;

深度强化学习模型训练目标是逼近状态—动作值函数Q(St,At),在每一训练步,输入一个采样状态St,模型预测输出当前状态下各个基站切换动作a对应的Q值;执行动作后环境给出反馈的实际即时奖励;计算Q(St,At)预测值和实际奖励的误差,通过损失函数反向传播更新模型参数;

在训练完成深度强化学习模型后,输入需要决策的时序状态,模型输出该状态下,不同基站切换动作对应的Q值,根据Q值最大选择基站,推演形成整个序列,计算序列总奖励,输出总奖励最大的最终序列,作为最优切换决策;

所述根据所述最优切换决策序列分时段发起移动设备的切换请求的过程包括:

S501、最优切换决策序列按照时间顺序给出了从时刻t1到tN,移动设备在每个时间点

需要切换到的目标基站;网络侧得到最优切换序列后,转换为信令交互过程;

S502、在时刻 t_1 ,网络侧向移动设备发送切换命令,通知设备切换到对应目标基站;移动设备收到切换命令后,向网络发送切换准备就绪信令;网络侧向源基站发送资源释放信令,向目标基站发送资源准备信令;

S503、源基站响应,释放移动设备占用资源;目标基站响应,预留资源;网络侧向移动设备发送确认信令,授权其执行切换资源重配置流程;移动设备与目标基站进行随机接入,完成切换;

S504、按照时间顺序,每到最优切换序列给出的时刻,重复执行S501-S503,完成切换,到时刻 t_N 全部切换完成。

2. 一种基于5G的手机数据传输系统,其基于权利要求1所述的一种基于5G的手机数据传输方法实现,其特征在于,包括:收集接收模块,用于采集设备业务参数、设备数据和基站数据;

对抗模型构造模块,用于根据设备业务参数计算收益值,根据设备数据和收益值构建设备加权函数;根据基站数据构建基站加权函数;根据设备加权函数和基站加权函数构建移动设备与 n 个5G基站之间的博弈资源对抗模型;

求解模块,用于利用强化学习中的Q学习方法对博弈资源对抗模型迭代计算求解;得到Q值矩阵;

状态构建模块,用于采集历史移动设备数据;根据历史移动设备数据建立用于预测联合状态的深度序列学习模型,根据预测的联合状态构建移动设备时间序列状态;

决策实施模块,用于联立Q值矩阵和移动设备时间序列状态,构建时序Q值矩阵;将时序Q值矩阵采用深度强化学习方法获取最优切换决策序列,根据所述最优切换决策序列分时段发起移动设备的切换请求。

一种基于5G的手机数据传输方法和系统

技术领域

[0001] 本发明涉及数据传输技术领域,更具体地说,本发明涉及一种基于5G的手机数据传输方法和系统。

背景技术

[0002] 申请公开号为CN116527572A的专利公开了一种基于5G的多路径数据传输系统,包括:主控模块:作为整个系统的核心,控制系统的运行;5G模组:搭载至少两个5G模块,与主控模块进行通信,通过多路径传输控制协议,聚合两条5G链路,实现5G多路径数据的传输;数据交互模块:与外部设备构成局域网,并与主控模块进行通信,实现数据的采集;人机交互模块:与主控模块进行通信,实现系统的控制和交流。本发明实现了MPTCP和5G技术的融合,成功将不同运营商的5G通信链路聚合,解决了现有5G网络目前覆盖性能较差、上行速率受限以及基站分布不均匀,导致不能满足日益增长的高速通信和基于物联网的应用需求问题。

[0003] 申请公开号为CN115551093A的专利公开了一种5G基站的低时延数据传输方法,包括步骤:S1、将5G基站与用户终端建立连;S2、在5G基站侧配置与用户终端对应的多个实体;S3、配置允许的最大传输时延参数T;S4、根据最大传输时延参数T进行数据传输。本发明通过设置最大时延参数来控制数据传输的时间,如果发现超时,则立即删除该超时数据,并统计超时传输的次数,重新配置资源调度的优先等级,这样,可以有效减少空口资源的浪费,提高数据传输的时效性,进一步降低时延。

[0004] 目前,随着5G网络的快速建设,大量新业务和应用涌现,这给网络的稳定运维带来了前所未有的挑战,现有的网络切换机制存在明显的缺陷,主要体现在:在移动性预测上,仍停留在传统的静态分析,无法实现设备个体精确建模,预测结果存在大量偏差,在业务感知上,现网络对新业务理解不够,无法根据业务场景实时调整手段,另外,各类网络调度措施之间缺乏协同,无法做到统筹兼顾,难以实现用户体验和网络收益的双赢;例如,在高铁场景中,因线路骤变,若无法准确预测旅客移动轨迹,将导致频繁错切、信令大量耗费的严重问题,再如针对新兴的AR/VR超高清视频业务,若网络无法感知识别业务形式,仍将其作为普通数据流对待,继而导致抖动大、延时的失败体验,此外,在城市商圈区域,各类网络主体(运营商/第三方服务商等)可能针对自己侧采取本地优化,这将造成区域网络整体效率大幅下降的局面这些问题的存在,严重制约了运营商提升5G网络运维水平的能力,也无法满足用户日益增长的网络体验需求。

[0005] 鉴于此,本发明提出一种基于5G的手机数据传输方法和系统以解决上述问题。

发明内容

[0006] 为了克服现有技术的上述缺陷,为实现上述目的,本发明提供如下技术方案:一种基于5G的手机数据传输方法,包括:

[0007] S1、采集设备业务参数、设备数据和基站数据;

[0008] S2、根据设备业务参数计算收益值,根据设备数据和收益值构建设备加权函数;根据基站数据构建基站加权函数;根据设备加权函数和基站加权函数构建移动设备与n个5G基站之间的博弈资源对抗模型;

[0009] S3、利用强化学习中的Q学习方法对博弈资源对抗模型迭代计算求解;得到Q值矩阵;

[0010] S4、采集历史移动设备数据;根据历史移动设备数据建立用于预测联合状态的深度序列学习模型,根据预测的联合状态构建移动设备时间序列状态;

[0011] S5、联立Q值矩阵和移动设备时间序列状态,构建时序Q值矩阵;将时序Q值矩阵采用深度强化学习方法获取最优切换决策序列,根据所述最优切换决策序列分时段发起移动设备的切换请求。

[0012] 进一步地,所述设备业务参数包括数据业务平均吞吐量、单位数据吞吐量收入、语音业务平均持续时间、单位时间语音收入、视频业务平均码率和单位视频码率收入;

[0013] 所述设备数据包括平均时延和移动性开销;所述基站数据包括基站负载提升值和平均传输功率;

[0014] 所述收益值的计算方式包括:

[0015] 收益值 $R = T_data \times R_T + T_voice \times R_V + B_video \times R_B$;

[0016] 其中, T_data 为数据业务平均吞吐量; R_T 为单位数据吞吐量收入; T_voice 为语音业务平均持续时间; R_V 为单位时间语音收入; B_video 为视频业务平均码率; R_B 为单位视频码率收入;

[0017] 所述平均时延 D 是一个预定义的业务质量统计参数,由网络侧的质量监控系统实时统计汇总,直接调用;

[0018] 测算移动设备切换前后的位置变化,即为移动性开销 C ;

[0019] 获取移动设备在切换前后的位置坐标;通过移动设备自身的GPS或定位模块获取切换前后的位置坐标,并上报网络侧;网络侧收到位置坐标后,按照坐标系的计算方法获得两位置之间的距离 ds ;查询移动设备的类型参数,获取其平均移动速度 V ;移动时间开销=距离 ds /平均速度 V ;移动时间开销即为移动性开销 C ;

[0020] 构建设备加权函数 $U = \omega 1 \times R + \omega 2 \times D + \omega 3 \times C$;

[0021] 其中, $\omega 1$ 、 $\omega 2$ 和 $\omega 3$ 为设备权重参数;

[0022] 所述博弈资源对抗模型的构建方式包括:

[0023] 获取参与博弈决策的每个移动设备,确定移动设备的唯一标识ID;获取每个移动设备当前所运行的业务类型;将每个移动设备定义设备动作空间,设备动作空间包括移动设备可申请切换的 L 个相邻基站;

[0024] 获取移动设备周边参与博弈决策的每个基站,确定基站的唯一标识ID;获取每个基站的网络频谱资源信息和基站参数;将每个基站定义基站动作空间,基站动作空间包括接受切换请求和拒绝切换请求;

[0025] 移动设备接入后,计算基站负载提升值 ΔL ;

[0026] 基站负载提升值 ΔL 的计算过程包括:

[0027] 获取基站的当前业务数 N 和移动设备的业务类型;根据业务类型查询对应类业务

的平均占用带宽BW_语音、BW_数据和BW_视频；

[0028] 计算基站当前总占用带宽BW_总= $N \times BW_{\text{语音}} + N \times BW_{\text{数据}} + N \times BW_{\text{视频}}$ ；

[0029] 移动设备接入后，总业务数变为 $N+1$ ，则总占用带宽为BW_新总= $(N+1) \times BW_{\text{语音}} + (N+1) \times BW_{\text{数据}} + (N+1) \times BW_{\text{视频}}$ ；带宽提升量BW_提升=BW_新总-BW_总；

[0030] 带宽提升量与基站总带宽计算负载提升 $\Delta L = BW_{\text{提升}} / \text{基站总带宽}$ ；

[0031] 查表获取移动设备的平均传输功率P；

[0032] 构建基站加权函数 $V = \alpha 1 \times \Delta L + \alpha 2 \times P$ ；

[0033] 其中， $\alpha 1$ 和 $\alpha 2$ 为基站权重参数；

[0034] 结合设备加权函数和基站加权函数获取博弈资源对抗模型；

[0035] 博弈资源对抗模型 $J = \max \sum U - \sum V$ 。

[0036] 进一步地，所述Q值矩阵的获取过程包括：

[0037] S301、定义博弈资源对抗模型中的状态空间S和动作空间A；其中，状态空间S表示场景数量；动作空间A表示每种状态下可供移动设备或基站选择执行的不同行为数量；初始化一个存储状态—动作值函数的Q(s,a)矩阵，其中Q(s,a)矩阵的大小为 $|S| \times |A|$ ，Q(s,a)矩阵的行表示状态S，列表示动作a；矩阵中的每个元素Q(s,a)表示在状态S下选择执行动作a对应的价值，初始化全部矩阵元素为0；

[0038] S302、设定迭代终止条件，迭代终止条件为最大迭代数M或Q矩阵元素值变化小于设置阈值p；设定Q学习算法的参数，参数包括学习率 α ，折扣因子 γ ，在 ϵ 贪婪策略中探索概率 ϵ ；

[0039] S303、基于 ϵ 贪婪策略以固定概率 ϵ 选择当前Q值最大的动作a，同时以 $1-\epsilon$ 概率随机选择一个动作a；执行所选择的动作a后，环境状态从S转移到S'，并产生一个即时收益r；收集执行动作a后环境的反馈，反馈包括新的状态S'和即时收益r；

[0040] S304、更新Q(s,a)；

[0041] 更新公式为 $Q(s,a) \leftarrow Q(s,a) + \alpha \times (r + \gamma \times \text{Max}(Qs') - Q(s,a))$ ；

[0042] 式中，S为环境状态；a为在状态S下可选择执行的动作；r为即时收益；S'为执行完动作后的新的环境状态； α 为学习率； γ 为折扣因子；Max(Qs')为下一状态S'下所有可能的动作对应的Q值最大值； $\leftarrow$ 为更新符号；

[0043] S305、判断是否满足终止条件，若不满足则继续更新Q(s,a)；若满足则得到状态空间下的Q值矩阵。

[0044] 进一步地，所述历史移动设备数据包括历史定位轨迹数据和历史业务数据；

[0045] 历史定位轨迹数据包括设备经纬度坐标及其对应的设备定位时间戳；

[0046] 历史业务数据包括业务类型、业务流量和网络连接数及其对应的时间戳；

[0047] 联合状态为未来f天的定位轨迹数据和业务数据；

[0048] 深度序列学习模型为联立定位预测模型和参数预测模型获取；

[0049] 定位预测模型获取方式包括：

[0050] 将历史定位轨迹数据作为测试数据集,并进行曲线拟合;

[0051] 曲线拟合过程包括:

[0052] 用线性多项式拟合曲线、二次多项式拟合曲线和三次多项式分别进行拟合曲线;线性多项式拟合曲线为 $y = ax + b$;二次多项式拟合曲线 $y = ax^2 + bx + c$;三次多项式拟合曲线 $y = ax^3 + bx^2 + cx + d$;式中, x 为设备定位时间戳; y 为设备经纬度坐标; a 、 b 、 c 和 d 为待求的参数,均通过曲线拟合得出;

[0053] 计算线性多项式、二次多项式和三次多项式拟合曲线的预测点位误差;预测点位误差 $E = \sum \frac{|\text{实际值} - \text{预测值}|}{h}$;

[0054] 其中,实际值为测试数据集中的真实值;预测值为拟合曲线计算的预测值; h 为测试数据集大小;

[0055] 选取多项式拟合曲线、二次多项式拟合曲线和三次多项式拟合曲线中预测点位误差最小的拟合曲线,即最佳拟合的曲线;最佳拟合的曲线即为定位预测模型。

[0056] 进一步地,参数预测模型为ARIMA时间序列预测模型,其获取方式包括:

[0057] 采集最近 r 天内的历史业务数据作为历史样本,历史样本中包括4个字段;4个字段分别为时间戳、业务类型、业务流量和网络连接数;将历史样本按照时间顺序构建输入样本矩阵,每一行代表一个采样点位;

[0058] 将输入样本矩阵的数据进行预处理;将24小时的全天数据依照分时段划分为 g 个时段区间,统计计算每个时段区间内各类业务数据所占比重,得到不同业务类别在各个时间段内的分布统计量;用于相应调整预测结果;

[0059] 将网络连接数的时间序列数据建立ARIMA自回归滑动平均模型;进行平稳性检验,对不满足稳定条件的序列进行差分处理,直至达到稳定条件;

[0060] 确定时间序列的AR和MA阶数,配置ARIMA(p, d, q)模型的参数;

[0061] 利用配置好的参数将ARIMA自回归滑动平均模型训练和优化,输入训练数据为预处理的输入样本矩阵,训练目标为预测未来 f 天内每小时的网络连接数;获得未来 f 天内网络连接数的时间序列;训练完成的ARIMA自回归滑动平均模型即为参数预测模型。

[0062] 进一步地,所述联立定位预测模型和参数预测模型获取深度序列学习模型的方式包括:

[0063] 将过去 r 天的历史定位轨迹数据和历史业务数据作为历史训练集;将预测得到的未来 f 天内设备经纬度坐标及其对应的定位时间戳、业务类型、业务流量和网络连接数及其对应的时间戳;利用时间戳进行时间对齐,并作为未来数据集;

[0064] 构建卷积神经网络LSTM模型;输入层为历史训练集和未来数据集拼接而成的三维矩阵;

[0065] 卷积神经网络LSTM模型包括卷积层、池化层、LSTM层和全连接层;编译卷积神经网络LSTM模型,配置优化器和损失函数;

[0066] 优化器选用Adam算法;Adam算法基于优化随机梯度下降,自适应地对学习率进行调整;

[0067] 损失函数选用均方误差 $MSE = \frac{1}{Z} \sum_{i=1}^Z (Y_{pred,i} - Y_{true,i})^2$;

[0068] 其中, Z 为样本数量; $Y_{pred,i}$ 为模型预测得到的第 i 个样本输出; $Y_{true,i}$ 为第 i 个样本真实的标签输出;

[0069] **MSE** 函数计算预测值和真实值之间的均方差, 惩罚得到的损失数值越小模型效果越好; 以历史训练集为卷积神经网络LSTM模型输入, 训练目标是预测出未来数据集;

[0070] 将历史训练集按照批大小分批输入卷积神经网络LSTM模型, 卷积神经网络LSTM模型输出一批预测结果; 计算这一批预测结果和未来数据集批之间的均方误差 **MSE** 函数值; 利用反向传播算法, 将 **MSE** 损失误差沿着模型计算图反方向回传, 更新每层的卷积神经网络LSTM模型的参数;

[0071] 当训练达到预定的迭代轮数后, 获得最终训练好的深度序列学习模型; 使用训练好的深度序列学习模型进行预测; 输入给定的过去 r 天历史数据, 深度序列学习模型预测输出未来 f 天的联合状态。

[0072] 进一步地, 所述移动设备时间序列状态的构建方式包括:

[0073] 根据预测的联合状态, 提取未来 f 天内的时间戳、经纬度坐标、业务类型、业务流量和网络连接数; 将提取的数据重新组织成 v 个时间序列, 时间序列状态包括时间戳序列、经纬度坐标时间序列、业务类型时间序列、业务流量时间序列和网络连接数时间序列; 即构建出移动设备在未来 f 天内的移动设备时间序列状态;

[0074] 所述时序 **Q** 值矩阵的构建方式包括:

[0075] 从移动设备时间序列状态中, 提取所有时刻的状态参数, 构成状态空间 S , 状态参数包括移动设备当前所运行的业务类型、移动设备的位置坐标和运动速度; 业务类型包括语音业务、视频业务和数据业务;

[0076] 从 **Q** 值矩阵中, 提取移动设备所有可选择的基站切换动作, 构成动作空间 A ; 初始化存储状态—动作值的时序 **Q** (St, At) 矩阵; 时序 **Q** (St, At) 矩阵的行表示不同时刻 t 的设备状态 St , 时序 **Q** (St, At) 矩阵表示 L 个相邻基站作为基站切换动作 At ; 矩阵元素 **Q** (St, At) 表示在状态 St 下选择动作 At 的价值, 初始化为 0。

[0077] 进一步地, 所述时序 **Q** 值矩阵采用深度强化学习方法获取最优切换决策序列的过程包括:

[0078] 构建深度强化学习模型为端到端的深度神经网络, 输入为设备的状态 St , 输出为当前状态下各个基站切换动作的 **Q** 值;

[0079] 深度强化学习模型内部结构包括卷积层、LSTM循环层和全连接层, 卷积层负责提取输入状态的时空特征表示, LSTM层负责挖掘状态序列的时序关联特征, 全连接层完成状态动作值函数 **Q** (St, At) 的估计;

[0080] 令马尔可夫决策过程 (MDP) 的五元组中即时奖励为 r_t ; 通过采取时序上的折扣累积方法定义总奖励 **TY**, 公式为: $TY = \sum_{t=0}^{tr} \gamma^t r_t$;

[0081] 其中, γ 是数值为0-1的折扣因子; tr 为决策序列终止时刻; t 表示的是当前的时间点, 从该时刻 t 开始计算一个序列中的累积折扣收益; 模型的目标是配置参数使得总奖励 TY 最大化;

[0082] 深度强化学习模型训练目标是逼近状态—动作值函数 $Q(S_t, A_t)$, 在每一训练步, 输入一个采样状态 S_t , 模型预测输出当前状态下各个基站切换动作 a 对应的 Q 值; 执行动作后环境给出反馈的实际即时奖励; 计算 $Q(S_t, A_t)$ 预测值和实际奖励的误差, 通过损失函数反向传播更新模型参数;

[0083] 在训练完成深度强化学习模型后, 输入需要决策的时序状态, 模型输出该状态下, 不同基站切换动作对应的 Q 值, 根据 Q 值最大选择基站, 推演形成整个序列, 计算序列总奖励, 输出总奖励最大的最终序列, 作为最优切换决策。

[0084] 进一步地, 所述根据所述最优切换决策序列分时段发起移动设备的切换请求的过程包括:

[0085] S501、最优切换决策序列按照时间顺序给出了从时刻 t_1 到 t_N , 移动设备在每个时间点需要切换到的目标基站; 网络侧得到最优切换序列后, 转换为信令交互过程;

[0086] S502、在时刻 t_1 , 网络侧向移动设备发送切换命令, 通知设备切换到对应目标基站; 移动设备收到切换命令后, 向网络发送切换准备就绪信令; 网络侧向源基站发送资源释放信令, 向目标基站发送资源准备信令;

[0087] S503、源基站响应, 释放移动设备占用资源; 目标基站响应, 预留资源; 网络侧向移动设备发送确认信令, 授权其执行切换资源重配置流程; 移动设备与目标基站进行随机接入, 完成切换;

[0088] S504、按照时间顺序, 每到最优切换序列给出的时刻, 重复执行S501-S503, 完成切换, 到时刻 t_N 全部切换完成。

[0089] 一种基于5G的手机数据传输系统, 其基于所述的一种基于5G的手机数据传输方法实现, 包括: 收集接收模块, 用于采集设备业务参数、设备数据和基站数据;

[0090] 对抗模型构造模块, 用于根据设备业务参数计算收益值, 根据设备数据和收益值构建设备加权函数; 根据基站数据构建基站加权函数; 根据设备加权函数和基站加权函数构建移动设备与 n 个5G基站之间的博弈资源对抗模型;

[0091] 求解模块, 用于利用强化学习中的 Q 学习方法对博弈资源对抗模型迭代计算求解; 得到 Q 值矩阵;

[0092] 状态构建模块, 用于采集历史移动设备数据; 根据历史移动设备数据建立用于预测联合状态的深度序列学习模型, 根据预测的联合状态构建移动设备时间序列状态;

[0093] 决策实施模块, 用于联立 Q 值矩阵和移动设备时间序列状态, 构建时序 Q 值矩阵; 将时序 Q 值矩阵采用深度强化学习方法获取最优切换决策序列, 根据所述最优切换决策序列分时段发起移动设备的切换请求。

[0094] 本发明一种基于5G的手机数据传输方法和系统的技术效果和优点:

[0095] 实现了智能感知和业务关联的移动网络切换新机制; 通过构建序列学习模型, 实现移动设备个体级别的高精度移动性预测, 准确得到用户未来的服务需求和网络连接趋

势,大幅减少预测偏差;通过建立移动设备与网络资源之间的关联认知模型,进行面向业务场景的切换控制,按照业务优先级进行基站动态切换,极大提升用户感知体验;本实施例利用强化学习算法,在保证单个用户效果提升的同时,实现了整个网络全局收益的最大化,获得考虑各方博弈的最优网络切换均衡解;融合个体建模、业务感知和全局优化为一体实现了稳定的5G网络运维。

附图说明

[0096] 图1为本发明的一种基于5G的手机数据传输方法示意图;

[0097] 图2为本发明的一种基于5G的手机数据传输系统示意图;

[0098] 图3为本发明的电子设备示意图;

[0099] 图4为本发明的存储介质示意图。

具体实施方式

[0100] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0101] 实施例1

[0102] 请参阅图1所示,本实施例所述一种基于5G的手机数据传输方法,包括:

[0103] S1、采集设备业务参数、设备数据和基站数据;

[0104] S2、根据设备业务参数计算收益值,根据设备数据和收益值构建设备加权函数;根据基站数据构建基站加权函数;根据设备加权函数和基站加权函数构建移动设备与n个5G基站之间的博弈资源对抗模型;

[0105] S3、利用强化学习中的Q学习方法对博弈资源对抗模型迭代计算求解;得到Q值矩阵;

[0106] S4、采集历史移动设备数据;根据历史移动设备数据建立用于预测联合状态的深度序列学习模型,根据预测的联合状态构建移动设备时间序列状态;

[0107] S5、联立Q值矩阵和移动设备时间序列状态,构建时序Q值矩阵;将时序Q值矩阵采用深度强化学习方法获取最优切换决策序列,根据所述最优切换决策序列分时段发起移动设备的切换请求;

[0108] 所述设备业务参数包括数据业务平均吞吐量、单位数据吞吐量收入、语音业务平均持续时间、单位时间语音收入、视频业务平均码率和单位视频码率收入;

[0109] 所述设备数据包括平均时延和移动性开销;所述基站数据包括基站负载提升值和平均传输功率;

[0110] 所述收益值的计算方式包括:

[0111] 收益值 $R = T_data \times R_T + T_voice \times R_V + B_video \times R_B$;

[0112] 其中, T_data 为数据业务平均吞吐量; R_T 为单位数据吞吐量收入; T_voice 为语音业务平均持续时间; R_V 为单位时间语音收入; B_video 为视频业务平均码率; R_B

为单位视频码率收入；

[0113] 所述平均时延 D 是一个预定义的业务质量统计参数,由网络侧的质量监控系统实时统计汇总,直接调用；

[0114] 测算移动设备切换前后的位置变化,即为移动性开销 C ；

[0115] 获取移动设备在切换前后的位置坐标；通过移动设备自身的GPS或定位模块获取切换前后的位置坐标,并上报网络侧；网络侧收到位置坐标后,按照坐标系的计算方法获得两位置之间的距离 ds ；查询移动设备的类型参数,获取其平均移动速度 V ；移动时间开销=距离 ds /平均速度 V ；移动时间开销即为移动性开销 C ；

[0116] 构建设备加权函数 $U = \omega 1 \times R + \omega 2 \times D + \omega 3 \times C$ ；

[0117] 其中, $\omega 1$ 、 $\omega 2$ 和 $\omega 3$ 为设备权重参数；其中, $\omega 1$ 、 $\omega 2$ 和 $\omega 3$ 的获取是通过回归分析和最小二乘法来拟合历史数据实现的；

[0118] 进一步的,所述博弈资源对抗模型的构建方式包括：

[0119] 获取参与博弈决策的每个移动设备,确定移动设备的唯一标识ID；获取每个移动设备当前所运行的业务类型,确定移动设备当前所运行是语音/视频/数据业务；将每个移动设备定义设备动作空间,设备动作空间包括移动设备可申请切换的 L 个相邻基站；

[0120] 需要说明的是, T_data 数据业务平均吞吐量的获取方式为收集移动设备在各基站的历史业务数据；统计计算数据业务流量时间占比；数据吞吐量=数据流量/数据业务占比时间；

[0121] R_T 单位数据吞吐量收入的获取方式为查阅运营商计费政策；数据业务收费按流量计费,即1KB流量的收费价格；单位数据吞吐量收入=每KB流量收费价格/1024；

[0122] T_voice 语音业务平均持续时间的获取方式为收集语音业务建立时长数据,统计计算全部语音业务的平均持续时间；

[0123] R_V 单位时间语音收入的获取方式为查阅运营商计费政策；语音业务按时间计费,即每分钟通话时间的收费价格；

[0124] B_video 视频业务平均码率的获取方式为收集移动设备视频业务的码率大小数据；统计计算全部视频业务的平均码率；

[0125] R_B 单位视频码率收入的获取方式为查阅运营商计费政策,视频业务按流量计费,即每MB视频流量的收费价格；单位码率收入=视频收费价格/码率大小；

[0126] 需要解释的是,所述坐标系的计算方法包括：

[0127] 假设移动设备切换前坐标为 $(x1, y1)$,切换后坐标为 $(x2, y2)$ ；则两坐标之间的距离计算公式为 $ds = \sqrt{((x1 - x2)^2 + (y1 - y2)^2)}$ ；就是基于笛卡尔坐标系的两点间距离计算公式,网络侧获取到设备的两位置坐标后,可据此计算直线距离；

[0128] 获取移动设备周边参与博弈决策的每个基站,确定基站的唯一标识ID；获取每个基站的网络频谱资源信息和基站参数；将每个基站定义基站动作空间,基站动作空间包括接受切换请求和拒绝切换请求；

[0129] 移动设备接入后,计算基站负载提升值 ΔL ；

[0130] 基站负载提升值 ΔL 的计算过程包括：

- [0131] 获取基站的当前业务数 N 和移动设备的业务类型;根据业务类型查询对应类业务的平均占用带宽 $BW_{\text{语音}}$ 、 $BW_{\text{数据}}$ 和 $BW_{\text{视频}}$;
- [0132] 计算基站当前总占用带宽 $BW_{\text{总}}=N \times BW_{\text{语音}}+N \times BW_{\text{数据}}+N \times BW_{\text{视频}}$;
- [0133] 移动设备接入后,总业务数变为 $N+1$,则总占用带宽为 $BW_{\text{新总}}=(N+1) \times BW_{\text{语音}}+(N+1) \times BW_{\text{数据}}+(N+1) \times BW_{\text{视频}}$;带宽提升量 $BW_{\text{提升}}=BW_{\text{新总}}-BW_{\text{总}}$;
- [0134] 带宽提升量与基站总带宽计算负载提升 $\Delta L=BW_{\text{提升}}/\text{基站总带宽}$;
- [0135] 查表获取移动设备的平均传输功率 P ;
- [0136] 构建基站加权函数 $V = \alpha 1 \times \Delta L + \alpha 2 \times P$;
- [0137] 其中, $\alpha 1$ 和 $\alpha 2$ 为基站权重参数;
- [0138] 结合设备加权函数和基站加权函数获取博弈资源对抗模型;
- [0139] 博弈资源对抗模型 $J = \max \sum U - \sum V$;
- [0140] 博弈资源对抗模型即为最大化所有移动设备的加权函数之和减去所有基站的加权函数之和;
- [0141] 进一步的,所述 Q 值矩阵的获取过程包括:
- [0142] S301、定义博弈资源对抗模型中的状态空间 S 和动作空间 A ;其中,状态空间 S 表示场景数量;动作空间 A 表示每种状态下可供移动设备或基站选择执行的不同行为数量;初始化一个存储状态—动作值函数的 $Q(s, a)$ 矩阵,其中 $Q(s, a)$ 矩阵的大小为 $|S| \times |A|$, $Q(s, a)$ 矩阵的行表示状态 S ,列表示动作 a ;矩阵中的每个元素 $Q(s, a)$ 表示在状态 S 下选择执行动作 a 对应的价值,初始化全部矩阵元素为0;
- [0143] S302、设定迭代终止条件,迭代终止条件为最大迭代数 M 或 Q 矩阵元素值变化小于设置阈值 p ;同时设定 Q 学习算法需要的参数,参数包括学习率 α ,折扣因子 γ , ϵ 贪婪策略中探索概率 ϵ ;其中,学习率 α 表示更新 Q 矩阵时当前信息和历史信息的权重比例;折扣因子 γ 表示评估当前收益和未来收益的重要性; ϵ 贪婪策略以 ϵ 概率随机选择动作,以 $1-\epsilon$ 概率选择 Q 值最大动作;
- [0144] 需要说明的是,最大迭代数 M 可根据实际情况设定;
- [0145] 阈值 p 的设置方式包括:
- [0146] 先使用较大 p 值,快速预训练,确定训练次数级别;然后从中选择一定比例,如设置为预训练次数的10%左右;以上次数作为阈值下限;上限可以设置的更大一些;在上下限范围中试验,选取一个较优的 p ;
- [0147] S303、基于 ϵ 贪婪策略以固定概率 ϵ 选择当前 Q 值最大的动作 a ,同时以 $1-\epsilon$ 概率随机选择一个动作 a ;执行所选择的动作 a 后,环境状态从 S 转移到 S' ,并产生一个即时收益 r ;收集执行动作 a 后环境的反馈,反馈包括新的状态 S' 和即时收益 r ;
- [0148] S304、更新 $Q(s, a)$;
- [0149] 更新公式为 $Q(s, a) \leftarrow Q(s, a) + \alpha \times (r + \gamma \times \text{Max}(Qs') - Q(s, a))$;
- [0150] 式中, S 为环境状态;对博弈问题来说,指博弈场景中参与博弈各方(移动用户、基站)的状态参数的总体描述; a 为在状态 S 下可选择执行的动作;博弈问题中指参与博弈各

方的不同决策选择; $\mathbf{r}$ 为即时收益, 选择某动作后环境产生的立即的收益; 对博弈问题, 指执行一决策的直接结果; $\mathbf{s}'$ 为执行完动作后的新的环境状态, 一个动作会改变博弈场景, 带来新的状态; α 为学习率, 更新 Q 时当前值和历史值的权重比例; α 越大, 越重视新信息; 一般取 0-1 区间数值; γ 为折扣因子, 考虑当前收益和未来收益的重要性程度; γ 取值越大, 表明未来收益更重要, 一般取 0-1 区间数值; $\text{Max}(Q_{s'})$ 为下一状态 $\mathbf{s}'$ 下所有可能的动作对应的 Q 值最大值; $\leftarrow$ 为更新符号;

[0151] S305、判断是否满足终止条件, 若不满足则继续更新 $Q(\mathbf{s}, \mathbf{a})$; 若满足则得到状态空间下的 Q 值矩阵;

[0152] 需要说明的是, Q 学习是一个值函数逼近算法, 其基本思想是通过试错学习一个行为价值函数 (Q 函数), 用于评估在给定状态 $\mathbf{S}$ 下采取动作 $\mathbf{a}$ 的长期收益价值; 其目标是找到一个策略 π , 该策略可以最大化长期收益;

[0153] 进一步的, 所述历史移动设备数据包括历史定位轨迹数据和历史业务数据;

[0154] 历史定位轨迹数据包括设备经纬度坐标及其对应的设备定位时间戳;

[0155] 历史业务数据包括业务类型、业务流量和网络连接数及其对应的时间戳;

[0156] 需要说明的是, 移动设备通过自身的 GPS 模块或基站定位模块, 周期性获取设备在基站覆盖区域内的精确经纬度坐标; 同时通过设备本地时钟, 获取坐标采集时所对应的时间戳; 将采集到的经纬度坐标数据和对应时间戳在本地存储, 并定期上报到网络侧数据库;

[0157] 利用网络侧核心网网元, 例如, HSS 存储每台移动设备的业务控制信息; 可统计提取每台设备在服务期内的语音/视频/数据业务开通和关闭的业务类型、持续时长、流量大小等参数; 同时绑定业务发生时间的时间戳; 最终形成设备级别的业务类型、流量大小、网络连接数与时间戳的对应历史业务数据

[0158] 联合状态为未来 f 天的定位轨迹数据和业务数据;

[0159] 深度序列学习模型为联立定位预测模型和参数预测模型获取;

[0160] 定位预测模型获取方式包括:

[0161] 将历史定位轨迹数据作为测试数据集, 并进行曲线拟合;

[0162] 曲线拟合过程包括:

[0163] 用线性多项式拟合曲线、二次多项式拟合曲线和三次多项式分别进行拟合曲线; 线性多项式拟合曲线为 $\mathbf{y} = \mathbf{ax} + \mathbf{b}$; 二次多项式拟合曲线 $\mathbf{y} = \mathbf{ax}^2 + \mathbf{bx} + \mathbf{c}$; 三次多项式拟合曲线 $\mathbf{y} = \mathbf{ax}^3 + \mathbf{bx}^2 + \mathbf{cx} + \mathbf{d}$; 式中, $\mathbf{x}$ 为设备定位时间戳; $\mathbf{y}$ 为设备经纬度坐标; $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{c}$ 和 $\mathbf{d}$ 为待求的参数, 均通过曲线拟合得出;

[0164] 计算线性多项式、二次多项式和三次多项式拟合曲线的预测点位误差; 预测点位

$$\text{误差} E = \sum \frac{|\text{实际值} - \text{预测值}|}{h};$$

[0165] 其中, 实际值为测试数据集中的真实值; 预测值为拟合曲线计算的预测值; h 为测试数据集大小;

[0166] 选取多项式拟合曲线、二次多项式拟合曲线和三次多项式拟合曲线中预测点位误

差最小的拟合曲线,即最佳拟合的曲线;最佳拟合的曲线即为定位预测模型,利用定位预测模型预测未来 f 天内的设备经纬度坐标及其对应的定位时间戳;

[0167] 参数预测模型为ARIMA时间序列预测模型,其获取方式包括:

[0168] 采集最近 r 天内的历史业务数据作为历史样本,历史样本中包括4个字段;4个字段分别为时间戳、业务类型、业务流量和网络连接数;将历史样本按照时间顺序构建输入样本矩阵,每一行代表一个采样点位;

[0169] 将输入样本矩阵的数据进行预处理;将24小时的全天数据依照分时段划分为 g 个时段区间,例如,将24小时划分为6个时间段,统计计算每个时段区间内各类业务数据所占比重,得到不同业务类别在各个时间段内的分布统计量;用于相应调整预测结果;

[0170] 将网络连接数的时间序列数据建立ARIMA自回归滑动平均模型;进行平稳性检验,对不满足稳定条件的序列进行差分处理,直至达到稳定条件;

[0171] 确定时间序列的AR和MA阶数,配置ARIMA(p, d, q)模型的参数;

[0172] 需要说明的是,将网络连接数的时间序列数据建立ARIMA自回归滑动平均模型的过程包括:

[0173] 对网络连接数的时间序列数据进行平稳性检验,判断是否满足平稳条件。检验方法可以采用ADF单位根检验法;当原始时间序列不满足平稳条件时,需要进行差分处理,目的是去除时间序列中的随机性和非平稳性;进行差分操作直至时间序列满足平稳条件,得到平稳序列确定时间序列的自回归参数的阶数 p 和滑动平均参数的阶数 q ;可以通过部分自相关图PACF和自相关图ACF的尾部情况来确定 p 和 q 的值;

[0174] 配置ARIMA(p, d, q)模型的参数;模型的参数包括 p 阶自回归参数, q 阶滑动平均参数, d 为进行了几次差分后的结果;利用最大似然估计法估计模型的参数;这样就建立了ARIMA(p, d, q)模型,可以用于时间序列分析和预测;

[0175] 利用配置好的参数将ARIMA自回归滑动平均模型训练和优化,输入训练数据为预处理的输入样本矩阵,训练目标为预测未来 f 天内每小时的网络连接数;获得未来 f 天内网络连接数的时间序列;训练完成的ARIMA自回归滑动平均模型即为参数预测模型;

[0176] 利用参数预测模型预测得到的未来 f 天网络连接数的时间序列,依据不同业务类别在各个时间段内的分布统计量进行校正;即对每一时段的网络连接数预测值乘以该时段内各业务比例,获得未来 f 天内各类业务分布预测结果;各类业务分布预测结果包括业务类型、业务流量和网络连接数及其对应的时间戳;

[0177] 进一步的,所述联立定位预测模型和参数预测模型获取深度序列学习模型的方式包括:

[0178] 将过去 r 天的历史定位轨迹数据和历史业务数据作为历史训练集;将预测得到的未来 f 天内设备经纬度坐标及其对应的定位时间戳、业务类型、业务流量和网络连接数及其对应的时间戳;利用时间戳进行时间对齐,并作为未来数据集;

[0179] 构建卷积神经网络LSTM模型;输入层为历史训练集和未来数据集拼接而成的三维矩阵,第一维和第二维表示样本 id 和时间步,第三维表示特征维度;

[0180] 卷积神经网络LSTM模型包括卷积层、池化层、LSTM层和全连接层;卷积层用于自动提取输入数据的局部特征;池化层用于减少特征维度;LSTM层用于提取时间序列数据的长短期特征;全连接层输出预测结果;

[0181] 编译卷积神经网络LSTM模型,配置优化器和损失函数;优化器选用Adam算法;Adam算法基于优化随机梯度下降,自适应地对学习率进行调整;

[0182] 损失函数选用均方误差 **MSE** 函数;

$$[0183] \quad \text{MSE} = \frac{1}{Z} \sum_{i=1}^Z (Y_{pred,i} - Y_{true,i})^2;$$

[0184] 其中, Z 为样本数量; $Y_{pred,i}$ 为模型预测得到的第 i 个样本输出; $Y_{true,i}$ 为第 i 个样本真实的标签输出;

[0185] **MSE** 函数计算预测值和真实值之间的均方差,惩罚得到的损失数值越小模型效果越好;

[0186] 以历史训练集为卷积神经网络LSTM模型输入,训练目标是预测出未来数据集;

[0187] 将历史训练集按照批大小分批输入卷积神经网络LSTM模型,卷积神经网络LSTM模型输出一批预测结果;计算这一批预测结果和未来数据集批之间的均方误差 **MSE** 函数值;

[0188] 利用反向传播算法,将 **MSE** 损失误差沿着模型计算图反方向回传,更新每层的卷积神经网络LSTM模型的参数;

[0189] 当训练达到预定的迭代轮数后,获得最终训练好的模型;预定的迭代轮数可以根据实际情况设定;

[0190] 使用训练好的深度序列学习模型进行预测;输入给定的过去 r 天历史数据,深度序列学习模型预测输出未来 f 天的联合状态;

[0191] 进一步的,所述移动设备时间序列状态的构建方式包括:

[0192] 根据预测的联合状态,提取未来 f 天内的时间戳、经纬度坐标、业务类型、业务流量和网络连接数;将提取的数据重新组织成 v 个时间序列,时间序列状态包括时间戳序列、经纬度坐标时间序列、业务类型时间序列、业务流量时间序列和网络连接数时间序列;即构建出移动设备在未来 f 天内的移动设备时间序列状态;

[0193] 进一步的,所述时序 **Q** 值矩阵的构建方式包括:

[0194] 从移动设备时间序列状态中,提取所有时刻的状态参数,构成状态空间 S ,状态参数包括移动设备当前所运行的业务类型、移动设备的位置坐标和运动速度;业务类型包括语音业务、视频业务和数据业务;

[0195] 从 **Q** 值矩阵中,提取移动设备所有可选择的基站切换动作,构成动作空间 A ;初始化存储状态—动作值的时序 **Q** (St, At) 矩阵;时序 **Q** (St, At) 矩阵的行表示不同时刻 t 的设备状态 St ,时序 **Q** (St, At) 矩阵表示 L 个相邻基站作为基站切换动作 At ;矩阵元素 **Q** (St, At) 表示在状态 St 下选择动作 At 的价值,初始化为0;

[0196] 进一步的,所述时序 **Q** 值矩阵采用深度强化学习方法获取最优切换决策序列的过程包括:

[0197] 构建深度强化学习模型为端到端的深度神经网络,输入为设备的状态 St ,输出为当前状态下各个基站切换动作的 **Q** 值;

[0198] 深度强化学习模型内部结构包括卷积层、LSTM循环层和全连接层,卷积层负责提

取输入状态的时空特征表示,LSTM层负责挖掘状态序列的时序关联特征,全连接层完成状态动作值函数 $Q(S_t, A_t)$ 的估计;

[0199] 令马尔可夫决策过程(MDP)的五元组中即时奖励为 r_t ;通过采取时序上的折扣累积方法定义总奖励TY,公式为: $TY = \sum_{t=0}^{tr} \gamma^t r_t$;

[0200] 其中, γ 是数值为0-1的折扣因子; tr 为决策序列终止时刻; t 表示的是当前的时间点,从该时刻 t 开始计算一个序列中的累积折扣收益;模型的目标是配置参数使得总奖励TY最大化;

[0201] 深度强化学习模型训练目标是逼近状态-动作值函数 $Q(S_t, A_t)$,在每一训练步,输入一个采样状态 S_t ,模型预测输出当前状态下各个基站切换动作 a 对应的Q值;执行动作后环境给出反馈的实际即时奖励;计算 $Q(S_t, A_t)$ 预测值和实际奖励的误差,通过损失函数反向传播更新模型参数;

[0202] 在训练完成深度强化学习模型后,输入需要决策的时序状态,模型输出该状态下,不同基站切换动作对应的Q值,根据Q值最大选择基站,推演形成整个序列,计算序列总奖励,输出总奖励最大的最终序列,作为最优切换决策;

[0203] 进一步的,所述根据所述最优切换决策序列分时段发起移动设备的切换请求的过程包括:

[0204] S501、最优切换决策序列按照时间顺序给出了从时刻 t_1 到 t_N ,移动设备在每个时间点需要切换到的目标基站;网络侧得到最优切换序列后,转换为信令交互过程;

[0205] S502、在时刻 t_1 ,网络侧向移动设备发送切换命令,通知设备切换到对应目标基站;移动设备收到切换命令后,向网络发送切换准备就绪信令;网络侧向源基站发送资源释放信令,向目标基站发送资源准备信令;

[0206] S503、源基站响应,释放移动设备占用资源;目标基站响应,预留资源;网络侧向移动设备发送确认信令,授权其执行切换资源重配置流程;移动设备与目标基站进行随机接入,完成切换;

[0207] S504、按照时间顺序,每到最优切换序列给出的时刻,重复执行S501-S503,完成切换,到时刻 t_N 全部切换完成;

[0208] 本实施例,实现了智能感知和业务关联的移动网络切换新机制;通过构建序列学习模型,实现移动设备个体级别的高精度移动性预测,准确得到用户未来的服务需求和网络连接趋势,大幅减少预测偏差;同时,通过建立移动设备与网络资源之间的关联认知模型,进行面向业务场景的切换控制,按照业务优先级进行基站动态切换,极大提升用户感知体验;本实施例利用强化学习算法,在保证单个用户效果提升的同时,实现了整个网络全局收益的最大化,获得考虑各方博弈的最优网络切换均衡解;融合个体建模、业务感知和全局优化为一体实现了稳定的5G网络运维。

[0209] 实施例2

[0210] 请参阅图2所示,本实施例未详细叙述部分见实施例1描述内容,提供一种基于5G的手机数据传输系统,包括:收集接收模块,用于采集设备业务参数、设备数据和基站数据;

[0211] 对抗模型构造模块,用于根据设备业务参数计算收益值,根据设备数据和收益值构建设备加权函数;根据基站数据构建基站加权函数;根据设备加权函数和基站加权函数构建移动设备与n个5G基站之间的博弈资源对抗模型;

[0212] 求解模块,用于利用强化学习中的Q学习方法对博弈资源对抗模型迭代计算求解;得到Q值矩阵;

[0213] 状态构建模块,用于采集历史移动设备数据;根据历史移动设备数据建立用于预测联合状态的深度序列学习模型,根据预测的联合状态构建移动设备时间序列状态;

[0214] 决策实施模块,用于联立Q值矩阵和移动设备时间序列状态,构建时序Q值矩阵;将时序Q值矩阵采用深度强化学习方法获取最优切换决策序列,根据所述最优切换决策序列分时段发起移动设备的切换请求;各个模块之间通过有线和/或无线的方式进行连接,实现模块间的数据传输。

[0215] 实施例3

[0216] 请参阅图3所示,根据本申请的又一方面还提供了电子设备。该电子设备可包括一个或多个处理器以及一个或多个存储器。其中,存储器中存储有计算机可读代码,计算机可读代码当由一个或多个处理器运行时,可以执行如上所述的一种基于5G的手机数据传输方法。

[0217] 根据本申请实施方式的方法或系统也可以借助于图3所示的电子设备的架构来实现。如图3所示,电子设备可包括输入设备、一个或多个运算器、一个或多个存储器、一个或多个控制器和输出设备等。电子设备中的存储器,可存储本申请提供的一种基于5G的手机数据传输方法。当然,图3所示的架构只是示例性的,在实现不同的设备时,根据实际需要,可以省略图3示出的电子设备中的一个或多个组件。

[0218] 实施例4

[0219] 请参阅图4所示,是根据本申请一个实施方式的计算机可读存储介质600。计算机可读存储介质600上存储有计算机可读指令。当计算机可读指令由处理器运行时,可执行参照以上附图描述的根据本申请实施方式的一种。存储介质600包括但不限于例如易失性存储器和/或非易失性存储器。易失性存储器例如可包括随机存取存储器(RAM)和高速缓冲存储器(cache)等。非易失性存储器例如可包括只读存储器(ROM)、硬盘、闪存等。

[0220] 另外,根据本申请的实施方式,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本申请提供了非暂时性机器可读存储介质,所述非暂时性机器可读存储介质存储有机器可读指令,所述机器可读指令能够由处理器运行以执行与本申请提供的方法步骤对应的指令,一种基于5G的手机数据传输方法。在该计算机程序被中央处理单元(CPU)执行时,执行本申请的方法中限定的上述功能。

[0221] 上述实施例,可以全部或部分地通过软件、硬件、固件或其他任意组合来实现。当使用软件实现时,上述实施例可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令或计算机程序。在计算机上加载或执行所述计算机指令或计算机程序时,全部或部分地产生按照本发明实施例所述的流程或功能。所述计算机可以为通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存

储介质传输,例如,所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线网络或无线网络方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集合的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质(例如,软盘、硬盘、磁带)、光介质(例如,DVD)、或者半导体介质。半导体介质可以是固态硬盘。

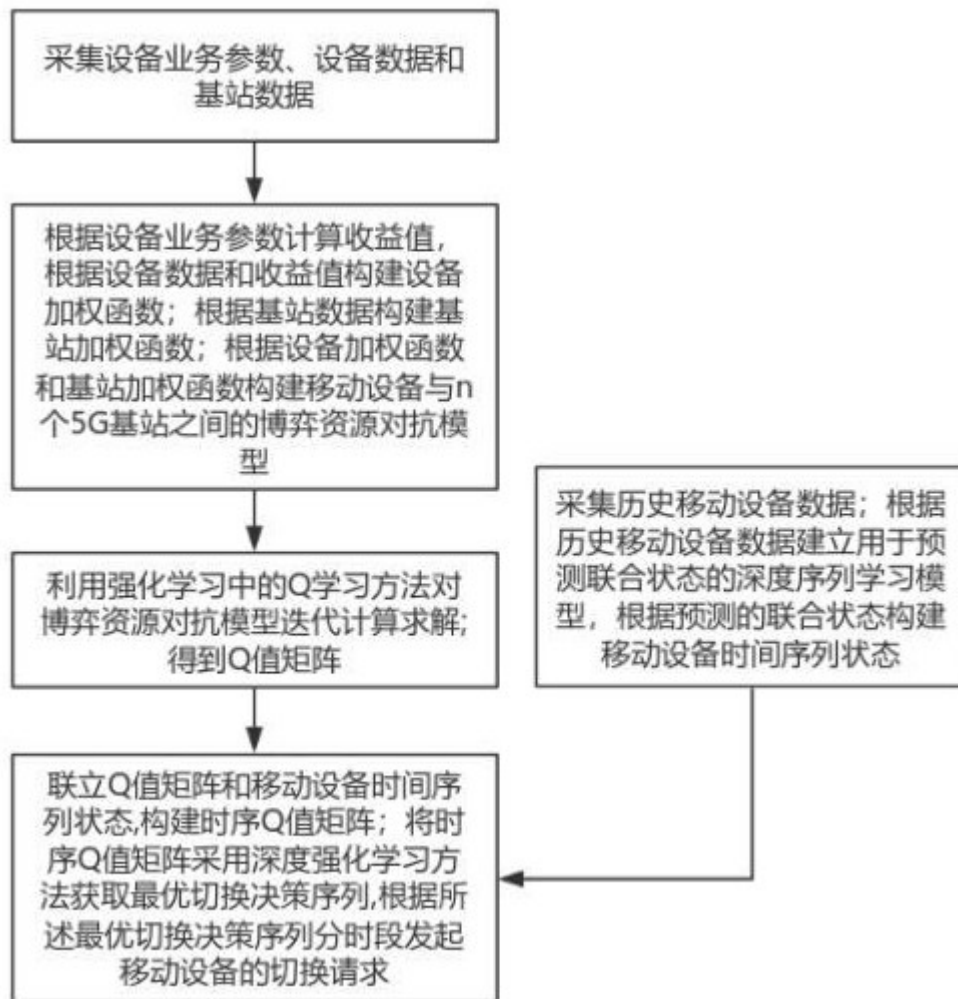


图 1

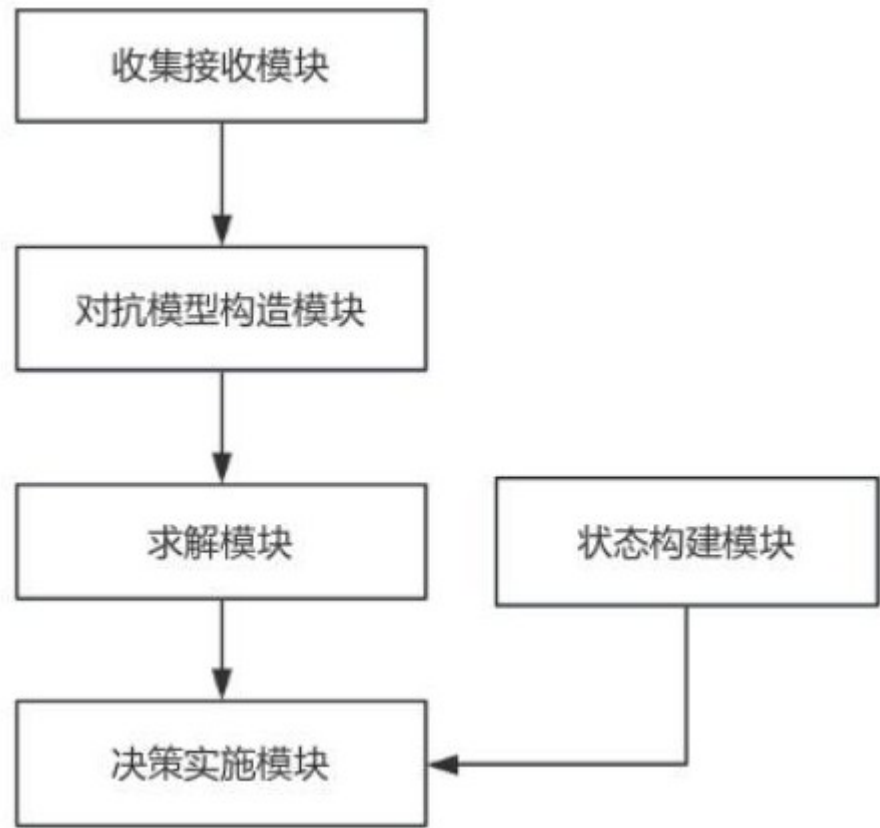


图 2

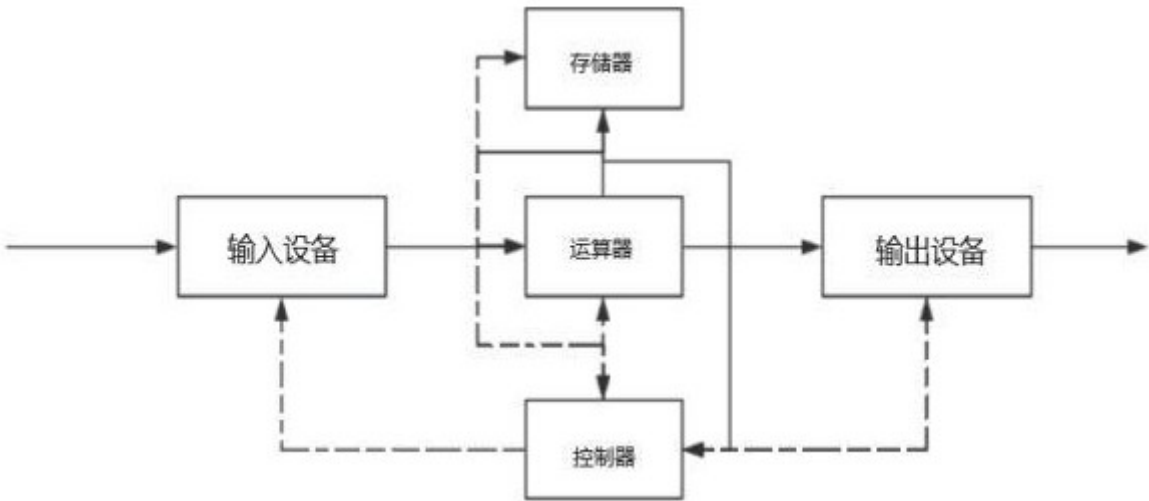


图 3

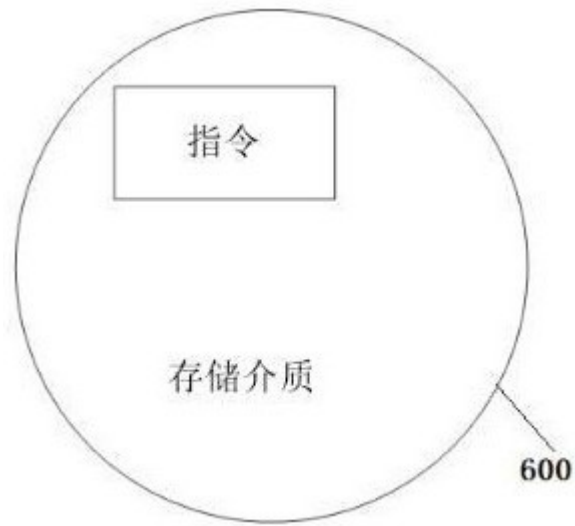


图 4