



(12)发明专利申请

(10)申请公布号 CN 111629380 A

(43)申请公布日 2020.09.04

(21)申请号 202010385842.9

H04W 16/22(2009.01)

(22)申请日 2020.05.09

G06N 3/08(2006.01)

(71)申请人 中国科学院沈阳自动化研究所

G06N 3/06(2006.01)

地址 110016 辽宁省沈阳市沈河区南塔街114号

G06N 3/04(2006.01)

申请人 沈阳智能机器人国家研究院有限公司

沈阳智能机器人创新中心有限公司

(72)发明人 于海斌 刘晓宇 许驰 曾鹏

金曦 夏长清

(74)专利代理机构 沈阳科苑专利商标代理有限公司 21002

代理人 王倩

(51)Int.Cl.

H04W 16/10(2009.01)

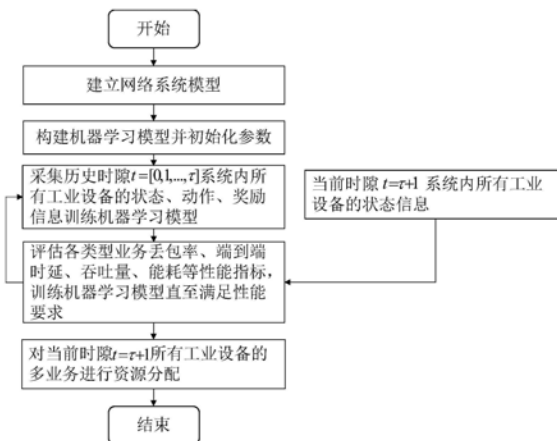
权利要求书3页 说明书7页 附图3页

(54)发明名称

面向高并发多业务工业5G网络的动态资源分配方法

(57)摘要

本发明涉及工业5G网络技术,具体地说,它是一种面向高并发多业务工业5G网络的动态资源分配方法。本发明包括以下步骤:建立网络系统模型,构建面向高并发多业务的工业5G网络动态资源分配的机器学习模型;采集工业5G网络内所有工业设备在不同时隙的状态、动作、奖励信息,训练机器学习模型;循环评估不同业务的网络性能指标,训练机器学习模型直至满足性能要求;以当前时隙工业5G网络内所有工业设备的状态信息作为机器学习模型输入,对所有工业设备多个不同类型并发业务进行资源分配。本发明解决了工业5G网络中大规模异构工业设备并发通信过程中控制命令、工业音视频、感知数据等多种类型业务对实时性、可靠性、吞吐量需求各异造成的资源冲突问题。



1. 面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,对于包括工业基站、计算节点和多个工业设备的工业5G网络,通过训练机器学习模型,对工业5G网络中所有工业设备多个不同类型并发业务进行资源分配,包括以下步骤:

1) 建立用于多工业设备并发多业务工业5G网络的动态资源分配的机器学习模型;

2) 采集工业5G网络内所有工业设备在不同时隙的状态、动作、奖励信息,训练机器学习模型;

3) 采集当前时隙工业5G网络内所有工业设备的状态信息作为机器学习模型输入,根据训练好的机器学习模型,对所有工业设备多个不同类型并发业务进行资源分配。

2. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,所述工业5G网络包括:1台工业基站,1个计算节点,以及N个工业设备;

所述计算节点,用于建立并训练机器学习模型;

所述工业基站,与计算节点通过网络相连接,从计算节点下载训练后的机器学习模型,用于调度不同工业设备及其传输业务的资源分配;

所述工业设备,包括工业控制类设备、工业音视频采集播放类设备、工业感知监测类设备,用于产生不同传输要求的工业业务数据,以保障工业生产的安全有效运行。

3. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,对于工业5G网络,建立工业5G网络模型,包括:工业5G网络覆盖范围及其范围内的工业设备数量N、资源块数量E、业务类型数量J、业务优先级P。

4. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,所述业务类型包括:工业控制类业务、工业音视频类业务、工业过程感知类业务。

5. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,所述建立用于多工业设备并发多业务工业5G网络的动态资源分配的机器学习模型,包括以下步骤:

建立评估网络和目标网络两个结构相同参数不同的深度神经网络模型,神经网络参数  $\text{params} = [L_{in}, L_{rnn}, L_{fc}, L_{out}, w, b]$ , 其中,  $L_{in}$  表示输入层神经元个数,  $L_{rnn}$  表示循环神经网络层神经元个数,  $L_{fc}$  表示全连接层神经元个数,  $L_{out}$  表示输出层神经元个数,  $w$  表示权重,  $b$  表示偏置;

其中,评估网络用于获得工业设备  $n$  ( $n \in N$ ) 当前状态向量  $s_n$  的动作向量  $a_n$  的估值函数  $Q(s_n, a_n)$ ; 目标网络用于选择工业设备  $n$  的下一个状态向量  $s'_n$  的最大动作向量  $a'_n$  的估值函数  $\max_{a'_n} Q(s'_n, a'_n)$ ;

利用强化学习的Q-learning更新方法  $Q(s_n, a_n) = Q(s_n, a_n) + \alpha (r_n + \gamma \max_{a'_n} Q(s'_n, a'_n) - Q(s_n, a_n))$  更新评估网络参数  $w, b$ ,  $\alpha$  表示学习速率,  $\gamma$  表示折扣比例,  $r_n$  表示工业设备  $n$  在当前状态  $s_n$  执行动作向量  $a_n$  获得的奖励;

目标网络的初始化参数与评估网络的初始化参数相同,评估网络的参数  $w, b$  每次迭代训练机器学习模型后进行更新,目标网络的参数  $w, b$  每迭代训练机器学习模型  $K$  次后更新一次。

6. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,所述工业设备的状态、动作、奖励信息包括:

工业设备 $n$  ( $n \in N$ ) 在时隙 $t$  ( $t \in T$ ) 的状态向量

$s_n(t) = [power_n^j(t), num_n^j(t), g_n^e(t), p_n^j(t), ack_n(t)]$ , 其中  $power_n^j(t)$  表示工业设备 $n$ 在时隙 $t$ 传输不同类型业务 $j$  ( $j \in J$ ) 所需的发送功率;  $num_n^j(t)$  表示工业设备 $n$ 在时隙 $t$ 准备发送的不同类型业务 $j$ 的数据包的数量;  $g_n^e(t)$  表示工业设备 $n$ 在时隙 $t$ 的各个资源块 $e$  ( $e \in E$ ) 上的增益;  $p_n^j(t)$  表示工业设备 $n$ 在时隙 $t$ 时不同类型业务 $j$ 的优先级 $p$  ( $p \in P$ );  $ack_n(t)$  表示工业设备 $n$ 在时隙 $t$ 结束业务是否发送成功;

工业设备 $n$ 在时隙 $t$ 的动作向量 $a_n(t) = [\overline{power}_n^j(t), \widetilde{num}_n^j(t), \widetilde{g}_n^e(t)]$ , 其中  $\widetilde{g}_n^e(t)$  是工业设备 $n$ 在时隙 $t$ 被分配的资源块 $e$ 的增益;  $\widetilde{num}_n^j(t)$  是工业设备 $n$ 在时隙 $t$ 选择传输的业务 $j$ 的数据包的数量;  $\overline{power}_n^j(t)$  是工业设备 $n$ 在时隙 $t$ 传输业务 $j$ 的发送功率;

工业设备 $n$ 在时隙 $t$ 的奖励向量 $r_n(t) = [r_n(t)]$ , 其中 $r_n(t)$  表示工业设备 $n$ 在时隙 $t$ 结束时获得的奖励。

7. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法, 其特征在于, 所述训练机器学习模型, 包括以下步骤:

a. 将工业设备 $n$ 时隙 $t$ 状态向量 $s_n(t) = [power_n^j(t), num_n^j(t), g_n^e(t), p_n^j(t), ack_n(t)]$  输入评估网络, 设定概率 $\epsilon$ , 根据 $\epsilon$ -greedy算法选择动作向量: 以概率 $\epsilon$ 随机选择动作向量, 或者以概率 $1-\epsilon$ 选择获得估值函数最大即  $\max_{a_n(t)} Q(s_n(t), a_n(t))$  的动作向量;

b. 根据动作向量 $a_n(t)$  计算获得的奖励 $r_n(t)$ ;

c. 根据工业设备 $n$ 时隙 $t$ 的状态向量 $s_n(t)$  和动作向量 $a_n(t)$ , 获得工业设备 $n$ 下一时隙 $t+1$ 的状态向量 $s'_n(t+1)$ , 将 $s'_n(t+1)$  输入目标网络获得  $\max_{a'_n(t+1)} Q(s'_n(t+1), a'_n(t+1))$  以及目标估值函数  $Q_{target} = r_n(t) + \gamma \max_{a'_n(t+1)} Q(s'_n(t+1), a'_n(t+1))$ , 存储  $\langle s_n(t), a_n(t), r_n(t), s'_n(t+1) \rangle$  作为经验池经验; 每个时隙的  $\langle s_n(t), a_n(t), r_n(t), s'_n(t+1) \rangle$  作为一个经验;

d. 从经验池中使用经验回放法随机抽取 $M$ 个经验, 计算每个经验的  $Q_{target}$  和  $\max_{a_n(t)} Q(s_n(t), a_n(t))$ , 并根据均方误差损失函数

$$Loss(\theta(t)) = \frac{\sum_{i=1}^M (Q_{target}(\theta(t)) - \max_{a_n(t)} Q(s_n(t), a_n(t), \theta(t)))^2}{M} \text{ 和梯度下降 } \theta_n(t+1) = \theta_n(t) - \eta \frac{\partial Loss(\theta(t))}{\partial (\theta(t))}$$

更新评估网络参数 $w, b$ , 其中 $\eta$ 表示神经网络学习速率,  $\theta(t)$  表示时隙 $t$ 的神经网络超参数; 重复迭代 $K$ 次评估网络后, 复制评估网络参数 $w, b$ 给目标网络参数 $w, b$ ;

e. 重复迭代a~d直至均方误差损失函数收敛, 此时得到的评估网络作为训练后的机器学习模型。

8. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法, 其特征在于, 得到训练的机器学习模型后, 对该机器学习模型进行优化:

采集工业5G网络内所有工业设备的历史时隙的状态信息, 通过该机器学习模型得到资源分配结果, 当该分配结果的网络性能符合要求时, 即对于不同类型业务, 丢包率、端到端

时延、吞吐量、能耗均小于对应业务的网络性能指标,则该机器学习模型作为最终的训练后的机器学习模型,用于进行最终的并发业务资源分配;

否则,重复权利要求1中的步骤1)-2),直到机器学习模型符合要求为止。

9. 根据权利要求8所述的面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,所述网络性能指标包括:

丢包率,为单位时间内,同一类型业务待发送数据包与已发送的数据包的差与待发送数据包的比值;

端到端时延,为同一类型业务的数据包传输所需的传播时延、传输时延、排队时延和硬件时延之和;其中,传播时延为电磁波从一个工业设备发送端到另一个工业设备接收端所经历的时延;传输时延为从数据包的第一个比特被发送到最后一个比特被发送所经历的时延;排队时延为数据包到达工业设备到离开工业设备所经历的时延;硬件时延为工业设备的硬件性能造成的时延;

吞吐量,为单位时间内同一类型业务成功传输的数据包的比特数;

能耗,为不同类型业务传输所需的能量消耗之和。

10. 根据权利要求1所述的面向高并发多业务工业5G网络的动态资源分配方法,其特征在于,所述步骤3包括:

收集当前时隙  $t$  工业 5 G 网络内的所有  $N$  个工业设备的状态向量

$\sum_{n=1}^N s_n(t) = [s_1(t), s_2(t), \dots, s_{N-1}(t), s_N(t)]$  输入至训练好的机器学习模型,得到输出动作向量

$\sum_{n=1}^N a_n(t) = [a_1(t), a_2(t), \dots, a_{N-1}(t), a_N(t)];$

根据得到的输出动作向量,工业基站为不同工业设备的不同业务分配资源。

## 面向高并发多业务工业5G网络的动态资源分配方法

### 技术领域

[0001] 本发明提供面向高并发多业务工业5G网络的动态资源分配方法,考虑工业5G网络中大规模异构工业设备并发通信过程中控制命令、工业音视频、感知数据等多种类型业务对实时性、可靠性、吞吐量需求各异造成的资源冲突问题,特别涉及多业务的丢包率、端到端时延、吞吐量和能耗约束,属于工业5G网络技术领域。

### 背景技术

[0002] 随着工业4.0的发展,大量的分布式工业设备互联互通,分布式工业设备间使用工业无线网络通信,产生了海量的具有不同传输要求的业务。工业5G网络以其高可靠低时延(Ultra-reliable and Low Latency Communications,URLLC)、移动增强带宽(Enhanced Mobile Broadband,eMBB)、大规模机器类型通信(massive Machine Type of Communication,mMTC)的典型通信场景应用于高并发多业务工业场景,既能保证工业设备的并发接入,又能满足各类业务不同的传输要求,成为工业无线网络的通信使能技术。

[0003] 将时域、频域传输资源组成资源块,通过对资源块的合理分配能够允许大规模工业设备的并发接入,保证不同业务的传输要求。然而对于大规模工业设备间通信的工业场景,设备数量和业务类型是时变的,传统的基于已知的系统模型的资源分配算法,难以获得准确的系统模型。工业设备的业务产生是时变的,不同类型的业务具有不同的传输需求,即URLLC业务需要传输的高实时高可靠,eMBB业务需要链路的高吞吐量,mMTC业务需要保证海量设备接入的稳定性。对于高并发多业务的大规模工业生产过程,这不仅导致准确的系统模型难以获得,还会造成状态空间爆炸的问题。利用机器学习求解高并发多业务的工业5G网络动态资源分配能够有效地解决建模难和状态空间爆炸问题。

### 发明内容

[0004] 为实现上述发明目的,本发明的目的在于提供面向高并发多业务工业5G网络的动态资源分配方法,能够实现特定丢包率、端到端时延、吞吐量和能耗约束下的资源分配。

[0005] 本发明采用如下技术方案:面向高并发多业务工业5G网络的动态资源分配方法,对于包括工业基站、计算节点和多个工业设备的工业5G网络,通过训练机器学习模型,对工业5G网络中所有工业设备多个不同类型并发业务进行资源分配,包括以下步骤:

[0006] 1) 建立用于多工业设备并发多业务工业5G网络的动态资源分配的机器学习模型;

[0007] 2) 采集工业5G网络内所有工业设备在不同时隙的状态、动作、奖励信息,训练机器学习模型;

[0008] 3) 采集当前时隙工业5G网络内所有工业设备的状态信息作为机器学习模型输入,根据训练好的机器学习模型,对所有工业设备多个不同类型并发业务进行资源分配。

[0009] 所述工业5G网络包括:1台工业基站,1个计算节点,以及N个工业设备;

[0010] 所述计算节点,用于建立并训练机器学习模型;

[0011] 所述工业基站,与计算节点通过网络相连接,从计算节点下载训练后的机器学习

模型,用于调度不同工业设备及其传输业务的资源分配;

[0012] 所述工业设备,包括工业控制类设备、工业音视频采集播放类设备、工业感知监测类设备,用于产生不同传输要求的工业业务数据,以保障工业生产的安全有效运行。

[0013] 对于工业5G网络,建立工业5G网络模型,包括:工业5G网络覆盖范围及其范围内的工业设备数量N、资源块数量E、业务类型数量J、业务优先级P。

[0014] 所述业务类型包括:工业控制类业务、工业音视频类业务、工业过程感知类业务。

[0015] 所述建立用于多工业设备并发多业务工业5G网络的动态资源分配的机器学习模型,包括以下步骤:

[0016] 建立评估网络和目标网络两个结构相同参数不同的深度神经网络模型,神经网络参数 $params = [L_{in}, L_{rnn}, L_{fc}, L_{out}, w, b]$ ,其中, $L_{in}$ 表示输入层神经元个数, $L_{rnn}$ 表示循环神经网络层神经元个数, $L_{fc}$ 表示全连接层神经元个数, $L_{out}$ 表示输出层神经元个数, $w$ 表示权重, $b$ 表示偏置;

[0017] 其中,评估网络用于获得工业设备 $n (n \in N)$ 当前状态向量 $s_n$ 的动作向量 $a_n$ 的估值函数 $Q(s_n, a_n)$ ;目标网络用于选择工业设备 $n$ 的下一个状态向量 $s'_n$ 的最大动作向量 $a'_n$ 的估值函数 $\max_{a'_n} Q(s'_n, a'_n)$ ;

[0018] 利用强化学习的Q-learning更新方法

$Q(s_n, a_n) = Q(s_n, a_n) + \alpha (r_n + \gamma \max_{a'_n} Q(s'_n, a'_n) - Q(s_n, a_n))$ 更新评估网络参数 $w, b$ , $\alpha$ 表示学习速率, $\gamma$

表示折扣比例, $r_n$ 表示工业设备 $n$ 在当前状态 $s_n$ 执行动作向量 $a_n$ 获得的奖励;

[0019] 目标网络的初始化参数与评估网络的初始化参数相同,评估网络的参数 $w, b$ 每次迭代训练机器学习模型后进行更新,目标网络的参数 $w, b$ 每迭代训练机器学习模型K次后更新一次。

[0020] 所述工业设备的状态、动作、奖励信息包括:

[0021] 工业设备 $n (n \in N)$ 在时隙 $t (t \in T)$ 的状态向量

$s_n(t) = [power_n^j(t), num_n^j(t), g_n^e(t), p_n^j(t), ack_n(t)]$ ,其中 $power_n^j(t)$ 表示工业设备 $n$ 在时隙 $t$ 传输不同类型业务 $j (j \in J)$ 所需的发送功率; $num_n^j(t)$ 表示工业设备 $n$ 在时隙 $t$ 准备发送的不同类型业务 $j$ 的数据包的数量; $g_n^e(t)$ 表示工业设备 $n$ 在时隙 $t$ 的各个资源块 $e (e \in E)$ 上的增益; $p_n^j(t)$ 表示工业设备 $n$ 在时隙 $t$ 时不同类型业务 $j$ 的优先级 $p (p \in P)$ ; $ack_n(t)$ 表示工业设备 $n$ 在时隙 $t$ 结束业务是否发送成功;

[0022] 工业设备 $n$ 在时隙 $t$ 的动作向量 $a_n(t) = [\widetilde{power}_n^j(t), \widetilde{num}_n^j(t), \widetilde{g}_n^e(t)]$ ,其中 $\widetilde{g}_n^e(t)$ 是工业设备 $n$ 在时隙 $t$ 被分配的资源块 $e$ 的增益; $\widetilde{num}_n^j(t)$ 是工业设备 $n$ 在时隙 $t$ 选择传输的业务 $j$ 的数据包的数量; $\widetilde{power}_n^j(t)$ 是工业设备 $n$ 在时隙 $t$ 传输业务 $j$ 的发送功率;

[0023] 工业设备 $n$ 在时隙 $t$ 的奖励向量 $r_n(t) = [r_n(t)]$ ,其中 $r_n(t)$ 表示工业设备 $n$ 在时隙 $t$ 结束时获得的奖励。

[0024] 所述训练机器学习模型,包括以下步骤:

[0025] a. 将工业设备 $n$ 时隙 $t$ 状态向量 $s_n(t) = [power_n^j(t), num_n^j(t), g_n^e(t), p_n^j(t), ack_n(t)]$ 输入评

估网络,设定概率 $\varepsilon$ ,根据 $\varepsilon$ -greedy算法选择动作向量,以概率 $\varepsilon$ 随机选择动作向量,或者以概率 $1-\varepsilon$ 选择获得估值函数最大(即 $\max_{a_n(t)} Q(s_n(t), a_n(t))$ )的动作向量;

[0026] b. 根据动作向量 $a_n(t)$ 计算获得的奖励 $r_n(t)$ ;

[0027] c. 根据工业设备 $n$ 时隙 $t$ 的状态向量 $s_n(t)$ 和动作向量 $a_n(t)$ ,获得工业设备 $n$ 下一时隙 $t+1$ 的状态向量 $s'_n(t+1)$ ,将 $s'_n(t+1)$ 输入目标网络获得 $\max_{a'_n(t+1)} Q(s'_n(t+1), a'_n(t+1))$ 以及目标估值函数 $Q_{target} = r_n(t) + \gamma \max_{a'_n(t+1)} Q(s'_n(t+1), a'_n(t+1))$ ,存储 $\langle s_n(t), a_n(t), r_n(t), s'_n(t+1) \rangle$ 作为经验池经验;每个时隙的 $\langle s_n(t), a_n(t), r_n(t), s'_n(t+1) \rangle$ 作为一个经验;

[0028] d. 从经验池中使用经验回放法随机抽取 $M$ 个经验,计算每个经验的 $Q_{target}$ 和 $\max_{a_n(t)} Q(s_n(t), a_n(t))$ ,并根据均方误差损失函数

$$Loss(\theta(t)) = \frac{\sum_{i=1}^M (Q_{target}(\theta(t)) - \max_{a_n(t)} Q(s_n(t), a_n(t), \theta(t)))^2}{M} \text{ 和梯度下降 } \theta_n(t+1) = \theta_n(t) - \eta \frac{\partial Loss(\theta(t))}{\partial (\theta(t))}$$

更新评估网络参数 $w, b$ ,其中 $\eta$ 表示神经网络学习速率, $\theta(t)$ 表示时隙 $t$ 的神经网络超参数;重复迭代 $K$ 次评估网络后,复制评估网络参数 $w, b$ 给目标网络参数 $w, b$ ;

[0029] e. 重复迭代a~d直至均方误差损失函数收敛,此时得到的评估网络作为训练后的机器学习模型。

[0030] 得到训练的机器学习模型后,对该机器学习模型进行优化:

[0031] 采集工业5G网络内所有工业设备的历史时隙的状态信息,通过该机器学习模型得到资源分配结果,当该分配结果的网络性能符合要求时,即对于不同类型业务,丢包率、端到端时延、吞吐量、能耗均小于对应业务的网络性能指标,则该机器学习模型作为最终的训练后的机器学习模型,用于进行最终的并发业务资源分配;

[0032] 否则,重复步骤1)-2),直到机器学习模型符合要求为止。

[0033] 所述网络性能指标包括:

[0034] 丢包率,为单位时间内,同一类型业务待发送数据包与已发送的数据包的差与待发送数据包的比值;

[0035] 端到端时延,为同一类型业务的数据包传输所需的传播时延、传输时延、排队时延和硬件时延之和;其中,传播时延为电磁波从一个工业设备发送端到另一个工业设备接收端所经历的时延;传输时延为从数据包的第一个比特被发送到最后一个比特被发送所经历的时延;排队时延为数据包到达工业设备到离开工业设备所经历的时延;硬件时延为工业设备的硬件性能造成的时延;

[0036] 吞吐量,为单位时间内同一类型业务成功传输的数据包的比特数;

[0037] 能耗,为不同类型业务传输所需的能量消耗之和。

[0038] 所述步骤3包括:

[0039] 收集当前时隙 $t$ 工业5G网络内的所有 $N$ 个工业设备的状态向量

$$\sum_{n=1}^N s_n(t) = [s_1(t), s_2(t), \dots, s_{N-1}(t), s_N(t)] \text{ 输入至训练好的机器学习模型,得到输出动作向量}$$

$$\sum_{n=1}^N a_n(t) = [a_1(t), a_2(t), \dots, a_{N-1}(t), a_N(t)];$$

[0040] 根据得到的输出动作向量,工业基站为不同工业设备的不同业务分配资源。

[0041] 本发明具有以下有益效果及优点:

[0042] 1、本发明面向含有工业控制、工业音视频、工业过程感知等多种类型业务的工业5G网络,针对其大规模分布式工业设备并发通信时资源分配建模难以及传统算法空间爆炸的问题,将多类型工业业务映射为高可靠低时延(URLLC)、增强移动带宽(eMBB)、大规模机器类型通信(mMTC),并使用基于机器学习方法建模、训练,能够有效地为不同工业设备的不同类型业务进行动态资源分配。

[0043] 2、本发明具有较强的通用性和实用性,能够自适应地处理工业设备、业务类型和传输资源的动态变化,能够有效地保证高并发多业务工业5G网络的动态资源分配,实现特定丢包率、端到端时延、吞吐量和能耗约束下的稳定传输,提高系统安全性和稳定性。

### 附图说明

[0044] 图1是本发明方法流程图;

[0045] 图2是系统模型图;

[0046] 图3是资源块示意图;

[0047] 图4是机器学习模型架构图。

### 具体实施方式

[0048] 下面结合附图对本发明进行详细说明。

[0049] 本发明主要包括以下实现过程,如图1,包括以下步骤:

[0050] 步骤1:建立网络系统模型,确定系统内工业设备数量、业务类型、业务数量、业务优先级以及网络资源块数量;

[0051] 步骤2:构建高并发多业务工业5G网络的动态资源分配的机器学习模型,并初始化参数;

[0052] 步骤3:采集工业5G网络内所有工业设备在不同时隙的状态、动作、奖励信息,训练机器学习模型;

[0053] 步骤4:循环评估不同业务的丢包率、端到端时延、吞吐量、能耗等指标,训练机器学习模型直至满足性能要求;

[0054] 步骤5:以当前时隙工业5G网络内所有工业设备的状态信息作为机器学习模型输入,对所有工业设备多个不同类型并发业务进行资源分配。该实施例是按照如图1所示的流程实施的,具体步骤如下:

[0055] 1、建立网络系统模型,确定系统内工业设备数量、业务类型、业务数量、业务优先级以及网络资源块数量,如图2,包括:

[0056] (1)工业5G网络包括:1台工业基站,1个计算节点,以及N个工业设备;其中,计算节点可以是移动边缘计算节点、雾计算节点或云计算节点,计算节点用于训练机器学习模型;工业基站与计算节点通过网络相连接,从计算节点下载更新训练的机器学习模型,用于调度不同工业设备及其传输业务的资源分配;工业设备,包括工业控制类设备、工业音视频采



集播放类设备、工业感知监测类设备,用于产生不同传输要求的工业业务数据,以保障工业生产的安全有效运行;

[0057] (2) 确定工业5G网络覆盖范围及其范围内工业设备数量N、资源块数量E、业务优先级P、业务数量J,系统模型主要包括两类情况:工业设备数量N小于资源块数量E,工业设备数量N大于等于资源块数量E:

[0058] a. 当工业设备数量N小于资源块数量E时,所有工业设备满足信道容量约束的业务都能分配到资源块并满足性能要求;

[0059] b. 当工业设备数量N大于等于资源块数量E时,根据本发明的资源分配方法,只有部分工业设备满足信道容量约束的业务能够分配到资源块并满足性能要求;

[0060] (3) 工业5G网络中在线工业设备数量动态变化,其变化范围为 $[0, N]$ ,资源块是传输的时域、频域所构成的资源块,如图3,业务优先级在优先级 $[0, P]$ 范围内动态变化的, $p$  ( $p \in P$ ) 值越小优先级越高,业务传输所需实时性、可靠性越高,业务类型数量在 $[0, J]$ 范围内动态变化;

[0061] (4) 业务类型可划分为:工业控制类业务、工业音视频类业务、工业过程感知类业务其中,工业控制类业务需要严格的实时性、可靠性传输要求,用于传输工业控制数据,映射为高可靠低时延 (URLLC) 业务;工业音视频类业务要求在截止日期前能够送达,用于传输工业音视频数据,映射为移动增强带宽 (eMBB) 业务;工业过程感知类业务对传输的实时性、可靠性要求最低,需要维持大规模的并发接入,用于传输工业过程感知监测数据,映射为大规模机器接入 (mMTC) 业务。

[0062] 2、构建高并发多业务工业5G网络的动态资源分配的机器学习模型,并初始化参数,如图4,包括:

[0063] (1) 建立基于深度强化学习的多工业设备多业务的工业5G网络动态资源分配机器学习模型,使用深度神经网络估计网络系统模型,使用强化学习求解动态资源分配;

[0064] (2) 构建评估网络和目标网络两个网络结构相同参数不同的深度神经网络,深度神经网络参数 $params = [L_{in}, L_{rnn}, L_{fc}, L_{out}, w, b]$ ,其中, $L_{in}$ 表示输入层神经元个数, $L_{rnn}$ 表示循环神经网络层神经元个数, $L_{fc}$ 表示全连接层神经元个数, $L_{out}$ 表示输出层神经元个数, $w$ 表示权重, $b$ 表示偏置。

[0065] 其中评估网络用于获得工业设备 $n$  ( $n \in N$ ) 当前状态向量 $s_n$ 的动作向量 $a_n$ 的估值函数 $Q(s_n, a_n)$ ;目标网络通过选择工业设备 $n$  ( $n \in N$ ) 的下一个状态向量 $s'_n$ 的最大动作 $a'_n$ 的估值函数 $\max_{a'_n} Q(s'_n, a'_n)$ ,利用强化学习中的Q-learning更新方法

$Q(s_n, a_n) = Q(s_n, a_n) + \alpha (r_n + \gamma \max_{a'_n} Q(s'_n, a'_n) - Q(s_n, a_n))$ 更新评估网络参数, $\alpha$ 表示学习速率, $\gamma$ 表示折扣比例, $r_n$ 表示工业设备 $n$ 在当前状态 $s_n$ 执行动作向量 $a_n$ 获得的奖励。

[0066] (3) 目标网络的初始化参数与评估网络的初始化参数相同,评估网络的参数 $w$ 、 $b$ 每次迭代训练机器学习模型后进行更新,目标网络的参数 $w$ 、 $b$ 每迭代训练机器学习模型K次后更新一次。

[0067] 3、采集工业5G网络内所有工业设备在不同时隙的状态、动作、奖励信息,训练机器学习模型,其中,工业设备的状态信息、动作信息、奖励信息,包括:

[0068] (1) 工业设备 $n$  ( $n \in N$ ) 在时隙 $t$  ( $t \in T$ ) 的状态向量

$s_n(t)=[power_n^j(t), num_n^j(t), g_n^e(t), p_n^j(t), ack_n(t)]$ , 其中  $power_n^j(t)$  表示工业设备n在时隙t时传输不同类型业务j所需的发送功率, 是大小为J的向量, 即

$power_n^j(t) \in [power_n^1(t), power_n^2(t), \dots, power_n^J(t)]$ , 需要满足能耗约束;  $num_n^j(t)$  表示工业设备n在时隙t准备发送的不同类型业务j的数据包的数量, 是大小为J的向量, 即  $num_n^j(t) \in [num_n^1(t), num_n^2(t), \dots, num_n^J(t)]$ , 要满足吞吐量约束;  $g_n^e(t)$  表示工业设备n在时隙t的各个资源块e的增益, 是大小为E的向量, 即  $g_n^e(t) \in [g_n^1(t), g_n^2(t), \dots, g_n^E(t)]$ ;  $p_n^j(t)$  表示工业设备n在时隙t时不同类型业务j的优先级  $p (p \in P)$ , 是动态变化的;  $ack_n(t)$  表示工业设备n在时隙t结束业务是否发送成功, 如果  $ack_n(t) = 0$ , 表示工业设备n在时隙t发送业务失败; 如果  $ack_n(t) = 1$ , 表示工业设备n在时隙t发送业务成功;

[0069] (2) 工业设备n在时隙t的动作向量  $a_n(t) = [\widetilde{power}_n^j(t), \widetilde{num}_n^j(t), \widetilde{g}_n^e(t)]$ , 其中  $\widetilde{g}_n^e(t)$  是工业设备n在时隙t被分配的资源块e的增益;  $\widetilde{num}_n^j(t)$  是工业设备n在时隙t选择传输的业务j的数据包的数量;  $\widetilde{power}_n^j(t)$  是工业设备n在时隙t传输业务j的发送功率;

[0070] (3) 工业设备n在时隙t的奖励向量  $r_n(t) = [r_n(t)]$ , 其中  $r_n(t)$  表示工业设备n在时隙t结束获得的奖励, 奖励值与业务传输是否成功、当前时隙的数据包的业务类型及业务优先级相关。如果工业设备n在时隙t发送业务失败, 工业设备n获得的奖励为负值; 如果工业设备n在时隙t发送业务成功, 发送业务的优先级和实时性、可靠性传输要求越高, 工业设备n获得的奖励越高, 为正值。

[0071] 4、采集工业5G网络内工业设备在不同时隙的状态、动作、奖励信息, 训练机器学习模型, 其中, 训练过程包括:

[0072] a. 将工业设备n时隙t状态向量  $s_n(t) = [power_n^j(t), num_n^j(t), g_n^e(t), p_n^j(t), ack_n(t)]$  输入评估网络, 设定概率  $\epsilon$ , 根据  $\epsilon$ -greedy 算法选择动作向量, 以概率  $\epsilon$  随机选择动作向量, 或者以概率  $1-\epsilon$  选择获得估值函数最大 (即  $\max_{a_n(t)} Q(s_n(t), a_n(t))$ ) 的动作向量;

[0073] b. 根据动作向量  $a_n(t)$  计算获得的奖励  $r_n(t)$ ;

[0074] c. 根据工业设备n在时隙t的状态向量  $s_n(t)$  和动作向量  $a_n(t)$ , 获得工业设备n下一时隙t+1的状态向量  $s'_n(t+1)$ , 将  $s'_n(t+1)$  输入目标网络获得  $\max_{a'_n(t+1)} Q(s'_n(t+1), a'_n(t+1))$  以及目标估值函数  $Q_{target} = r_n(t) + \gamma \max_{a'_n(t+1)} Q(s'_n(t+1), a'_n(t+1))$ , 存储  $\langle s_n(t), a_n(t), r_n(t), s'_n(t+1) \rangle$  作为经验池经验; 每个时隙的  $\langle s_n(t), a_n(t), r_n(t), s'_n(t+1) \rangle$  作为一个经验;

[0075] d. 从经验池中使用经验回放法随机抽取M个经验, 计算每个经验的  $Q_{target}$  和  $\max_{a_n(t)} Q(s_n(t), a_n(t))$ , 并根据均方误差损失函数

$$Loss(\theta(t)) = \frac{\sum_{i=1}^M (Q_{target}(\theta(t)) - \max_{a_n(t)} Q(s_n(t), a_n(t), \theta(t)))^2}{M} \text{ 和梯度下降 } \theta_n(t+1) = \theta_n(t) - \eta \frac{\partial Loss(\theta(t))}{\partial(\theta(t))}$$

更新评估网络参数, 其中  $\eta$  表示神经网络学习速率,  $\theta(t)$  表示时隙t的神经网络超参数。重复

迭代K次评估网络后,复制评估网络参数给目标网络;

[0076] e.重复迭代a~d直至均方差损失函数收敛。

[0077] 5、循环评估不同业务的丢包率、端到端时延、吞吐量、能耗等指标,训练神经网络模型直至满足性能要求包括:

[0078] (1) 丢包率定义为单位时间内,同一类型业务待发送数据包与已发送的数据包的差与待发送数据包之比,用于衡量业务传输可靠性,丢包率越低,传输可靠性越高,反之传输可靠性越低;

[0079] (2) 端到端时延定义为同一类型业务的传输数据包所需的传播时延、传输时延、排队时延和硬件时延之和,用于衡量业务传输实时性。其中传播时延为电磁波从一个工业设备发送端到另一个工业设备接收端所经历的时延,与传输介质相关;传输时延为从数据包的第一个比特被发送到最后一个比特被发送所经历的时延,与链路带宽相关;排队时延为数据包到达工业设备到离开工业设备所经历的时延,与资源分配算法相关;硬件时延为工业设备的硬件性能造成的时延,与硬件性能相关;

[0080] (3) 吞吐量定义为单位时间内同一类型业务的成功传输的数据包的比特数,用于衡量资源块能够提供的最大传输业务量,对于要求传输速率高的业务,高吞吐量能够保证业务的有效传输;

[0081] (4) 能耗定义为不同类型业务传输所需的能量消耗之和,当工业设备分配到资源块发送业务,不同类型业务传输所需能耗不同,用于衡量能耗约束;

[0082] (5) URLLC业务要求高可靠低时延传输,需要满足严格的端到端时延和丢包率约束;eMBB业务要求传输的高带宽,需要高吞吐量传输保证;mMTC业务要求海量设备稳定接入,需要限制设备能耗减少对URLLC和eMBB业务的干扰;

[0083] (6) 判断丢包率、端到端时延、吞吐量、能耗,是否满足具体系统模型下不同类型业务的性能要求,如果满足,模型训练完成,否则,继续训练模型直至满足性能要求。

[0084] 6、以当前时隙工业5G网络内所有工业设备的状态信息作为神经网络模型输入,对所有工业设备多个不同类型并发业务进行资源分配包括:

[0085] (1) 收集当前时隙 $t$ 工业5G网络内的所有 $N$ 个工业设备的状态向量

$\sum_{n=1}^N s_n(t) = [s_1(t), s_2(t), \dots, s_{N-1}(t), s_N(t)]$  输入至训练好的机器学习模型,得到输出动作向量

$\sum_{n=1}^N a_n(t) = [a_1(t), a_2(t), \dots, a_{N-1}(t), a_N(t)];$

[0086] (2) 根据得到的输出动作向量,工业基站集中为不同工业设备的不同业务分配资源。



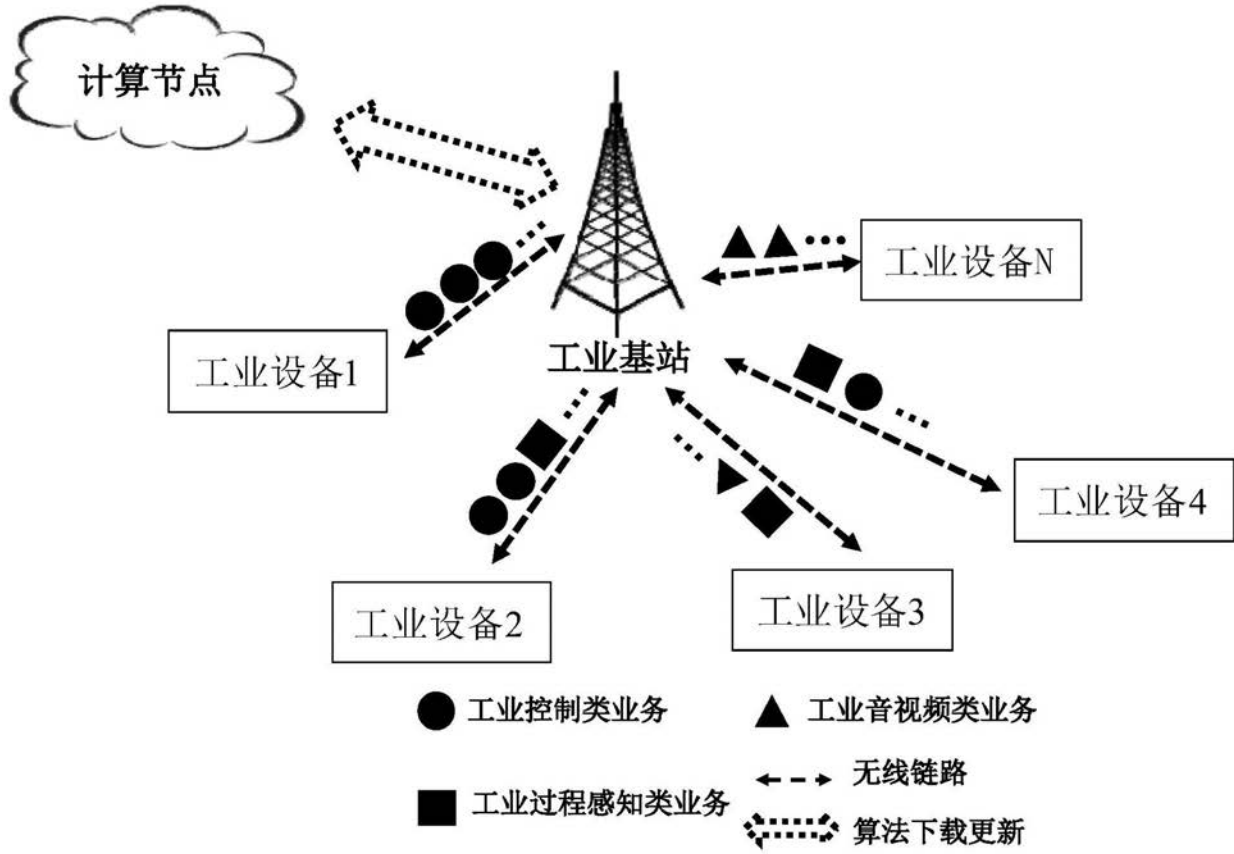


图2

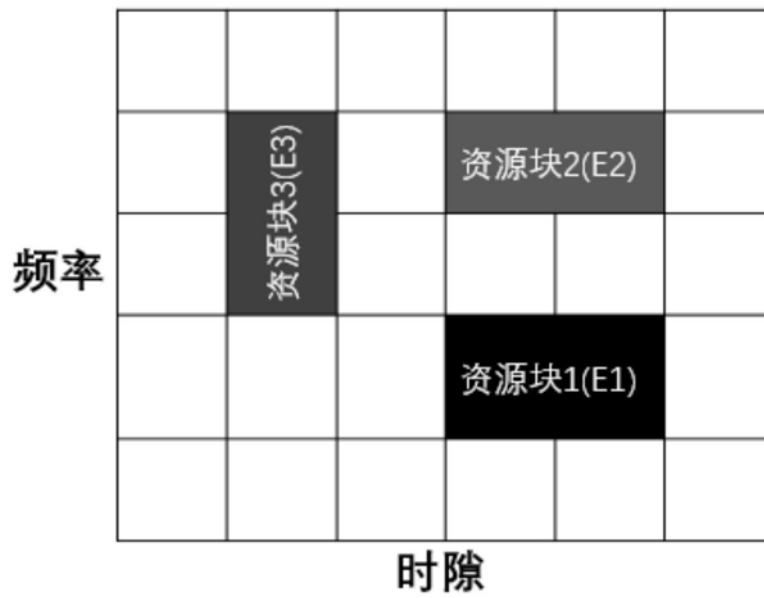


图3

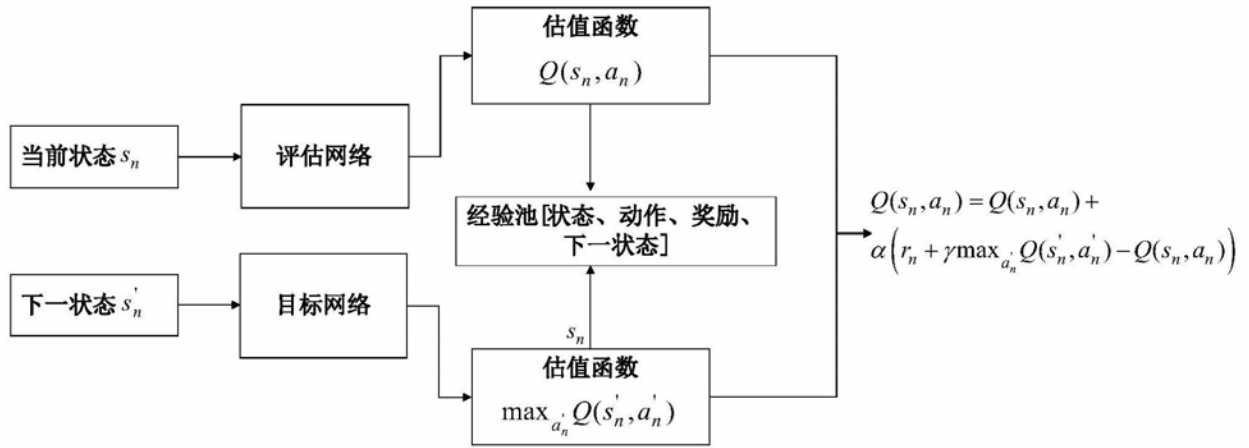


图4