(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 112202848 B
(45)授权公告日 2021.11.30

(21)申请号 202010968137.1

(22)申请日 2020.09.15

(65)同一申请的已公布的文献号
申请公布号 CN 112202848 A

(43)申请公布日 2021.01.08

(73)专利权人 中国科学院计算技术研究所
地址 100080 北京市海淀区中关村科学院南路6号

(72)发明人 刘建敏 王琪 徐勇军 何晨涛 徐亦达

(74)专利代理机构 北京律诚同业知识产权代理有限公司 11006
代理人 祁建国

(51)Int.Cl.
H04L 29/08(2006.01)
H04L 12/721(2013.01)
H04L 12/727(2013.01)
H04L 12/729(2013.01)

(56)对比文件
CN 111343608 A,2020.06.26
CN 111065105 A,2020.04.24
CN 109726866 A,2019.05.07
CN 110906935 A,2020.03.24
CN 111432433 A,2020.07.17
US 2018374356 A1,2018.12.27
US 10691127 B2,2020.06.23
陈思宇."基于智能推理的移动边缘计算资源分配方法研究".《中国优秀硕士学位论文全文数据库(电子期刊)信息科技辑》.2020,
Xianglong Zhou;Yun Lin."Dynamic Channel Allocation for Multi-UAVs: A Deep Reinforcement Learning Approach".《2019 IEEE Global Communications Conference (GLOBECOM)》.2020,

审查员 曹洪菠

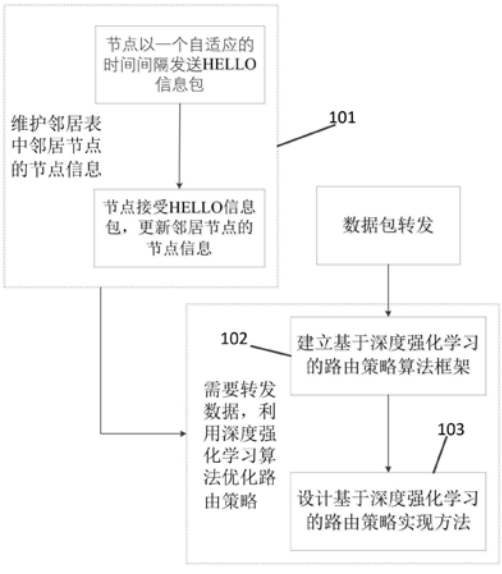
权利要求书3页 说明书11页 附图5页

(54)发明名称

基于深度强化学习的无人系统网络自适应路由方法和系统

(57)摘要

本发明提出一种基于深度强化学习的无人系统网络自适应路由方法,旨在解决现有技术中节点的高速移动、频繁变化的网络拓扑,无法提供自适应路由策略的技术问题。所述方法包括:所有节点以一个自适应的时间间隔发送HELLO信息包;任一节点收到其邻居节点发送的HELLO信息包后,更新该节点的邻居表中该邻居节点的节点信息;建立基于深度强化学习的路由策略算法框架;设计基于深度强化学习的路由策略实现方法。本发明具备良好的模型泛化能力,能泛化于具有不同网络规模和不同节点移动速度的网络上,使得本发明更适用于具有动态变化的无人系统网络。



CN 112202848 B

1.一种基于深度强化学习的无人系统网络自适应路由方法,其特征在于,包括:

步骤1、以无人系统网络中的每一个无人装置作为节点,所有节点以一个自适应的时间间隔发送HELLO信息包;任一节点收到其邻居节点发送的HELLO信息包后,更新该节点的邻居表中该邻居节点的节点信息;

步骤2、将该无人系统网络中所有节点以及由所有节点形成所有链路作为系统环境,该无人系统网络中每个节点从系统环境中获取当前时刻的环境状态,并执行行为作用于系统环境,系统环境根据该执行行为反馈给节点奖励值,其中该环境状态包括当前节点和当前节点的所有邻居节点的链路状态;

步骤3、无人系统网络中节点i根据其环境状态,利用深度Q网络(Deep Q-learning network,DQN)计算当前节点所有邻居节点的Q值,当前节点执行一个行为a_t,以最大Q值的邻居节点作为下一跳节点进行数据包的路由;

该步骤2包括:

在当前时刻t下,节点i所观察到的环境状态s_t为:

$s_t=\{C_{i,1},...,C_{i,j},...,C_{i,M}\}$,其中$C_{i,j}$是由该节点i和该节点i的邻居j所形成的链路$l_{i,j}$的状态,M为该节点i拥有的邻居节点数量;

基于该节点i的邻居表中该邻居节点j的信息,计算$C_{i,j}$:

$C_{i,j}=\{ct_{i,j},PER_{i,j},e_j,d_{j,des},d_{min}\}$,$ct_{i,j}$是链路$l_{i,j}$的期望连接时间,$PER_{i,j}$是链路$l_{i,j}$的包的错误率,$e_j$是该节点i的邻居节点j的剩余电量,$d_{j,des}$是该节点i的邻居节点j与该目标节点des间的距离,d_{min}是该节点i的2跳邻居节点k与该目标节点des的最小距离;

节点通过选择一个优化的邻居节点作为下一跳节点来完成行为a_t;

系统环境给予节点的奖励值r_t为:

当该节点i的邻居节点j是该目标节点des,$r_t=R_{max}$,R_{max}是预设最大奖励值;

当该节点i的所有邻居节点与该目标节点des的距离均大于该节点i与该目标节点des的距离,$r_t=-R_{max}$;

否则,$r_t=RD_{i,j}$,$RD_{i,j}=\frac{d_{i,des}}{d_{j,des}}(1-PER_{i,j})$。

2.如权利要求1所述的基于深度强化学习的无人系统网络自适应路由方法,其特征在于,该步骤1包括:所有节点以一个自适应的时间间隔发送HELLO信息包,其中自适应的时间间隔方法如下:

$$T_i=\min\{T_{\min}*\frac{v_{\max}}{v_{avg}^i},T_{\max}\}$$

其中,T_{min}和T_{max}分别是预设最短和最长时间间隔,v_{max}是节点i预设置的最大移动速度,v_{avg}^i为该节点i的平均速度。

3.如权利要求1所述的基于深度强化学习的无人系统网络自适应路由方法,其特征在于,该步骤3包括:

收集节点i与环境交互的经验(s_t,a_t,r_t,s_{t+1}),并将该经验存储到经验回放存储器;从该经验回放存储器中随机采样部分经验以及最小化预先设置的损失函数,更新该深度Q网

络的参数，该损失函数：$L(\theta)=\sum_{(s_t,a_t,r_t,s_{t+1})}(Q_{target}-q(s_t,a_t;\theta_t))$，其中$Q_{target}=r_t+\gamma\max_{a'}q(s_{t+1},a';\theta_t)$，θ表示所述DQN的网络参数，q($s_t$,$a_t$;$\theta_t$)表示将环境状态$s_t$输入所述DQN后，输出在该环境状态$s_t$下选择行为$a_t$后获得累积奖励值，$\theta_t$为t时刻DQN的网络参数，a′表示在环境状态$s_{t+1}$下节点所采取的行为，$\max_{a'}q(s_{t+1},a';\theta_t)$表示在环境状态$s_{t+1}$下的最优累积奖励值，γ表示折扣因子，0≤γ≤1；

一旦该深度Q网络的参数被更新，将更新后的参数发送给该无人系统网络中每个节点。

4.如权利要求1或2所述的基于深度强化学习的无人系统网络自适应路由方法，其特征在于，该邻居表中邻居节点的节点信息包括：邻居节点的移动速度、位置坐标和剩余的电量。

5.一种基于深度强化学习的无人系统网络自适应路由系统，其特征在于，包括：

以无人系统网络中的每一个无人装置作为节点，所有节点以一个自适应的时间间隔发送HELLO信息包；任一节点收到其邻居节点发送的HELLO信息包后，更新该节点的邻居表中该邻居节点的节点信息；

将该无人系统网络中所有节点以及由所有节点形成所有链路作为系统环境，该无人系统网络中每个节点从系统环境中获取当前时刻的环境状态，并执行行为作用于系统环境，系统环境根据该执行行为反馈给节点奖励值，其中该环境状态包括当前节点和当前节点的所有邻居节点的链路状态；

无人系统网络中节点i根据其环境状态，利用深度Q网络(Deep Q-learning network，DQN)计算当前节点所有邻居节点的Q值，当前节点执行一个行为a_t，以最大Q值的邻居节点作为下一跳节点进行数据包的路由；

其中，在当前时刻t下，节点i所观察到的环境状态s_t为：

$s_t=\{C_{i,1},\ldots,C_{i,j},\ldots,C_{i,M}\}$，其中$C_{i,j}$是由该节点i和该节点i的邻居j所形成的链路$l_{i,j}$的状态，M为该节点i拥有的邻居节点数量；

基于该节点i的邻居表中该邻居节点j的信息，计算$C_{i,j}$：

$C_{i,j}=\{ct_{i,j},PER_{i,j},e_j,d_{j,des},d_{min}\}$，$ct_{i,j}$是链路$l_{i,j}$的期望连接时间，$PER_{i,j}$是链路$l_{i,j}$的包的错误率，$e_j$是该节点i的邻居节点j的剩余电量，$d_{j,des}$是该节点i的邻居节点j与该目标节点des间的距离，d_{min}是该节点i的2跳邻居节点k与该目标节点des的最小距离；

节点通过选择一个优化的邻居节点作为下一跳节点来完成行为a_t；

系统环境给予节点的奖励值r_t为：

当该节点i的邻居节点j是该目标节点des，$r_t=R_{max}$，R_{max}是预设最大奖励值；

当该节点i的所有邻居节点与该目标节点des的距离均大于该节点i与该目标节点des的距离，$r_t=-R_{max}$；

否则，$r_t=RD_{i,j}$，$RD_{i,j}=\frac{d_{i,des}}{d_{j,des}}(1-PER_{i,j})$。

6.如权利要求5所述的基于深度强化学习的无人系统网络自适应路由系统，其特征在于，所有节点以一个自适应的时间间隔发送HELLO信息包，其中自适应的时间间隔系统如下：

$$T_i = \min\{T_{\min} * \frac{v_{\max}}{v_{avg}^i}, T_{\max}\}$$

其中，T_{min}和T_{max}分别是预设最短和最长时间间隔，v_{max}是节点i预设置的最大移动速度，v_{avg}^i为该节点i的平均速度。

7.如权利要求5所述的基于深度强化学习的无人系统网络自适应路由系统，其特征在于，具体包括：

收集节点i与环境交互的经验(s_t, a_t, r_t, s_{t+1})，并将该经验存储到经验回放存储器；从该经验回放存储器中随机采样部分经验以及最小化预先设置的损失函数，更新该深度Q网络的参数，该损失函数：$L(\theta) = \sum_{(s_t,a_t,r_t,s_{t+1})}(Q_{target} - q(s_t, a_t; \theta_t))$，其中$Q_{target} = r_i + \gamma \max_{a'} q(s_{t+1}, a'; \theta_t)$，θ表示所述DQN的网络参数，$q(s_t, a_t; \theta_t)$表示将环境状态$s_t$输入所述DQN后，输出在该环境状态$s_t$下选择行为$a_t$后获得累积奖励值，$\theta_t$为t时刻DQN的网络参数，a′表示在环境状态$s_{t+1}$下节点所采取的行为，$\max_{a'} q(s_{t+1}, a'; \theta_t)$表示在环境状态$s_{t+1}$下的最优累积奖励值，γ表示折扣因子，$0 \leqslant \gamma \leqslant 1$；

一旦该深度Q网络的参数被更新，将更新后的参数发送给该无人系统网络中每个节点。

8.如权利要求5或6所述的基于深度强化学习的无人系统网络自适应路由系统，其特征在于，该邻居表中邻居节点的节点信息包括：邻居节点的移动速度、位置坐标和剩余的电量。

基于深度强化学习的无人系统网络自适应路由方法和系统

技术领域

[0001] 本发明涉及一种基于深度强化学习的无人系统网络自适应路由方法，属于信息技术领域。

背景技术

[0002] 无人系统(Unmanned System)是由若干必要的数据处理单元、传感器、自动控制单元、通信系统组成，无需人为介入即可自主完成特定任务的机器或装置，这些无人机器或装置可以是无人机、无人车、地面机器人、水下机器人、水面机器人和卫星等。

[0003] 无人系统网络是由无人系统通过以自组织形式或基于网络基础设施所建立的网络。其中，以自组织形式建立的无人系统网络可以充分发挥无人系统的感知能力和较强的计算能力，并可有效地适应网络的变化。本发明将重点围绕无人系统自组织网络(下文简称无人系统网络)展开。

[0004] 在无人系统中，由于节点高速移动，无线链路不稳定，网络环境不确定等因素，导致的移动机器人网络拓扑频繁变化。在具有高频繁变化的网络拓扑的无人系统网络中，数据无法沿固定的路径传输，路由策略必须根据网络的变化，自适应地调节。因此，设计一个自适应且可靠的路由协议，是无人系统网络通信领域重要挑战之一。现有的基于拓扑的路由协议由于维护和重建路由路径而趋向于增加路由开销，不适用于无人系统网络。基于地理位置的路由协议是减少路由开销的主要选择之一，但由于缺乏对动态环境变化的理解，这些协议限制了路由路径的选择，所以基于地理位置的路由协议也不适用于无人系统网络。

[0005] 近年来，已有研究工作利用强化学习优化无人系统网络中的数据转发策略。这些研究工作往往将节点视为网络环境状态，而忽视了链路状态变化。然而在无人系统网络中，由于间歇性和不稳定的无线链路，链路状态频繁变化，进而影响数据转发策略。这些研究工作由于无法感知链路状态变化，因而对网络环境变化的适应性较低。此外，在这些研究工作中，节点以固定的时间间隔交换HELLO信息包。较长的时间间隔会导致邻居表中的邻居信息没有及时更新而过时，同时，较短的时间间隔也不能保证邻居信息被实时地更新，因为HELLO信息包可能会与数据包发生冲突而丢失。在这种具有低准确性的邻居信息情况下，实现可靠性数据转发是非常困难的。因此，这些研究工作无法提供可靠的数据转发。

发明内容

[0006] 针对现有技术的不足，本发明的目的在于提出一种基于深度强化学习的无人系统网络自适应路由方法和系统，以解决现有技术中移动无人系统网络中由于受节点高速移动、无线链路不稳定、移动机器人网络拓扑频繁变化的影响，无法提供自适应且可靠路由决策的技术问题。

[0007] 针对现有技术的不足，本发明提出一种基于深度强化学习的无人系统网络自适应路由方法，包括：

[0008] 步骤1、以无人系统网络中的每一个无人装置作为节点，所有节点以一个自适应的时间间隔发送HELLO信息包；任一节点收到其邻居节点发送的HELLO信息包后，更新该节点的邻居表中该邻居节点的节点信息；

[0009] 步骤2、将该无人系统网络中所有节点以及由所有节点形成所有链路作为系统环境，该无人系统网络中每个节点从系统环境中获取当前时刻的环境状态，并执行行为作用于系统环境，系统环境根据该执行行为反馈给节点奖励值，其中该环境状态包括当前节点和当前节点的所有邻居节点的链路状态；

[0010] 步骤3、无人系统网络中节点i根据其环境状态，利用深度Q网络(Deep Q-learning network,DQN)计算当前节点所有邻居节点的Q值，当前节点执行一个行为a_t，以最大Q值的邻居节点作为下一跳节点进行数据包的路由。

[0011] 所述的基于深度强化学习的无人系统网络自适应路由方法，该步骤1包括：所有节点以一个自适应的时间间隔发送HELLO信息包，其中自适应的时间间隔方法如下：

[0012] $$T_i = \min\{T_{\min} * \frac{v_{\max}}{v_{avg}^i}, T_{\max}\}$$

[0013] 其中，T_{min}和T_{max}分别是预设最短和最长时间间隔，v_{max}是节点i预设置的最大移动速度，v_{avg}^i为该节点i的平均速度。

[0014] 所述的基于深度强化学习的无人系统网络自适应路由方法，该步骤2包括：

[0015] 在当前时刻t下，节点i所观察到的环境状态s_t为：

[0016] $s_t=\{C_{i,1},\ldots,C_{i,j},\ldots,C_{i,M}\}$，其中$C_{i,j}$是由该节点i和该节点i的邻居j所形成的链路$l_{i,j}$的状态，M为该节点i拥有的邻居节点数量；

[0017] 基于该节点i的邻居表中该邻居节点j的信息，计算$C_{i,j}$：

[0018] $C_{i,j}=\{ct_{i,j},PER_{i,j},e_j,d_{j,des},d_{min}\}$，$ct_{i,j}$是链路$l_{i,j}$的期望连接时间，$PER_{i,j}$是链路$l_{i,j}$的包的错误率，$e_j$是该节点i的邻居节点j的剩余电量，$d_{j,des}$是该节点i的邻居节点j与该目标节点des间的距离，d_{min}是该节点i的2跳邻居节点k与该目标节点des的最小距离；

[0019] 节点通过选择一个优化的邻居节点作为下一跳节点来完成行为a_t；

[0020] 系统环境给予节点的奖励值r_t为：

[0021] 当该节点i的邻居节点j是该目标节点des，$r_t=R_{max}$，R_{max}是预设最大奖励值；

[0022] 当该节点i的所有邻居节点与该目标节点des的距离均大于该节点i与该目标节点des的距离，$r_t=-R_{max}$；

[0023] 否则，$r_t=RD_{i,j}$，$RD_{i,j}=\frac{d_{i,des}}{d_{j,des}}(1-PER_{i,j})$。

[0024] 所述的基于深度强化学习的无人系统网络自适应路由方法，该步骤3包括：

[0025] 收集节点i与环境交互的经验(s_t,a_t,r_t,s_{t+1})，并将该经验存储到经验回放存储器；从该经验回放存储器中随机采样部分经验以及最小化预先设置的损失函数，更新该深度Q网络的参数，该损失函数：$L(\theta)=\sum_{(s_t,a_t,r_t,s_{t+1})}(Q_{target}-q(s_t,a_t;\theta_t))$，其中$Q_{target}=r_i+\gamma\max_{a'}q(s_{t+1},a';\theta_t)$，θ表示所述DQN的网络参数，$q(s_t,a_t;\theta_t)$表示将环境状态$s_t$输入所述DQN后，输出在该环境状态$s_t$下选择行为$a_t$后获得累积奖励值，a′表示在环境状态$s_{t+1}$

下节点所采取的行为，$\max_{a'} q(s_{t+1}, a'; \theta_t)$表示在环境状态$s_{t+1}$下的最优累积奖励值，γ表示折扣因子，0≤γ≤1；

[0026] 一旦该深度Q网络的参数被更新，将更新后的参数发送给该无人系统网络中每个节点。

[0027] 所述的基于深度强化学习的无人系统网络自适应路由方法，该邻居表中邻居节点的节点信息包括：邻居节点的移动速度、位置坐标和剩余的电量。

[0028] 本发明还提供了一种基于深度强化学习的无人系统网络自适应路由系统，包括：

[0029] 以无人系统网络中的每一个无人装置作为节点，所有节点以一个自适应的时间间隔发送HELLO信息包；任一节点收到其邻居节点发送的HELLO信息包后，更新该节点的邻居表中该邻居节点的节点信息；

[0030] 将该无人系统网络中所有节点以及由所有节点形成所有链路作为系统环境，该无人系统网络中每个节点从系统环境中获取当前时刻的环境状态，并执行行为作用于系统环境，系统环境根据该执行行为反馈给节点奖励值，其中该环境状态包括当前节点和当前节点的所有邻居节点的链路状态；

[0031] 无人系统网络中节点i根据其环境状态，利用深度Q网络(Deep Q-learning network，DQN)计算当前节点所有邻居节点的Q值，当前节点执行一个行为a_t，以最大Q值的邻居节点作为下一跳节点进行数据包的路由。

[0032] 所述的基于深度强化学习的无人系统网络自适应路由系统，所有节点以一个自适应的时间间隔发送HELLO信息包，其中自适应的时间间隔系统如下：

[0033] $$T_i = \min\{T_{\min} * \frac{v_{\max}}{v_{avg}^i}, T_{\max}\}$$

[0034] 其中，T_{min}和T_{max}分别是预设最短和最长时间间隔，v_{max}是节点i预设置的最大移动速度，v_{avg}^i为该节点i的平均速度。

[0035] 所述的基于深度强化学习的无人系统网络自适应路由系统，具体包括：

[0036] 在当前时刻t下，节点i所观察到的环境状态s_t为：

[0037] $s_t = \{C_{i,1}, \ldots, C_{i,j}, \ldots, C_{i,M}\}$，其中$C_{i,j}$是由该节点i和该节点i的邻居j所形成的链路$l_{i,j}$的状态，M为该节点i拥有的邻居节点数量；

[0038] 基于该节点i的邻居表中该邻居节点j的信息，计算$C_{i,j}$：

[0039] $C_{i,j} = \{ct_{i,j}, PER_{i,j}, e_j, d_{j,des}, d_{min}\}$，$ct_{i,j}$是链路$l_{i,j}$的期望连接时间，$PER_{i,j}$是链路$l_{i,j}$的包的错误率，$e_j$是该节点i的邻居节点j的剩余电量，$d_{j,des}$是该节点i的邻居节点j与该目标节点des间的距离，d_{min}是该节点i的2跳邻居节点k与该目标节点des的最小距离；

[0040] 节点通过选择一个优化的邻居节点作为下一跳节点来完成行为a_t；

[0041] 系统环境给予节点的奖励值r_t为：

[0042] 当该节点i的邻居节点j是该目标节点des，$r_t = R_{max}$，R_{max}是预设最大奖励值；

[0043] 当该节点i的所有邻居节点与该目标节点des的距离均大于该节点i与该目标节点des的距离，$r_t = -R_{max}$；

[0044] 否则，$r_t = RD_{i,j}$，$RD_{i,j} = \frac{d_{i,des}}{d_{j,des}}(1 - PER_{i,j})$。

[0045] 所述的基于深度强化学习的无人系统网络自适应路由系统，具体包括：

[0046] 收集节点i与环境交互的经验(s_t, a_t, r_t, s_{t+1})，并将该经验存储到经验回放存储器；从该经验回放存储器中随机采样部分经验以及最小化预先设置的损失函数，更新该深度Q网络的参数，该损失函数：$L(\theta) = \sum_{(s_t,a_t,r_t,s_{t+1})}(Q_{target} - q(s_t,a_t;\theta_t))$，其中$Q_{target} = r_t + \gamma \max_{a'} q(s_{t+1},a';\theta_t)$，θ表示所述DQN的网络参数，$q(s_t, a_t; \theta_t)$表示将环境状态$s_t$输入所述DQN后，输出在该环境状态$s_t$下选择行为$a_t$后获得累积奖励值，a′表示在环境状态$s_{t+1}$下节点所采取的行为，$\max_{a'} q(s_{t+1},a';\theta_t)$表示在环境状态$s_{t+1}$下的最优累积奖励值，γ表示折扣因子，$0 \leq \gamma \leq 1$；

[0047] 一旦该深度Q网络的参数被更新，将更新后的参数发送给该无人系统网络中每个节点。

[0048] 所述的基于深度强化学习的无人系统网络自适应路由系统，该邻居表中邻居节点的节点信息包括：邻居节点的移动速度、位置坐标和剩余的电量。

[0049] 本发明与现有技术相比，具有以下优点：

[0050] 1.由于本发明创新性地提出了利用深度强化学习自适应地优化路由策略方法，与现有技术相比，本发明可以自主地在动态的无人系统网络中优化策略，以适应高动态变化的网络环境。此外，本发明具备良好的模型泛化能力，能泛化于具有不同网络规模和不同节点移动速度下的网络，这是一个非常重要的特征去适应动态无人系统网络。

[0051] 2.由于本发明在优化路由策略时考虑了链路状态，包括包的错误率，链路的期望连接时间，邻居节点的剩余能量以及邻居节点与目标之间的距离，与现有技术相比，本发明可以感知到链路状态的变化并且可以更好的推理网络环境变化，以做出更合适的路由策略。

[0052] 3.由于本发明提出了自适应调节HELLO信息包时间间隔方案，通过根据节点的平均移动速度自适应地调节HELLO信息包时间间隔，与现有技术相比，每个节点以不同的时间间隔自适应地发送HELLO信息包，有效地减少了HELLO信息包与数据包的冲突且改善了邻居表中邻居信息的准确性，从而提供可靠的数据转发。

[0053] 4.本发明实现了分布式路由决策机制，基于深度Q网络DQN的路由策略在每个节点上分布式执行，而DQN被一个预先设置的优化器集中式训练，进而简化了路由实施并且改善了DQN训练的稳定性。

附图说明

[0054] 图1是本发明方法实施例框架原理图；

[0055] 图2是本发明方法实施例基于深度强化学习的路由策略实现框架；

[0056] 图3至图8是本发明实例的仿真实验结果图。

具体实施方式

[0057] 为了解决上述技术问题，本发明的所采用的技术方案是：

[0058] 以无人系统网络中的无人机器或装置作为节点，所有节点以一个自适应的时间间隔发送HELLO信息包；任一节点收到其邻居节点发送的HELLO信息包后，更新该节点的邻居表中该邻居节点的节点信息；

[0059] 建立基于深度强化学习的路由策略算法框架；

[0060] 设计基于深度强化学习的路由策略实现方法。

[0061] 进一步地，该节点i发送HELLO信息包的时间间隔计算方法如下：

[0062] $$T_i = \min\{T_{\min} * \frac{v_{\max}}{v_{avg}^i}, T_{\max}\}$$

[0063] 其中，T_{min}和T_{max}分别是预设定的最短和最长时间间隔。v_{max}是该节点i预设置的最大移动速度，v_{avg}^i为该节点i的平均速度。

[0064] 进一步地，基于深度强化学习的路由策略算法框架：

[0065] (1)无人系统网络中的每个节点视为深度强化学习的智能体；

[0066] (2)抽象环境为无人系统网络包括网络中所有的节点以及由所有节点形成的所有链路；

[0067] (3)抽象环境状态为由该节点i和该节点i的所有邻居节点所形成的链路的状态。

[0068] (4)深度强化学习智能体从环境中获取当前时刻t的环境状态s_t，并执行行为a_t作用于环境，环境将反馈于深度强化学习智能体一个奖励值r_t，以实现深度强化学习智能体与环境的交互。

[0069] 进一步地，在当前时刻t下，该节点i所观察到的环境状态s_t为：

[0070] $s_t=\{C_{i,1},...,C_{i,j},...,C_{i,M}\}$，其中$C_{i,j}$是一个向量，其用于特征由该节点i和该节点的邻居j所形成的链路$l_{i,j}$的状态。

[0071] 进一步地，基于该节点i的邻居表中该邻居节点j的信息，计算$C_{i,j}$：

[0072] $C_{i,j}=\{ct_{i,j},PER_{i,j},e_j,d_{j,des},d_{min}\}$，$ct_{i,j}$是链路$l_{i,j}$的期望连接时间，即从当前时刻t直到该节点i与该节点的邻居j之间的距离达到最大通信距离的持续时间，$PER_{i,j}$是链路$l_{i,j}$的包的错误率，e_j是该节点i的邻居节点j的剩余电量，$d_{j,des}$是该节点i的邻居节点j与该目标节点des间的距离，d_{min}是该节点i的2跳邻居节点k与该目标节点des的最小距离。

[0073] 进一步地，深度强化学习智能体通过选择一个优化的邻居节点作为下一跳节点来完成行为a_t。

[0074] 进一步地，环境给予深度强化学习智能体的奖励值r_t为：

[0075] 当该节点i的邻居节点j是该目标节点des，$r_t=R_{max}$；

[0076] 当该节点i的所有邻居节点与该目标节点des的距离均大于该节点i与该目标节点des的距离，$r_t=-R_{max}$；

[0077] 否则，$r_t=RD_{i,j}$，$RD_{i,j}=\frac{d_{i,des}}{d_{j,des}}(1-PER_{i,j})$其中，$R_{max}$是预设定的最大奖励值。

[0078] 进一步地，基于深度强化学习的路由策略实现方法：基于深度Q网络(Deep Q-learningNetwork,DQN)的路由决策在每个节点上分布式执行，同时，DQN被一个预先设置的

优化器集中式训练。

[0079] (1)在分布式执行过程中,该节点i根据其在当前时刻t所观察到的环境状态s_t,利用DQN计算该节点i的所有邻居节点的Q值,该节点i执行一个行为a_t,以最大Q值对应的邻居节点作为下一跳节点进行数据包的路由。一个行为a_t执行后,该节点i获得一个奖励值r_t。一个预先设置的优化器收集该节点i与环境交互的经验(s_t,a_t,r_t,s_{t+1}),并将该经验存储到一个预先设置的经验回放存储器中。

[0080] (2)在集中式训练过程中,一个预先设置的优化器从预先设置的经验回放存储器中随机采样小批量经验来更新DQN的参数,通过最小化一个预先设置的损失函数:$L(\theta)=\sum_{(s_t,a_t,r_t,s_{t+1})}(Q_{target}-q(s_t,a_t;\theta_t))$,其中$Q_{target}=r_i+\gamma\max_{a'}q(s_{t+1},a';\theta_t)$,θ表示所述DQN的网络参数,$q(s_t,a_t;\theta_t)$表示将环境状态$s_t$输入所述DQN后,输出在该环境状态$s_t$下选择行为$a_t$后获得累积奖励值,a′表示在环境状态$s_{t+1}$下节点所采取的行为,$\max_{a'}q(s_{t+1},a';\theta_t)$表示在环境状态$s_{t+1}$下的最优累积奖励值,γ表示折扣因子,$0\leqslant\gamma\leqslant1$。

[0081] 一旦DQN的参数被更新,集中优化器会将更新后的DQN参数发送给无人系统网络中的每个节点。每个节点利用所收到的DQN参数更新该节点的DQN参数。

[0082] 为让本发明的上述特征和效果能阐述的更明确易懂,下文特举实施例,并配合说明书附图作详细说明如下。

[0083] 下面结合附图和具体实施例,对本发明进一步的详细描述。

[0084] 本发明具体实施方式提供了一种基于强化学习算法的无人系统网络的路由方法,本发明方法实施例的原理框架如图1所示,主要包括如下步骤:

[0085] 步骤101:以无人系统网络中的无人机器或装置作为节点,假设每个节点都会以一个自适应的时间间隔发送HELLO信息包,HELLO信息包的报文中包括节点自身的ID,节点的移动速度(v_x,v_y)、位置坐标(x,y)和剩余的电量e;每个节点维护一张邻居表用于存储邻居节点的移动速度、位置坐标和剩余的电量;

[0086] 具体地,节点i发送HELLO信息包的时间间隔计算方法如下:

[0087] $$T_i=\min\{T_{\min}*\frac{v_{\max}}{v_{avg}^i},T_{\max}\}$$

[0088] 其中,T_{min}是最短时间间隔,T_{min}=30ms,T_{max}是最长时间间隔,T_{max}=50ms。v_{max}是该节点i的最大移动速度,v_{max}=50m/s,v^i_{avg}为该节点i的平均速度。

[0089] 步骤102:建立基于深度强化学习的路由策略算法框架,包括深度强化学习的智能体和环境两大模块,并设计各个模块交互的内容;

[0090] (1)无人系统网络中的每个节点视为深度强化学习的智能体;

[0091] (2)抽象环境为无人系统网络包括网络中所有的节点以及由所有节点形成的所有链路;

[0092] (3)抽象环境状态为由该节点i和该节点i的所有邻居节点所形成的链路的状态。

[0093] (4)深度强化学习智能体从环境中获取当前环境状态s_t,并执行行为a_t作用于环境,环境将反馈于深度强化学习智能体一个奖励值r_t,以实现深度强化学习智能体与环境的交互。

[0094] 在当前时刻t下,该节点i所观察到的环境状态s_t为:$s_t=\{C_{i,1},\ldots,C_{i,j},\ldots,C_{i,M}\}$,

其中$C_{i,j}$是一个向量，其用于特征由该节点i和该节点的邻居j所形成的链路$l_{i,j}$的状态。$C_{i,j}$被计算基于该节点i的邻居表中该邻居节点j的信息：$C_{i,j}=\{ct_{i,j},PER_{i,j},e_j,d_{j,des},d_{min}\}$，$ct_{i,j}$是链路$l_{i,j}$的期望连接时间，即从当前时刻$t_1$直到该节点i与该节点的邻居j之间的距离达到最大通信距离的持续时间。假设在时刻t_1，该节点i的位置为(x_i,y_i)，速度为(v_x^i,v_y^i)，该节点i的邻居节点j的位置为(x_j,y_j)，速度为(v_x^j,v_y^j)，在时刻t_1，该节点i与该节点的邻居节点j的距离$d_{i,j}(t_1)$为：

[0095] $$d_{i,j}(t_1)=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}$$

[0096] 假设从时刻t_1到时刻t_2($t_2=t_1+\tau$)，该节点i和该节点的邻居节点j的速度未改变，那么τ时刻后，该节点i与该节点的邻居节点j的距离$d_{i,j}(t_1+\tau)$为：

[0097] $$d_{i,j}(t_1+\tau)=\sqrt{((x_i+v_x^i\tau)-(x_j+v_x^j\tau))^2+((y_i+v_y^i\tau)-(y_j+v_y^j\tau))^2}$$

[0098] 假设该节点i与该节点j的通信半径为R，当$d_{i,j}(t_1+\tau)>R$，该节点i与该节点j之间的链路$l_{i,j}$就会断开，因此我们可以通过$d_{i,j}(t_1+\tau)=R$求解该节点i与该节点j之间的链路$l_{i,j}$的期望连接时间$ct_{i,j}$，此时$ct_{i,j}=\tau$。

[0099] 假设链路$l_{i,j}$的包的错误率$PER_{i,j}$可以提前从网络环境中获得，e_j是该节点i的邻居节点j的剩余电量，$d_{j,des}$是该节点i的邻居节点j与该目标节点des间的距离，d_{min}是该节点i的2跳邻居节点k与该目标节点des的最小距离。

[0100] 深度强化学习智能体通过选择一个优化的邻居节点j作为下一跳节点来完成行为a_t。执行行为a_t后，环境将给予深度强化学习智能体一个奖励值r_t：

[0101] 当该节点i的邻居节点j是该目标节点des，给予智能体一个最大的奖励值，即，$r_t=R_{max}$，$R_{max}=2$；

[0102] 当该节点i的所有邻居节点与该目标节点des的距离均大于该节点i与该目标节点des的距离，给予智能体一个最小的奖励值，以避免路由空洞问题，即，$r_t=-R_{max}$；

[0103] 否则，在其他情况下，奖励值被计算根据节点与目标节点间的距离以及链路的质量：$r_t=RD_{i,j}$，$RD_{i,j}=\frac{d_{i,des}}{d_{j,des}}(1-PER_{i,j})$

[0104] 步骤103：设计基于深度强化学习的路由策略实现方法，如图2所示，基于深度强化学习的路由策略的实现具体包括基于深度Q网络DQN的路由策略在每个节点上分布式执行，以及利用一个预先设置的优化器集中式训练DQN。

[0105] (1)在分布式执行过程中，该节点i根据其所观察到的环境状态s_t，利用DQN计算该节点i的所有邻居节点的Q值，该节点i执行一个行为a_t，以最大Q值对应的邻居节点作为下一跳节点进行数据包的路由。一个行为a_t执行后，该节点i获得一个奖励值r_t。一个预先设置的优化器收集该节点i与环境交互的经验(s_t,a_t,r_t,s_{t+1})，并将该经验存储到一个预先设置的经验回放存储器M中。

[0106] (2)在集中式训练过程中，一个预先设置的优化器从预先设置的经验回放存储器M中随机采样小批量经验来更新DQN的参数，通过最小化一个预先设置的损失函数：$L(\theta)=\sum_{(s_t,a_t,r_t,s_{t+1})}(Q_{target}-q(s_t,a_t;\theta_t))$，其中$Q_{target}=r_t+\gamma\max_a q(s_{t+1},a';\theta_t)$，θ表

示所述DQN的网络参数，$q(s_t,a_t;\theta_t)$表示将环境状态s_t输入所述DQN后，输出在该环境状态s_t下选择行为a_t后获得累积奖励值，a′表示在环境状态s_{t+1}下节点所采取的行为，$\max_{a} q(s_{t+1},a';\theta_t)$表示在环境状态$s_{t+1}$下的最优累积奖励值，γ表示折扣因子，γ＝0.9。

[0107] 一旦DQN的参数被更新，集中优化器会将更新后的DQN参数θ_{t+1}发送给无人系统网络中的每个节点。每个节点利用所收到的DQN参数更新该节点的DQN参数。

[0108] 下面将通过具体的实例对本发明所述的一种基于深度强化学习的无人系统网络自适应路由方法进行仿真实验并给予说明。

[0109] 本实例在无线网络模拟器WSNet环境中仿真实验，实例中，节点分布在1000m x 1000m的区域内，其他节点随机分布。表1描述了下面路由协议对比实验共同参数的详细信息。

[0110] 表1参数配置表

[0111]

参数	配置
通信区域(Area Size)	1000m x 1000m
MAC 层协议(MAC protocol)	IEEE 802.11 dcf
无线传播模型(Radio propagation)	propagation_range, rang:300m
通信干扰模型(Interferences)	interferences_orthogonal
调制方式(Modulation)	modulation_bpsk
天线模型(Antenna)	antenna_omnidirectionnal
节点移动模型(Mobility)	Gaussian Markov mobility model
电量消耗模型(Battery)	energy_linear

[0112]

数据包大小(Data Size)	127 Bytes
包的错误率(PER)范围	[0, 0.2]

[0113] 在本实例中，采用IEEE 802.11dcfMAC协议和antenna_omnidirectionnal天线模型协议，每个节点利用propagation_range模型进行通信，且通信范围为300m，同时，利用energy_linear模型(节点发送和接受一个数据包消耗1单位能量(焦耳:J))，进行电量消耗的评估。实验中，仅有源节点在发送数据，目的节点接收数据，而其他节点对收到的数据进行转发。除了目的节点，其他节点均采用高斯移动模型移动。

[0114] 本实验中，将本发明实例与现有的QGeo路由协议(QGeo:Q-Learning based Geographic Ad-Hoc Routing Protocol for Unmanned Robotic Networks,Jung W S,2017)和GPSR路由协议(GPSR:Greedyperimeter stateless routing for wireless networks)进行了比较，并从端到端平均时延和数据包到达率，这2个性能指标对本发明所述的一种基于无人系统网络的自适应路由方法进行评估。在分析实验结果之前，先对本实验所涉及的2个性能指标进行简单的说明：

[0115] 端到端平均时延:数据包从源节点S成功到达目的节点D的平均时延；

[0116] 能耗:我们用目的节点收到一个数据包需要每个节点转发和接受的平均数据包数

来近似能耗，即能耗等于每个节点平均转发和接受的总包数除以目的节点收到的包数。

[0117] 首先，我们在不同节点移动速度下比较本发明实例与现有的QGeo路由协议和GPSR路由协议。图3显示了在节点数为25的情况下，数据包到达率与节点移动速度的关系。可以看出，随着节点移动速度的增大，数据包到达率降低。本发明具有更高的数据包到达率，且相比于现有的QGeo路由协议和GPSR路由协议，数据包到达率分别增加了16％和25％。GPSR路由协议通过利用局部信息，尝试发现最近邻的邻居来转发数据包。由于缺乏全局的路径信息，导致了低的数据包到达率。相比于GPSR路由协议，QGeo路由协议通过利用Q-learning可以引导更高的数据包到达率，但是在高动态场景中，由于缺乏对链路状态变化的理解，导致数据包到达率降低。相反，本发明在路由决策时考虑了链路状态包括链路质量、链路的期望连接时间、节点的剩余电量以及节点与目的节点间的距离，本发明可以很好地捕获链路的变化以至于可以做出更好的路由决策，引导了高的数据包到达率。

[0118] 图4显示了在节点数为25的情况下，能耗与节点移动速度的关系。可以看出，随着节点移动速度的增大，能耗增大。本发明具有更低的能耗，相比于现有的QGeo路由协议和GPSR路由协议，能耗减少了16％和28％。由于本发明通过使用深度强化学习方法可以发现更可靠的路由路径，导致了更少的数据重传和电量利用效率。此外，本发明提出了一种自适应HELLO消息间隔方法，该方法减少了节点发送不必要的HELLO消息包的概率，进一步地提高了电量利用效率。

[0119] 其次，我们在不同网络规模下比较本发明实例与现有的QGeo路由协议和GPSR路由协议。

[0120] 图5显示了在节点移动速度的范围为20～30m/s下，数据包到达率与节点数量的关系。可以看出，随着节点数量的增加，数据包的到达率也在增加。这是因为当节点数较多时，更多可靠的节点可以被选择去转发数据包。在不同网络规模下，本发明的数据包到达率高于现有的QGeo路由协议。相比于QGeo路由协议和GPSR路由协议，本发明的数据包到达率增加了18％和27％，即使在具有10个节点的低密度网络中，本发明的数据包到达率是82％，然而现有的QGeo路由协议和GPSR路由协议仅有68％和61％的到达率。

[0121] 图6显示了在节点移动速度的范围为20～30m/s下，能耗与节点数量的关系。可以看出，本发明具有更高的电量利用效率，相比于现有的QGeo路由协议和GPSR路由协议，在不同的网络规模下能耗平均减少了14％和23％。

[0122] 最后，我们验证了本发明在不同节点移动速度下和网络规模下的泛化能力。为了验证在不同移动速度下的泛化能力，在节点移动速度为30m/s下，我们首先为本发明实例训练一个DQN模型，定义为$train_{v=30}$。同时我们为现有的QGeo方法优化一个查询表，定义为$opt_{v=30}$。然后我们使用训练好的DQN模型和优化好的查询表来测试在其他节点移动速度下的路由性能，我们将这些测试结果定义为$(train_{v=30},test_{v=i},i=10,20,\cdots,100)$。最后，我们将这些结果与在相同移动速度下的训练和测试结果(定义为$train_{v=i},test_{v=i},i=10,20,\cdots,100$)进行比较。图7显示了本发明在不同移动速度下的泛化能力，可以看出，在本发明事例中，$(train_{v=30},test_{v=i},i=10,20,\cdots,100)$结果与$(train_{v=i},test_{v=i},i=10,20,\cdots,100)$结果较为吻合，这验证了本发明方法在不同节点移动速度下的泛化能力。然而在现有的QGeo路由协议中，$(train_{v=30},test_{v=i},i=10,20,\cdots,100)$结果与$(train_{v=i},test_{v=i},i=10,20,\cdots,100)$结果差距较大，这说明现有的QGeo路由协议在不同的节点移动速度下，不具

备泛化能力。

[0123] 为了验证在不同网络规模下的泛化能力，在节点数为20下，我们首先为本发明实例训练一个DQN模型，定义为$train_{N=20}$。同时我们为现有的QGeo方法优化一个查询表，定义为$opt_{N=20}$。然后我们使用训练好的DQN模型和优化好的查询表来测试在其他网络规模下的路由性能，我们将这些测试结果定义为($train_{N=20}$,$test_{N=i}$,i＝10,15,…,50)。最后，我们将这些结果与在相同网络规模下的训练和测试结果(定义为($train_{N=i}$,$test_{N=i}$,i＝10,15,…,50)进行比较。图8显示了本发明在不同网络规模下的泛化能力，可以看出，在本发明事例中，($train_{N=20}$,$test_{N=i}$,i＝10,15,…,50)结果与($train_{N=i}$,$test_{N=i}$,i＝10,15,…,50)结果较为吻合，这验证了本发明方法在不同网络规模下的泛化能力。然而在现有的QGeo路由协议中，($train_{N=i}$,$test_{N=i}$,i＝10,15,…,50)结果与($train_{N=i}$,$test_{N=i}$,i＝10,15,…,50)结果差距较大，这说明现有的QGeo路由协议在不同的网络规模下，不具备泛化能力。

[0124] 本实例的实验结果说明了本发明所述的基于深度强化学习的无人系统网络自适应路由方法较现有路由协议有更高的数据包到达率和更低的能耗。

[0125] 以下为与上述方法实施例对应的系统实施例，本实施系统可与上述实施方式互相配合实施。上述施方式中提到的相关技术细节在本实施系统中依然有效，为了减少重复，这里不再赘述。相应地，本实施系统中提到的相关技术细节也可应用在上述实施方式中。

[0126] 本发明还提供了一种基于深度强化学习的无人系统网络自适应路由系统，包括：

[0127] 以无人系统网络中的每一个无人装置作为节点，所有节点以一个自适应的时间间隔发送HELLO信息包；任一节点收到其邻居节点发送的HELLO信息包后，更新该节点的邻居表中该邻居节点的节点信息；

[0128] 将该无人系统网络中所有节点以及由所有节点形成所有链路作为系统环境，该无人系统网络中每个节点从系统环境中获取当前时刻的环境状态，并执行行为作用于系统环境，系统环境根据该执行行为反馈给节点奖励值，其中该环境状态包括当前节点和当前节点的所有邻居节点的链路状态；

[0129] 无人系统网络中节点i根据其环境状态，利用深度Q网络(Deep Q-learning network,DQN)计算当前节点所有邻居节点的Q值，当前节点执行一个行为a_t，以最大Q值的邻居节点作为下一跳节点进行数据包的路由。

[0130] 所述的基于深度强化学习的无人系统网络自适应路由系统，所有节点以一个自适应的时间间隔发送HELLO信息包，其中自适应的时间间隔系统如下：

[0131] $$T_i = \min\{T_{\min} * \frac{v_{\max}}{v_{avg}^i}, T_{\max}\}$$

[0132] 其中，T_{min}和T_{max}分别是预设最短和最长时间间隔，v_{max}是节点i预设置的最大移动速度，v_{avg}^i为该节点i的平均速度。

[0133] 所述的基于深度强化学习的无人系统网络自适应路由系统，具体包括：

[0134] 在当前时刻t下，节点i所观察到的环境状态s_t为：

[0135] s_t＝{$C_{i,1}$,...,$C_{i,j}$,...,$C_{i,M}$}，其中$C_{i,j}$是由该节点i和该节点i的邻居j所形成的链路$l_{i,j}$的状态，M为该节点i拥有的邻居节点数量；

[0136] 基于该节点i的邻居表中该邻居节点j的信息，计算$C_{i,j}$：

[0137] $C_{i,j}=\{ct_{i,j},PER_{i,j},e_j,d_{j,des},d_{min}\}$，$ct_{i,j}$是链路$l_{i,j}$的期望连接时间，$PER_{i,j}$是链路$l_{i,j}$的包的错误率，$e_j$是该节点i的邻居节点j的剩余电量，$d_{j,des}$是该节点i的邻居节点j与该目标节点des间的距离，d_{min}是该节点i的2跳邻居节点k与该目标节点des的最小距离；

[0138] 节点通过选择一个优化的邻居节点作为下一跳节点来完成行为a_t；

[0139] 系统环境给予节点的奖励值r_t为：

[0140] 当该节点i的邻居节点j是该目标节点des，$r_t=R_{max}$，R_{max}是预设最大奖励值；

[0141] 当该节点i的所有邻居节点与该目标节点des的距离均大于该节点i与该目标节点des的距离，$r_t=-R_{max}$；

[0142] 否则，$r_t=RD_{i,j}$，$RD_{i,j}=\frac{d_{i,des}}{d_{j,des}}(1-PER_{i,j})$。

[0143] 所述的基于深度强化学习的无人系统网络自适应路由系统，具体包括：

[0144] 收集节点i与环境交互的经验(s_t,a_t,r_t,s_{t+1})，并将该经验存储到经验回放存储器；从该经验回放存储器中随机采样部分经验以及最小化预先设置的损失函数，更新该深度Q网络的参数，该损失函数：$L(\theta)=\sum_{(s_t,a_t,r_t,s_{t+1})}(Q_{target}-q(s_t,a_t;\theta_t))$，其中$Q_{target}=r_t+\gamma\max_{a'}q(s_{t+1},a';\theta_t)$，θ表示所述DQN的网络参数，$q(s_t,a_t;\theta_t)$表示将环境状态$s_t$输入所述DQN后，输出在该环境状态$s_t$下选择行为$a_t$后获得累积奖励值，a′表示在环境状态$s_{t+1}$下节点所采取的行为，$\max_{a'}q(s_{t+1},a';\theta_t)$表示在环境状态$s_{t+1}$下的最优累积奖励值，γ表示折扣因子，$0\leqslant\gamma\leqslant1$；

[0145] 一旦该深度Q网络的参数被更新，将更新后的参数发送给该无人系统网络中每个节点。

[0146] 所述的基于深度强化学习的无人系统网络自适应路由系统，该邻居表中邻居节点的节点信息包括：邻居节点的移动速度、位置坐标和剩余的电量。

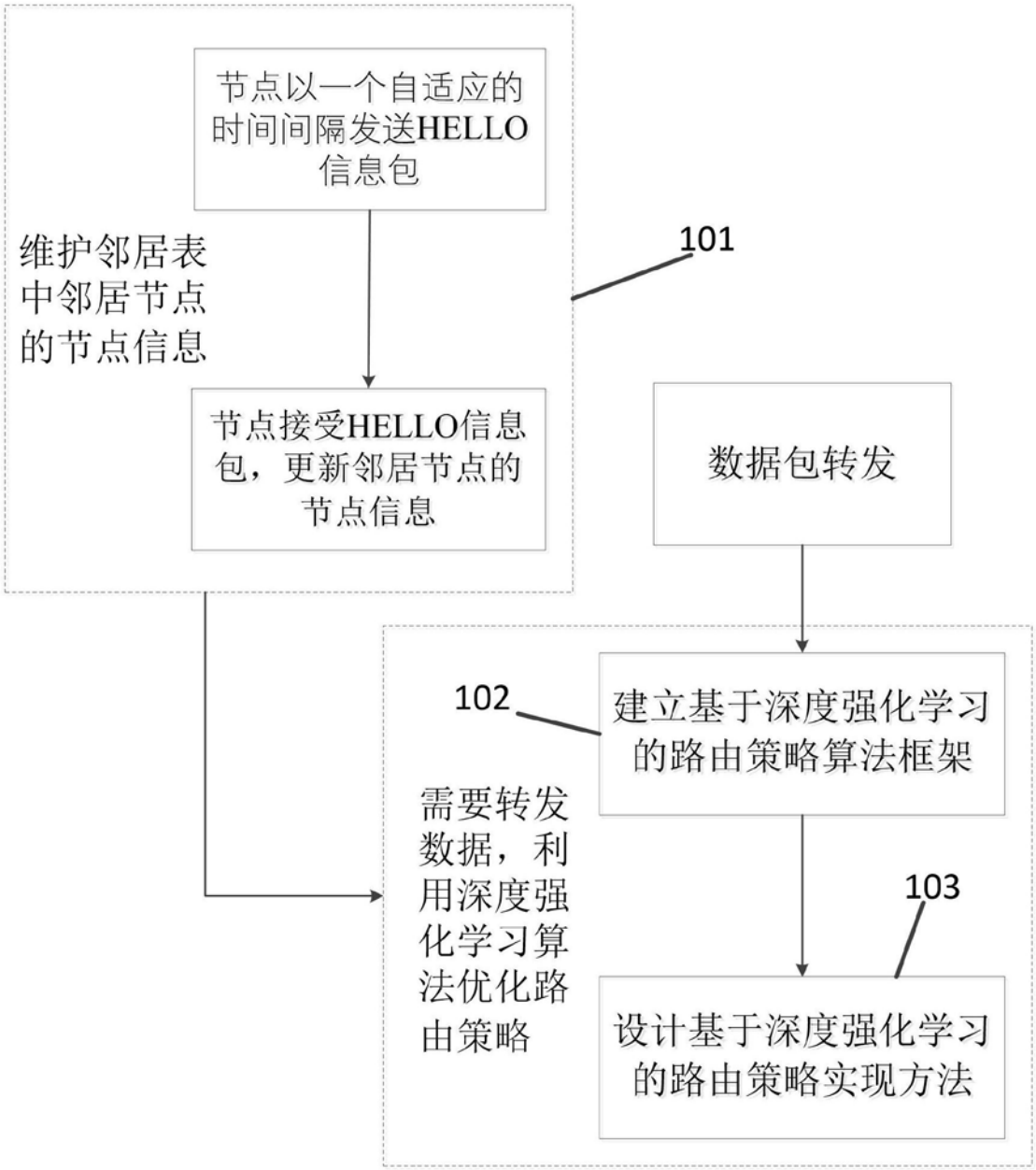

节点以一个自适应的时间间隔发送HELLO信息包
维护邻居表中邻居节点的节点信息
101
节点接受HELLO信息包，更新邻居节点的节点信息
数据包转发
102
建立基于深度强化学习的路由策略算法框架
需要转发数据，利用深度强化学习算法优化路由策略
103
设计基于深度强化学习的路由策略实现方法

图1

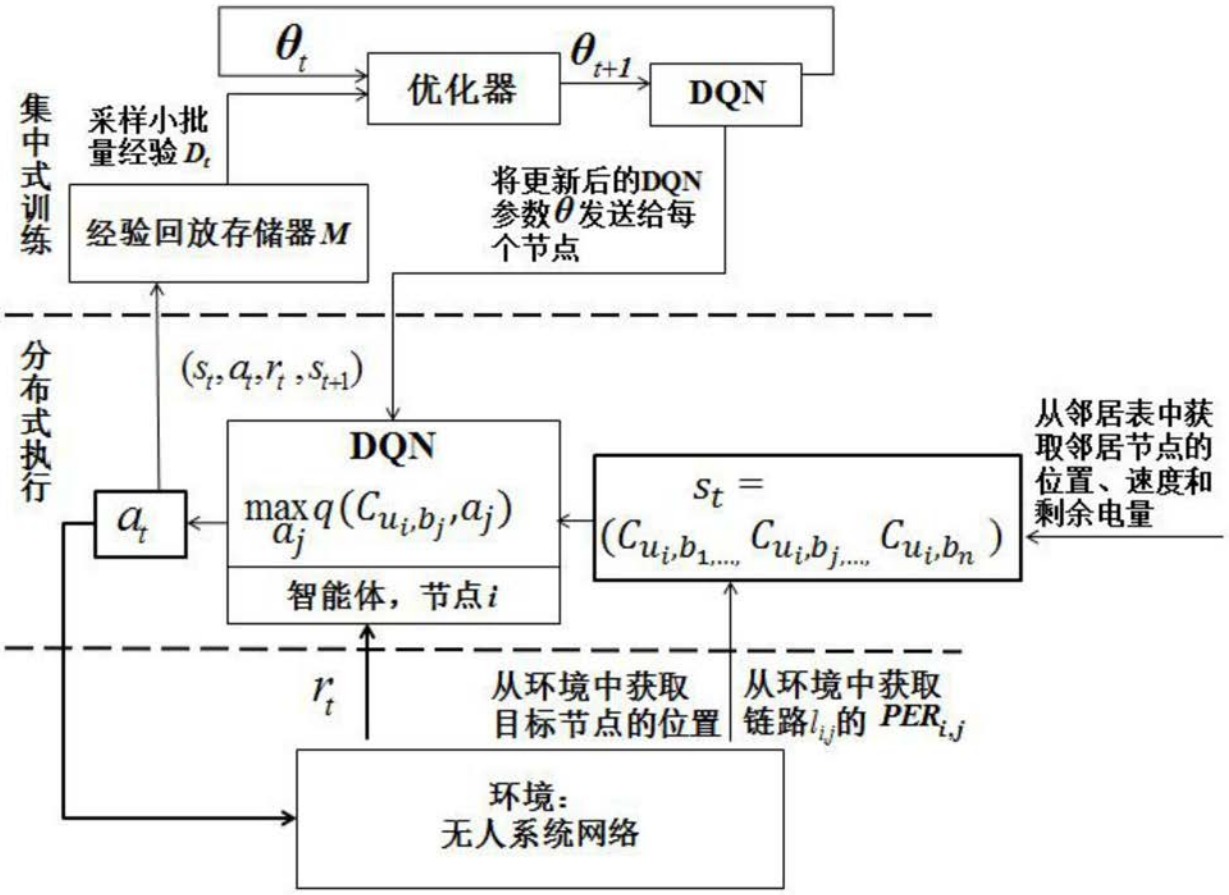

θ_t
优化器
θ_{t+1}
DQN
集中式训练
采样小批量经验D_t
经验回放存储器M
将更新后的DQN参数θ发送给每个节点
分布式执行
(s_t, a_t, r_t, s_{t+1})
DQN
$\max_{a_j} q(C_{u_i,b_j}, a_j)$
智能体，节点i
a_t
$s_t = (C_{u_i,b_1,\ldots}, C_{u_i,b_j,\ldots}, C_{u_i,b_n})$
从邻居表中获取邻居节点的位置、速度和剩余电量
r_t
从环境中获取目标节点的位置
从环境中获取链路l_{ij}的$PER_{i,j}$
环境：
无人系统网络

图2

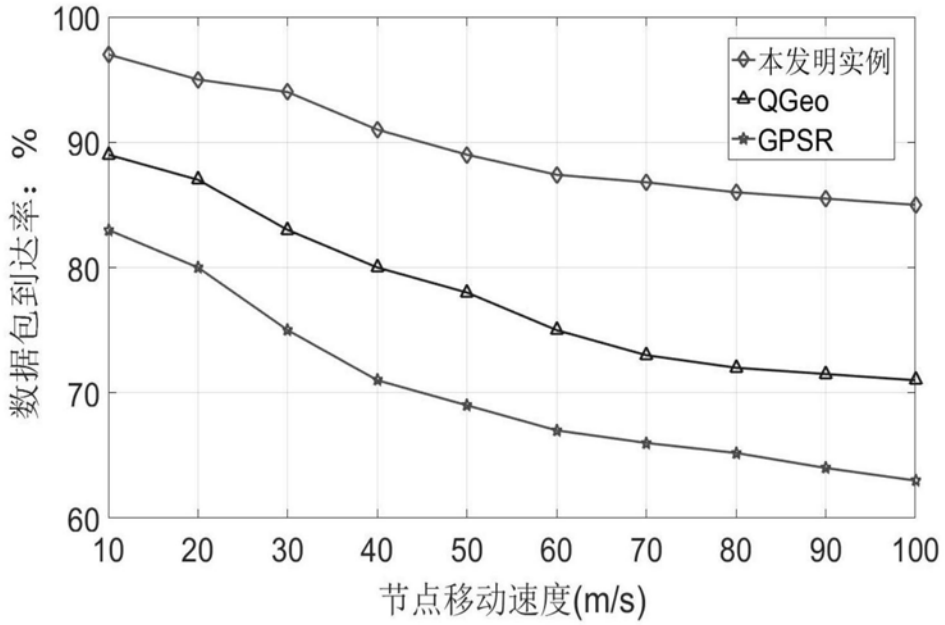

本发明实例
QGeo
GPSR
数据包到达率：%
节点移动速度(m/s)
100
90
80
70
60
10
20
30
40
50
60
70
80
90
100

图3

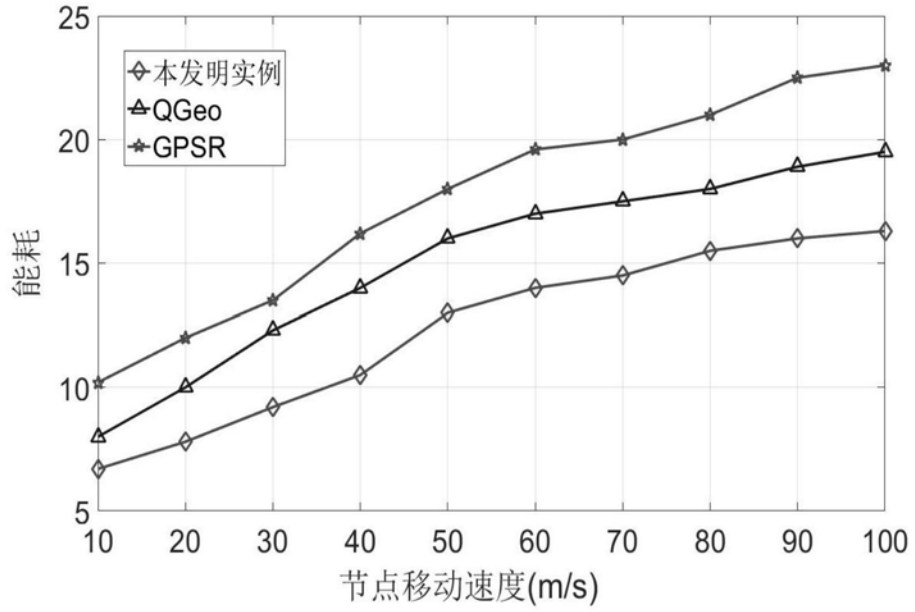

本发明实例
QGeo
GPSR
能耗
节点移动速度(m/s)
25
20
15
10
5
10
20
30
40
50
60
70
80
90
100

图4

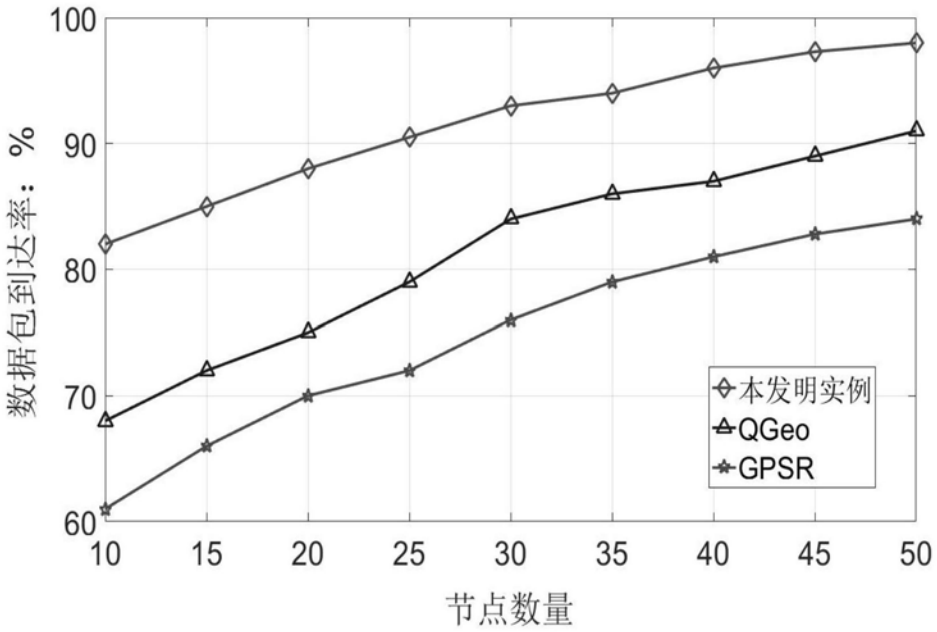

本发明实例
QGeo
GPSR
数据包到达率：%
节点数量
100
90
80
70
60
10
15
20
25
30
35
40
45
50

图5

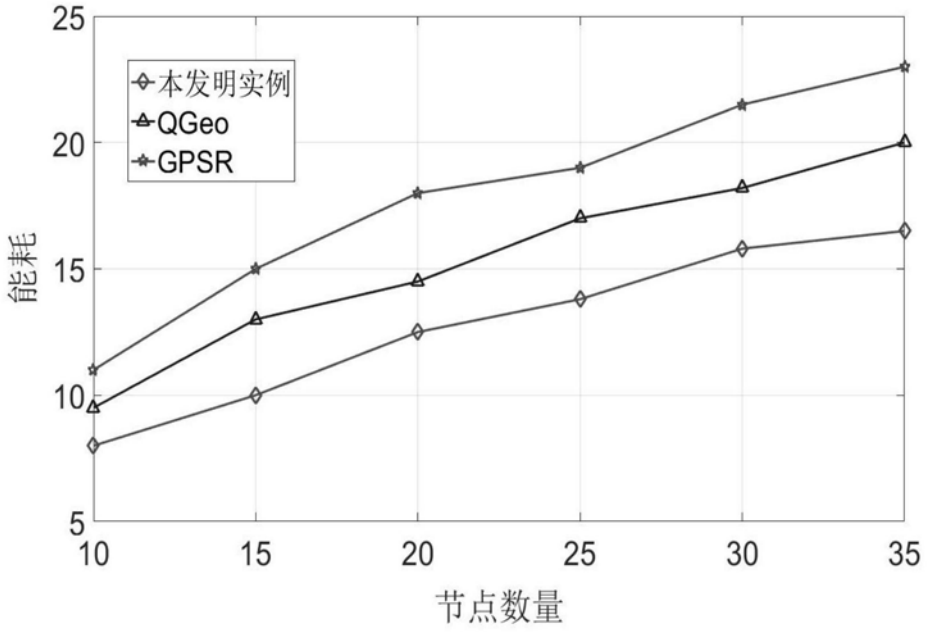

本发明实例
QGeo
GPSR
能耗
节点数量
25
20
15
10
5
10
15
20
25
30
35

图6

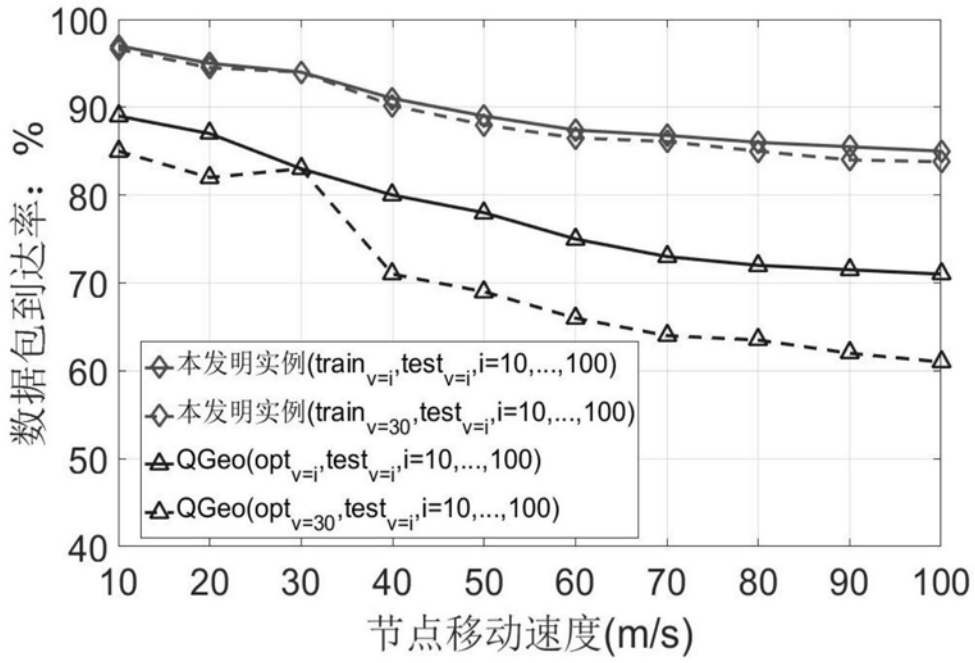

本发明实例(train$_{v=i}$,test$_{v=i}$,i=10,...,100)
本发明实例(train$_{v=30}$,test$_{v=i}$,i=10,...,100)
QGeo(opt$_{v=i}$,test$_{v=i}$,i=10,...,100)
QGeo(opt$_{v=30}$,test$_{v=i}$,i=10,...,100)
数据包到达率：%
节点移动速度(m/s)
100
90
80
70
60
50
40
10
20
30
40
50
60
70
80
90
100

图7

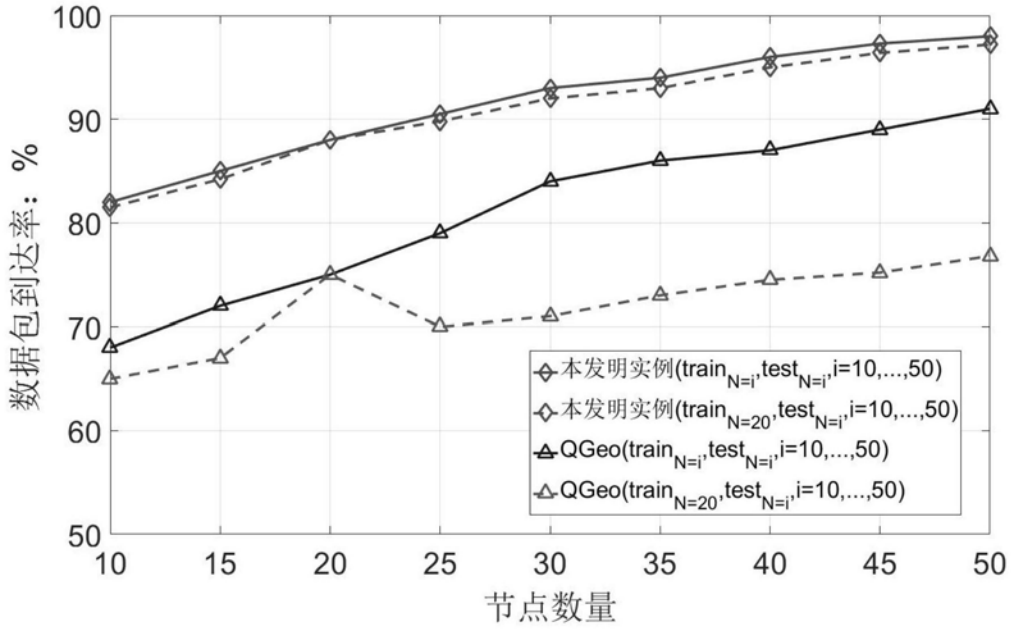

100
90
80
70
60
50
10
15
20
25
30
35
40
45
50
数据包到达率：%
节点数量
本发明实例(train$_{N=i}$,test$_{N=i}$,i=10,...,50)
本发明实例(train$_{N=20}$,test$_{N=i}$,i=10,...,50)
QGeo(train$_{N=i}$,test$_{N=i}$,i=10,...,50)
QGeo(train$_{N=20}$,test$_{N=i}$,i=10,...,50)

图8