 DOI: https://doi.org/10.12349/iser.v7i6.10797

Resource Allocation Algorithm for NBIoT/LoRa Heterogeneous Networks Based on Reinforcement Learning

Jie Ding

Wenhua College, Hongshan District, Wuhan City, Wuhan, Hubei, 430000, China

Abstract

With the rapid development of Internet of Things (IoT) technology, NBIoT and LoRa, as two mainstream narrowband IoT communication technologies, have been widely applied in smart cities, industrial monitoring, agricultural sensing, and other fields. However, a single communication technology is difficult to meet the diverse business requirements in different scenarios, and the integration of heterogeneous networks has become an important direction for improving the overall performance of the IoT. This paper addresses the resource allocation problem in NBIoT/LoRa heterogeneous networks and proposes a dynamic resource allocation algorithm based on deep reinforcement learning. The algorithm models the network resource allocation as a Markov decision process and uses a deep Q-network to jointly optimize the spectrum resources, power resources, and time slot resources in the heterogeneous network. By designing a reasonable state space, action space, and reward function, the algorithm can adaptively learn the optimal resource allocation strategy in a complex network environment, achieving performance improvements in system throughput, energy efficiency, and user fairness, and providing an effective technical solution for resource management in heterogeneous IoT networks.

Keywords

NBIoT; LoRa; heterogeneous network; resource allocation; deep reinforcement learning

基于强化学习的 NBIoT/LoRa 异构网络资源分配算法

丁洁

武汉市洪山区文华学院，中国 · 湖北 武汉 430000

摘　要

随着物联网技术的发展，NBIoT与LoRa是目前两种常见的窄带物联网通信技术，在智慧城市的各个领域如智能交通、智能照明等以及工业监测、农业生产中被大量使用。但是由于每一种通信技术都有其局限性而不能完全胜任各种各样的应用，因此将不同的通信技术结合在一起形成异构网络可以更好地发挥各自优势从而提高整个物联网系统的性能。本文针对NBIoT与LoRa异构网络中的资源调度问题提出了一种基于深度强化学习的方法来解决。把网络资源调度看作一个马尔科夫决策问题，用深度Q网络在异构网络中对频谱资源、功率资源以及时隙资源进行优化。根据实际情况定义合适的状态空间、动作空间以及奖励函数使算法可以在复杂环境中自动找到最优资源配置方案，在系统吞吐率、能效比和用户体验上都有所改善，对于异构物联网网络中的资源管理具有积极意义。

关键词

NBIoT；LoRa；异构网络；资源分配；深度强化学习

1 引言

物联网是新一代信息技术的重要组成部分，正在改变人们的生活方式以及工作方式。根据有关资料表明，全球物联网连接设备的数量呈爆炸式增长的趋势，预计到 2025 年会达到数十亿甚至更多。在此背景下，低功耗广域网络出现，而在这其中 NBIoT（NarrowbandInternetofThings）以及 LoRa（LongRange）是最为典型的两类。NBIoT 利用蜂窝网络进行通信，有较大的覆盖范围，安全性能良好并且支持现有的蜂窝网络；LoRa 使用未授权的频段，部署简单，价格便宜，易于组建网络。这两种技术有着很大的区别，在覆盖范围、传输速度、能耗上都有所不同，因此适用于不同类型的应用场景。

2 NBIoT/LoRa 异构网络概述

2.1 异构网络架构

NBIoT/LoRa 混合网络采取分层式融合结构，分为核心层、聚合层及接入层三层结构，核心层用于对数据进行收集、分析、传递，完成对上层应用平台的交互；聚合层设置混合网关装置，可以同时接入 NBIoT 和 LoRa 两种协议并进行相互转化；接入层包括众多感测节点，在不同业务场景以及

【作者简介】丁洁（1986–），女，中国河南漯河人，硕士，工程师，从事光通信、电信研究。

不同的覆盖区域下接入相应的 NBIoT 或 LoRa 网络。

在接入层中，NBIoT 终端通过蜂窝基站接入核心网，采用 OFDMA 技术进行多址接入，支持上行和下行非对称传输。NBIoT 系统工作在授权频段，具有较好的抗干扰能力和服务质量保障机制。LoRa 终端则通过 LoRa 网关接入网络，采用基于 CSS（ChirpSpreadSpectrum）的调制技术，通过调整扩频因子、编码率和带宽等参数实现不同的传输速率和覆盖范围。LoRa 网络采用 ALOHA 或时隙 ALOHA 的随机接入方式，工作在 ISM 非授权频段。

异构网络中的资源类型主要包括频谱资源、功率资源和时隙资源。NBIoT 网络的频谱资源以物理资源块为单位进行分配，每个资源块包含若干子载波和时隙；LoRa 网络的频谱资源以信道为单位，不同扩频因子在同一信道上可实现准正交传输。功率资源方面，NBIoT 终端的发射功率受到基站调度和功率控制机制的限制，LoRa 终端则根据链路预算和覆盖需求自主选择发射功率。

2.2 资源分配问题分析

异构网络资源调度的根本任务就是在保证各种服务等级需求的基础上使整个系统的性能最大化，因此，在进行资源调度时应当兼顾以下限制因素以及优化准则。

在约束条件上，首先要满足各类终端的基础业务要求，如最底传输速率、最高时延接受范围以及最大错误率等等；其次，NBIoT 终端发射功率不能超过基站下发的最大限值，LoRa终端发射功率要满足国家法律法规对发射功率的要求。另外频率资源的发放应注意不要有相同频率的相互影响，做到不同网路能相互兼容。时隙资源的分配也要注意不要对同网络下不同类型的设备造成互相影响，也要注意不同网路之间进行相互配合。

关于优化目标，在系统吞吐量上来说它是评价网络数据传输能力的一项重要指标，表征着单位时间内传输成功的总数据数量；能量效率则是一个表征单位能量下可传输的数据量大小的一个参数，对电池供电的物联网设备来说很重要；用户公平性体现的是各个设备间的资源分配是否合理，防止某些设备因为信道环境恶劣或者过于偏僻导致长时间无服务的情况出现；再者就是系统的时延也是一个重要的性能指标，尤其是对于那些需要低时延的应用场景。

因为 NBIoT 与 LoRa 两种网络的技术体制不同，资源种类不同以及它们所受到的限制不同，所以资源分配问题具有多目标、多约束、高维数的特点。传统固定的资源分配算法已不足以应用于动态变化的网络环境中，在复杂的异构环境下也很难使用基于模型化的优化算法来解决资源分配的问题。为此要采用智能化的资源分配算法，让系统可以依据当前的网络环境的变化而动态性的改变自己的资源分配方案。

3 基于强化学习的资源分配算法设计

3.1 问题建模

本文将 NBIoT/LoRa 异构网络资源分配问题建模为马尔可夫决策过程（MDP），通过强化学习算法求解最优资源分配策略。马尔可夫决策过程由状态空间、动作空间、状态转移概率和奖励函数四个要素构成。

状态空间的设计需要充分反映异构网络的运行状态。本文定义系统状态为各终端的信道增益、剩余能量、队列长度以及网络负载等信息的集合。具体而言，对于每个 NBIoT 终端，状态包括其与基站之间的信道增益、当前队列中的待传数据量、电池剩余能量以及已分配的资源块索引；对于每个 LoRa 终端，状态包括其与网关之间的路径损耗、当前扩频因子配置、队列长度和发射功率等级。此外，系统状态还包含网络层面的信息，如当前接入的终端总数、各信道的干扰水平和时隙利用率等。

动作空间定义了资源分配算法在每个决策时刻可执行的操作集合。对于 NBIoT 网络，动作包括为各终端分配资源块、设置调制编码方案和确定发射功率等级。对于 LoRa 网络，动作包括为各终端分配扩频因子、编码率和传输信道，以及调整发射功率。由于两种网络的资源类型不同，动作空间采用联合编码方式表示，每个动作向量包含所有终端的资源分配决策。

3.2 深度 Q 网络算法

异构网络的状态空间和动作空间维度很高，传统 Q 学习算法遇到了维数灾难的问题不能很好地解决，因此本篇文章使用了深度 Q 网络（DQN）的方法，使用深度神经网络来拟合 Q 值函数，以此来克服高维状态空间下的决策问题。

深度 Q 网络的基本思路就是用一个深度神经网络来近似 Q 函数，在给定状态下输出各个动作对应的 Q 值估算，一般是一个输入层、多个隐藏层、输出层组成的模型，在隐含层内采用完全连接的形式并且加入非线性的激活函数使网络可以更好的逼近复杂的状态 - 动作映射关系，输入层接受异构网络当前时刻的状态信息即各终端的信道增益，队列长度，剩余能量及网络负荷等高维特征向量，输出层即为所有的可能采取的资源分配行为所对应的 Q 值预测，每一个输出端口就代表着一个具体的行为其未来长期获得的累计奖励期望。为了提升算法训练性能及收敛效果，使用了经验回放和目标网络的方法，经验回放会把智能体和环境之间的交互过程中的转移示例以四元组的形式存放在经验池中，四元组表示当下的状态，采取的行动，立即得到的奖赏还有下一时刻的状态，训练的时候通过经验池中随机抽取出一个小批量的样本用于更新网络模型，从而消除了样本间的因果联系，不会因为出现连续相似样本而影响到梯度的计算，极大提高了数据的有效性和学习的稳定性。

3.3 算法流程

我们提出的算法运行过程分为初始化、交互学习与策略执行三个环节，在初始化环节中建立深层 Q 网络以及目标网络并初始化其参数，在初始化过程中设置一些超参数如经验库大小、批次数目、学习率等等，在此同时还需要初始

化异构网络环境即终端位置分布、信道模型与业务流量模型。在交互学习过程中，算法在每个时间片初始时刻获得最新的系统状态，依据当前策略挑选并作出行为，观察作出行为之后立即产生的奖赏以及下一时段的状态，将转移经验保存到经验库中。

4 算法特性与适用性分析

4.1 算法特性分析

所提基于深度强化学习的资源分配算法具有以下显著特性。

自适应性为算法的主要亮点之一。因其利用与环境不断互动优化自身策略，算法可以做到对不断变化的网络环境进行自我调节以达到较好的效果；而在终端数目增多时，终端分布广泛、不同类型的业务、信道状态等发生改变的情况下，算法也可以做出相应地调整而不需要人为操作，从而使其始终维持在一个相对较高的水平上，这对于物联网网络来说是非常重要的，物联网网络一般会存在大量的终端设备分布分散、业务多种多样以及信道情况复杂等问题，固定下来的分配策略很难应对瞬息万变的情况。

多目标优化特点在于奖励函数的构造。以吞吐率、节能性能以及公平性等多个优化指标为基础构建一个整体的奖励函数平台，在训练的时候可以自动地找到各个指标之间的平衡点，相较于传统的单一指标优化方式更加贴近实际网络运行需求，不会为了提高某一项性能指标而减少其他方面的性能，同时权重系数设置可以让网络运营商根据实际情况自行调节优化侧重点。

4.2 适用场景分析

本文提出的算法可应用于所有常见的 NBIoT 与 LoRa 混用网布局方案中。

智慧城市场景下，在城市区域内部署了广泛分布着大量的传感器节点。例如智能电表、环境监测设备、车流量感应器等等，这些节点间的数据上报速率以及对可靠性的需求程度以及供电情况都有着较大的区别。所以 NBIoT 适用于放置在重要基础设施类监控节点的地方，而 LoRa 则是适用于放置在数据量较少，上报频率较低的一些普通的环境监测节点处；本文所提出的方法可以针对不同的节点其业务特性以及当前网络状况来动态地去进行频谱及功率分配，从而使得两种网络之间达到优势资源共享的效果。

针对工业物联网应用场景下的厂内部署着一整套传感器系统来进行设备健康状况监控、环境指标采集及警报提醒等一系列工作，部分重要设备的数据上传要求极高可靠性和极低时延，应该走 NBIoT 通道来传输，大面积采集温度湿度光强等环境信息可以利用 LoRa 技术以较低的成本进行全覆盖，提出的方案可以区别对待不同类型业务的服务质量要求，做到差异化调度，提高工业物联网整体运行性能。

4.3 实现考虑

在实际部署所提算法时，需要考虑以下实现层面的问题。

计算资源方面，深度学习中的训练与推断都需要一定的计算能力，训练阶段，因为有大量的样本遍历及参数调整，推荐在拥有 GPU 加速能力的边缘服务器或者云计算平台上做离线训练，在推断阶段，神经网络的前向计算量并不大，可以部署在网络端的边缘计算节点处进行实时刻度推断来适应对资源分配及时性的需求。

通信消耗方面，算法要运行就需要得到各个设备的状态信息，对于 NBIoT 网络来说，设备状态的信息是通过目前的信令来传送的，在 LoRa 网络中则需要在协议框架内添加状态信息上报的字段或是频率。为了减少通信消耗，可以选择事件触发机制，即只有在设备发生重大改变的时候才会发送状态信息，也可以对状态信息做一定的压缩处理以节约上传的数据量。

5 结论与展望

本文针对 NB-IoT、LoRa 异构网络下资源调度的问题，提出了基于深度强化学习的动态资源调度方法。把资源调度问题抽象成一个 MDP 模型，在这个模型中定义了包括信道状况，能况以及队列状况等元素在内的多维的状态空间，定义了包括频谱资源，功率资源和时隙资源在内的联合动作集，并设置了综合考虑了吞吐量、能量效率、公平性的目标函数。使用深度 Q 网络来逼近 Q 函数值，用经验回放和目标网络相结合的方法实现对高维复杂状态空间的有效探索从而达到更好的策略学习效果。

算法具有自主性、多目标性和可伸缩性等诸多优点，在智慧城市、工业互联网以及农业互联网等多个典型应用中均适用该算法，算法在不断与环境中交互的过程中自主探索出最佳资源配置方案，不需要构建具体的系统模型来模拟这个过程；可以应对各种各样的复杂的网络环境。但是在实际应用过程中必须兼顾计算资源，通信代价及系统的安全性等多个方面的实现因素才能保证算法能正常工作下去。

参考文献

[1] 魏泽豪,徐少毅.面向工业5G异构网络保证时延确定性的资源管理[J].北京交通大学学报,2025,49(06):183192.

[2] 贺园园,谢健骊,李翠然.面向用户体验质量的超密集异构网络资源分配算法[J].计算机应用与软件,2025,42(10):147155.

[3] 魏梦霞.室内可见光异构网络中资源分配算法研究[D].西安理工大学,2025.DOI:10.27398/d.cnki.gxalu.2025.000871.

[4] 蒋潇逸.元宇宙业务体验保障的无线网络资源分配方法[D].北京邮电大学,2025.DOI:10.26969/d.cnki.gbydu.2025.002537.

[5] 曹树强.异构网络中的无线时频资源协同技术研究及实现[D].北京邮电大学,2025.DOI:10.26969/d.cnki.gbydu.2025.001691.