

氏 名 Thi Ha Ly DINH

学位(専攻分野) 博士(情報学)

学位記番号 総研大甲第 2359 号

学位授与の日付 2022 年 9 月 28 日

学位授与の要件 複合科学研究科 情報学専攻  
学位規則第6条第1項該当

学位論文題目 Reinforcement Learning-based Methods for Wireless Access  
Optimization and Multi-Interface Connectivity

論文審査委員 主 査 金子 めぐみ  
情報学専攻 准教授  
計 宇生  
情報学専攻 教授  
福田 健介  
情報学専攻 准教授  
阿部 俊二  
情報学専攻 准教授  
栗本 崇  
情報学専攻 准教授  
Lila Boukhatem  
フランス・パリサクレ大学 准教授

## **Summary of Doctoral Thesis**

Name in full    Thi Ha Ly DINH

Title    Reinforcement Learning-based Methods for Wireless Access Optimization and Multi-Interface Connectivity

Recently, 5G wireless networks introduced three new use cases: Enhanced Mobile Broadband (eMBB) for high data rate transmissions, Ultra-Reliable Low Latency Communications (URLLC) enabling high reliability and low latency of connections, and Massive Machine Type Communications (mMTC) for supporting low data rate communications of a massive number of devices. Given the forecasted exponential growth of mobile data traffic and the unprecedented diversification of applications with the spread of the Internet of Things (IoT) applications, Beyond 5G (B5G) and 6G wireless networks will be facing more daunting challenges of Quality of Service (QoS) provision as compared to 5G. Towards this end, future networks are expected to leverage these promising directions: the joint exploitation of a wide range of spectrum from Sub-6GHz to mmWaves for multi-interface connectivity, AI-enabled network entities and energy efficient Deep Learning (DL).

In such a context, this thesis investigates the fundamental issues of wireless access design, namely user association and interference management at user side, and develops new radio resource allocation optimization methods at Access Point (AP) side to enhance global network performances while satisfying individual user QoS constraints. Unlike most of existing studies, we focus on the situation where both user devices and APs are equipped by multiple wireless interfaces, and by DL capabilities. For that, wireless access optimization methods to support multiple applications/interfaces simultaneously with heterogeneous types of QoS and levels, requested by each user device, are investigated. Namely, depending on the QoS requirements and the state of the dynamically varying wireless environment, each application should be served on the most suitable wireless interface at a given time, to offer the utmost user satisfaction to the maximum number of users over the whole network.

Namely, we first consider the problem of user-to-multiple APs association, where a user requesting several applications with different QoS constraints can be served by multiple APs simultaneously, in current Sub-6GHz wireless system. We propose two distributed user-to-multiple APs association methods that leverage Reinforcement Learning (RL), namely Q-Learning (QL) at each user device, enabling each user to optimize its own association decision while aiming at global network optimization. Then, to cope with large-scale networks, we extend this initial QL-based association method by making use of Deep RL (DRL) tools such as DQN and Double DQN (DDQN)

based on Deep Neural Networks (DNN). Based on that, in the Sub-6GHz/mmWave integrated networks envisioned for B5G/6G, we handle the issue of joint user-to-multiple APs association and beamforming by proposing a scheme where DNN-enabled user devices optimize their requests (APs, interfaces), while APs perform a greedy-based beamforming to select their best sets of users. The goal is to maximize the system throughput while satisfying the users' QoS requirements and APs' load constraints.

Running such DL functionalities generally requires tremendous energy consumption, which may be prohibitive for battery-limited user devices. Indeed, a large amount of energy is consumed not only for DNN computations using massive data, but also to access, read and write data in the device memory. As Energy Efficiency (EE) will become one of the major Key Performance Indicators (KPI) in B5G/6G system, we also investigate the EE issue of DQN-based method at the user device. In particular, unlike existing works, we conduct a comprehensive analysis of the energy consumption for both computation and data access by DNN. Based on that, to obtain higher EE and better cope with dynamic environments, we enhance our proposed DQN-based user association and beamforming scheme by proposing an adaptive  $\epsilon$ -greedy strategy which enables the user to explore whenever notable changes of its surrounding environment are detected. Moreover, to further improve the network performance, we design a beamforming method based on Branch-and-Bound algorithm at AP side, taking into account the features of mmWave bands. The trade-off between achievable network performances and energy costs at user side is then investigated.

Finally, since realizing extreme reliability is another of the major milestones paving the way towards B5G, we also consider the issue of reliability enhancement for mMTC use case under Sub6GHz/mmWave integrated systems. For that, we design a method based on the Risk-Averse Averaged Q-Learning (RAQL) framework, whereby each AP avoids to transmit on interfaces with high risks of violating devices' Packet Loss Rate (PLR) targets, based on limited feedback from their associated devices.

We assess our proposed methods through numerical evaluations over various network settings. These results show that the proposed approach enables all users to associate to multiple APs/interfaces distributively and efficiently, while satisfying their heterogeneous QoS requirements and enhancing the long-term global sum-rate. Moreover, the proposed algorithms are also shown to outperform benchmark methods, both in terms of global sum-rate and application outage probabilities. In particular, the proposed methods enable to cope with dynamic environments and to strike a balanced trade-off between network sum-rate, QoS satisfaction of diverse applications, as well as user energy consumption. In case of reliability enhancement, the proposed RAQL-based method can significantly improve network performance by increasing the global successful packet delivery rate while reducing individual PLRs as compared to baseline algorithms.

## 博士論文審査結果

Name in Full

氏 名

Thi Ha Ly DINH

Title

論文題目

Reinforcement Learning-based Methods for Wireless Access  
Optimization and Multi-Interface Connectivity

本学位論文は、強化学習を活用した無線アクセス法及びマルチ無線インターフェース接続法の設計に関するものである。全 7 つの章から構成されている。周波数逼迫問題が悪化する中、Beyond 5G (B5G)で要求される厳しい無線通信品質の達成に向け、強化学習技術や数値最適化を活用し、高い通信性能と無線資源利用効率を達成できる無線アクセス設計を目的としている。

第 1 章では研究の背景と目的、第 2 章では強化学習や数値最適化に基づく、既存無線アクセス法や無線資源割当て法に関する関連研究を紹介し、それらの問題点を説明している。

第 3 章では、低い周波数帯であるサブ 6 GHz 帯のための、強化学習に基づくマルチ無線アクセス法の設計について、問題の定式化・提案法および計算機評価について述べている。従来無線アクセス法とは異なり、深層強化学習を活用した提案法では、各ユーザがローカルな情報のみで、多様なアプリケーションの性能要求を最大限に達成できるアクセスポイント(AP)の組み合わせを学習し、分散型なマルチ AP 選択法を実現している。計算機評価により、提案法は従来の無線 AP 選択法と比べて、システムの総伝送速度とアウトゲージ確率を改善していることを明らかにしている。

第 4 章では、B5G で普及していくサブ 6 GHz とミリ波帯の統合的ネットワークを対象にし、ユーザ側での深層強化学習を活用したマルチ AP・無線インターフェース接続法と、AP 側でのユーザ選択最適化について述べている。更に、高い周波数帯であるミリ波帯では距離や障害物による伝搬損失が著しいため、そのモデル化を強化学習と数値最適化に取り入れ、ビームフォーミング最適化を提案している。計算機評価により、提案法は厳しいモバイル環境でも、従来法と比べて高い通信品質を達成することを示している。

第 5 章では、深層強化学習を用いたマルチ AP・インターフェース接続法におけるエネルギー利用効率に注目している。従来研究とは異なり、深層強化学習における計算処理及びデータ移動による消費電力を考慮し、無線アクセス法の総合的な消費電力を導出している。この解析により、提案法はエネルギー利用効率を考慮し、探索する AP 数を設定することが可能となる。計算機評価により、提案法は各ユーザの消費エネルギーと無線通信特性の優れたトレードオフを達成することを明らかにしている。従来法と比べてエネルギー利用効率を向上し、提案法の有効性を確認している。

第 6 章ではサブ 6 GHz・ミリ波帯の統合的ネットワークの信頼性向上を目指し、リスクアバースな強化学習を活用した無線アクセス最適化において、問題の定式化・提案法及び計算機評価について述べている。既存法とは異なり、提案法では瞬時チャネル情報は未知といった厳しい現実条件の下で、要求信頼性レベルを達成している。計算機評価により、

従来法と比べて提案法はシステム全体の信頼性を改善することを明らかにしている。

最後に第 7 章では研究結果をまとめ、未解決課題と今後の方向性を提示した。

公開発表会では博士論文の章立てに従って発表が行われた。その後行われた論文審査会及び口述試験では、審査委員からの質疑に対して適切に回答がなされた。

質疑応答後に審査委員会を開催し、審査委員で議論を行った。審査委員会では、出願者の博士研究の学術レベルが十分高く、実用的な意義も有していることが評価された。

以上を要するに本学位論文は、今後の厳しい通信要求やモバイル環境への対応に向け、無線アクセスやインターフェース接続の最適化における諸問題の解決方法とその有効性を示したものであり、情報通信研究分野の発展に寄与しているという点で学術的価値が大きい。また、本学位論文の成果は、学術雑誌論文 2 件、情報学専攻の定めるトップ国際会議論文 2 件、及び国際会議論文 2 件（フルペーパー査読付き）として発表され、社会的な評価も得ている。

以上の理由により、審査委員会は、本学位論文が学位の授与に値すると判断した。