

氏 名 LE Tien Thanh

学位(専攻分野) 博士(情報学)

学位記番号 総研大甲第 2576 号

学位授与の日付 2025 年 3 月 24 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 Distributed Learning Techniques for Dynamic Access  
Control in massive Machine-Type Communication

論文審査委員 主 査 計 宇生  
情報学コース 教授  
金子 めぐみ  
情報学コース 教授  
福田 健介  
情報学コース 教授  
栗本 崇  
情報学コース 教授  
John C.S. Lui  
香港中文大学 教授

# Summary of Doctoral Thesis

Name in Full : Le Tien Thanh

Title : Distributed Learning Techniques for Dynamic Access Control in massive Machine-Type Communication

Massive Machine-Type Communication (mMTC) is vital for supporting numerous Internet of Things (IoT) devices in applications such as smart cities, smart factories, and environmental sensing. The number of connected devices is projected to reach 34.7 billion by 2028, a significant increase from the current number. However, available wireless resources are restricted by the scarcity of radio frequencies, and certain frequency bands must be reserved for human communication. Consequently, the limited wireless resources must be shared efficiently among an ever-increasing number of connecting IoT devices. This dissertation focuses on enhancing multiple access control protocols to address the new scaling challenges introduced by mMTC.

Traditional centralized resource allocation methods lead to significant signaling overhead, while uncoordinated resource contention schemes, which refer to the situation where multiple devices try to access the same set of wireless resources without coordination, fail to fully utilize the available bandwidth. We explore distributed learning techniques to improve uncoordinated resource contention in mMTC networks, proposing two tailored multi-agent learning methods for two well-known mMTC system models, namely TinyQMIX and MFTTS.

TinyQMIX is designed for dynamic channel selection using grant-free random access. It leverages the distributed decision-making capabilities of multi-agent deep reinforcement learning to optimize channel usage. In this context, we categorize devices into two groups: those in a regular state and those in an event-triggered state. Devices in a regular state only need to send simple status updates at low frequencies. In contrast, those in an event-triggered state must convey more critical information to a centralized controller due to its higher value, resulting in an increased transmission rate. We consider a dynamic traffic arrival scenario where the state of each device changes, hence the distribution of traffic for each device also varies over time.

This work introduces a practical and efficient cooperative multi-agent deep reinforcement learning algorithm to determine the optimal cooperative channel selection policy for each mMTC device, ensuring a high average global success transmission rate. Additionally, we modify QMIX, a popular cooperative multi-agent deep reinforcement learning algorithm, to reduce its computational complexity, making it easier to implement on IoT devices. TinyQMIX effectively minimizes transmission delays by enabling devices to independently select optimal channels based on local observations. Our findings indicate that an effective distributed resource allocation method can outperform centralized approaches, primarily due to the reduced costs associated with information exchange with a central controller. Furthermore, our results support the hypothesis that reinforcement learning can effectively generalize and adapt to new changes when trained on diverse patterns. Our proposed method achieves performance levels closely aligned with theoretical performance bounds while maintaining a low computational footprint suitable for typical IoT devices.

For TinyQMIX, we assume that only a single device can transmit data on one channel at a time, leading to a high probability of data collisions. Then we address this issue using a recent technique called unsourced random access, a method that allows all available channels to form a shared codebook, enabling multiple devices to use the same set of channels without causing data collisions. Additionally, standard multi-agent learning techniques like QMIX do not scale well as

the number of users increases to thousands, requiring decomposing thousands of devices into smaller groups. Therefore, we propose a novel approach called Mean-field Transfer Thomson Sampling (MFTTS) to solve the distributed power control problem in the hybrid Non-Orthogonal Multiple Access (NOMA) systems in the context of mMTC. MFTTS focuses on distributed power allocation in unsourced random access scenarios by employing a decentralized approach using Thompson Sampling, a well-known multi-armed bandit algorithm, to select the optimal power level for each mMTC device. We demonstrate that as the number of devices in the system increases, the behavior resembles an average system of ordinary differential equations (ODEs). Our proposed method, MFTTS, initializes local power selection models based on the convergence behavior of approximated large-scale networks, which can be obtained by solving the system of ODEs. MFTTS addresses key requirements of mMTC devices by being decentralized and lightweight while adaptively minimizing power consumption. We prove and quantify the conditions necessary for sub-linear regret in the system dynamics, indicating that the system can learn to improve its power allocation decisions.

In summary, we explored distributed (multi-agent) learning techniques for the wireless resource allocation problem in mMTC networks. Distributed learning is well-suited for this problem because knowledge about the network state is inherently distributed among many devices. For instance, channel conditions and traffic demands are localized within each device. Gathering this information centrally at the base station consumes more wireless channel resources than deciding resource selection locally at each device. We proposed TinyQMIX and MFTTS for two widely-used mMTC system models. While TinyQMIX can adapt better to changes, MFTTS can scale to accommodate thousands of players or agents.

Results of the Doctoral Thesis Defense

## 博士論文審査結果

Name in Full

氏 名 LE Tien Thanh

T i t l e

論文題目 Distributed Learning Techniques for Dynamic Access Control in massive Machine-Type Communication

本学位論文は英語で書かれており、「Distributed Learning Techniques for Dynamic Access Control in massive Machine-Type Communication (和訳: 分散学習技術を用いた多数同時接続通信のための動的接続制御)」と題し、全5章から構成されている。

第1章と第2章では研究の背景と関連研究について紹介した。IoT システムでの大規模応用が期待されている多数同時接続通信 (mMTC) システムのトラヒックは、上りリンク方向が主で、パケット長が短く、かつ通信チャネル数に対して接続デバイス数が圧倒的に多いことが特徴である。大量な IoT デバイスを同時に接続するための遅延を低減させ、通信スループットを向上させるために、マルチエージェント分散学習の手法を利用して、制御のオーバーヘッドが小さく、分散デバイス間の自律的協調を可能にする接続制御メカニズムが必要であると説明した。

第3章では、mMTC システムにおける動的チャネル選択の研究内容について紹介した。研究では、デバイスによるチャネル選択の問題を分散的部分観測可能なマルコフ決定過程として表現し、マルチエージェント深層強化学習モデルを利用した接続制御方法を提示した。大規模分散システムで課題となる制御のオーバーヘッドを回避し、計算能力が限られた IoT デバイス上でも分散実行可能にするために、集中訓練・分散推論が可能な軽量の学習モデルを採用した。分散学習のエージェント間の自律的協調による提案手法では、基地局による集中的なチャネル割当のオーバーヘッドを回避しながら、集中制御による上限性能に近い性能を有し、トラヒックパターンの変化にも一定程度適応可能であることをシミュレーションの結果で示した。

第4章では、より多くのデバイスを収容できる、電力と符号化の混合非直交多元接続システムについて、デバイスが自律的に電力レベルを選択するための研究内容について紹介した。研究では、復号化のための最小電力要求を満足しながら、平均送信電力を最小化する問題をマルチプレイヤ多腕バンデット問題に帰着し、そのための学習アルゴリズムを提示した。トンプソンサンプリング法を利用して、多腕バンデットの探索と利用のバランスを図り、平均場近似を利用して、デバイス数が数千以上規模のシステムにおける均衡状態を導出することで学習の収束過程を短縮させることに成功した。また、システム規模が大きい場合に均衡状態に収束できることを理論的に証明した。性能評価では提案手法が既存の分散制御方式よりも高い性能を有し、集中制御システムの性能により近い性能を有すると同時に、計算遅延と制御のオーバーヘッドが大幅に削減できることを示した。

最後に第5章で結論をまとめ、研究の適用範囲に関する議論を行い、将来の方向性を提示した。

公开发表会では博士論文の章立てに従って発表が行われ、その後に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に高く、研究内容が多数同時接続通信システムの性能向上のための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、大規模 IoT システムのための多数同時接続通信において、その特有な通信トラヒックの特徴に適した分散アクセス制御と資源割当の課題について、マルチエージェント強化学習とマルチプレイヤ多腕バンデットの分散学習の手法を利用した解決方法とその有効性を示したものである。本学位論文の成果として、国際会議論文 1 件とジャーナル論文 1 件が発表され、学術コミュニティからの評価も得ている。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。