

審査の結果の要旨

氏 名 シュレストマリ サソット

本論文は「Reinforcement Learning-based Optimization in Energy Harvesting Wireless Sensor Nodes」と題し、7つの章から構成される。近年、あらゆるものがネットワークにつながるInternet of Things (IoT) が急速に実現されてきており、中でも太陽光などから発電する環境発電で動作する無線接続されたセンサノードが広く普及することが予想されている。しかし、給電能力が周囲の環境に依存して変動するため、動作時間をできるだけ長く、しかも給電不足による動作の中断を生じさせないためには、センサノードの動作を環境に応じて動的に最適化することが必須となる。本論文は、強化学習 (reinforcement learning) を用いてその動的最適化を実現することを目的としている。

第1章「Introduction」では、本論文の背景と目的、貢献、及び構成について述べている。まずenergy neutral operation (ENO) つまり、発電量と電力消費量をバランスさせながら動作時間を最大化する運用が重要であることを述べた後、ENOを強化学習で実現する際の課題と、その課題に対する本論文の貢献について述べている。

第2章「Literature Overview」では、これまでの関連研究についてまとめており、ENOを実現する解析的手法と学習に基づく手法、および、目的関数が複数存在する場合の最適化手法について、これまでの研究を体系立てて整理し述べている。

第3章「Theory」では、本論文がその実現を目標とするENOの理論的な定式化を述べた後、本論文がセンサノードの動作最適化手法として着目する強化学習の理論について、Markov Decision Processとの関係、ならびにQ-value, Q学習, SARSA, Deep Q Network, などについてその特徴を述べている。また目的関数が複数存在する場合の強化学習である、Multi-Objective Markov Decision Processの形式的な定義も示している。

第4章「General framework for adaptive policies」では、太陽光パネルとバッテリーを有する一般的なセンサノードにおいて、稼働時間の割合を調整することでENOを実現する問題に対し、設置環境に依存した個別最適化をすることなく、強化学習により動作の最適化を実現する汎用的なフレームワークを提案している。評価では、天候の異なる2か所、具体的には稚内と東京の気象データを用いることで、提案するフレームワークが環境への優れた適応性を有することを示している。また、バッテリーの劣化や太陽光パネルの大きさを半減させた場合についても評価を行い、提案手法が多様な環境に対し動作の最適化を汎用的に実現可能であることも示している。

第5章「Coordinated Exploration with Distributed Reinforcement Learning」では、深層強

化学習(Deep Reinforcement Learning)を用いたセンサノードの動作最適化を論じている。深層強化学習はその学習能力は高いが、計算量が多く学習するまでに時間を要し、計算資源に制約があるセンサノード等のIoTシステムでは、学習が完了するまで安定した運用ができないという問題がある。この問題に対し、分散環境下で学習結果を共有する Distributed Reinforcement Learning (DiRL)を適用すると、探索が並列に行われるため学習時間は短くなるが、すべての学習結果を共有するため優先すべき局所的な学習が反映されるまでには時間を要する。センサノードが多様な環境に分散配置される状況では、局所的な学習を早期に反映させることが安定な動作の実現には必要である。一方、探索空間を分割し局所的に学習すると、学習結果が共有されないため最終的な品質が低下する。そこで、異なる環境のセンサノードによる探索を動的に協調させる新しい手法を提案し、従来の深層強化学習に比べ、学習に要する時間を 1/50 に短縮できることを示している。

第6章「Multi-objective Framework for ENO」では、複数の目的をもつセンサノードの動作最適化を実現する強化学習について論じている。センサノードはセンシング、処理、通信などの異なるタスクを実行しており、ユーザ要求を満たしながらこれらの処理のエネルギー消費量を最適化する必要がある。多目的強化学習 (Multi-Objective Reinforcement Learning: MORL) を適用することが考えられるが、従来のMORLアルゴリズムは計算コストが大きいという課題がある。そこで、本章では、ユーザが望む複数の性能要件の中でその優先度に応じたトレードオフを考慮でき、しかも学習コストを低減できる新しいMORLのフレームワークを提案し、その有効性を示している。

第7章「Conclusion」では、本論文が提案する手法とその貢献を要約している。また、残された課題を整理し、今後必要となる研究を論じている。

以上を要するに、本論文は環境発電で動作する無線接続されたセンサノードに対し、強化学習を用いてその動作を最適化する方法を提案し、有効性を明らかにしたものであり、その成果は情報理工学の発展に大きく貢献すると考えられる。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。