

2022年度 卒業論文

マルチエージェント協調巡回問題における
エネルギー消費抑制手法の提案

松本 航平

早稲田大学 基幹理工学部
情報理工学科

学籍番号 1W193102

提出日 2023/

指導教授 菅原 俊治

目次

| | | |
|----------|--|----------|
| 1 | 序論 | 1 |
| 2 | 関連研究 | 2 |
| 3 | モデルの定義 | 3 |
| 3.1 | 環境 | 3 |
| 3.2 | エージェント | 3 |
| 3.3 | 評価指標 | 3 |
| 4 | 準備 | 3 |
| 4.1 | Adaptive meta target decision strategy (AMTDS) | 3 |
| 4.1.1 | 目標決定戦略 | 3 |
| 4.1.2 | 経路生成戦略 | 6 |
| 4.2 | AMTDS with learning of dirt accumulation (AMTDS/LD) | 6 |
| 4.3 | AMTDS for energy saving and cleanliness (AMTDS/ESC) | 6 |
| 4.3.1 | 要求充足の判断 | 6 |
| 4.3.2 | 自己重要度評価 | 6 |
| 4.3.3 | 帰還動作 (Homing) | 6 |
| 4.3.4 | 待機動作 (Pausing) | 6 |
| 5 | 提案手法 | 6 |
| 5.1 | AMTDS for energy saving under the requirement (AMTDS/ER) | 6 |
| 5.1.1 | Homing と Pausing の組み合わせ | 6 |
| 5.1.2 | 補正係数 K の導入 | 6 |
| 5.1.3 | 未来のイベント発生量の予測 | 6 |
| 5.2 | AMTDS for energy saving under the requirement with learning of event probabilities (AMTDS/ERL) | 6 |
| 5.2.1 | イベント発生量の予測に使用するノードの範囲の変更 | 6 |
| 5.2.2 | 補正係数 K^i の更新方法の変更 | 6 |
| 6 | 評価実験 | 6 |
| 6.1 | 実験環境 | 6 |
| 6.2 | AMTDS/ER についての実験結果 | 6 |
| 6.2.1 | 実験 1: 性能評価 | 6 |
| 6.2.2 | 実験 2: 環境の変化による性能の違い | 6 |
| 6.2.3 | 実験 3: エージェント数減少による性能の変化 | 6 |
| 6.2.4 | 実験 4: K^i の降順にエージェント数を減少させたときの性能の変化 | 6 |
| 6.3 | AMTDS/ERC についての実験結果 | 6 |
| 6.3.1 | 実験 5: 性能評価 | 6 |
| 6.3.2 | 実験 6: 環境の変化による性能の違い | 6 |
| 6.3.3 | 実験 7: エージェント数減少による性能の変化 | 6 |

| | | |
|-------|---------------------------------------|---|
| 6.3.4 | 実験 8: K^i の降順にエージェント数を減少させたときの性能の変化 | 6 |
| 7 | 結論 | 6 |

概要

後で追記する

1 序論

後で追記する

近年、ロボット技術が発達し、巡回パトロールや清掃などといった、人間が日常的に繰り返す作業や、災害地や原子力発電所、宇宙などでの作業を複数の自律的ロボットで代替する動きが加速している。このような複数のロボットが協調して共通の作業を行う問題は、ロボットを自律的に動作するエージェントとしてモデル化し、マルチエージェント協調巡回問題 (*multi-agent cooperative patrolling problem*, MACPP) と呼ばれる。MACPP の研究では、複数のエージェントが協力・協調することで、与えられた環境において効率的かつ効果的に巡回を行うための方法・アルゴリズムを見出すことを目的としている。

巡回効率だけを追求した高度な行動や学習は、確かに巡回の効率は向上するものの、必要以上にエネルギーを消費する可能性があり、これはMACPPの極めて大きな課題となっている。特に、本研究で想定する自律型エージェントは独自のバッテリーを持ち、頻繁な充電を強いられることになる。一方、アプリケーションによっては、巡回作業に対する品質要求があり、それを超えることは必ずしも期待されているわけではない。例えば、清掃問題では、ある程度環境がきれいになっていれば十分であり、過度の巡回作業はかえって単位エネルギー当たりの作業効率を低下させる。また、環境が複雑で大規模な場合、どれだけのエージェントが必要かを事前に判断することができないこともあり、エージェントが少ないと品質要求を満たせず、多すぎるとエネルギーの浪費に繋がる。

マルチエージェントシステムにおける協調の観点からエネルギー効率に注目した研究 [1, 2, 3] がある。例えば、[2] では、マルチロボットにおける探索問題において、移動ロボットの総運動エネルギーを節約するために、運動時間を短縮する分散型協調手法を提案している。しかし、これらの研究は、タスクのために効率的に移動・作業し、結果として総消費エネルギーを削減する方法を目指したものである。一方、本手法では、他のエージェントが要求されるタスクの質を満たすことができれば、あるエージェントが自らの判断でしばらく停止したり、システムから退出したりすることを実現しようとするものである。また、[4, 5] では、エージェントが要求を満たせば自律的に一時停止してエネルギーを節約し、再び充電して探索に戻るという手法を提案している、しかし、その方法は不十分であり、エージェントは依然として不必要に環境内を巡回していることが分かった。

そこで本研究では、[5] の手法を拡張し、その後の行動によって起こりうる貢献を自律的に予測し、環境の現状を推定することで要求品質の全体的な達成度を把握しながら、より効率的に要求品質の充足とエネルギー消費の削減を両立する手法を提案する。主な違いは、エージェントが休止・充電している間の巡回タスクの進捗は、他のエージェントの行動に依存し、エージェントごとに異なるため、個々のエージェントから見た省エネルギー行動の自律学習を導入しているという点である。

さらに、この学習が進むとエージェントは、短い休止時間で巡回する Busy グループと、比較的長い時間休止してエネルギーを消費しない Energy save グループに分かれることを発見した。そのため、後者のグループのエージェントは、品質要求を満たしたまま、順次、活動を停止することができる。実験の結果、提案手法により、従来手法と比べて大幅にエネルギー消費量を削減することができることが分かった。また、Busy グループに属するエージェントの数は、環境条件によって変化することが分かった。その後、品質要求を満たしつつ、Energy save グループに属するエージェントを順次停止することで、活動しているエージェント数を削減することができた。

2 関連研究

後で、杉山さんの研究も追加する

MACPP とその応用に関する研究は数多く行われている [6, 7, 8, 9, 10]。例えば、[9] では、*edge Markov evolution graphs* を用いた動的環境変化のモデルを提案した。また、[10] では、巡回問題をバイズ適応型遷移分離部分観測可能マルコフ決定過程として定式化し、モンテカルロ木探索法を拡張した分散型オンライン学習・計画アルゴリズムを提案している。さらに、[11] では、人口ニューラルネットワークを用いて、個人の idleness から共用の idleness を予測する方法を提案している。しかし、これらの研究では、バッテリーの放電やエネルギー消費による定期的な充電を無視し、効率性のみを考慮したものである。[12] では、Q 学習によって決定される計画戦略のもと、複数のエージェントが定期的に充電しながら協調して巡回する AMTDS(*adaptive meta-target decision strategy*) という手法を提案している。さらに、[ここから杉山さんの研究](#)。しかし、この研究もエネルギー消費量の削減を考慮せず、巡回の効率化を目的とした学習のみを行うものである。

一方、[1, 2, 3, 1, 5, 13] では一部、省エネルギーに着目している。[1] では、消防ロボット以外にも複数のサブロボットを導入し、消火作業における総作業時間の延長を図った。また、[13] では、協調型群ロボットに対して、採餌などの連続タスクを解決するためのエネルギー配慮型分散タスク配分アルゴリズムを提案し、高効率なミッションの実現を実現した。しかし、これらの研究では、本研究とは異なり、動作時間の延長も考慮されている。これに対して、[5] では、AMTDS を巡回の品質要求に合わせて拡張し、省エネルギー化を図った。しかし、この手法によるエネルギー削減は不十分であり、エージェントの活動にはまだ不要な行動が含まれている。そこで、本研究では、個々の視点からの学習を導入することで、品質要求を満たしながら、さらにエネルギー消費を削減した。さらに、複数のエージェントを停止する手法を提案し、消費エネルギーの削減を図った。

3 モデルの定義

3.1 環境

3.2 エージェント

3.3 評価指標

4 準備

4.1 Adaptive meta target decision strategy (AMTDS)

4.1.1 目標決定戦略

Random selection (R)

環境全体のノード集合 V からランダムに v_{tar}^i を選ぶ.

Probabilistic greedy selection (PGS)

環境全体のノード集合 V 内のノード v におけるイベント発生量の推定値 $EL_t^i(v)$ の上位 N_g 個のノードから, ランダムに 1 つ v_{tar}^i を選ぶ. この際に, 学習や訪問をする v_{tar}^i の偏りを防ぐため, N_g 番目のノードと $EL_t^i(v)$ の値が同じノードが存在する場合, そのノードをすべて含めた後, その中から v_{tar}^i をランダムに選んでいる.

Prioritizing unvisited interval (PI)

環境全体のノード集合 V 内のノード v における訪問間隔 $I_t^i(v)$ の上位 N_i 個のノードから, ランダムに 1 つ v_{tar}^i を選ぶ. この際に, 学習や訪問をする v_{tar}^i の偏りを防ぐため, N_i 番目のノードと $I_t^i(v)$ の値が同じノードが存在する場合, そのノードをすべて含めた後, その中から v_{tar}^i をランダムに選んでいる.

Balanced neighbor-preferential selection (BNPS)

近隣のノードにイベント発生量が多いとエージェントが判断したとき, 近隣を優先的に巡回する. v_{tar}^i の決定時にエージェントの現在地 v_t^i との距離が d_{rad} 以下のノード集合を近領域 V_{area}^i とする. ここで, V_{area}^i における 1 ステップあたりのイベント処理量の期待値 EV_t^i は以下の式で求められる.

$$EV_t^i = \frac{\sum_{v \in V_{area}^i} EL_t^i(v)}{|V_{area}^i|} \quad (1)$$

エージェント i は近領域内のイベントを処理するか判断するための閾値 $EV_{threshold}$ と EV_t^i の値を比較し, $EV_t^i > EV_{threshold}$ の間は PGS によって近領域内から v_{tar}^i を選ぶ. その後, $EV_t^i \leq EV_{threshold}$ となった場合, 環境全体を対象とし, PGS で v_{tar}^i を選ぶ. 環境全体から v_{tar}^i を選択した後, V_{area}^i を更新する. 更新後の V_{area}^i の 1 ステップあたりのイベント処理量の期待値を EV_{t+1}^i とし, $EV_{threshold}$ の値を以下の式に従って更新する.

$$EV_{threshold} \leftarrow EV_{threshold} + \alpha(EV_{t+1}^i - EV_{threshold}) \quad (2)$$

ここで, $\alpha (0 < \alpha < 1)$ は学習率である. また, $EV_{threshold}$ の初期値は初めに V_{area}^i を設定した際の EV_t^i の値である.

表 1: エージェントに関するパラメータ

| 種類 | パラメーター | 値 |
|---------|----------------|-----|
| エージェント数 | —A— | 20 |
| バッテリー | B_{max}^i | 900 |
| | B_{drain}^i | 1 |
| | k_{charge}^i | 3 |
| 経路生成戦略 | d_{myopia} | 10 |
| | k_{att} | 1.0 |
| | k_{rover} | 1.2 |

表 2: 目標決定戦略のパラメーター

| 目標決定戦略 | パラメーター | 値 |
|----------|---------------|------|
| PGS | N_g | 5 |
| PI | N_i | 5 |
| BNPS | α | 0.1 |
| | d_{rad} | 15 |
| AMTDS | α | 0.1 |
| | ε | 0.05 |
| AMTDS/LD | β | 0.1 |

表 3: エネルギー節約行動に関するパラメーター

| 種類 | パラメーター | 値 |
|-----------|--------------|---------------|
| 自己重要度評価 | T_s | 20 |
| | T_l | 50 |
| | T_f | 10 |
| Homing | T_{check} | 100 |
| | k_{homing} | $\frac{1}{3}$ |
| Pausing | T_{basic} | 100 |
| AMTDS/ERL | T_{hp} | 500,000 |

4.1.2 経路生成戦略

4.2 AMTDS with learning of dirt accumulation (AMTDS/LD)

4.3 AMTDS for energy saving and cleanliness (AMTDS/ESC)

4.3.1 要求充足の判断

4.3.2 自己重要度評価

4.3.3 帰還動作 (Homing)

4.3.4 待機動作 (Pausing)

5 提案手法

5.1 AMTDS for energy saving under the requirement (AMTDS/ER)

5.1.1 Homing と Pausing の組み合わせ

5.1.2 補正係数 K の導入

5.1.3 未来のイベント発生量の予測

5.2 AMTDS for energy saving under the requirement with learning of event probabilities (AMTDS/ERL)

5.2.1 イベント発生量の予測に使用するノードの範囲の変更

5.2.2 補正係数 K^i の更新方法の変更

6 評価実験

6.1 実験環境

6.2 AMTDS/ER についての実験結果

6.2.1 実験 1: 性能評価

6.2.2 実験 2: 環境の変化による性能の違い

6.2.3 実験 3: エージェント数減少による性能の変化

6.2.4 実験 4: K^i の降順にエージェント数を減少させたときの性能の変化

6.3 AMTDS/ERC についての実験結果

6.3.1 実験 5: 性能評価

6.3.2 実験 6: 環境の変化による性能の違い

6.3.3 実験 7: エージェント数減少による性能の変化

6.3.4 実験 8: K^i の降順にエージェント数を減少させたときの性能の変化

7 結論

参考文献

- [1] Jeongwan Kim, J. Eric Dietz, and Eric T Matson. Modeling of a multi-robot energy saving system to increase operating time of a firefighting robot. In *2016 IEEE Symposium on Technologies for Homeland Security (HST)*, pp. 1–6, 2016.
- [2] Abdenour Benkrid, Abdelaziz Benallegue, and Noura Achour. Multi-robot coordination for energy-efficient exploration. *Journal of Control, Automation and Electrical Systems*, Vol. 30, No. 6, pp. 911–920, 2019.
- [3] Gennaro Notomista, Siddharth Mayya, Yousef Emam, Christopher Kroninger, Addison Bohannon, Seth Hutchinson, and Magnus Egerstedt. A resilient and energy-aware task allocation framework for heterogeneous multirobot systems. *IEEE Transactions on Robotics*, Vol. 38, No. 1, pp. 159–179, 2022.
- [4] Lingying Wu and Toshiharu Sugawara. Strategies for Energy-Aware Multi-agent Continuous Cooperative Patrolling Problems Subject to Requirements. In Matteo Baldoni, Mehdi Dastani, Beishui Liao, Yuko Sakurai, and Rym Zalila Wenkstern, editors, *PRIMA 2019: Principles and Practice of Multi-Agent Systems*, pp. 585–593, Cham, 2019. Springer International Publishing.
- [5] Lingying Wu, Ayumi Sugiyama, and Toshiharu Sugawara. Energy-efficient strategies for multi-agent continuous cooperative patrolling problems. *Procedia Computer Science*, Vol. 159, pp. 465–474, 2019.
- [6] Katsuya Hattori and Toshiharu Sugawara. Effective Area Partitioning in a Multi-Agent Patrolling Domain for Better Efficiency. In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART*, pp. 281–288. INSTICC, SciTePress, 2021.
- [7] Gleb K Tevyashov, Mark V Mamchenko, Andrey N Migachev, Rinat R Galin, Konstantin A Kulagin, Petr M Trefilov, Rodion O Onisimov, and Nikolay V Goloburdin. Algorithm for multi-drone path planning and coverage of agricultural fields. In *Agriculture Digitalization and Organic Production*, pp. 299–310. Springer, 2022.
- [8] Bernát Wiandt and Vilmos Simon. Autonomous graph partitioning for multi-agent patrolling problems. In *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 261–268. IEEE, 2018.
- [9] Mehdi Othmani-Guibourg, Amal El Fallah-Seghrouchni, Jean-Loup Farges, and Maria Potop-Butucaru. Multi-agent patrolling in dynamic environments. In *2017 IEEE international conference on agents (ICA)*, pp. 72–77. IEEE, 2017.
- [10] Xin Zhou, Weiping Wang, Tao Wang, Yonglin Lei, and Fangcheng Zhong. Bayesian reinforcement learning for multi-robot decentralized patrolling in uncertain environ-

- ments. *IEEE Transactions on Vehicular Technology*, Vol. 68, No. 12, pp. 11691–11703, 2019.
- [11] Mehdi Othmani-Guibourg, Amal El Fallah-Seghrouchni, and Jean-Loup Farges. Decentralized multi-agent patrolling strategies using global idleness estimation. In *International Conference on Principles and Practice of Multi-Agent Systems*, pp. 603–611. Springer, 2018.
- [12] Keisuke Yoneda, Chihiro Kato, and Toshiharu Sugawara. Autonomous learning of target decision strategies without communications for continuous coordinated cleaning tasks. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, Vol. 2, pp. 216–223, 2013.
- [13] Ehsan Latif, Yikang Gui, Aiman Munir, and Ramvivas Parasuraman. Energy-aware multi-robot task allocation in persistent tasks, 2021.