

卒業論文概要書

CD

2023年1月13日提出

学籍番号 1W193102

学科名	情報理工学科	氏名	松本 航平	指導 教員	菅原 俊治
研究 題目	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				

1 序論

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を複数の自律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、マルチエージェント協調巡回問題 (MACPP) と呼ばれる。MACPPでは、複数のエージェントが協力・協調することで、与えられた環境で効率的に巡回を行うことを目的とし、様々な研究が行われている [1, 2]。

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。

[2] では、エージェントが要求を満たせば自律的に充電基地への帰還 (Homing) や、充電基地での待機 (Pausing) を行い、エネルギーを節約する手法を提案している。本研究では、[2] 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品質要求の全体的な達成度を把握しながら、品質要求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。

2 モデルの定義

2.1 環境

エージェントの巡回環境をグラフ $G = (V, E)$ で表す。ここで、 $V = \{v_1, \dots, v_n\}$ はノード集合を表し、各ノードにエージェントやイベント、障害物が存在する。また、 E はノード v_i と v_j 間のエッジ $e_{i,j}$ である。本研究では、エッジの長さはすべて1とする。さらに、ステップを単位とする離散時間を導入する。したがって、エージェントは1ステップで障害物のない隣接ノードに移動することができる。

全てのノード $v \in V$ 上でイベントが発生し、そのイベント発生確率を $p(v)$ ($0 \leq p(v) \leq 1$) とする。毎時刻 t において、ノード v に蓄積されたイベント数 $L_t(v)$

は以下の式で更新される。

$$L_t(v) = \begin{cases} L_{t-1}(v) + 1 & (\text{確率 } p(v) \text{ のイベント発生時}) \\ L_{t-1}(v) & (\text{その他}) \end{cases}$$

時刻 t にエージェントがノード v を訪れた時に v 上のイベントは処理され、 $L_t(v) = 0$ となる。

2.2 エージェント

n 個のエージェントの集合を $A = \{1, \dots, n\}$ と表す。エージェント $i \in A$ はバッテリーを持ち、バッテリー容量が0になる前に充電基地に戻り充電し、充電完了後は再び環境を巡回する。

エージェントは目標決定戦略によって、目標ノード v_{tar}^i を決定する。その後、経路生成戦略によって、 v_{tar}^i までの経路を生成し、それに沿って行動を進める。これをバッテリー残量を参照しながら繰り返す。

2.3 評価指標

本研究では評価指標として以下の式で定められるイベント残存時間の総和 D_{t_s, t_e} と、エージェントの総エネルギー消費量 C_{t_s, t_e} を用いる。

$$D_{t_s, t_e} = \sum_{v \in V} \sum_{t=t_s+1}^{t_e} L_t(v)$$

$$C_{t_s, t_e} = \sum_{i \in A} \sum_{t=t_s+1}^{t_e} \mathcal{E}_t(i)$$

ここで、 $[t_s, t_e]$ ($t_s < t_e$) は時間間隔を表し、 $\mathcal{E}_t(i)$ は t におけるエージェント i の消費エネルギーを表す。したがって、 i が隣接ノードに移動したとき $\mathcal{E}_t(i) = 1$ 、それ以外は $\mathcal{E}_t(i) = 0$ となる。

また、[2] と同様に1stepにおけるイベント量の要求値 D_{req} を設定した。エージェントは以下の式を満たせるように協調を行う。

$$D_{t_s, t_e} \leq D_{req} \times (t_e - t_s)$$

本研究では、品質要求 (上式) を満たしつつ、 C_{t_s, t_e} をできるだけ小さくすることが目的である。以降、それぞれを $D(s), C(s)$ と表す。

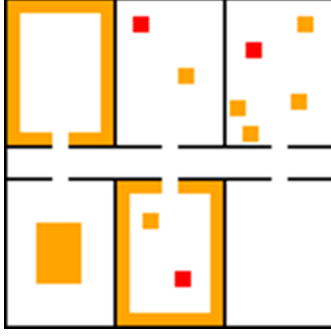


図 1: 実験環境

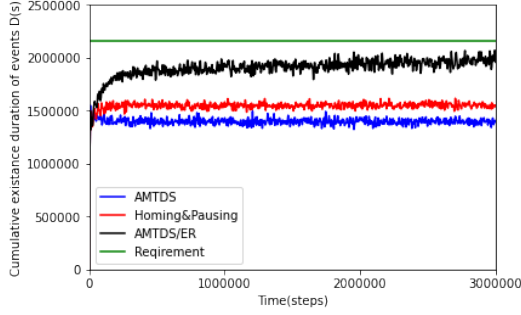


図 2: $D(s)$ の時間推移

3 提案手法

本研究では、エージェントがそれぞれの視点で将来の環境状態を予測する方法を学習し、充電基地について将来の環境状態を予測し、待機時間を動的に決定する手法として、AMTDS/ER を提案した。

まず、エネルギー節約行動の調整のため、学習パラメータ K^i を $\forall i \in A$ に導入する。Pausing 後にイベント量の推定値 $E^i(D_t)$ を用いて K^i は個々に次の式に従って更新される。

$$\begin{cases} K^i \leftarrow (1 - \alpha_k) \cdot K^i + \alpha_k \cdot \frac{D_{req}}{E^i(D_t)} K^i & (E^i(D_t) \leq D_{req}) \\ K^i \leftarrow K^i - \left(\frac{E^i(D_t)}{D_{req}} - 1 \right) & (E^i(D_t) > D_{req}) \end{cases}$$

次に、従来手法では Pausing は固定時間 S_{pause} ステップしか待機しなかったが、以下の式を満たす最大の $T (= \gamma_p \cdot S_{pause})$ まで待機するように変更した。ここで、 t_c は現在の時刻であり、 γ_p は可変長の変数である。さらに、Homing の後は必ず Pausing を行うようにした。

$$\sum_{v \in V} E^i(L_{t_c+T}(v)) \div K^i \leq D_{req}$$

4 実験

4.1 実験環境

実験環境は図 1 のような 101×101 の 2 次元グリッド環境である。黒線は壁を表す。図 1 の色に従って、

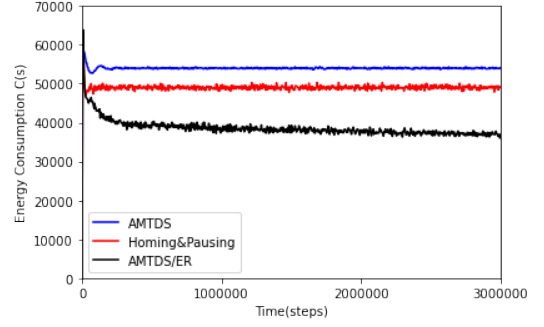


図 3: $C(s)$ の時間推移

$p(v)$ を次のように設定した。

$$p(v) = \begin{cases} 10^{-3} & (v \text{ が赤色の領域内の場合}) \\ 10^{-4} & (v \text{ がオレンジ色の領域内の場合}) \\ 10^{-6} & (v \text{ が白色の領域内の場合}) \end{cases}$$

エージェント数は 20 体である。 $D(s)$ と $C(s)$ の値は、独立した 50 回の試行による平均値である。

4.2 実験結果

従来手法である AMTDS や AMTDS/ESC と、提案手法である AMTDS/ER の $D(s)$ と $C(s)$ の時間推移を比較した結果を図 2, 3 に示す。なお、図 2 内の緑色の直線が品質要求値である。これらの図より、AMTDS/ER は従来手法よりも、品質要求を満たしながらエネルギー消費量を削減することができたといえる。

5 結論

本研究では、MACPP において、各エージェントがエネルギー節約行動を調整する手法を提案した。実験の結果、品質要求を満たしながらエネルギー消費量を削減できることを示した。

今後はエージェントが $p(v)$ を学習しながらエネルギー消費量の削減を図る必要がある。

参考文献

- [1] 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.
- [2] 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.