

強化学習を導入した津波避難行動シミュレーションの開発

東北大学工学部	学生会員	○中島 拓巳
東北大学災害科学国際研究所	非 会 員	モヤ ルイス
東北大学災害科学国際研究所	正 会 員	マス エリック
東北大学災害科学国際研究所	正 会 員	越村 俊一

1. はじめに

東日本大震災では津波からの逃げ遅れにより多くの犠牲者が出た。津波避難においては全員が短時間で避難できる方法が必要であり、適切な対策立案のため津波避難行動シミュレーションが行われている。既往の津波避難行動シミュレーションの多くは、避難経路に単純距離的な最短経路を仮定している。しかし選択経路が一つに集中することで混雑が生じ避難者の移動速度が遅くなることで、全体の避難完了時間が長くなる危険性がある。

本研究では混雑度を考慮して経路選択を行い全体の避難完了時間を短くすることを目的とし、強化学習手法の一つであるAction-Value法 (Sutton *et al.*, 2000)を用いた津波避難行動シミュレーションを開発した。混雑度を変更した複数のケースでシミュレーションを実行し、本手法の有効性を検証した。

2. モデルとアルゴリズム

(1) 環境の設定

シミュレーションを実行する環境として、東日本大震災前の仙台市荒浜地区のデータを参考に240のノードと312のリンクからなる無向ノードリンクネットワークを作成した。(図-1)それぞれのリンクの長さや避難所に指定するノードも震災前の仙台市荒浜地区を参考にして決定した。

(2) エージェントの設定

震災前の仙台市荒浜地区のデータを参考に2732人のエージェントを設定した。エージェントにはそれぞれ出発するノードが与えられており、シミュレーション開始時間に一斉に避難を開始しノードリンク上を移動しながら避難所に指定されているノードを目指す。エージェントの移動速度 v はあるノードから次のリンクに移る時にそれぞれ与えられ、移動速度はそのリンク密度 d によって決

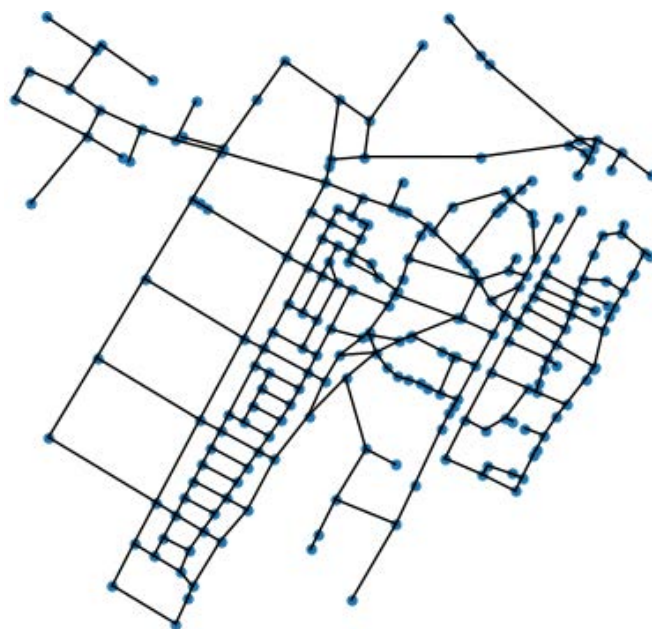


図-1 使用するノードリンクネットワーク
定する。リンク密度は、そのリンクの長さやリンク上にいるエージェントの数によって決定する。移動速度とリンク密度の関係式は (Takabatake *et al.*, 2017) を参考に作成し、式(1)に示す。

$$\begin{aligned} v &= 1.19 \quad (d < 0.5) \\ v &= -0.396d + 1.388 \quad (0.5 \leq d < 3.0) \\ v &= 0.2 \quad (3.0 \leq d) \end{aligned} \quad (1)$$

(3) 経路選択アルゴリズム

本研究ではエージェントが混雑を避け自律的に経路選択することを目的に、エージェントの行動価値を数値化した価値関数 Q を更新していく強化学習を導入する。強化学習手法の中でも報酬の与え方のみを決定することで使用できるAction-Value法を使用する。試行回数 $i+1$ と i の価値関数の関係式を式(2)に示す。価値関数はそれぞれのノードに設定されており一回の試行回数終了する度に更新される。また全ての価値関数の初期値は0.5とした。

$$Q_{i+1} = \frac{Q_i \times N_i + R}{N_i + n} \quad (2)$$

表-1 実行したシミュレーションケース

ケース番号	道路幅	経路選択方法
ケース1	2.0 m	ダイクストラ法
ケース2	2.0 m	強化学習アルゴリズム
ケース3	0.10 m	ダイクストラ法
ケース4	0.10 m	強化学習アルゴリズム

N_i は i 回目の試行回数まででそのノードが使用された総数であり、 n は $i+1$ 回目のみでのそのノードの使用数である。 R は $i+1$ 回目で得られた報酬の総数であり、報酬はそのノードを使用したエージェントが一回の試行回数のシミュレーション終了時間までに避難所のノードに達成できるかで決定する。避難達成できた場合の報酬値は1、達成できなかった場合は0とした。また一回の試行回数のシミュレーション時間は、東日本大震災時の津波到達時間を参考にし67 minとした。エージェントはあるノードに到着した時に次のノードを選択する。価値関数が大きいノードを選択する割合を90%とし、価値関数に関係なくランダムにノードを選択する割合を10%とした。

3. シミュレーションの結果と考察

(1) シミュレーションの条件

シミュレーションの条件は道路幅と経路選択方法を変更した4ケースである。(表-1)ダイクストラ法とは単純距離的な最短経路を選択する、数理的アルゴリズムである。強化学習アルゴリズムの試行回数は3000回とした。

(2) シミュレーションの結果と考察

強化学習アルゴリズムの試行回数ごとの避難達成人数の累積平均を図-2に示す。試行回数が大きくなるに従って避難達成者の累積平均も大きくなっていき、収束に向かう様子が確認できた。

次にそれぞれのケースにおける時間ごと避難達成人数を図-3に示す。ケース2とケース4では試行回数3000回目の結果を使用した。ケース1とケース2では全員が避難達成できた。ケース1はケース2に比べて早くに全員が避難達成していることから、混雑度が小さい場合にはダイクストラ法による経路選択が有効であると考えられる。またケース3とはケース4では全員が避難達成できなかった。これは混雑度が大きくなりエージェントの移動速度が小さくなったためであると考えられる。ケース4ではケース3に比べて避難達成人数は多くなった。これは強化学習アルゴリズムにより混雑度が小さくなるようにエージェントの選択経路が分散され、エージェントの移動速度が

ケース3に比べて小さくならなかったためであると考えられる。よって混雑度が大きい場合には強化学習アルゴリズムによる経路選択が有効であると考えられる。

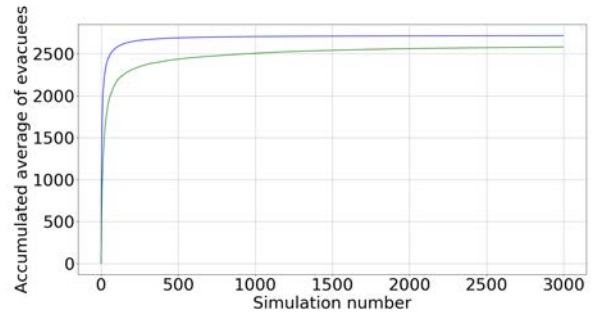


図-2 収束の様子、(青)ケース2、(緑)ケース4。

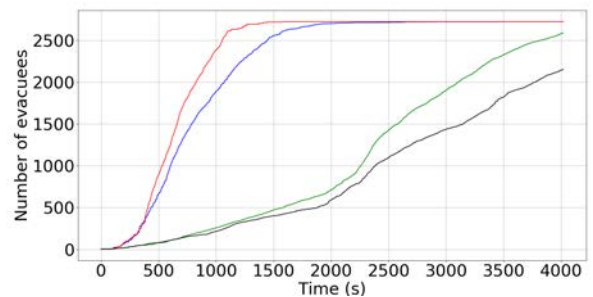


図-3 避難達成人数の比較、(赤)ケース1、(青)ケース2、(黒)ケース3、(緑)ケース4。

4. おわりに

本研究では混雑度を考慮して選択経路を分散させることを目的とし、強化学習を導入した津波避難シミュレーションを開発した。複数のケースでシミュレーションを実行し比較検討することで、強化学習アルゴリズムによる経路選択は混雑度が大きい場合に有効であることを定量的に示した。本研究で開発したアルゴリズムはエージェントの特性を増やすことや避難開始時間を変更することも可能である。今後は本研究で開発したシミュレーションを使用し、適切な津波避難対策の提言に役立てたい。

参 考 文 献

- Takabatake, T., Shibayama, T., Esteban, M., Ishii, H. and Hamano, G., Simulated tsunami evacuation behavior of local residents and visitors in Kamakura, Japan, *International Journal of Disaster Risk Reduction*, **23**, pp.1–14., 2017.
- Sutton, R.S. and Barto, A.G.(translated by Mikami, S. and Minagawa, M.), Reinforcemet Learning, *Morikita Publishing Co*, 2000.