

京都大学工学部 学生員 道頭 理緒奈
 京都大学防災研究所 フェロー 堀 智晴

1. 序論

2018年の西日本豪雨以降、自治体による治水整備や情報提供だけでなく、住民主体での避難・災害対応が被害を防止するために重要であるという認識が高まってきた。しかし、適切な時期に避難を判断し、行動を開始することは被災経験のない住民にとって決して容易ではない。大きな水害を経験することで防災や避難に対する意識が高まる事例がある一方、かえって経験が規模の異なる水害への対応を誤らせる事例もある。そこで本研究では、水害時に住民が様々な避難を経験し、身を守るための行動指針を獲得していく過程を、強化学習機能を付加したエージェントモデルを使って再現することを試みた。

2. 避難シミュレーションモデルの概要

従来、開発されてきた水害避難モデルでは、降雨流出氾濫(RRI)モデル[1]を用いて浸水深を計算している。避難行動場は、国土地理院作成の数値地図2500[2]を利用し、街路区間を表す線(アーク)と交差点部を表す点(ノード)で表現する。

モデル避難者は初期速度1.1(m/sec)で自宅から避難を開始し、浸水を避けながら、避難所の最寄りのノードまでアークの上を移動する。あるノードに到達するとそのノードに接続しているアークの中から次に進むべきアークを選択する。周りが浸水していて他に経路の選択肢がない場合、モデル避難者は浸水しているアークを通る。その場合には、浸水深の影響により式(1)に従って歩行速度が低減する。式中の v_t は時刻 t における歩行速度(m/sec)、 d_t は時刻 t におけるモデル避難者の位置の浸水深(m)を示す。

$$v_t = 1.1 \left(1 - \frac{d_t}{0.7} \right) \quad (1)$$

式に従えば浸水深が0.7mのとき速度は0(m/sec)となるが、浸水深が0.7mを超える場合には歩行不能と判定し、その時点で避難失敗とする。本研究では細野[3]の結

果を用い、モデル避難者は降雨開始直後から降雨ピークまで10分ずつ避難開始時刻をずらして避難を行った。

3. 強化学習による水害時避難行動指針の獲得

先ほど述べたモデルのシミュレーション結果をもとに、強化学習を用いてエージェントモデルに避難の判断について学習させることを考える。

今回は経験を再現するという観点から、代表的な強化学習の手法の一つであるQ学習[4]を採用した。これは、ある状況において実行可能な選択肢を、その状況でその選択肢を選んだ結果得た報酬によって点数をつけ評価するものである。この点数はQ値と呼ばれ、同じ状況でその選択肢が選ばれるたびに式(2)によって更新される。

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha(r + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)) \quad (2)$$

式中の s_t は時刻 t での状況、 a_t は状況 s_t において選ばれた選択肢、 $\max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$ は次の時刻 $t+1$ で選択可能な選択肢のQ値のうちで最大のものを意味する。三つの文字 r , α , γ はそれぞれ報酬、学習係数、割引率を表す。一般に、正の報酬が得られQ値が高くなった選択肢ほど、その状況において好ましい選択肢であるといえる。今回は状況 s_t としてモデル避難者の自宅近くの河川水位を、各状況において実行可能な選択肢 a_t として指定された避難所へ避難を開始するか、自宅に留まるかの二択を設定した。学習のフローチャートは図1の通りである。学習は以下の二種類の報酬の与え方に対して開始時間が二通りずつの合計四パターンで行った。

- 1) 選択後自宅が浸水しなければ留まったことに報酬+1、選択後避難所にたどり着ける時刻であれば避難したことに報酬+1、それ以外は報酬-1を与える場合。
- 2) 避難所にたどり着ける最後の時刻に避難を選択すれば避難に報酬+1、それ以前では留まれば報酬+1、それ以降ではどちらを選んでも報酬-1を与える場合。

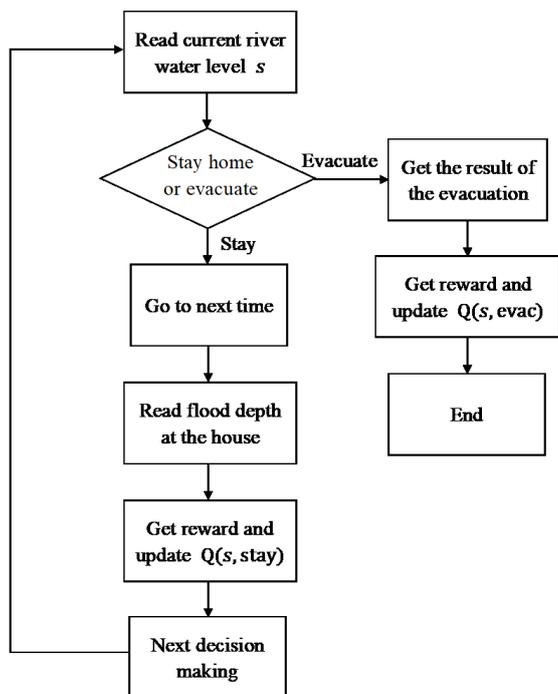


図1. 学習の手順

なお、上記の自宅が浸水する時刻、避難所にとどける最後の時刻はそれぞれ降雨開始後 670 分、580 分である。学習の開始時刻は降雨開始直後 (Case 1-1, 2-1) と、降雨開始後 270 分 (Case 1-2, 2-2) である。

4. 結果

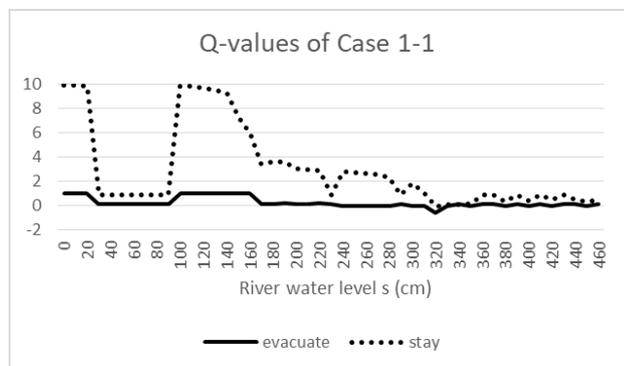


図2. Case 1-1 のQ値

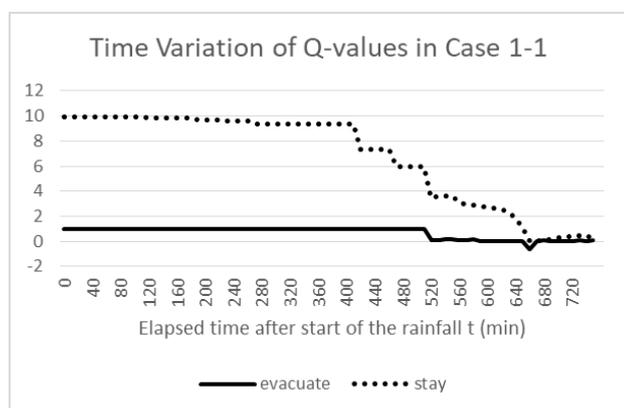


図3. Case 1-1 のQ値の時間変化

図2は1) の報酬の与え方をした際の、河川水位にQ値を表したものである。実線が避難、点線が自宅に留まることに対するQ値を表している。モデルにとって未体験の河川水位に対するQ値は初期値 (避難 0.1、自宅 0.9) のまま更新されず、低い水位のうちどちらのQ値も初期値の10倍ほどになっている。そしてまだ自宅が浸水せず避難も可能な安全であるはずのある水位を境に避難のQ値はほぼ0になり、自宅に留まるQ値は比較的緩やかに下降していく。

図3では降雨開始後の経過時間を横軸にとり、各時刻における河川水位に対応するQ値を表している。報酬が負になる時刻に向かって自宅に留まるQ値が下降していくことがわかる。

5. 結論

どのパターンにおいても、自宅に留まるQ値に関しては緩やかな値の変化が見られたが、避難のQ値は0か1に極端に分かれる結果となった。水害時に自宅に留まるべきかどうかは、自宅に留まるQ値に対して閾値を設定するなどしてエージェントに判断させることが可能であると思われる。一方、同様のことを避難のQ値をもとに判断させることは、今回の学習の結果からは困難であった。

参考文献

- [1] 降雨流出氾濫モデル (RRI モデル), ユネスコ後援機関 水災害・リスクマネジメント国際センター (ICHARM), 2019/02/13 accessed http://www.icharm.pwri.go.jp/research/rri/index_j.html
- [2] 数値地図 2500 (空間データ基盤), 国土交通省国土地理院, 2019/02/13 accessed <http://www.gsi.go.jp/geoinfo/dmap/dm2500sdf/index.html>
- [3] 細野時由: 人工知能技術を用いた水害時最適避難開始タイミング判断モデルの試作, 京都大学工学部地球工学科, 卒業論文, 2018
- [4] 牧野貴樹: 1.1 強化学習とは, これからの強化学習, 森北出版株式会社, 2016