



TOHOKU
UNIVERSITY

エージェントモデル及び強化 学習による災害時の避難誘 導システムの研究

Erick MAS (マス・エリック)

東北大学・災害科学国際研究所・広域被害把握研究分野(准教授)

タフ・サイバーフィジカルAI研究センター・HPC計算モジュール研究部門



TCPAI

Tough Cyberphysical AI Research Center - Tohoku University

交通渋滞



A traffic jam is seen as people evacuate after tsunami advisories were issued following an earthquake, in Iwaki, Fukushima prefecture, Japan, in this photo taken by Kyodo November 22, 2016. Mandatory credit Kyodo Kyodo/via REUTERS

Source: <https://www.reuters.com/article/us-japan-quake-evacuation/bitter-lessons-of-japans-2011-tsunami-put-to-use-with-latest-quake-idUSKBN13H0EC>

混雑



© Reuters

Tsunami warning evacuation in Indonesia. (photo credit: REUTERS)

Source: <https://www.jpost.com/International/Quakes-off-Indonesia-stir-panic-but-no-big-tsunami>



東北地方整備局

下閉伊郡山田町大沢
 R45大沢(241.8K)上り



渋滞回避（混雑を避ける）

最適避難経路

これまでの情報社会(4.0)



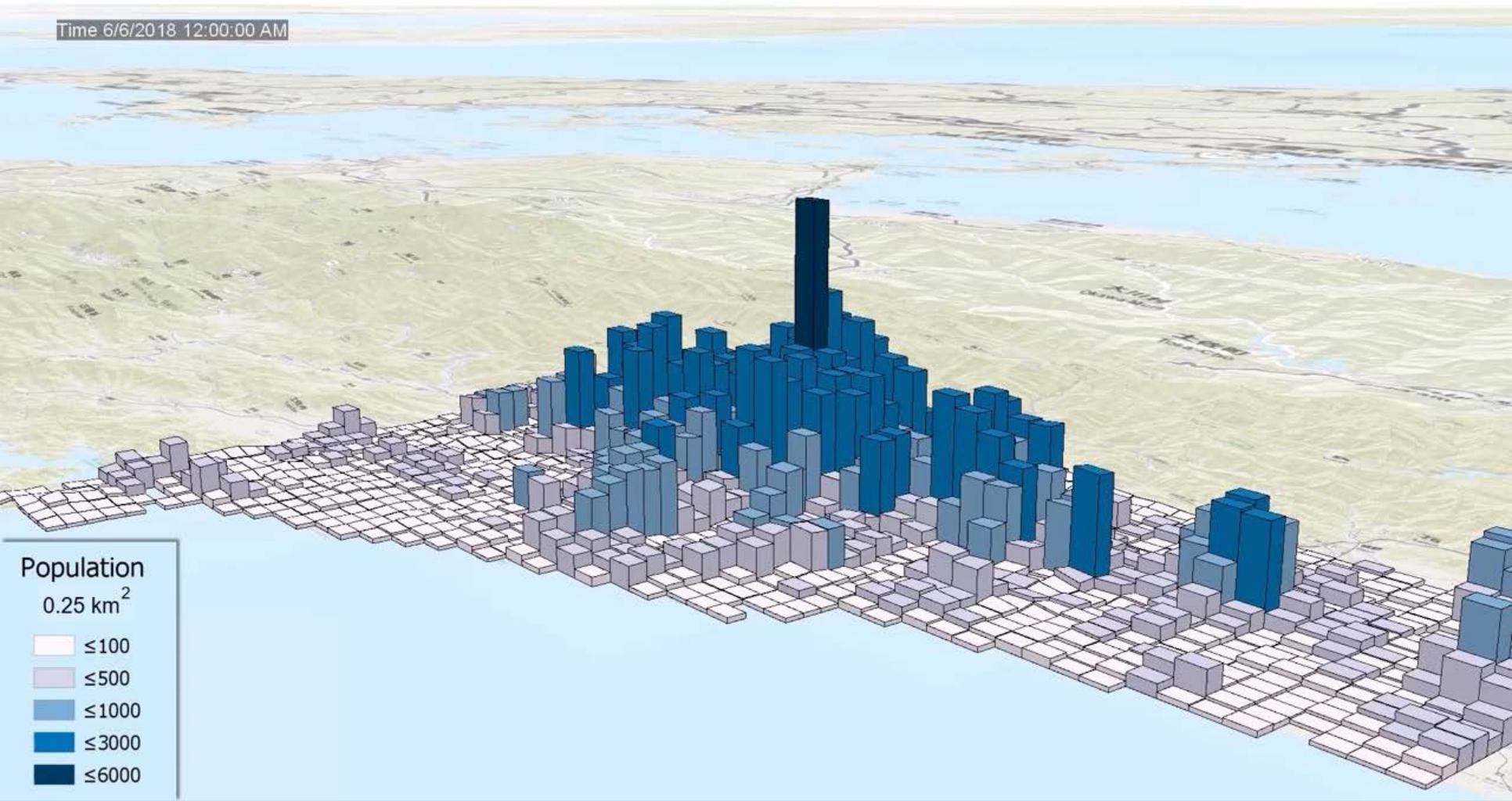
Society 5.0

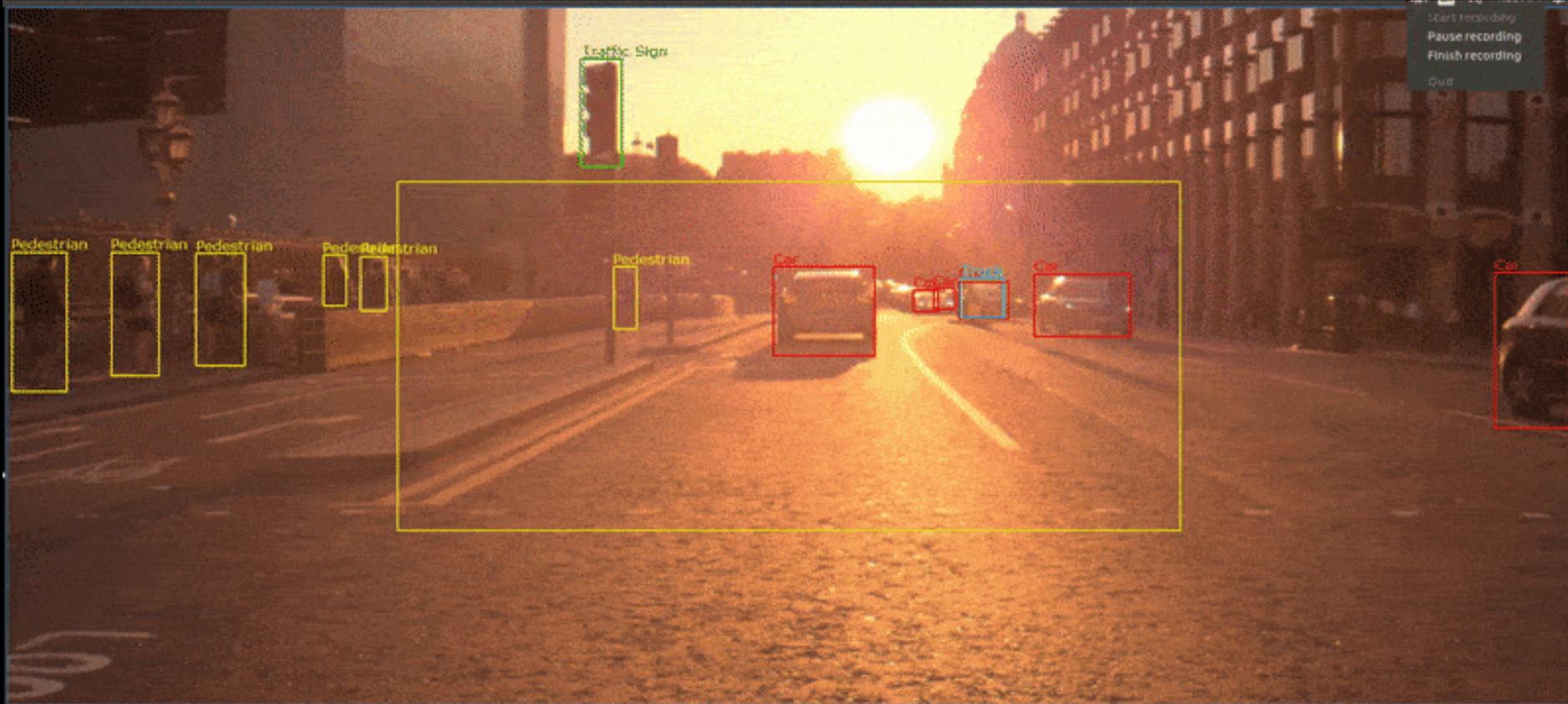


[内閣府作成]



Time 6/6/2018 12:00:00 AM





Real-time Neural Net Processing,
Westminster, London

www.academyofrobotics.co.uk

1. リアルタイム津波浸水予測と 強化学習による避難誘導最適化

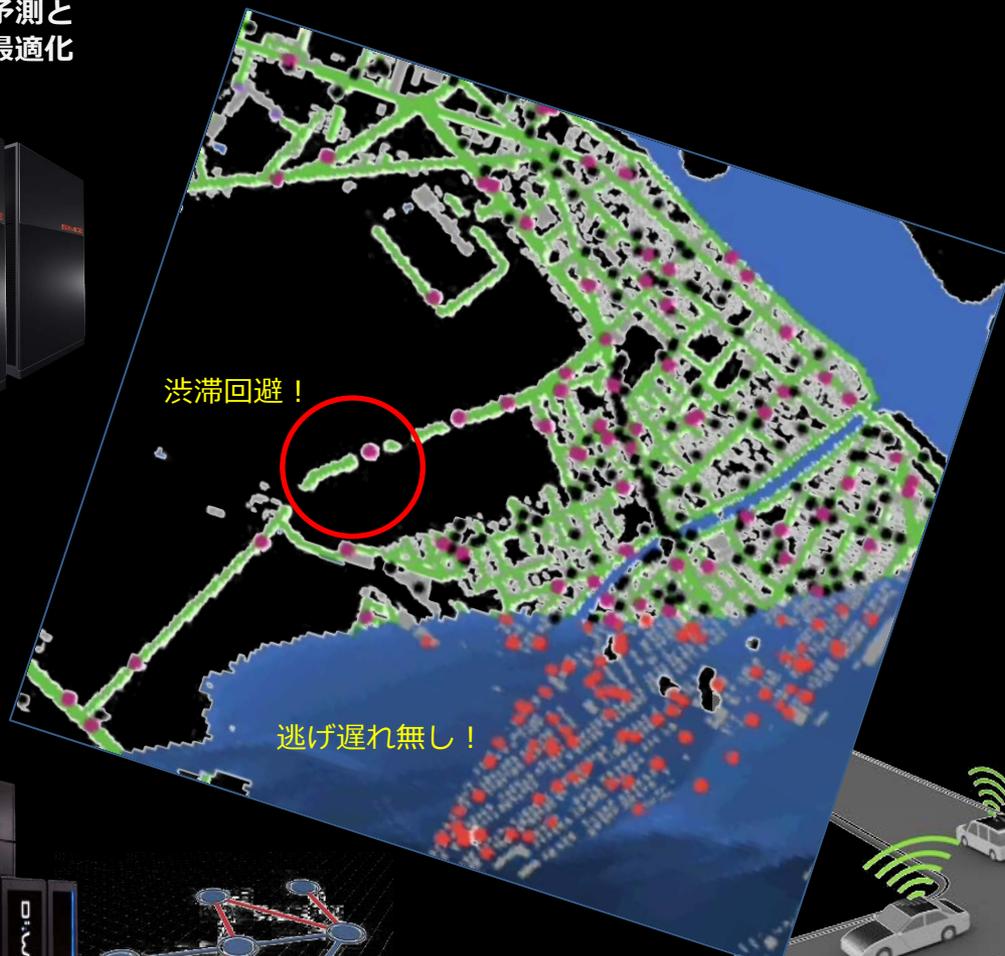


東北大SX-ACE/Aurora
(災害時運用可能)

2. 量子アニーリングによる 避難経路探索最適化



東北大
D-waveマシン



3. 人・車のリアルタイムモニタリングと避難誘導更新



4. 避難誘導システム (*)

(*)

(*) <https://technology-illustration.com/highresolution/027-illustration-technology.png>

避難経路選択の問題点

- 経路の全体的な被害状況を知らず
- 最適の道を選択されない時もある
 - 最適とは最短経路のみではない。さらに混雑していない、津波到達時間の余裕、高いところを向かう、などの意味を考慮する。



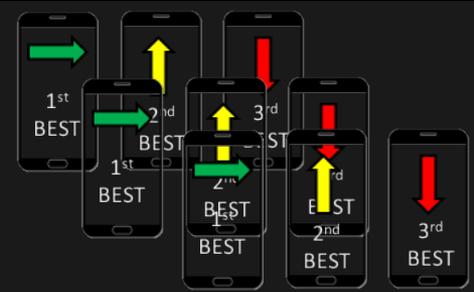
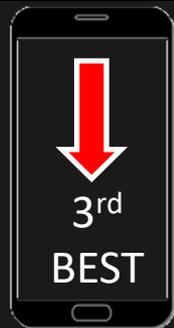
Input: 被害状況の想定あるいはリアルタイム情報

Input: 避難者の位置情報と目的地

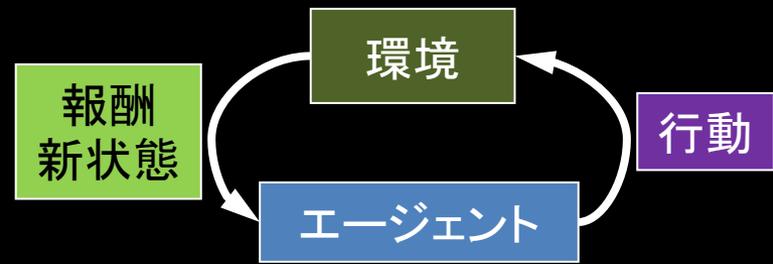
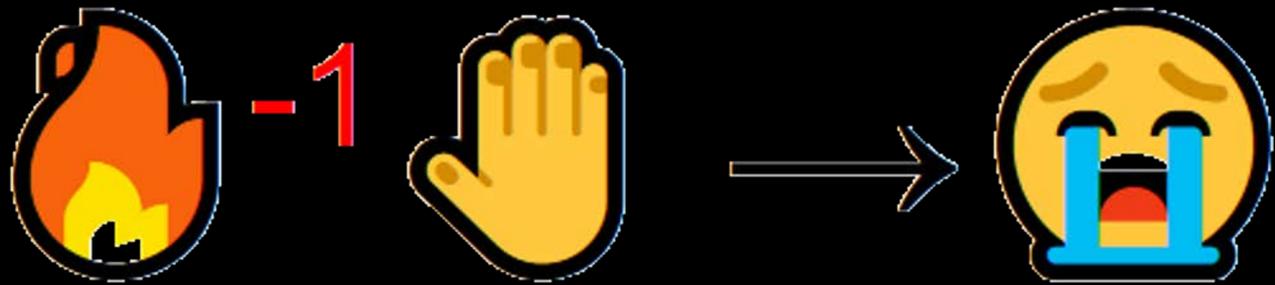
強化学習と次世代ベクトル機による経路選択問題を解く

+

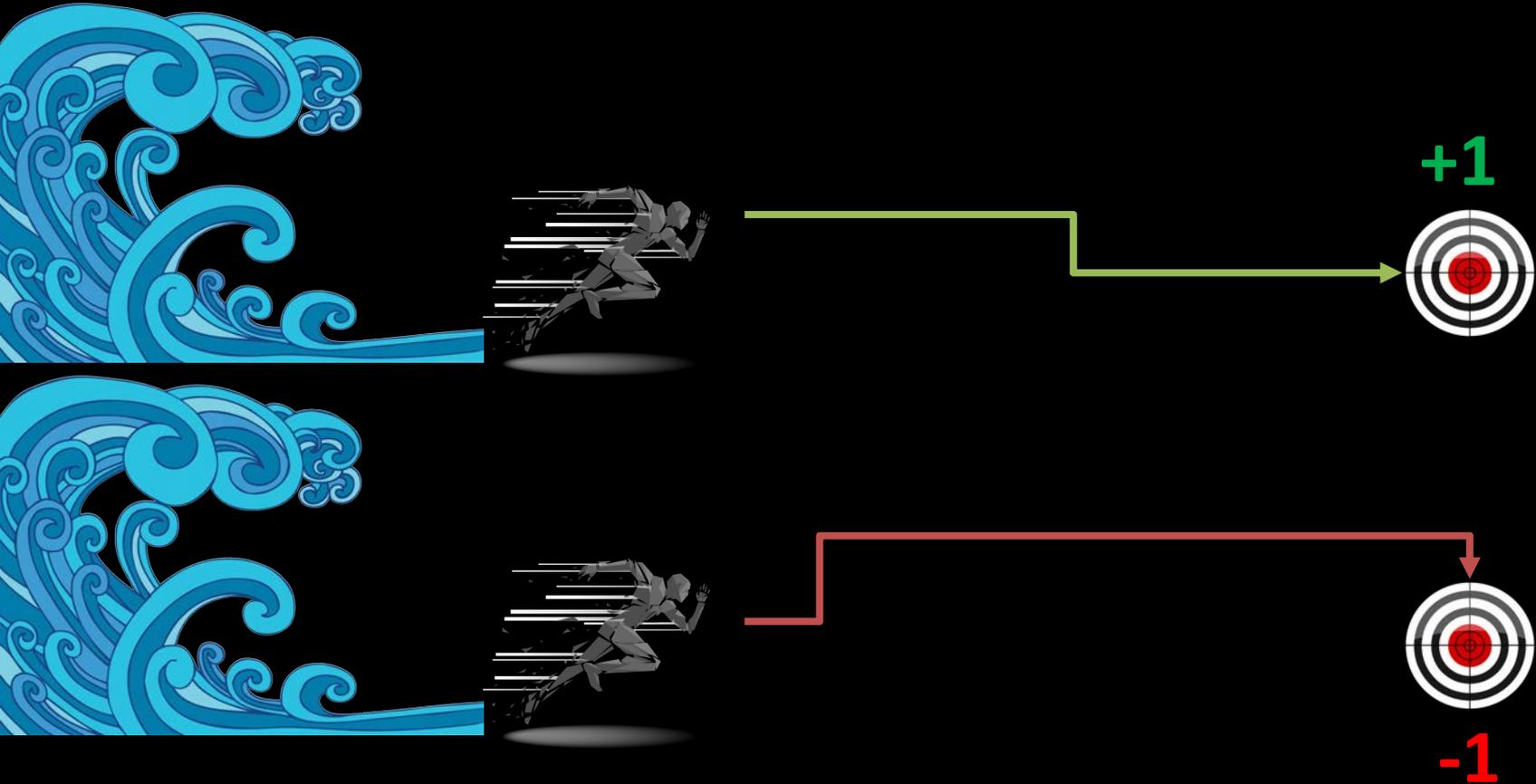
量子アニーリングによる避難者を複数の道に配布する



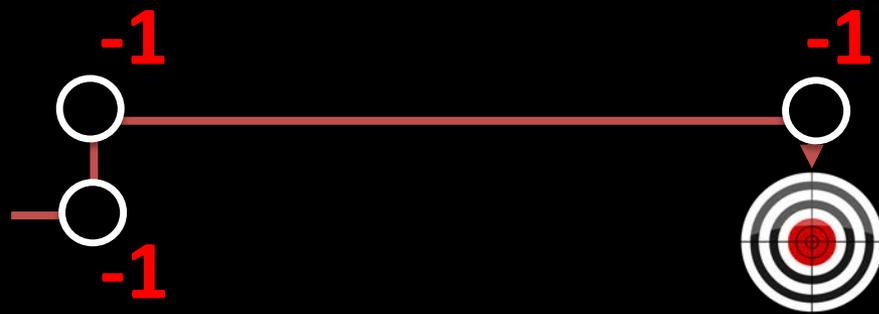
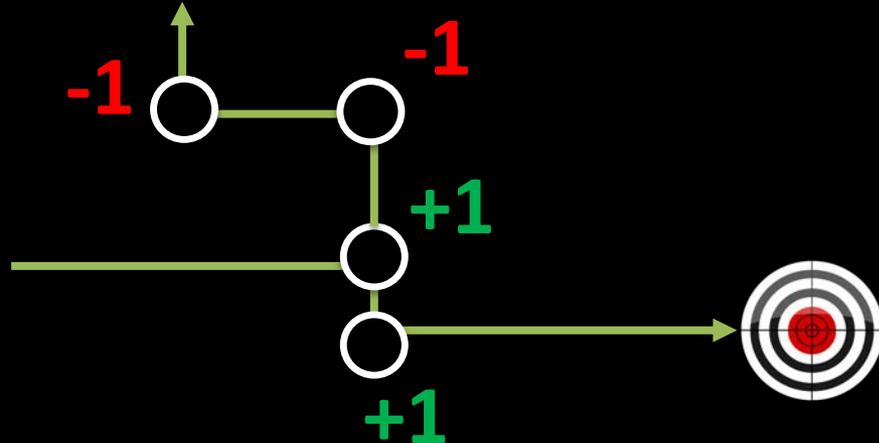
強化学習について



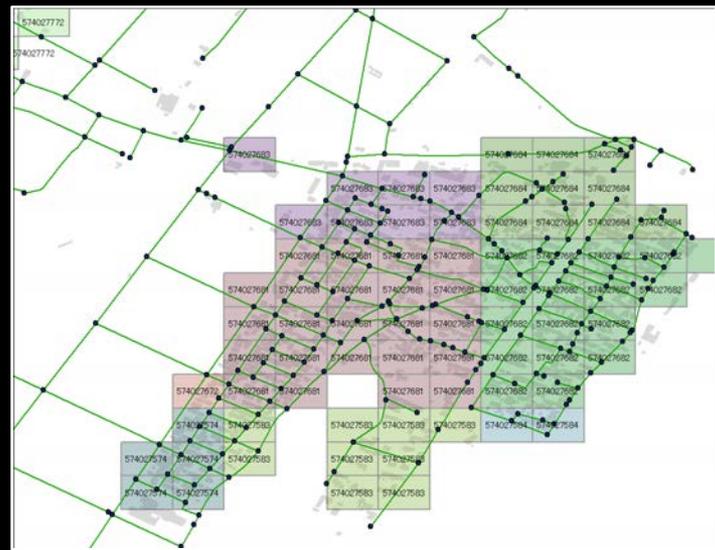
津波避難の場合



津波避難の場合

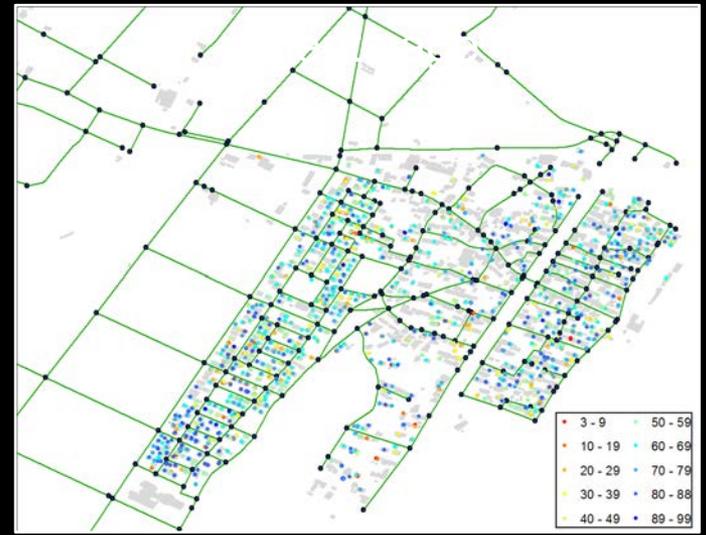
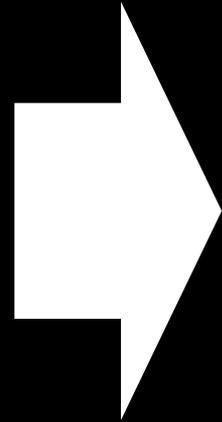


国勢調査



N = 2,723 人数

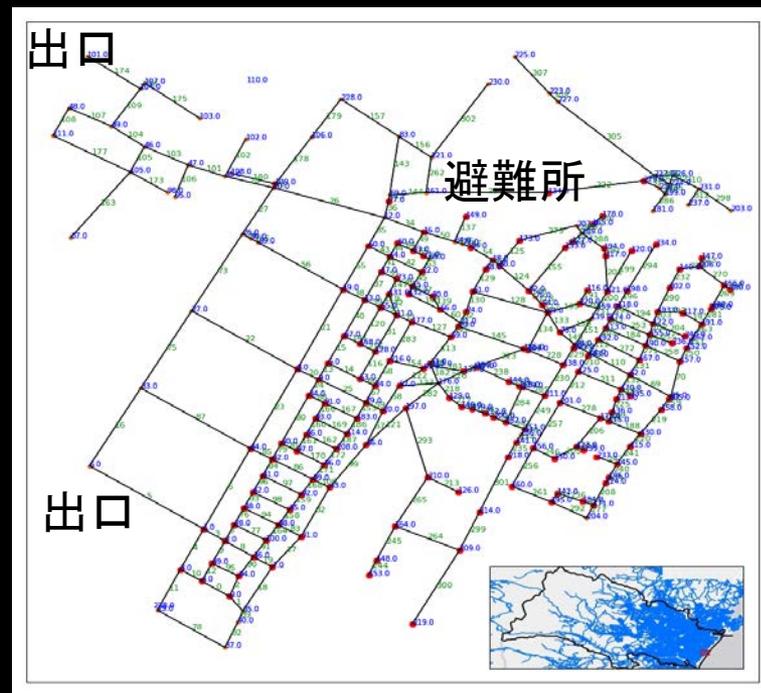
分解



ノード
(240個)

リンク
(312枚)

code	coord_x	coord_y	evacuation	# number	node1	node2	length
0	497729.4	4229597	0	0	0	4	70
1	497673.6	4229756	0	1	0	35	58
2	497719.6	4229728	0	2	0	94	52
3	497623.3	4229661	0	3	1	2	54
4	497669.2	4229633	0	4	1	3	108
5	497940.6	4230147	0	5	1	6	288
6	497427.4	4229905	1	6	1	44	215
7	497821.3	4229666	0	7	2	28	44
8	497875	4230132	0	8	2	36	72
...	9	2	39	60



リンクの接続

Code	Num. links	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
0	3	0	1	2	0	0	0	0	0	0	0
1	4	3	4	5	6	0	0	0	0	0	0
2	4	7	8	9	3	0	0	0	0	0	0
...

ノードの接続

Code	Num. nodes	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9
0	3	4	35	94	0	0	0	0	0	0	0
1	4	2	3	6	44	0	0	0	0	0	0
2	4	28	36	39	1	0	0	0	0	0	0
...

Links connected to node0

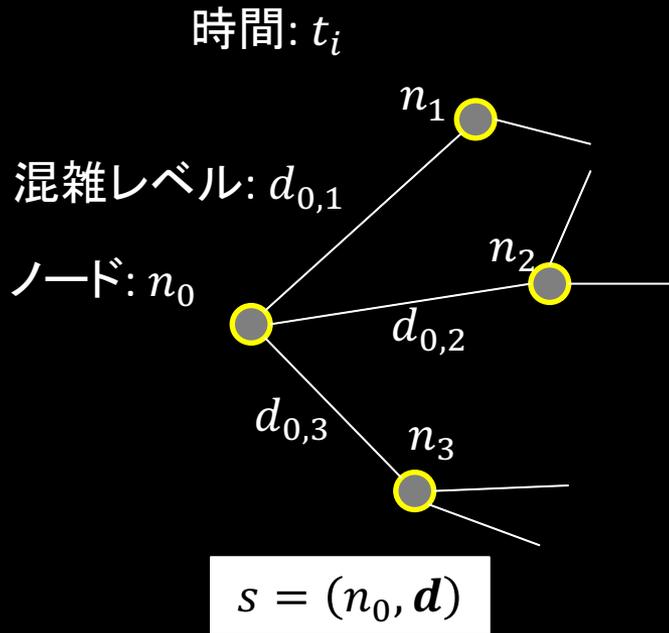
Number of links connected to node0

Node0 code

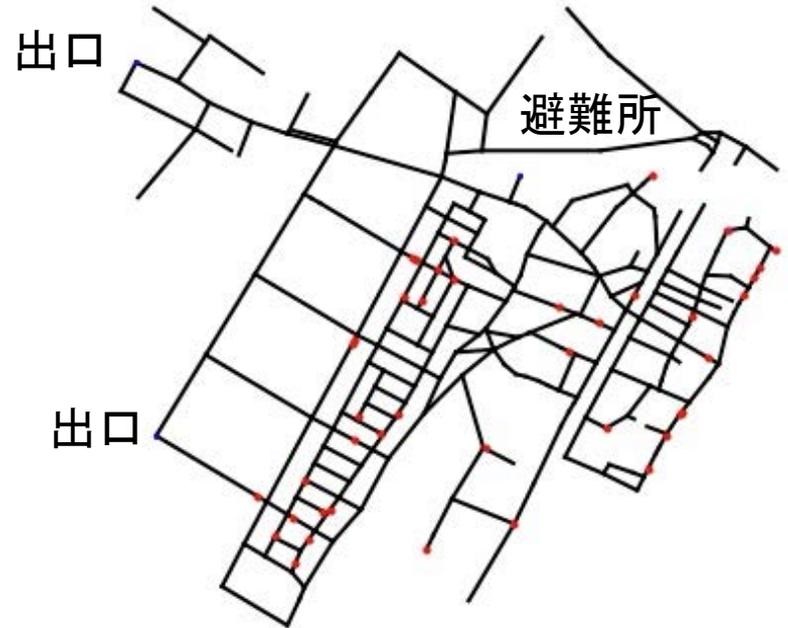
Nodes connected to node0

Number of nodes connected to node0

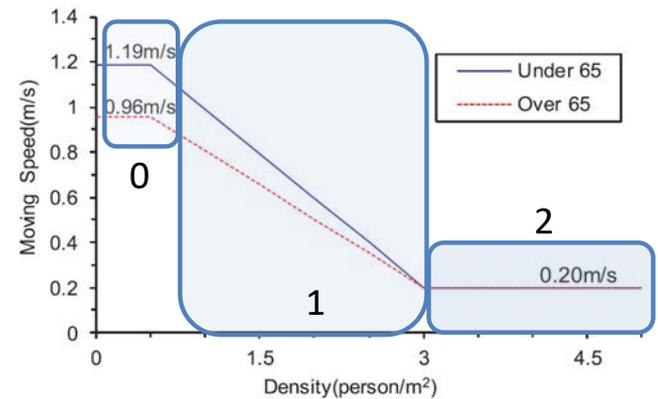
Node0 code



- (1) 1回目の試行 (ランダム誘導)
- (2) 結果の評価
- (3) 報酬を与える
- (4) 政策更新と保存
- (5) 2回目の試行 (現在の最高政策を入力)
- (6) 避難者に誘導 (数人へランダム誘導)
- (7) (2)へ戻ります



$t = 61.00$



政策の更新式

$$Q_{i+1} = \frac{Q_i \times N_i + R}{N_i + n}$$

N_i : i 番目の試行回数まででそのノードが使用された総数

n : $i+1$ 番目の試行回数のみでそのノードが使用された総数

R : $i+1$ 番目の試行回数のみで得られた報酬の総数



State code										Action-Value function										Number of observations									
...
36	0	0	1	0	0	0	0	0	0	0.86	0.99	0.99	0.99	0	0	0	0	0	0	130	79	68	62	0	0	0	0	0	0
2	0	1	0	1	0	0	0	0	0	0.98	0.97	0.98	0.96	0	0	0	0	0	0	26	19	21	14	0	0	0	0	0	0
100	0	1	0	0	0	0	0	0	0	0.99	0.79	0.96	0	0	0	0	0	0	0	189	468	175	0	0	0	0	0	0	0
36	0	1	0	0	0	0	0	0	0	0.92	0.66	0.97	0.93	0	0	0	0	0	0	18	297	15	23	0	0	0	0	0	0
36	0	1	1	0	0	0	0	0	0	0.99	0.99	1	1	0	0	0	0	0	0	186	176	169	159	0	0	0	0	0	0
36	0	1	1	1	0	0	0	0	0	0.83	0.5	0.75	0.83	0	0	0	0	0	0	3	1	2	3	0	0	0	0	0	0
7	0	0	1	0	0	0	0	0	0	0.88	0.75	0.5	0	0	0	0	0	0	0	4	2	1	0	0	0	0	0	0	0
94	0	1	0	0	0	0	0	0	0	0.98	0.98	0.99	0	0	0	0	0	0	0	21	30	36	0	0	0	0	0	0	0
36	1	1	1	0	0	0	0	0	0	0.99	0.99	0.99	0.99	0	0	0	0	0	0	81	74	81	53	0	0	0	0	0	0
100	0	1	1	0	0	0	0	0	0	0.5	0.88	0.75	0	0	0	0	0	0	0	1	4	2	0	0	0	0	0	0	0
88	1	0	0	0	0	0	0	0	0	0.75	0.5	0.5	0.75	0	0	0	0	0	0	2	1	1	2	0	0	0	0	0	0
28	1	0	1	0	0	0	0	0	0	0.83	0.9	0.93	0	0	0	0	0	0	0	3	5	7	0	0	0	0	0	0	0
38	0	0	1	0	0	0	0	0	0	0.92	0.92	0.94	0	0	0	0	0	0	0	6	6	8	0	0	0	0	0	0	0
28	0	1	1	0	0	0	0	0	0	0.99	0.98	0.99	0	0	0	0	0	0	0	88	84	72	0	0	0	0	0	0	0
100	1	1	0	0	0	0	0	0	0	1	0.99	1	0	0	0	0	0	0	0	114	124	116	0	0	0	0	0	0	0
28	1	記録状態							0	0	0.94	政策評価							0	9	8	使用された総数							
2	1	記録状態							0	0	0.98	政策評価							0	32	32	使用された総数							
100	1	1	1	0	0	0	0	0	0	0.75	0.75	0.5	0	0	0	0	0	0	0	2	2	1	0	0	0	0	0	0	0
28	0	1	0	0	0	0	0	0	0	0.99	0.98	0.99	0	0	0	0	0	0	0	74	89	70	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0.99	0.99	0.98	0.99	0	0	0	0	0	0	69	63	80	58	0	0	0	0	0	0
100	1	0	0	0	0	0	0	0	0	0.97	0.99	0.99	0	0	0	0	0	0	0	52	59	61	0	0	0	0	0	0	0
21	0	1	0	0	0	0	0	0	0	0.75	0.83	0.75	0	0	0	0	0	0	0	2	3	2	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0.96	0.79	0.97	0.95	0	0	0	0	0	0	14	12	15	11	0	0	0	0	0	0
1	2	0	0	0	0	0	0	0	0	0.81	0.64	0.98	0.8	0	0	0	0	0	0	8	7	31	339	0	0	0	0	0	0
220	1	0	0	0	0	0	0	0	0	0.5	0.9	0.75	0	0	0	0	0	0	0	1	5	2	0	0	0	0	0	0	0
63	0	0	1	0	0	0	0	0	0	0.75	0.5	0.5	0.5	0.833	0	0	0	0	0	2	1	1	1	3	0	0	0	0	0
2	0	0	0	2	0	0	0	0	0	0.97	0.74	0.96	0.95	0	0	0	0	0	0	18	21	12	336	0	0	0	0	0	0
1	2	0	0	1	0	0	0	0	0	0.5	0.83	0.92	0.75	0	0	0	0	0	0	1	3	6	2	0	0	0	0	0	0
44	1	0	0	0	0	0	0	0	0	0.76	0.95	0.89	0.99	0	0	0	0	0	0	1518	69	60	58	0	0	0	0	0	0
...

新状態を記録されます

毎試行後に更新されます

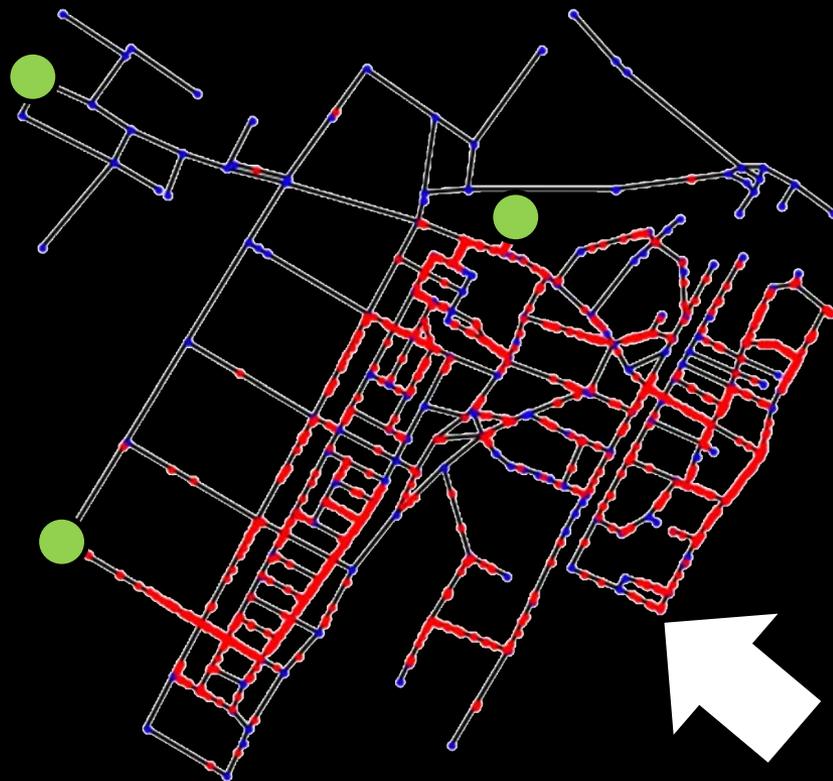
エージェントが任意の状態を経験し、任意の政策を実行した回数

津波到達時間: 67分

5分

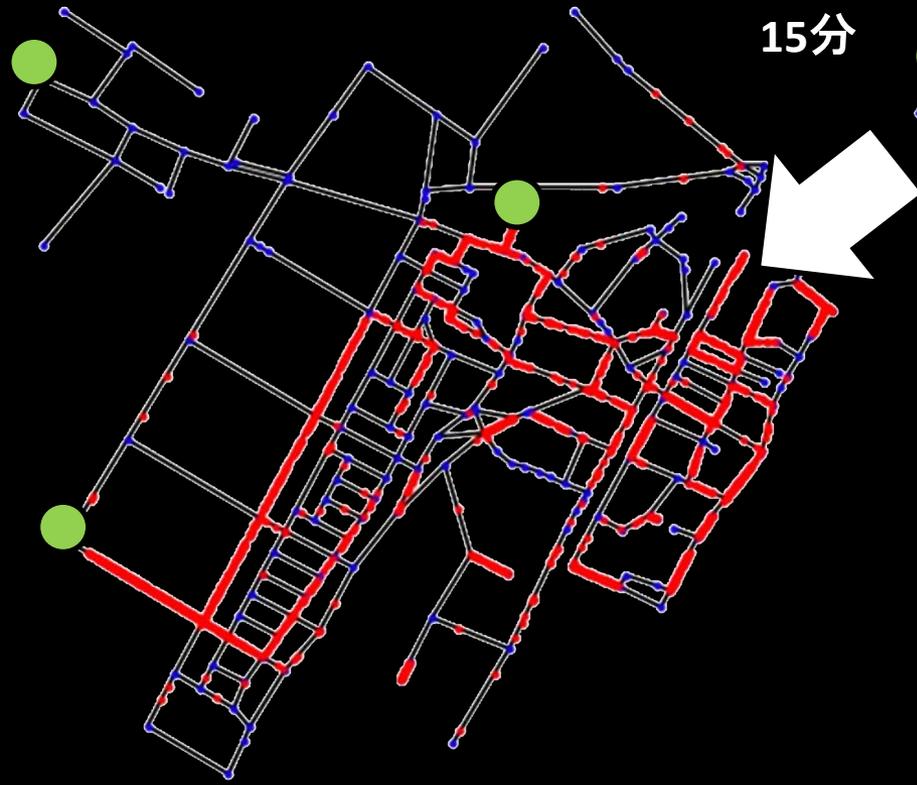


1回目の試行 (n = 1)

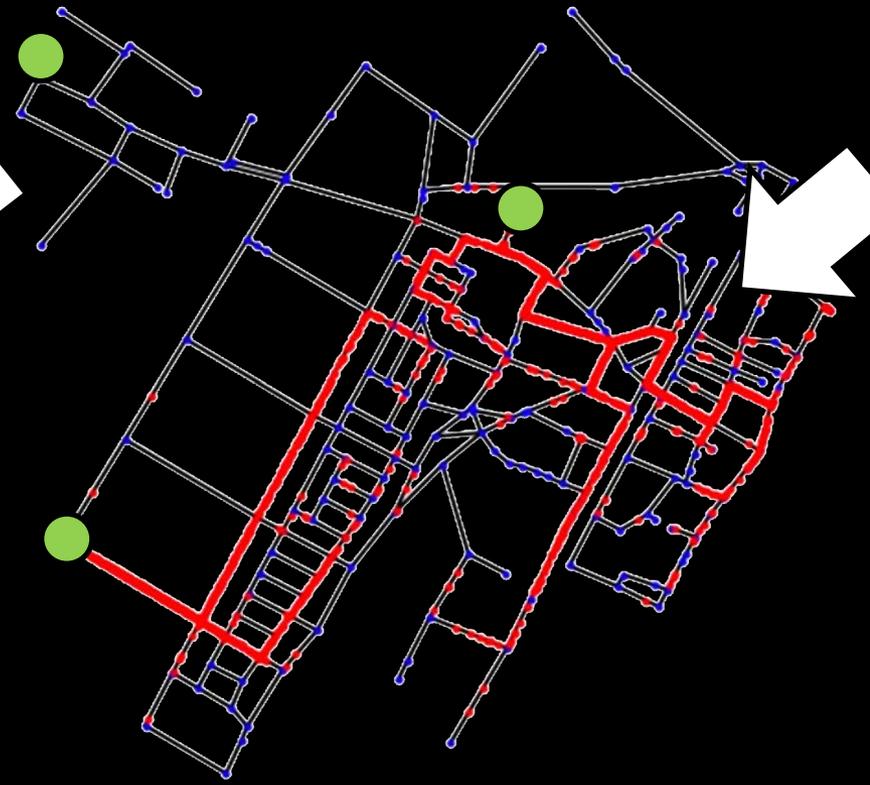


最終の試行 (n = 1000)

津波到達時間: 67分



1回目の試行 (n = 1)

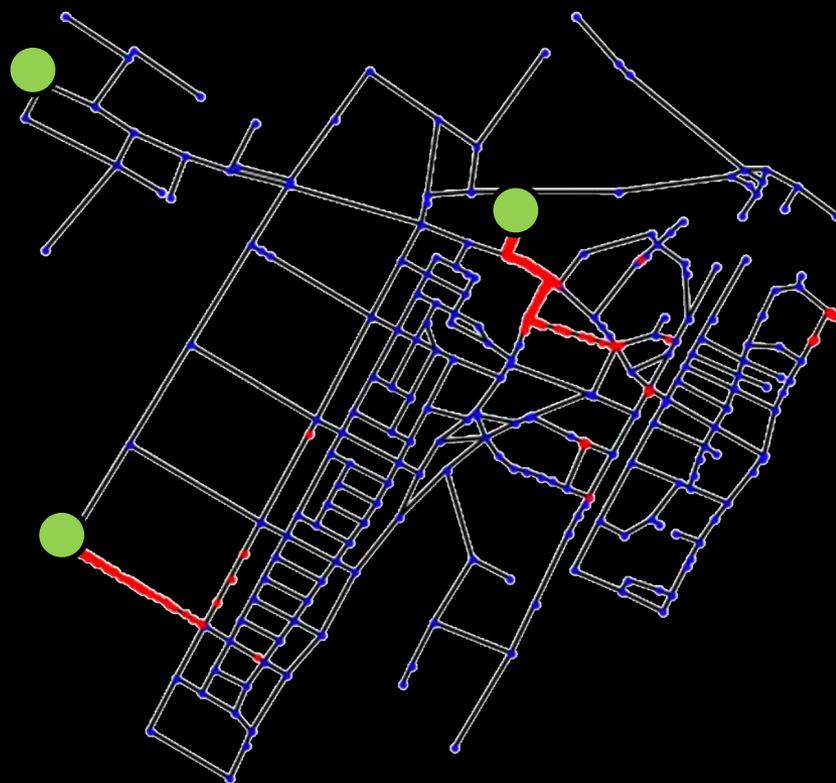


最終の試行 (n = 1000)

津波到達時間: 67分



1回目の試行 (n = 1)

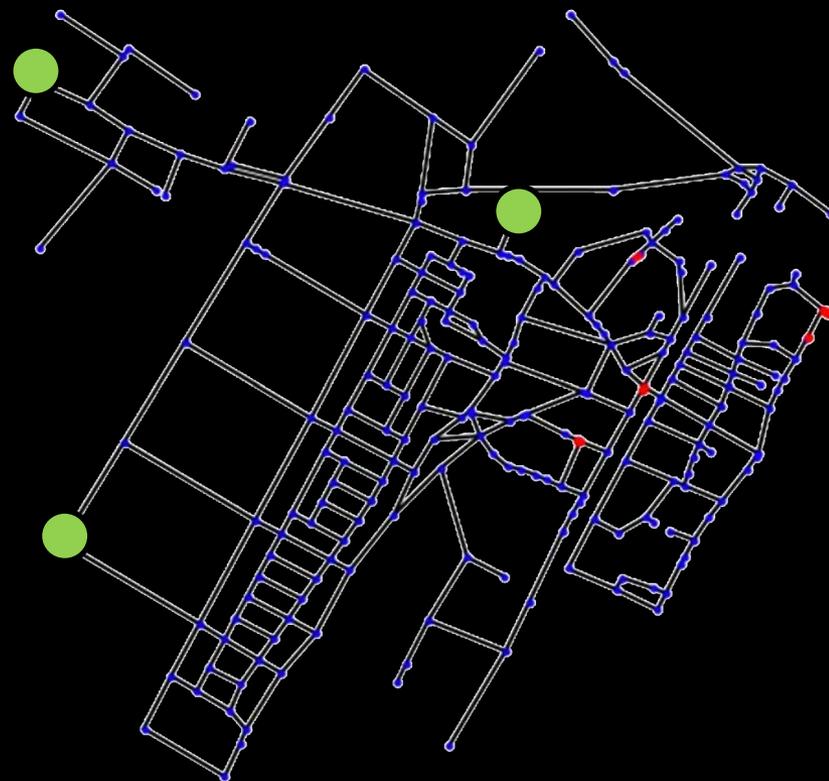


最終の試行 (n = 1000)

津波到達時間: 67分



1回目の試行 (n = 1)

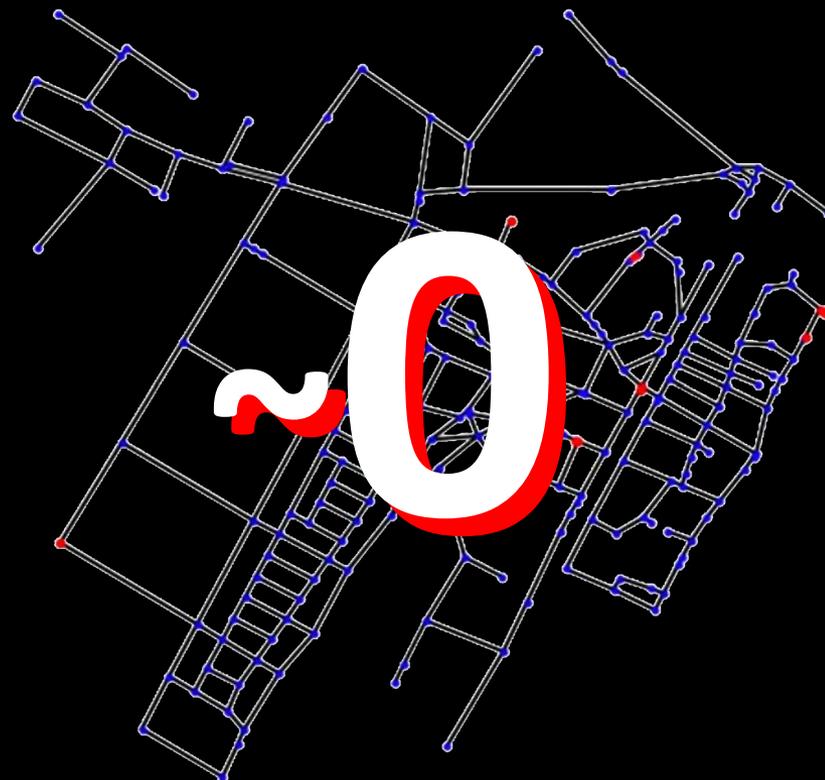


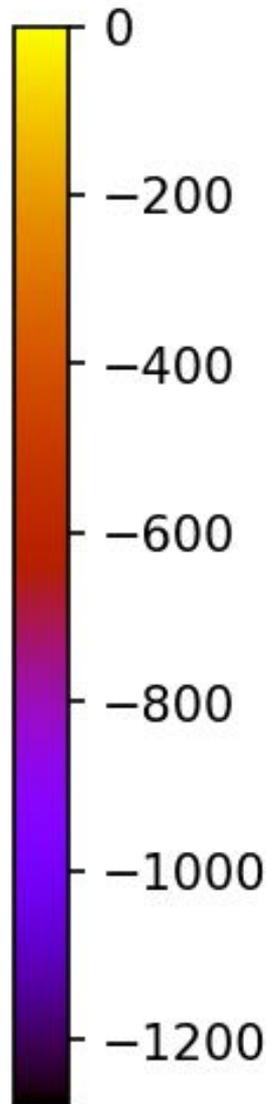
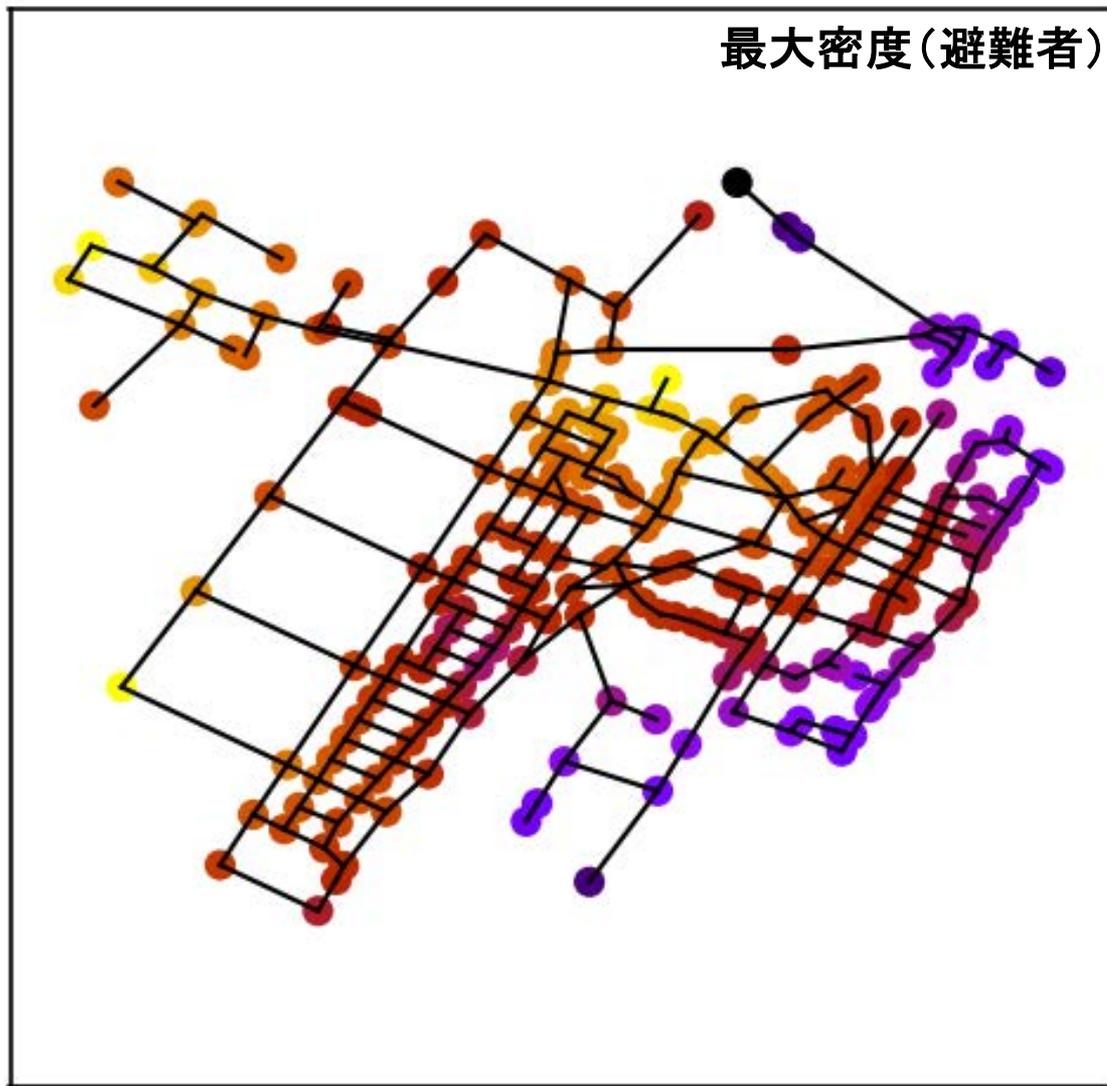
最終の試行 (n = 1000)

従来の避難モデル
(誘導なし)



強化学習の避難モデル
(誘導あり)





最大密度(避難者)
「重み」



経路探索最適化モデル

まとめ

- Society5.0とHPCは、リアルタイム津波避難シミュレーションの機会を提供する
- 交通渋滞の回避と最適避難経路を実装するために効果的な避難誘導システムが必要である
- 強化学習(RL)は効果的な避難誘導システムを設計するための適切な方法のようである
- HPC-RLとQAの組み合わせにより、避難誘導システムに堅牢性が提供されます



TCPAI
東北大学 タフ・サイバーフィジカルAI研究センター

ご清聴ありがとうございました