
Xiaoling Zeng, Kezhong Liu, Yuting Ma, Mozi Chen

Wuhan University of Technology, Wuhan, China
May 2023, MOST
Motivation

Emergency evacuation is critical following a ship accident, as passengers are required to escape the dynamic hazards and reach the muster station before the deadline.
Motivation

(a) A navigation scheme based on a road backbone may lead to heavy congestion.

(b) Unconstrained detours may increase the user’s exposure time to hazards.
Our Contribution

- We constructed a crowd movement data set of ship indoor evacuation via a simulation platform Anylogic for the DC-HEN training.
- We proposed a method for constructing a graph model with environmental structural features.
- We developed a hierarchical emergency navigation algorithm that combines the global reference path and local environmental information based on reinforcement learning technology.
Our Proposed Hierarchical Navigation System

Simulation model

Offline

Module 1: Feature graph construction

- Feature node extraction
- GVD feature node
- Roadmap generation
  \[ G' = (V', E') \]
- Look-up table construction

Start position
Exit position

Online

Module 2: Constrained global guidance

- Navigation model
- Global navigator
- Subgoal selection
- Navigation segment

Module 3: Adaptive local navigation

- RL-based planner
- Next node selection
- Update

Navigation command
Sensor data
Offline Feature Graph Construction
[Baruah et al. 2018]

(i) Process of feature node extraction

Input:
• Typical delay: $d_T(v'_i v'_j)$
• The worst-case delay: $d_W(v'_i v'_j)$
• Deadline

Output: a 3-tuple table $\text{Tab}[v'] = (d'_v, \pi'_v, \delta'_v)$

(ii) A simple example of look-up table to the exit
Decision-making Agent Development
[Van Hasselt et al. 2016]

**Local observation:** the set of locations of free space, walls, other users, hazards, and global guidance segments within the observation range respectively.

**Action space:** a discrete set of navigation decisions.

**Reward function:**
- a small negative reward at each time step to encourage the agent to reach the exit with less time compensation;
- a penalty of when the agent collides with walls or other users;
- a great penalty for exposure to hazards;
- a positive reward denoted as $N_t \times 10$ for following the global guidance;
- a great positive reward when the agent reaches the exit.
Hierarchical Emergency Navigation System
[Van Hasselt et al. 2016]

The target Q-value:

\[ Y_t^{DoubleDQN} = r_t + \gamma Q \left( s_{t+1}, \arg\max_{a_t} Q(s_{t+1}, a_t; \theta); \theta' \right) \]

The loss function:

\[ L(\theta) = \frac{1}{N_b} \sum_{i=1}^{N_b} \left[ Y_{t_i} - Q(s_{t_i}^i, a_{t_i}^i; \theta) \right]^2 \]
Simulation Setup

**Data set generation:** we simulate the evacuation process on a single deck using the visualization simulation platform Anylogic to generate the crowd movement data set.

**Ablation study:**

<table>
<thead>
<tr>
<th>$H_o \times W_o$</th>
<th>9 × 9</th>
<th>11 × 11</th>
<th>13 × 13</th>
<th>15 × 15</th>
<th>17 × 17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowd-50</td>
<td>1.12</td>
<td>1.10</td>
<td>1.08</td>
<td>1.08</td>
<td>1.07</td>
</tr>
<tr>
<td>Crowd-100</td>
<td>1.25</td>
<td>1.24</td>
<td>1.18</td>
<td>1.12</td>
<td>1.10</td>
</tr>
<tr>
<td>Crowd-150</td>
<td>1.44</td>
<td>1.37</td>
<td>1.25</td>
<td>1.15</td>
<td>1.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N_t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowd-50</td>
<td>1.28</td>
<td>1.19</td>
<td>1.12</td>
<td>1.08</td>
<td>1.06</td>
</tr>
<tr>
<td>Crowd-100</td>
<td>1.43</td>
<td>1.33</td>
<td>1.27</td>
<td>1.12</td>
<td>1.10</td>
</tr>
<tr>
<td>Crowd-150</td>
<td>1.47</td>
<td>1.36</td>
<td>1.28</td>
<td>1.15</td>
<td>1.15</td>
</tr>
</tbody>
</table>
## Compared Approaches


<table>
<thead>
<tr>
<th>Approach</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC-HEN</td>
<td>Hierarchical emergency navigation, our proposed approach</td>
</tr>
<tr>
<td>DDQN</td>
<td>Deep RL-based emergency navigation without global guidance</td>
</tr>
<tr>
<td>CANS</td>
<td>A congestion-adaptive method with potential map and hazard level map</td>
</tr>
<tr>
<td>ECSSN</td>
<td>Clogging-free and shortest-safe path navigation method</td>
</tr>
</tbody>
</table>
Experimental results

- DC-HEN's training curve rises faster than the DDQN method.
- DC-HEN's navigation success rate rises rapidly with the increase of training times and finally reaches 78.3%.
Experimental results

- The path planned by DC-HEN has a certain distance from the static obstacles, while the trajectory of ECSSN is close to the wall.
- The path stretch results for DC-HEN are similar to those of ECSSN in all cases.
Experimental results

- Congestion distribution

- Nodes involved in DC-HEN are at most participate in about 300 navigation paths, and the blue curve rapidly reaches 1.
Conclusions

- DC-HEN utilizes reinforcement learning and designs a novel reward function to provide congestion-relieved evacuation guidance for each user in real-time.
- DC-HEN has a higher success rate with 78.3%, relatively short average path stretch, and better congestion avoidance performance.

Future work

- Designing a multi-agent decision system that takes into account the allocation of limited life-saving resources.
- Incorporating users’ personalized preferences.
Thank you!

Xiaoling Zeng
xlzeng@whut.edu.cn

Kezhong Liu
kzliu@whut.edu.cn

Yuting Ma
278827@whut.edu.cn

Mozi Chen
chenmz@whut.edu.cn