An Integrative Approach of Social Dynamic Long Short-Term Memory and Deep Reinforcement Learning for Socially Aware Robot Navigation

Xuan Tung Truong and Trung Dung Ngo

***Abstract*— In this study, we propose an integrative approach of Long short-term memory (LSTM) networks and a deep reinforcement learning (DRL) technique for socially aware robot navigation in crowded and dynamic environments. The proposed system is an integration of two main stages: (1) socio-spatio-temporal characteristics of the humans are encoded by using the LSTM networks, and (2) the encoded social dynamic characteristics are then fed into the DRL algorithm in order to generate motion control commands for a mobile robot. We integrated the developed system onto the conventional mobile robot navigation system and verified it in a simulated environment. The simulation results show that the proposed socially aware robot navigation system enables the mobile robot to behave in socially acceptable manners.**

I. INTRODUCTION

The ability to autonomously navigate in crowded and dynamic environments is crucial to deploy a mobile robot into our daily-life settings. Several frameworks of socially aware navigation systems have been proposed to generate the socially acceptable behaviours for the robot, and ensure the human safety and comfort within the it's vicinity [1]. The conventional robot navigation frameworks can be roughly classified into two categories: (i) model-based methods e.g. [2] and (ii) learning-based approaches e.g. [3], [4]. In the former, social force models and velocity obstacles techniques are used to develop the navigation systems. In the later, machine learning algorithms are utilized to enable the robots to navigate safely in crowded and dynamic environments.

Although the existing navigation systems are capable of driving the mobile robot to navigate safely towards a given goal, they still suffer essential drawbacks in navigating in crowded and dynamic social environments because *only current human states are taken into account*. Therefore, the robot navigation systems may fail in social situations, as shown in Fig. 1. To overcome this drawback, we propose an integrative approach of LSTM networks [5] and a DRL algorithm. Inspired by the works presented by Alahi [6] and Everett [3], *the past, current, and predicted future states of humans and objects are taken into account to design the socially aware mobile robot navigation systems.*

Xuan Tung Truong and Trung Dung Ngo are with The More-Than-One Robotics Laboratory, University of Prince Edward Island, Canada. Xuan Tung Truong is also with the Le Quy Don Technical University, Hanoi,Vietnam. xuantung.truong@gmail.com, tngo@upei.ca

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.01-2018.10.

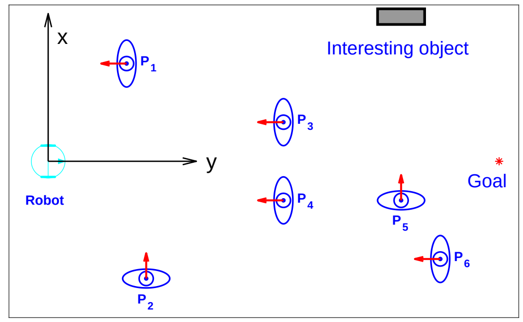


Fig. 1. An example of a crowded and dynamic environment consisting of a mobile robot, six moving people and one interesting object. The people P_3 and P_4 are forming a group. Person P_5 intends to interact with the interesting object. The mobile robot is requested to move towards the given goal while avoiding people, their social interaction, and object.

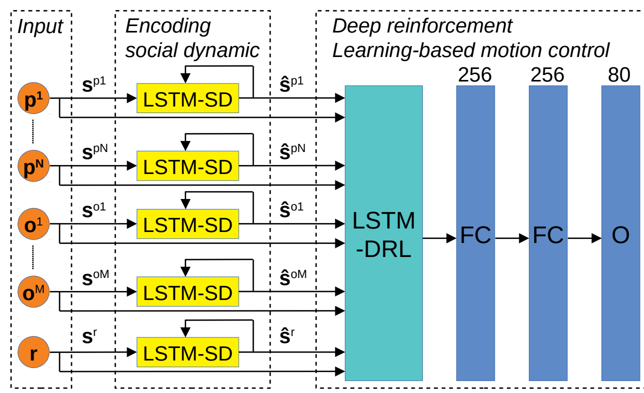


Fig. 2. The flowchart of the proposed socially aware navigation framework. The input of the framework is the state of the humans $\mathbf{p}^i$, interesting objects $\mathbf{o}^j$ and robot $\mathbf{r}$. The LSTM-SD networks are the LSTM networks that are used to encode the social dynamic characteristics of the humans. Whereas, the LSTM-DRL is utilized to encode the states and the predicted sub-states of the humans, interesting objects and robot. The input of the LSTM-SD networks is the sub-state of the humans $\mathbf{s}^{pi} = [x^{pi}, y^{pi}, \theta^{pi}]$, interesting object $\mathbf{s}^{oj} = [x^{oj}, y^{oj}, \theta^{oj}]$ and the robot $\mathbf{s}^r = [x^r, y^r, \theta^r]$, and the its output is the corresponding predicted sub-state $\hat{\mathbf{s}}^{pi}$, $\hat{\mathbf{s}}^{oj}$ and $\hat{\mathbf{s}}^r$, respectively. The input of the deep reinforcement learning-based motion planing system is their state and the predicted sub-state. Follow the LSTM-DRL is the two fully connected layers (FC) and the output layer (O).

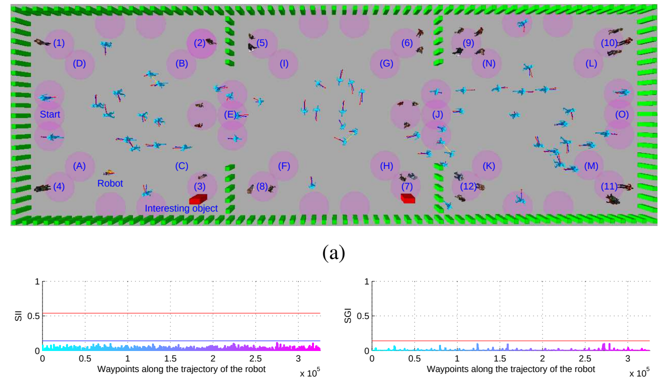
II. THE INTEGRATIVE APPROACH

To make the model independent of the global coordinate, we use a reference frame with the origin fixed at the robot position, the $x-axis$ toward the north, and the $y-axis$ toward the east, as seen in Fig. 1. Assume that, states of the robot in the local coordinate at time t are represented as $\mathbf{r}_t = [\theta_t^r, \dot{x}_t^r, \dot{y}_t^r, r_t^r, v_{pref}^r]^T$, where θ_t^r is the orientation, $(\dot{x}_t^r, \dot{y}_t^r)$ is the velocity, r_t^r is the radius, and v_{pref}^r is the preferable speed. We also assume that, there are N people appearing in the robot's vicinity at time t, $P = \{\mathbf{p}_t^1, \mathbf{p}_t^2, ..., \mathbf{p}_t^N\}$, where $\mathbf{p}_t^i$ is the i^{th} person. The states of person $\mathbf{p}_t^i$ are represented as $\mathbf{p}_t^i = [x_t^{pi}, y_t^{pi}, \theta_t^{pi}, \dot{x}_t^{pi}, \dot{y}_t^{pi}, r_t^{pi}]^T$, where (x_t^{pi}, y_t^{pi}) is the position, θ_t^{pi} is the head orientation, $(\dot{x}_t^{pi}, \dot{y}_t^{pi})$ is the velocity, and r_t^{pi} is the radius. There are M objects $O = \{\mathbf{o}_t^1, \mathbf{o}_t^2, ..., \mathbf{o}_t^M\}$, where $\mathbf{o}_t^j$ is the j^{th} interesting object. The states of object o_t^j are represented as $\mathbf{o}_t^j = [x_t^{oj}, y_t^{oj}, \theta_t^{oj}, \dot{x}_t^{oj}, \dot{y}_t^{oj}, r_t^{oj}]^T$, where (x_t^{oj}, y_t^{oj}) is the position, θ_t^{oj} is the orientation, $(\dot{x}_t^{oj}, \dot{y}_t^{oj})$ is the velocity, and r_t^{oj} is the radius.

We consider a social context of a mobile robot navigating towards a given goal (x_t^g, y_t^g) through a social environment in the presence of N humans and M objects. The robot should navigate towards the given goal timely while guaranteeing human safety and comfort. To do that, we propose an integrative approach of LSTM networks and DRL technique as shown in Fig. 2.

Encoding socio-spatio-temporal human characteristics: Given the sub-states of human $\mathbf{s}^{pi} = [x^{pi}, y^{pi}, \theta^{pi}]$, object $\mathbf{s}^{oj} = [x^{oj}, y^{oj}, \theta^{oj}]$ and robot $\mathbf{s}^r$ at time $t = 1, 2, ..., T_{obs}$, the corresponding future location $\hat{\mathbf{s}}^{pi}$, $\hat{\mathbf{s}}^{oj}$ at time $t = T_{obs+1}, T_{obs+2}, ..., T_{pred}$ should be predicted accordingly. To do that, we encode the social dynamics of the humans by using LSTM networks [5]. Particularly, we extend the model, proposed by Alahi [6], by adding the human head orientation into the LSTM model. In addition, using the same scheme of the social pooling layer [6], which incorporates the human–human interaction information into the model, we embed the human–object, human–robot interaction information into the LSTM networks. The future states of the humans and objects are predicted by using the LSTM-SD networks, as shown in Fig. 2.

Deep reinforcement learning for the robot motion planing system: We propose a DRL model enabling a robot to learn how to navigate in a crowded environment, as shown in Fig. 2. The current states of the humans, objects and their predicted sub-states are taken into account of the DRL model. Robot behaviours are identified through the network outputs. In addition, a number of humans and objects within the robot's vicinity are arbitrary, thus inspired by Everett [3] their states are sequentially fed into LSTM networks (named as LSTM-DRL in Fig. 2) by using descending order of their distances to the robot. The final state of the LSTM-DRL network is then used as the input of the two fully connected layer neural network. The output layer is the probability for each action in the discrete action space of the robot. The action space of the mobile robot consists of 80 discrete actions including: (i) 5 values of the linear velocity ranging from $[0, v_{pref}^r]$ and (ii) 16 values of the angular velocity spacing between $[-\pi/3, \pi/3]$. The preferable velocity of humans and the mobile robot is ranged from 0 to 1.5 [m/s] in simulation, which is close to the pedestrian's velocity in the real-world. The joint state of the mobile robot, humans, objects and their predicted sub-states is $\mathbf{s}_t^{por} = [\mathbf{p}^i, \mathbf{o}^i, \hat{\mathbf{s}}^{pi}, \hat{\mathbf{s}}^{oj}, \mathbf{r}]$.



(a)

(b) (c)

Fig. 3. The simulation environment (a) and the simulation results measured by the social individual indices (SII) and the social group indices (SGI).

III. Experiments

We have created a shopping mall-like scenario using the available software platform[1], as shown in Fig. 3(a). The proposed method is implemented using Python, PyTorch library and the robot operating system (ROS).

To evaluate the performance of the proposed navigation method, we utilized the human comfortable safety index, presented in [2] including social individual index (SII), social group index (SGI). The experimental results in Figs. 3(b) and 3(c) highlight that the mobile robot is capable of avoiding humans in socially acceptable manners while guaranteeing the human safety and comfort in dynamic social environments.

IV. Conclusions

We have presented a socially aware robot navigation system based on LSTM networks and a DRL technique. The simulation results confirm that the proposed method is capable of safely and socially driving the robot while guaranteeing human safety and comfort. In the future, we will implement the proposed method on our mobile robot platform and evaluate it in numerous social situations.

References

[1] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1726–1743, 2013.

[2] X. T. Truong and T. D. Ngo, "Toward socially aware robot navigation in dynamic and crowded environments: A proactive social motion model," *IEEE Transactions on Automation Science and Engineering*, vol. 14, pp. 1743–1760, October 2017.

[3] M. Everett, Y. F. Chen, and J. P. How, "Motion planning among dynamic, decision-making agents with deep reinforcement learning," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3052–3059, October 2018.

[4] C. Chen, Y. Liu, S. Kreiss, and A. Alahi, "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning," *arXiv:1809.08835v2*, 2018.

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, pp. 1735–1780, November 1997.

[6] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 961–971, June 2016.

[1] https://github.com/srl-freiburg