

Autonomous Navigation Using Deep Reinforcement Learning in ROS

Ganesh Khokare, G.H. Raisoni College of Engineering, Nagpur, India

 <https://orcid.org/0000-0002-1687-4699>

Shahrukh Sheikh, G.H. Raisoni College of Engineering, Nagpur, India

 <https://orcid.org/0000-0002-2877-6279>

ABSTRACT

For an autonomous robot to move safely in an environment where people are around and moving dynamically without knowing their goal position, it is required to set navigation rules and human behaviors. This problem is challenging with the highly stochastic behavior of people. Previous methods believe to provide features of human behavior, but these features vary from person to person. The method focuses on setting social norms that are telling the robot what not to do. With deep reinforcement learning, it has become possible to set a time-efficient navigation scheme that regulates social norms. The solution enables mobile robot full autonomy along with collision avoidance in people rich environment.

KEYWORDS

Automatic Path Finder, Collision Avoidance, Crowd Navigation, Deep Reinforcement Learning, Robotics

INTRODUCTION

Autonomous Navigation in a highly stochastic environment is bound to uncertainty with the prediction of future behavior in a multi-agent system. Sensor fusion is again a challenging task that involves the coupling of different sensors for object detection and smooth movement of the robot (Bai et al., 2015).

This paper deals with a strong foundation of mapping, localization, and path planning with a deep reinforcement learning method having multiple agents (Kim et al., 2014). Autonomous navigation problem having a set of challenges which includes an ever-changing environment, prediction of people's behavior while moving, avoiding obstacles, etc. While predicting people's behavior, it is essential to bound the scope with simple kinematics and impose rules for avoiding collision amongst the agents (Helbing et al., 1995).

LITERATURE SURVEY

Deep Reinforcement Learning in ROS

Deep reinforcement learning is a combination of deep learning and reinforcement learning. It allows robots to achieve a high degree of autonomy of exploration and navigation (Kim et al., 2016). There

DOI: 10.4018/IJAIML.20210701.oa4

This article, published as an Open Access article on April 23, 2021 in the gold Open Access journal, International Journal of Artificial Intelligence and Machine Learning (converted to gold Open Access on January 1, 2021), is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

are two approaches in deep reinforcement learning mainly (a) Model-based learning (b) Value-based learning. This work focuses on value-based learning, this value decides the next action be taken (Unhelkar et al., 2015).

Figure 1 illustrates the workflow of memory-based deep-reinforcement learning model, where the input is through the sensor database on the depth and range sensing (Gohane & Khekare, 2015). Along with it, autonomous exploration is based on Markov Decision Process (Trautman, 2015). Using DRL it learns to take the optimal decision at time t in a safe state (Kretzschmar et al., 2016). The decision will be taken according to the policy learned by the algorithm. It uses a socially aware collision avoidance algorithm using deep reinforcement learning (Berg et al., 2011).

Where a multi-agent collision avoidance problem is addressed as a sequential decision-making problem in the reinforcement learning framework (Khekare, 2014). Where the agent's position, velocity, and size is described by a set of states. Also inducing social norms for example the distance to the other agents is taken into consideration. Previous works have reported that the teleoperation of a robot bound to be cooperative and time-efficient (Khekare & Janardhan, 2017).

To solve the problem of collision avoidance, it performs multi-agent value training on a two-agent network, but the drawback of their system is, it can- not be scalable to more than two agents (Khekare et al., 2020). Whereas social norms of the agent's environment are taken into consideration by the SA-CADRL algorithm (Mehta et al., 2016). Whereas some papers discuss the social force model for pedestrian dynamics. It uses the Deep Q Network (DQN) which is a combination of deep learning and Q-learning thinking is drawing attention as a reinforcement learning algorithm (Berg et al., 2008).

METHODOLOGY

Our Strategy is divided into four parts including Mapping, Localization, Path Planning, and collision avoidance.

Mapping

Mapping is the process of building a map of the floor or the environment where the robot is in tented to move. The map is the representation of the environment created from sensor readings of the robot, here the LIDAR sensor is used based on lased guided navigation (Kim et al., 2018). For mapping turtle, bot navigation package in ROS is used, along with mapping package of gazebo simulation. To map the robot teleop is used so that the robot should the complete floor. Illustrated in Figure 2 and Figure 3.

Figure 1. Schematic of Deep reinforcement learning model

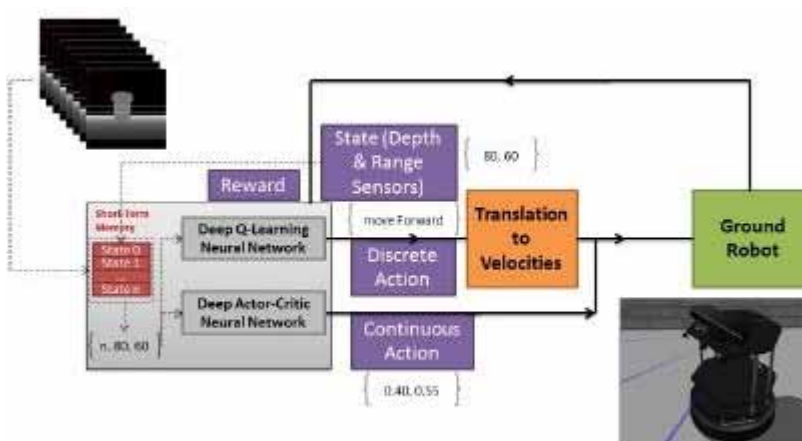


Figure 2. Mapping of the Robot with gmapping and ROS Navigation Stack

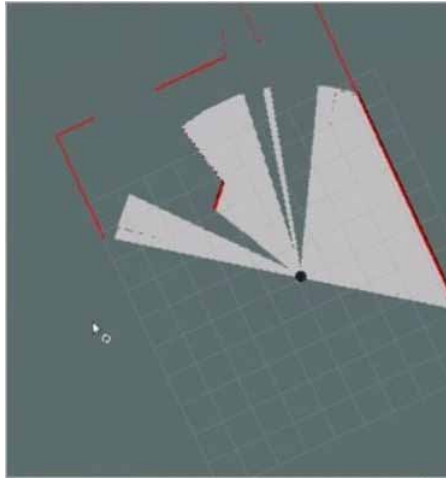
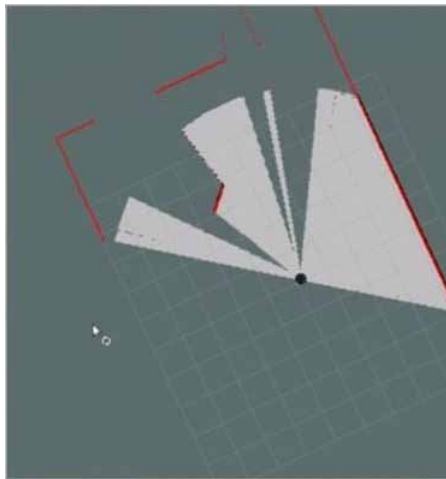


Figure 3. Map generation with Teleop Navigation in RViz



Localization

In Autonomous Navigation robot has to locate its position concerning the environment, so that it can easily move towards its goal position (Khekare et al., 2019). Localization is the method to find out where the robot is concerning the map. Gazebo AMCL package is used for localization along with turtle bot navigation stack, Illustrated in Figure 4. To view the localization of the robot, RViz visualization package of ROS is used (Chen et al., 2017).

Obstacle Avoidance

Obstacle avoidance is the method which updates real-time data from sensors if some sudden change in the environment happens or some new agent or object comes in, it considers it as an obstacle and tries to change its path (Fox et al., 1997). As illustrated in Figure 5, a new object is spawned into the environment as an obstacle and the robot is trying to avoid it.

Figure 4. Localization of the Robot with AMCL package and visualization in RViz of ROS

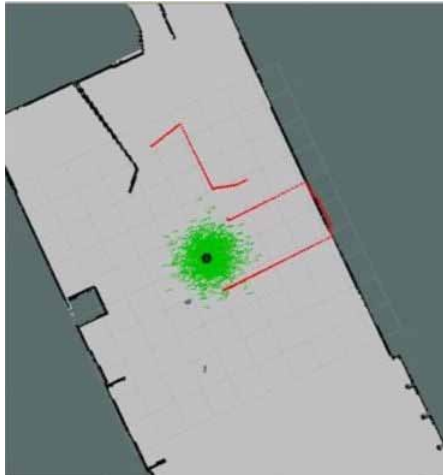
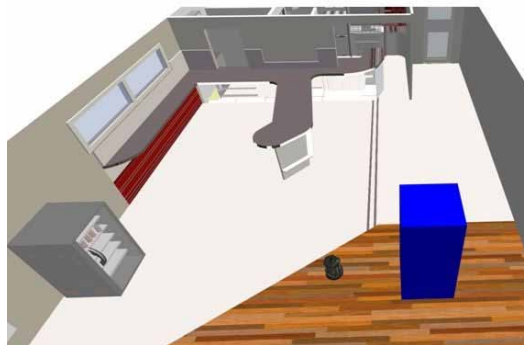


Figure 5. Obstacle avoidance of the Robot having obstacle spawn in environment



Path Planning

Deep reinforcement learning is used as a base algorithm for path planning inside the navigation stack of the ROS framework (Ferrer et al., 2013). Navigation stack is a set of ROS nodes and algorithm which are used to autonomously move the robot (Dooraki et al., 2018). It will take as input the current location of the robot, desired location (goal pose), the Odometry data of the robot (wheel encoders, IMU, GPS), and LIDAR. Illustrates the path planning by the robot in Figure 6.

The main purpose of the ROS distribution package is to move the robot from the starting position to the goal position, without causing collisions with the environment (Ferrer et al., 2014). The ROS Navigation package comes with the introduction of a few navigation-related algorithms that can easily facilitate the use of automation in mobile robots (Phillips et al., 2011). The user only needs to feed the target position of the robot and robot Odometry data from sensors such as wheels, IMU, and GPS, as well as other sensor data sources such as laser scanner data or 3D point cloud from sensors such as Kinect. Navigation Package Release will be the velocity commands that will move the robot to the given point as shown in Figure 7.

Figure 6. Visualization of Obstacle avoidance of the Robot

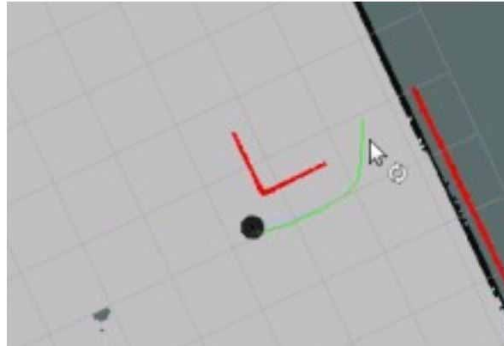
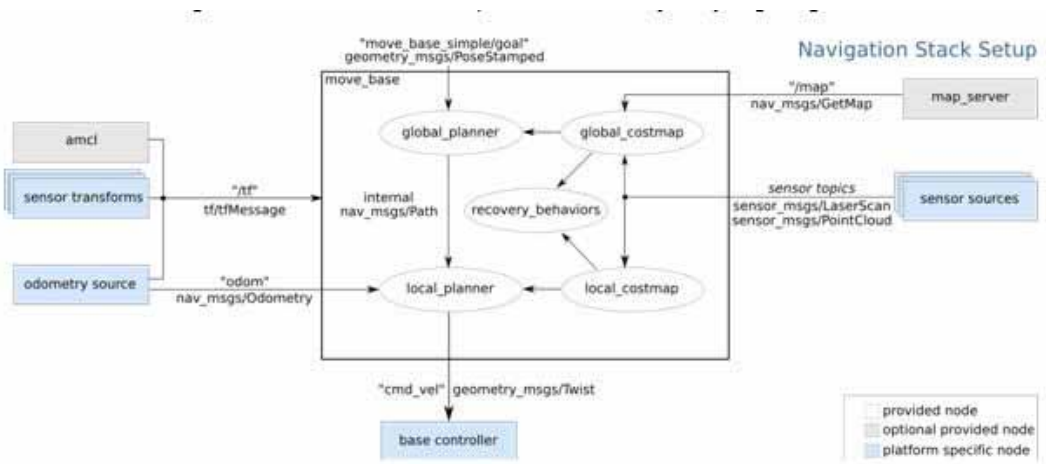


Figure 7. Navigation Stack setup and communication between different nodes of the Robot

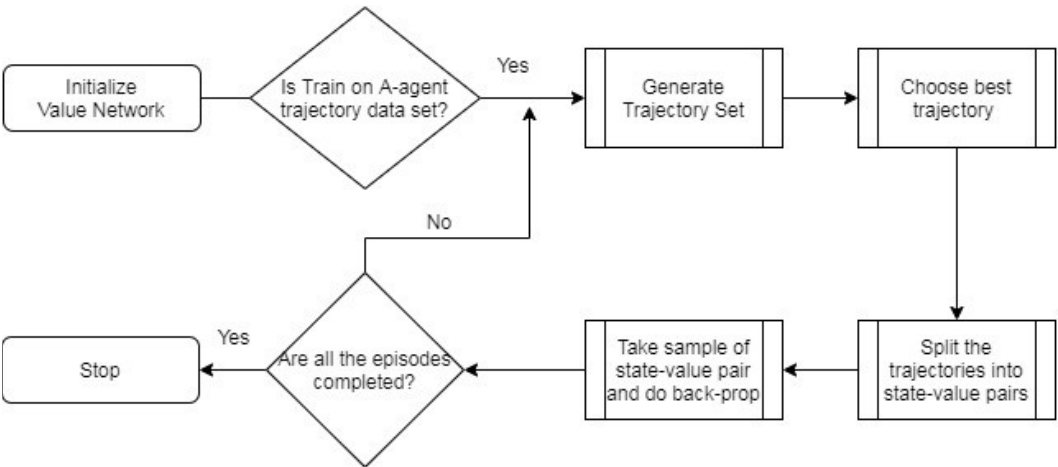


Proposed Path Planning Solution

The path planning algorithm framed as socially aware deep reinforcement value learning on the multi-agent environment (SADRVL). The drawback of CADRL is that the neural network cannot be scaled on more than two agents. SADRVL uses a shared weight policy by which it can get a multi-agent value network more than two agents. Weights are shared and scaled followed by max-pooling layer for feature aggregation. The Rectified Linear Unit (Relu) is being used as the non-linear activation function in the hidden layers.

The SADRVL algorithm initializes the value net by training on A-agents trajectory data set through NN-regression as illustrated in Figure 8. A value neural network provides policies and trajectories from random test cases. It transforms these trajectories into the state-value pair also account for them as experience set E, Eb. Now, it takes the portion of state-value pair sampled from E, Eb. And performs an update on value network using the back-propagation technique. This process is repeated on all the pre-specified number of episodes.

Figure 8. Proposed SADRVLN algorithm



RESULTS

Computational Details

The A-Agent SADRVLN network is trained in real-time access. Building the system with an i7 CPU with 4GB of GPU memory, Python's use of a five- agent SADRVLN policy takes an average of 15ms for each query receiving the appropriate action. The training took 6 hours to receive 2,000 training sessions. The five-agent system took longer to train because its space was much larger than that of the two-agent system in CADRL. The training process was repeated many times and all runs to the same policy - indicating the appropriate social goals for all test cases in the test set.

CONCLUSION

This paper presents the implementation of Autonomous Navigation using deep reinforcement learning in a ROS framework. The process navigating the robot autonomously in the highly stochastic environment starts from mapping the environment, localization of the robot concerning the environment, then which is the desired action to be taken to reach to the goal position avoiding the obstacles along the way for this step Deep reinforcement learning technique is used based on Markov decision process, an actor-critic method involved specifically our proposed socially-aware deep reinforcement value learning algorithm is used giving the most probable action based on multi-agent environment moving around without having any prior knowledge of their behavior. The approach is successfully demonstrated in the Gazebo simulation environment and RViz visualization tools.

In future works, planning to make the deep learning model light enough which can be run on cloud architectures is there also and on edge devices along with adding more capability to its prediction efficiency.

REFERENCES

- Bai, H., Cai, S., Ye, N., Hsu, D., & Lee, W. S. (2015). Intention-aware online POMDP planning for autonomous driving in a crowd. *IEEE International Conference on Robotics and Automation (ICRA)*, 454-460. doi:10.1109/ICRA.2015.7139219
- Berg, J. V. D., Guy, S. J., Lin, M., & Manocha, D. (2011). Reciprocal n-Body Collision Avoidance. In C. Pradalier, R. Siegwart, & G. Hirzinger (Eds.), *Robotics Research. Springer Tracts in Advanced Robotics* (Vol. 70). Springer.
- Berg, J. V. D., Lin, M. C., & Manocha, D. (2008). Reciprocal Velocity Obstacles for Real-Time Multi-Agent Navigation. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.
- Chen, Y. F., Everett, M., & How, M. (2017). *Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning*. .10.1109/ICRA.2017.7989037
- Dooraki, R. (2018). An End-to-End Deep Reinforcement Learning- Based Intelligent Agent Capable of Autonomous Exploration in Unknown Environments. *Sensors (Basel)*, 18(10), 3575. doi:10.3390/s18103575 PMID:30360397
- Ferrer, G., Garrell, A., & Sanfeliu, A. (2013). Robot companion: A social-force based approach with human awareness-navigation in crowded environments. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1688-1694. doi:10.1109/IROS.2013.6696576
- Ferrer, G., & Sanfeliu, A. (2014). *Behavior estimation for a complete framework for human motion prediction in crowded environments*. In *IEEE International Conference on Robotics and Automation*. ICRA.
- Fox, D., Burgard, W., & Thrun, S. (1997). The dynamic window approach to collision avoidance. *IEEE Robotics & Automation Magazine*, 4(1), 23-33. doi:10.1109/100.580977
- Gohane, S. P., & Khekare, G. S. (2015). Reconfiguration of industrial embedded system in WSN. *IEEE 9th International Conference on Intelligent Systems and Control (ISCO)*, 1-5. doi:10.1109/ISCO.2015.7282284
- Helbing & Molnar. (1995). *Social force model for pedestrian dynamics*. Academic Press.
- Khekare, G., Verma, P., Dhanre, U., Raut, S., & Sheikh, S. (2020). The Optimal Path Finding Algorithm Based on Reinforcement Learning. *International Journal of Software Science and Computational Intelligence*, 12(4), 1-23.
- Khekare, G. S. (2014). Design of emergency system for intelligent traffic system using VANET. *International Conference on Information Communication and Embedded Systems (ICICES2014)*, 1-7. doi:10.1109/ICICES.2014.7033910
- Khekare, G. S., Dhanre, U. T., Dhanre, G. T., & Yede, S. S. (2019). Design of Optimized and Innovative Remotely Operated Machine for Water Surface Garbage Assortment. *International Journal on Computer Science and Engineering*, 7(1), 113-117.
- Khekare, S., & Janardhan, S. (2017). Stability Analysis of A Vector Host Epidemic Model. *Asian Journal of Mathematics and Computer Research*, 21(3), 98-109.
- Kim, G. S., Liu, S., Lau, W., Lin, R., & Manocha, M. (2014). Predicting Pedestrian Trajectories Using Velocity-Space Reasoning. *The International Journal of Robotics Research*. Advance online publication. 34
- Kim, K., Kim, D., & Lee, J. (2018). Deep Learning Based on Smooth Driving for Autonomous Navigation. *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 616-621. doi:10.1109/AIM.2018.8452266
- Kim, P., & Pineau, J. (2016). Socially Adaptive Path Planning in Human Environments Using Inverse Reinforcement Learning. *International Journal of Social Robotics*, 8(1), 51-66. doi:10.1007/s12369-015-0310-2
- Kretzschmar, H., Spies, M., Sprunk, C., & Burgard, W. (2016). Socially compliant mobile robot navigation via inverse reinforcement learning. *The International Journal of Robotics Research*, 35(11), 1289-1307. doi:10.1177/0278364915619772
- Mehta, D., Ferrer, G., & Olson, E. (2016). Autonomous navigation in dynamic social environments using Multi-Policy Decision Making. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1190-1197. doi:10.1109/IROS.2016.7759200

Phillips, M., & Likhachev, M. (2011). SIPP: Safe interval path planning for dynamic environments. *IEEE International Conference on Robotics and Automation*, 5628-5635. doi:10.1109/ICRA.2011.5980306

Trautman, P., Ma, J., Murray, R. M., & Krause, A. (2015). Robot navigation in dense human crowds: Statistical models and experimental studies of human–robot cooperation. *The International Journal of Robotics Research*, 34(3), 335–356. doi:10.1177/0278364914557874

Unhelkar, V. V., P'erez-D'Arpino, C., Stirling, L., & Shah, J. A. (2015). Human-robot co-navigation using anticipatory indicators of human walking motion. *IEEE International Conference on Robotics and Automation (ICRA)*, 6183-6190.

Ganesh Khekare is pursuing a Ph.D. in Computer Science & Engineering from Bhagwant University, India. He pursued Master of Engineering from G H Raison College of Engineering, Nagpur in the year 2013, and Bachelor of Engineering from Priyadarshini College of Engineering, Nagpur in 2010. Currently working as an Assistant Professor at G. H. Raison College of Engineering, Nagpur since 2012. He is a member of ACM since 2014, a life member of the ISTE and IEEE, IAENG, IFERP, IERD member since 2018. He has published more than 15 research articles in reputed international journals and conferences including Thomson Reuters, IGI Global, Inderscience, Springer, IEEE, etc. He has filled one patent and one copyright. His main research work focuses on the Internet of Everything, Machine Learning, VANET, Image Processing, Smart Cities, Data Analytics, Intelligent Transportation System, etc. He has more than 9 years of teaching experience and 4 years of research experience.

Shahrukh Sheikh is working as an Assistant Professor at G H Raison College of Engineering, India. He received his Master of Technology degree in robotics from Amrita School of Engineering, Amrita University, India. Also received his Bachelors of engineering degree in Computer Science from Nagpur University in 2016. His Areas of research includes Computer Vision, Robotics, Deep learning.