2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)

Navigation in Multi Robot System using Cooperative Learning : A Survey

Priyanka Singh*, Ritu Tiwari†, Mahua Bhattacharya‡
*†‡ *Robotics & Intelligent System Design Lab,Indian Institute of Information Technology & Management,Gwalior,India*
*priyankascpmt@gmail.com,†tiwariritu2@gmail.com,‡mb@iiitm.ac.in

***Abstract*—Multi-Robot system became an important research area in the Robotics and Artificial Intelligence. In the presence of obstacles, uncertainty and incomplete information Multi-Robot system is used to accomplish the task in more efficient way as compared to single robot system. Navigation of robots in its surrounding is essential to avoid unacceptable situation such as avoiding obstacles, trajectory planning. Cooperative behavior of multi-robots provides efficient way to avoid collision. It is a matter of concern that how a group of mobile robots should be controlled so that they can move in a cohesive way toward a single direction. For the problem of cooperation among robots, flocking strategy is a good solution. To learn cooperation among robots, various machine learning strategies have been developed. One of the machine learning techniques is called reinforcement learning. It is a very challenging issue in the area of robotics and artificial intelligence. Combining cooperative and learning strategy collision avoidance can be improved. Survey of multi-agent reinforcement learning and flocking control is presented in this paper.**

***Index Terms*—Navigation, Multi Robot, Reinforcement Learning, Flocking Strategy, Collision Avoidance, Predator**

I. Introduction

There are several advantages of Multi-Robot System (MRS) over single robot system. A collection of robots can be used to perform the difficult task in an unknown environment, partial information, computation and distributed control which can not be performed effectively by single robot system. MRS can be efficiently used in many applications such as spacecraft, mobile robots, Unmanned aerial vehicles (UAVs), autonomous underwater vehicles (AUVs) and in other robot applications. Thus, MRS has been given a lot of attention in last decade [1], [2].

Although a lot of development has been done in MRS but there have many challenging issues are still present. These issues include cooperation control, path planning, collision avoidance, task allocation, communication among robots, coordination, navigation etc. Robot navigation is an important issue in the field of robotics. Navigation can be defined as an ability of a robot to find out its location and then find the desired path towards its goal. Robot or any other mobility device need a mapping of its environment and capability to represent that to navigate in its environment.

For problem related to robot navigation [3], a number of algorithms have been developed till now. The aim of navigation is to provide collision free path towards the desired goal point by guiding the robots. In the presence of unwanted obstacles, it is difficult to design a fast and efficient method for navigation of mobile agents.

Concept of navigation can be explained with three fundamental terms

- Self-localization
- Path finding
- Map-building

Robot localization is the ability of robot to determine its location and direction within the reference frame. *Path finding* is extension of localization that is used to determine collision free path from robots current position to the goal position, both position should have the same frame of reference. *Map building* may be any representation which describes a position in the reference frame of robots.

The primary requirement of any mobile autonomous robot is obstacle avoidance. In general Robot has many sensors mounted on it which acquire information about its environment. To detect obstacles different types of sensors may be used such as IR sensors, range finder, ultrasonic sensors. One of the methods that can be used for the purpose of avoiding collision [4] in multi robot system is cooperative behavior among the robots. Cooperative behavior is moving of robots in group towards desired goal avoiding undesired situation in between the path.

When a group of robots are located for any type of mission then there is a possibility that enemy may attack robot team. Robot network may be broken with the attack by opponent. So in these situations team of robots should have the capability to avoid opponent or predators. It is prudent that crew of robot should avoid predator, and also it should maintain the network structure and coordination among robots. From the biology, it can be seen that when a group of animal, fish of school and flock of birds move together, they produce overload on opponent or predator.

978-1-5090-0082-1/16/$31.00 ©2016 IEEE

Figure 1. Flock Behavior [5], [6], [7], [8]

There are many advantages of mobile robots such as redundancy and flexibility. Team of mobile robots sometimes fulfill the task that would be difficult for a single robot. Collective movement of a team of the mobile robot is to control mobile robots and make them able to move in a group towards desired target. Collective movement is not only for moving from one place to another, but it also perform a complex task. Cooperation in multi agent (MAS) has been focused a lot during last decade. Till now many methods have been proposed for this area such as genetic algorithm(GA), neural networks and reinforcement learning etc.

The remaining part of this paper has been structured as follows. Section II presents a literature review of cooperative strategy and reinforcement learning. In section III research gaps are mentioned and at last, in section IV, a precise conclusion of survey paper is given.

II. Literature Review

In this section, the brief introduction of related literature is given. This work involves collective movement and learning of team of robots to avoid an unwanted situation. Many flocking strategies are used for controlling the group behavior of robots. One of the learning techniques that are called reinforcement learning is used to learn robot from its environment. So work on both the strategies is presented separately.

A. Flocking strategies

Flocking is a collection of cooperating agents that have a common objective. An Example of these agents is a group of birds, fish school, ants, bees, crowd, etc. Member of these flocks is known as alpha agent. Flocking algorithms have the property of organizing, healing and configuring in distributed system. Scientists from different areas like physics, social science, computer science have been attracted by the evolution of flocking behavior. Various artificial intelligence techniques, computational functions and geometric strategy are used by flocking strategy to achieve the global objective by controlling the local parameter.

Reynolds gave three rules related to flocking which is known as ***Flocking rules of Reynolds*** [9].

Three flocking rules of Reynolds are given as:

1) **Flock Centering:** Neighborhood of agents is maintained
2) **Collision Avoidance:** collision with neighboring agents and other obstacles should be avoided.
3) **Velocity Matching:** velocity of agent should be matched with all the other agents.

Distribution of global and local knowledge among the different agents play an important role in controlling the flocking behavior of multi-agent system. Depending on the nonuniform distribution of knowledge in multi-agent system various structures have been presented.

- Leader Follower Structure
- Virtual Leader Structure
- Pinning Structure
- Soft Control Structure

1. Leader Follower Structure. In this structure two kinds of agents are present. First one is Leader agent and second is Follower agent. Leader agents have some extra features than the Follower agents and they can work independently. Leader can be controlled externally. There are some rules by which Follower can differentiate Leader. Follower agents are supposed to chase Leader agent while preserving the group behavior. In this structure, one or more than one agent can be elected as a leader who has knowledge of path towards the target. Remaining agents are Follower agents, and they do not have global information about the desired target, but even they can communicate with each other. This structure is used in many multi-robot algorithms in which Leader have information of desired path, and they have more facility as compared to Follower agents [10], [11].

In flocking strategy members of multi-robot system communicate with each other to produce information. It may be the case that some robots may get crashed in between and they will stop functioning and never recovered. Remaining robots will not differentiate active robot in waiting condition from the damaged robot. Thus the desired formation will be not be maintained while navigation. So it is important to develop a procedure that can discriminate working robots in active state to crashed robots. Yoshida et al. [12] have given a method in which working robots can be selected in waiting state from the team of mobile agents. Authors have only taken the case of crash in the initial stage i.e. when a robot gets damaged from starting point and makes no motion.

2. Virtual Leader Structure. Virtual leader [13] is not real agent it is a moving reference point that is used to guide other robots in the flock. In [14] it is described that even if some agents have the information about target then also flocking will be stable. In this structure also, multiple virtual leaders can be selected. The concept of multiple virtual leaders is discussed in [15]. All follower agents should have knowledge of leader agent in pure virtual leader structure [16].

3. Pinning structure . For stable flock structure, it is not necessary to have direct control of each agent in a dynamic network. So knowledge about a leader to every agent is not needed. Only some of agents can be informed about the leader, and these agents are known as informed agents. It is given in [17] that even by informing only some agents flocking behavior will be satisfactory.

4. Soft Control structure. In this structure special type of robots are used which can be managed externally. These robots are treated as ordinary agents by all the other robots [18]. Local rules will be same for this structure. For complex network soft control structure is an efficient method.

Canepa and Potop-Butucaru [19] have given a flocking strategy in which member of the flock will elect a leader on the basis of probability. After that robots adjust their position in formation and formation will move ahead. But in this algorithm formation will only move in one direction. Leaders are selected on the probabilistic method but once selected it will not be changed. In the case of faulty robots, this may be a significant limitation.

Target tracking is one of the important tasks in mobile sensor networks. To control a mobile sensor network for the purpose of target tracking flocking strategy can be used. There exist some limitations in the existing work when there is a change in the environment. When mobile robot system moves across small space surrounded by obstacles, there arises some change in the network. These changes are

- Connectivity may be lost due to split in the network.
- Configuration of network may be changed totally.
- Due to low speed, network may have poor performance

To deal with these problems an adaptive flocking control strategy is designed.

In this paper [20] authors discussed a method in which each agent determines networks parameter in decentralized fashion so that obstacle avoidance can be done and performance of tracking, connectivity, configuration can be enhanced. In the flocking strategy network size can be reduced to zero to deal with the environmental change so connectivity and similar configuration of the network remains unchanged and that is why neighbors of each robot can be unaltered. So this algorithm allows to reduce the complexity of the network by keeping the same topology during path finding process. Due to wide application of flocking strategy in the different area it has received considerable attention.

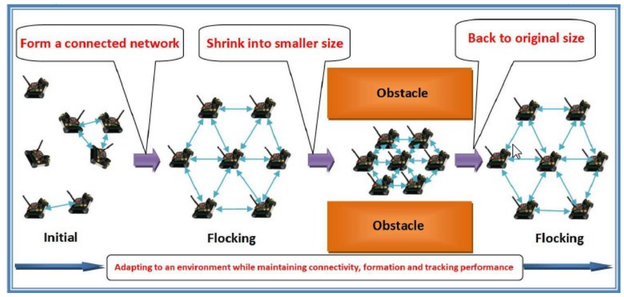

Figure 2. Adaptive flocking control strategy [20]

B. Reinforcement Learning

Littman [21] proposed the concept of reinforcement learning in 1994. It is a learning technique in which agents interact with the surrounding environment and uses trial and error technique to obtain a policy. This learning technique is self-adjusting in nature and has online learning property. In multi-robot system to learn cooperative behavior among agents reinforcement learning technique is used. It is difficult to anticipate all the cases a robot may experience and define all the action optimally in advance while designing MRSs. A robot that grasp by its action does not need a teacher who presents situation and advice suitable action that should be taken. When robot encounters any situation while experiencing surrounding of a robot, it simply tries different action and select most appropriate action depending on reinforcement and performance feedback signal. Reinforcement learning has the advantage of improving its behavior continuously and adopting new environment. So control of the robot network and knowledge of environment become two important issues in MRSs. The purpose of reinforcement learning is know how to react to an unexpected situation in dynamic environment from past events by maximizing reward and minimizing the cost function.

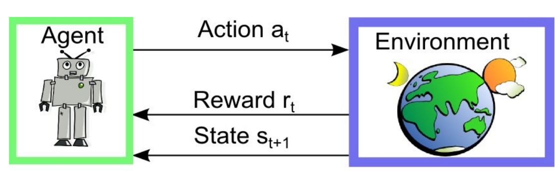

Figure 3. Reinforcement Learning setup [22]

In [23] author proposed reactive mobile robots in which reinforcement learning is used. Complex and continuous events and actions are described in the given structure. This structure is based on two concepts. In the first one reactive component is arranged in a group of module. In the second concept it is planned to find out which section of action space should be given more attention for each situation. Mobile robot should be able to deal with continuous situations. If actions of robot are taken discrete in nature then it may be possible that some reasonable solution may be neglected and unreasonable solution may be adopted. Planning is not done in advance in pure reactive system. Actions are selected immediately depending on the current situation. There are some sensors inbuilt in the reactive system, so reading of sensor is used to take appropriate action.

Reinforcement learning based on the vision is proposed in [24] to obtain cooperative behavior in the dynamic environment. This method is used effectively in robot soccer game that is initiated by Robocup. Team members of each group work together to obtain common goal against the opponent. Designing robot with visual information which have capability of learning and ability to execute task is a very challenging issue. Visual information and actions are tightly coupled to each other and are not separable. As humans are not able to see without eye this implies action in robot are significantly affected by visual information and vice versa.

Learning agent should estimate suitable state vectors for learning to be successful. However learning agent can not acquire all the information that is necessary for estimation because of limitation of sensors sensing range capability. Learning agent can acquire all the data observed and get the mapping between observed agents and behavior of learner.

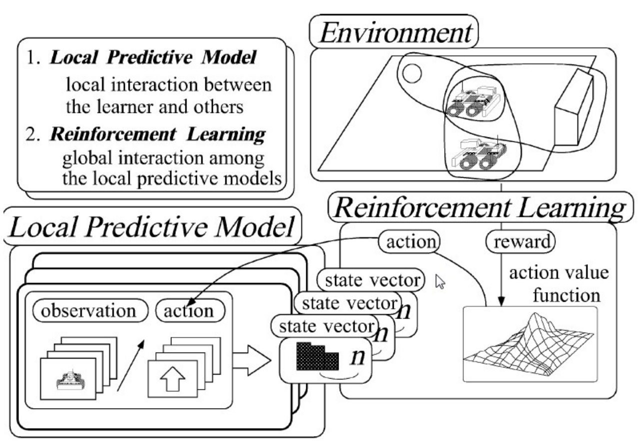

Figure 4. Learning Architecture [24]

Learning architecture of each robot is shown in the fig. Combining the output of sensors and also its action local predictive model is constructed by each learning agents. It requires vector states with the help of which it can forecast future state vectors in the dynamic environment. Then it learns cooperation on the basis of the state vector that is obtained by local predictive model. This method is two phase learning process and explanation is given as: group of robots coordinate with each other and local predictive model is constructed by learning agent considering all these interaction. It is very difficult to obtain all the sequences and so the suitable model since size of state vector becomes very large. So for all the other robots, learning robot first obtain the local predictive model and through post reinforcement learning process higher interaction among all the robots are obtained.

C. Collision Avoidance

Navigation in robots is an important issue in robotics. Depending on the surrounding environment of robots, navigation algorithms can be classified as local or global. Navigation is known as global navigation when surrounding of robots is known, and collision free path is selected in advance. Navigation is known as local navigation when surrounding of robots is not known and collision free path is also unknown. So in local navigation sensors are used to find out path which avoids collision. To deal with the navigation problem in robots, a number of robot navigation algorithms have been designed till now by many researchers.

In many cases multiple robots operate in the same environment. When separation between robots are more then they work independent to each other, but when robots gets closer to each other there may be possibility of collision so all robots must coordinate each other to avoid collision and deadlock condition. For example, when two robots wants to pass through narrow space and both have opposite direction then one have to wait until other one passes.

Dynamic obstacle avoidance is an challenging issue for mobile robots. When undesired objects whether it is static or dynamic appear robot should have capability to avoid them without having collision and should returned to its original path after these obstacle passed. Dynamic obstacle avoidance is known as local or reactive obstacle avoidance.
Local navigation is an important issue. A robot must move in the direction of target avoiding all the unwanted objects in between the path. To achieve this goal usually robot uses sensors like SONAR sensors and rangefinder etc. It is difficult to move locally using vision only. But assigning this ability to robot is very challenging. A visual local navigation is proposed in [25] inspired by human navigation through vision.

R. Abiyev et al. [26] given a method in which uses classical and fuzzy based algorithm for collision avoidance in static environment. In this environment of simulation is menu driven where obstacles of normal size are drawn and start and end point of robots are also defined. Robot is reach to its desired target position without avoiding obstacles. Navigation algorithm and dynamic behavior of robots are important parameters in planning the trajectory where getting the knowledge about location of obstacles desired path towards goal can be determined. Basic concept used here is potential field method (PFM) which is a trajectory determination method.

TABLE 1. LITERATURE REVIEW BASED ON FLOCKING STRATEGY

Year	Author	Contribution
1987	Reynolds, Craig W	Given three rules of flocking
1997	T. Masuzawa and H. Fuji-wara	Given a method to elect the working agents from a team of mobile agents.
2006	Han, Jing ,Ming Li, and Lei Guoic	Soft control structure of flocking is given.
2007	Su, Housheng, Xiaofan Wang, and Zongli Lin	Discussed that flocking will be stable even informing few robots about leader.
2007	D. Canepa and M. G.Potop-Butucaru	Probability based leader selection method is proposed.
2009	Gu, Dongbing, and zongyao Wang	Explained Leader follower flocking structure
2009	Hung Manh La; Weihua Sheng	Adaptive flocking control strategy is given
2010	Wang, Xiaofan, Xiang Li, and Jinhu Lu	Explained Pinning structure of flocking

TABLE 2. LITERATURE REVIEW BASED ON REINFORCEMENT LEARNING

Year	Author	Contribution
1994	M.L. Littman	Proposed reinforcement learning
1995	Jos del R. Milln	Designed reactive mobile robot system using . reinforcement learning
1999	Minoru Asada, Eiji Uchibe, Koh Hosoda	Cooperative strategy via vision-based reinforcement learning

TABLE 3. LITERATURE REVIEW BASED ON COLLISION AVOIDANCE

Year	Author	Contribution
1997	Chia-Han Lin, Ling-Ling Wang	Obstacle avoidance using fuzzy algorithm.
2006	Wesley H. Huang,Brett R.Fajen, Jonathan R. Fink,William H. Warren	Collision avoidance using visual information
2010	R. Abiyev, D. Ibrahim, B. Erin	Combination of both classical and fuzzy algorithms for avoiding collision in static environment.

III. Research Gaps

MRSs have given significant concern during last decade. Robot should learn and adopt surrounding environment. So learning environment and controlling robot system becomes two important challenges. To gain more number of rewards agent should take actions which have been already tried in past but to find these action it should try new actions. Robot should exploit the previous information to gain rewards but to improve actions it should explore. So neither exploration nor exploitation can be done exclusively. In most of work it is assumed that robot which is in learning state have full knowledge of other robots. But it is not possible for robot to get full information about other robots especially for competitive games. Mostly all reinforcement learning strategies are value based methods. Thus if these techniques are applied for continuous domain it has to be approximated. So this creates a little gap in calculated and exact value. Maintaining cooperation and tracking efficiency is still an open challenge when size of network increases.

IV. Conclusion

In this paper navigation in multi robots maintaining the collective behavior among them is explained. Collective movement of group of robots can be maintained by using flocking strategy. Flocking is a kind of behavior of group of robots which are interacting with each other having common objective. Flocking is applied so group of robot can move in cohesion towards a common direction. One of the machine learning techniques that is reinforcement learning is used for multi-agent system that enables robot to learn coordination. Integrated form of reinforcement learning and flocking control is used for collision avoidance.

References

[1] M. J. Matarić, "Reinforcement learning in the multi-robot domain," in *Robot colonies*. Springer, 1997, pp. 73–83.

[2] C. F. Touzet, "Robot awareness in cooperative mobile robot learning," *Autonomous Robots*, vol. 8, no. 1, pp. 87–97, 2000.

[3] V. Di Gesù, B. Lenzitti, G. Lo Bosco, and D. Tegolo, "A distributed architecture for autonomous navigation of robots," in *Computer Architectures for Machine Perception, 2000. Proceedings. Fifth IEEE International Workshop on*. IEEE, 2000, pp. 190–194.

[4] C.-H. Lin and L.-L. Wang, "Intelligent collision avoidance by fuzzy logic control," *Robotics and Autonomous Systems*, vol. 20, no. 1, pp. 61–83, 1997.

[5] [Online]. Available: http://depositphotos.com/8295259/stock-photo-swarm-of-bees.html

[6] [Online]. Available: http://www.fws.gov/r5fedaid/vt.html

[7] [Online]. Available: http://picsant.com/12009814-tahiti-snorkeling-scuba-diving-fish.html

[8] [Online]. Available: http://www.birdsasart-blog.com/page/308/

[9] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," in *ACM Siggraph Computer Graphics*, vol. 21, no. 4. ACM, 1987, pp. 25–34.

[10] Y. Hanada, G. Lee, and N. Y. Chong, "Adaptive flocking of a swarm of robots based on local interactions," in *Swarm Intelligence Symposium, 2007. SIS 2007. IEEE*. IEEE, 2007, pp. 340–347.

[11] A. Kumar, S. Sharma, R. Tiwari, and S. Majumdar, "Area exploration by flocking of multi robot," *Procedia Engineering*, vol. 41, pp. 377–382, 2012.

[12] S. Carpin and L. E. Parker, "Cooperative leader following in a distributed multi-robot system," in *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, vol. 3. IEEE, 2002, pp. 2994–3001.

[13] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *Automatic Control, IEEE Transactions on*, vol. 51, no. 3, pp. 401–420, 2006.

[14] H. Su, X. Wang, and Z. Lin, "Flocking of multi-agents with a virtual leader," *Automatic Control, IEEE Transactions on*, vol. 54, no. 2, pp. 293–307, 2009.

[15] H. Su, X. Wang, and W. Yang, "Flocking in multi-agent systems with multiple virtual leaders," *Asian Journal of control*, vol. 10, no. 2, pp. 238–245, 2008.

[16] S. Souissi, Y. Yang, and X. Défago, "Fault-tolerant flocking in a k-bounded asynchronous system," in *Principles of Distributed Systems*. Springer, 2008, pp. 145–163.

[17] H. Su, X. Wang, and Z. Lin, "Flocking of multi-agents with a virtual leader part i: with a minority of informed agents," in *Decision and Control, 2007 46th IEEE Conference on*. IEEE, 2007, pp. 2937–2942.

[18] J. Han, M. Li, and L. Guo, "Soft control on collective behavior of a group of autonomous agents by a shill agent," *Journal of Systems Science and Complexity*, vol. 19, no. 1, pp. 54–62, 2006.

[19] D. Canepa and M. G. Potop-Butucaru, "Stabilizing flocking via leader election in robot networks," in *Stabilization, Safety, and Security of Distributed Systems*. Springer, 2007, pp. 52–66.

[20] H. M. La and W. Sheng, "Adaptive flocking control for dynamic target tracking in mobile sensor networks," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 4843–4848.

[21] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *Proceedings of the eleventh international conference on machine learning*, vol. 157, 1994, pp. 157–163.

[22] [Online]. Available: http://www.ias.tu-darmstadt.de/Workshops/NIPS2014

[23] J. d. R. Millán, "Reinforcement learning of goal-directed obstacle-avoiding reaction strategies in an autonomous mobile robot," *Robotics and Autonomous Systems*, vol. 15, no. 4, pp. 275–299, 1995.

[24] M. Asada, E. Uchibe, and K. Hosoda, "Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development," *Artificial Intelligence*, vol. 110, no. 2, pp. 275–292, 1999.

[25] W. H. Huang, B. R. Fajen, J. R. Fink, and W. H. Warren, "Visual navigation and obstacle avoidance using a steering potential function," *Robotics and Autonomous Systems*, vol. 54, no. 4, pp. 288–299, 2006.

[26] R. Abiyev, D. Ibrahim, and B. Erin, "Navigation of mobile robots in the presence of obstacles," *Advances in Engineering Software*, vol. 41, no. 10, pp. 1179–1186, 2010.

[27] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *Robotics, IEEE Transactions on*, vol. 21, no. 3, pp. 376–386, 2005.

[28] M. Cao, C. Yu, and B. Anderson, "Coordination with the leader in a robotic team without active communication," in *Control and Automation, 2009. MED'09. 17th Mediterranean Conference on*. IEEE, 2009, pp. 252–257.

[29] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *Robotics, IEEE Transactions on*, vol. 21, no. 3, pp. 376–386, 2005.