© https://fti-tn.net/publications Future Technologies and Innovations (FTI) Proceedings: 4th international conference on computer applications and information security (iccais'2021) / March 19 / 2021/ Tunisia: https://fti-tn.net/iccais-2021-list-of-papers

Robotics Wireless Sensors Network with Reinforcement learning

Sondes Dallali
Innov'com Laboratory, SUP' COM
University of Carthage
Sondes.Dallali@fsb.rnu.tn

Fatma Louati
Innov'com Laboratory, SUP' COM
University of Carthage
fatma.louati@enicar.univ-carthage.tn

Kaouther Sethom
Innov'com Laboratory, SUP' COM
University of Carthage
kaouthar.sethom@enicar.univ-carthage.tn

***Abstract*—Robotic Wireless Sensors Networks (RWSN) is an autonomous networked robots' system that aims to achieve certain sensing goal while meeting and maintaining certain communication performance requirements. Interact robot with machine learning is an optimal idea to maintain function for robots and facility interaction between sensors and base station (BS).Our works proposed a robotics Wireless sensors networks with reinforcement learning algorithm to search an optimal path for robots and minimize energy consumption for sensors.**

Keywords: RWSN, reinforcement learning, Artificial Intelligence, energy consumption

1. INTRODUCTION

Wireless sensors networks are a best solution to collect data in hazardous environment and can be used in a variety of application such as military surveillance health monitoring and agriculture, the sensors are limited in battery and capacities and cannot move and maintain connectivity they can distributed over a vast area such as in disaster area or harsh remote environment where it is difficult for people to function.

Mobile sensor node is required to accomplish many difficult tasks such as maintain connectivity collecting data to base station node replacement hole and partition recovery autonomous deployment and redeployment moving and detect obstacle.

Wireless sensor network (WSN) is one of the most promising technologies for some real-time applications because of its size, cost-effective and easily deployable nature. The job of WSN is to monitor a field of interest and gather certain information and transmit them to the base station for post data analysis. Some of the WSN applications consists of many sensor nodes. Therefore, managing such many nodes requires a scalable and efficient algorithm. In addition, due to the external causes or intended by the system designers, the WSNs may change dynamically. Therefore, it may affect network routing strategies, localization, delay, cross-layer design, coverage, QoS, link quality, fault detection, etc. Because of the highly dynamic nature, it may require depreciating dispensable redesign of the network, but the traditional approaches for the WSNs are explicitly programmed, and as a result, the network does not work properly for the dynamic environment.[9]

Machine Learning (ML) is the process that automatically improves or learns from the study or experience and acts without being explicitly programmed. ML was making our computing processes more efficient, reliable and cost-effective. ML produce models by analyzing even more complex data automatically, quickly and more accurately. It is mainly classified into supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. The strength of ML lies in their ability to provide generalized solutions through an architecture that can learn to improve its performance. Because of the interdisciplinary nature, it plays a pivotal role in various fields including engineering, medical, and computing.[10] Recent advances in ML have been applied to solve various issues in WSNs. Applying ML not only improves the performance of WSNs and limits the human intervention or reprograming [11].

The main contribution of this paper is to introduce some robots inside the sensor networks to reduce energy for nodes and maintain connectivity between sensors and increase lifetime for network. AML algorithm pilots the robots for predictive analysis or to autonomously make decisions in response to incoming events and can also be applied to historical data to identify patterns or anomalies in the Wireless sensor networks.

2. Proposed Solution

WSN is dedicated to monitor a field of interest and collect data and transmit them to the remote cloud through the nearest base station [11-13]. Therefore, managing such many nodes requires times energy, a scalable and efficient algorithm is a solution. In addition, due to the environments or causes intended by the system designers, the WSNs may change dynamically. Therefore, it may affect network routing strategies, localization, delay, cross-layer design [6], coverage, QoS, link quality, fault detection, etc. Because of the highly dynamic nature, it may require depreciating dispensable redesign of the network, but the traditional approaches for the WSNs are explicitly programmed, and as a result, the network does not work properly for the dynamic environment. Machine Learning (ML) seems to be the good solution to overcome this.

We propose to add some robots inside the WSNs and use AI to monitor the overall system. Sensors and robots will represent the body part that perform physical tasks, whereas the AI represents the brain that controls the physical entity according to the application requirements. The wireless communication network plays a vital rule in transmitting and sharing data between robots over the pervasive network, communicating with each other, and also with remote cloud. While robots will represent the middleware for data collection, AI shall be used for processing data intelligently for the architecture. The combination of robots, AI and sensors will result in robotic systems having high capabilities to perform complex tasks autonomously. In such systems, robots will facilitate high quality and high-speed data exchange amongst them and the network.

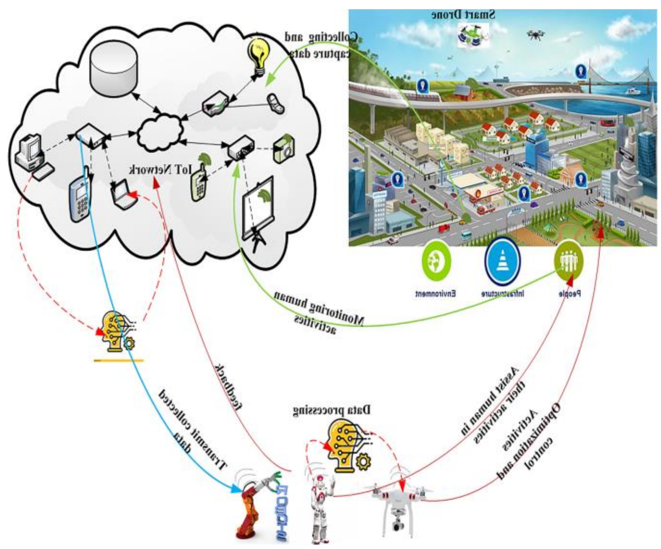


Fig.1 Synergy among ML, robots, IoT, and environment

We need some artificial intelligence to smartly monitor the system and the robots interaction with the sensors nodes: when/where to assist the WSNs; the best trajectory to follow for initial network discovery, emergency response to node's dead, AI will empower robots with some self-learning and self-reconfiguring abilities. It will help robots to move autonomously and extend robotic functions to perform any given task effectively and efficiently with better (or at least equal) skill than a human counterpart.

Moreover, AI may also be leveraged for Quality of Service (QoS) enhancement and data collection for energy efficiency. The most used metrics for robot communication systems are categorized and presented as energy metrics (related to data collection from surrounding environment), QoS metrics, mobility (related to collision prevention by assigning smooth trajectories for robots). AI techniques must successfully be applied for controlling and managing the above metrics for enhancing the effectiveness of robotic collaboration.

2.1 Network topology

First, we decide to use clustering in our wireless sensors network because this technique reduce packets transmissions and by the way save energy. To achieve this, we used a cluster head selection method proposed in our previous work [1] Fig3.To overcome connectivity problems between clusters and/or Base station, we introduce some robots in the system (one or two depending on network size) Figure 2.

The main role of robots is to:
-Discover the topology after the initial deployment of sensor nodes. This can help to decide if we need to add some nodes in the network for better coverage.
-Collect data from isolated nodes if any relay periodically information to the BS to save CHs' energy.
-Enhance the QoS for best communication inside the network if needed.

Each sensor node communicates its location and current energy level to the closest CH. Periodically, the robot following a pre-calculated path receives data (residual energy and positions) related to all sensors through cluster heads and transmit it to the BS.
The machine learning algorithm uses the collected data for the reselection for CH if necessary, to maintain connectivity It can send the robot on a particular position to recharge battery for some dead sensors and/or replace battery or even the node itself.

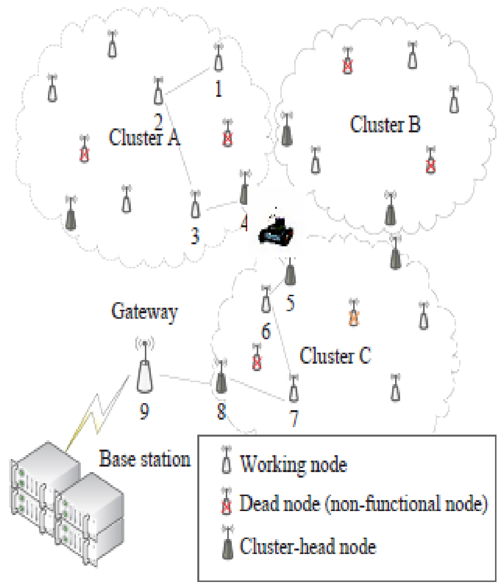


Fig.2 Robot collect data with closest CHs

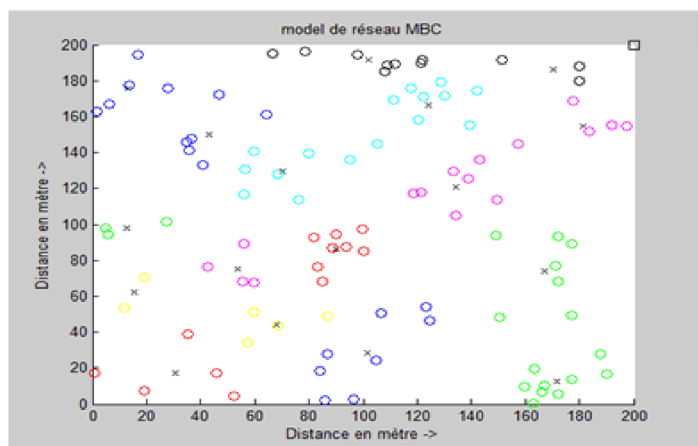


Fig.3Our network with 17CHs [1]

3.2Machine learning algorithm implementation

In our case the ML will help us in resolving the following problems:

- For target area coverage problem, deciding on additional nodes' placement to best cover the area.
- Robots' trajectory calculation for emergency intervention.

- CHs' recalculation for energy saving and node dead prevention.

- Sensor nodes may change their location due to some internal or external factors. Accurate localization is smooth and rapid with the help of ML algorithms.

- Routing data place a major role in improving the network lifetime. The dynamic behavior of sensor network requires dynamic routing mechanisms to enhance the system performance.

Different types of machine learning algorithms [8]exist in the literature such as supervised, unsupervised and reinforcement learning algorithms. We choose the last one because it is the most suited to our scenario. Reinforcement Learning (RL) is learning what to do and how to map situations to actions to maximize an environmental numerical reward signal. The learner is not told which actions to take, as in most forms of ML, but instead he must take his own action and a discovering the action which yield the most reward by keep on trying them.

Deep Q-Learning is one of important reinforcement learning is characterized by the temporal difference method which allows learning directly by experience and does not need to wait until the final result to calculate the value of each state like dynamic programming.
When a robot is moving in a discrete environment it chooses one of a set of definite behaviors at every time interval and assume that it is in markov state, the state changes to the probability s of different.

$$\Pr[S_{t+1}]=s'[s_t,a_t]=\Pr[a_t] \qquad (1)$$

At every time interval t a robot can get status s from the environment and then take action a_t it receives a stochastics prize r which depends on the state and behavior of the expected prize R_{st} to find the optimal policy that an entity wants to achieve:

$$R_{st}(a_i)=\mathrm{E}\{\sum_{i=0}^{\infty} \gamma^j\, r_{t+j}\} \ (2)$$

The discount factor means that the rewards received at t time intervals are less affected than the rewards currently received. The operational value function Va is calculated using the policy value function r and the policy value function V_p as shown in equation 3.The state value function for the expected prize when starting from state s and following the policy is expressed by the following equation:

$$V_a(S_t)=R_s(\mathrm{r}(S_t))+\gamma \sum_u Pxy[r(S_t)]V_p(S_t). \qquad (3)$$

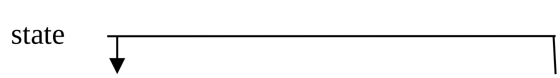
It is proved that there is at least one optimal policy as follows. The goal of Q-learning is to set an optimal policy without initial conditions.For the policy define the value as follows:

$$Q_P(S_t,a_t)=R_s(a_t)+\gamma\sum_u Pxy\,[r(S_t)]V_p(\mathrm{y}) \qquad (4)$$

Q(S_t,a_t) is the newly calculated $Q(S_{t-1},a_{t-1})$ and $Q(S_{t-1},a_{t-1})$ corresponding to the next state by the current $Q(S_{t-1},a_{t-1})$ value and the current $Q(S_{t-1},a_{t-1})$.

Fig 4 is the form of the proposed algorithm. The proposed algorithm uses empirical representation technique. The learning experience that occurs at each time step through multiple episodes to be stored in the dataset is called memory regeneration. The learning data samples are used for updating with a certain probability in the reconstructed memory each time. Data efficiency can be improved by reusing experience data and reducing correlations between samples.

Proposed algorithm:



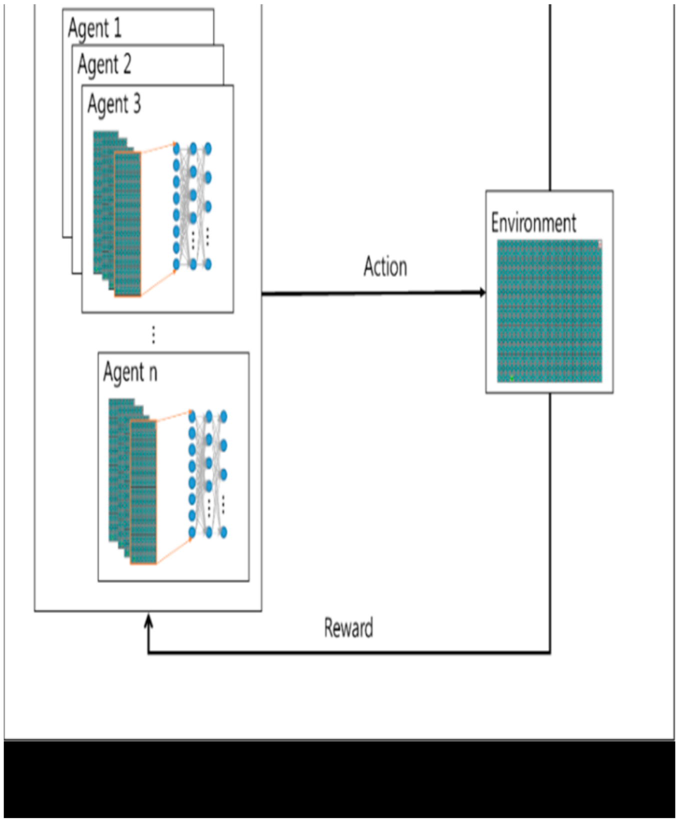


Fig.4: Proposed algorithm

The proposed algorithm uses empirical data according to the roles assigned individually. In learning how to set goal value by dividing several situations, different expectation value for each role before starting to learn is set. For example, assuming that the path planning should take less time in any case and that it is a success factor for each episode, the proposed algorithm designates an arbitrary position and number of obstacles in the map of a given size. This time is used as a compensation function of deep q algorithm.

The agent learns so that the compensation value always increases. If the searching time increases, the compensation value decreases, and learning is performed so that the search time does not increase.

In the preprocessing part, the features of the image are searched using input images, and these features are collected and learned. In this case, the Q value is learned for each robot assigned to each role, but the CNN value has the same input with a different expected value. Therefore, the Q value is shared when learning, and used for the learning machine. In order to optimize the updating of Q value, it is necessary to define an objective function as defined as error of target value and prediction value of Q value. Equation (5)
describes an objective function

$$L=1/2[\mathrm{r} + max_{a'}\mathrm{Q}(\mathrm{s}',\mathrm{a}') - \mathrm{Q}(\mathrm{s},\mathrm{a})]^2 \qquad (5)$$

The basic information for obtaining a loss function is a transition <s, a, r, s'>. Therefore , first, the Q-network forward pass is performed using the state as an input to obtain an action-value value for all.

After obtaining the environment return value <r, s'> for the action a, the action-value value for all action a' is obtained again using the state s'. Then, we get all the information to get the loss function, which updates the weight parameter so that the Q-value updates for the selected action converges; that is, as close as possible to the target value and the prediction value. Algorithm 1 is a function of compensation, which greatly increases compensation if the distance to the current target point is reduced before and decreases the compensation if the distance becomes longer and closer.

Algorithm 1 Reward Function
if distancet-1 >$distance_t$
then reward = 0.2
else
then reward = 0.2
if action == 0
then reward = reward + 0.1
else if action == 1 or action == 2
then reward = reward 0.1
else if action == 3 or action == 4
then reward = reward 0.2
if collision
then reward = -10
else if goal

```
then reward = 100
return reward
```

9. CONCLUSION

In this paper, we have introduced a new solution to interact sensors with robots and machine learning algorithm to make robots mobility more interactive. Future works will focus on simulation results with different mobility scenario for robots.

REFERENCES

[1] sondes dallali,"CHSM with multihop balanced clustering routing protocols," in ASET,2018

[2] Vaidehi Vijayakumar" Application of Machine Learning inWirelessSensorNetwork"Springer Nature Switzerland AG 2019 X.(S.) Shen et al. (eds.), *Encyclopedia of Wireless Networks*

[3] Uthman Baroudi] ,"Robot-Assisted Maintenance of Wireless Sensor Networks Using Wireless Energy Transfer,"IEEE SENSORS JOURNAL, VOL. 17, NO. 14, JULY 15, 2017

[4] S. H. Alsamhi , Oumohd. Samar Ansari "Convergence of Machine Learning and Robotics Communication in Collaborative Assembly: Mobility, Connectivity and Future Perspectives"Journal of Intelligent & Robotic Systems (2020) 98:541–566

[5] P. Rawat, K.D. Singh, H. Chaouchi , J.M. Bonnin , "Wireless sensor networks: a survey on recent developments and potential synergies", J. Supercomput 68 (1) (2014) 1–48 .

[6] D.K. Sah , T. Amgoth ," Parametric survey on cross-layer designs for wireless sensor networks", Comput Sci. Rev. 27 (2018) 112–134 .

[7] M.A. Alsheikh , S. Lin , D. Niyato , H.-P. Tan ," Machine learning in wireless sensor networks: algorithms, strategies, and applications", IEEE Commun. Surv. Tutorials 16 (4) (2014) 1996–2018

[8] Vaidehi Vijayakumar" Application of Machine Learning in Wireless Sensor Network" Springer Nature Switzerland AG 2019

[9] D. Praveen Kumar, Tarachand Amgoth∗ , Chandra Sekhara Rao Annavarapu" Machine learning algorithms for wireless sensor networks: A survey"Information Fusion 49 (2019) 1–25

[10] G. T. Sibley, M. H. Rahimi, and G. S. Sukhatme, "Robomote: a tiny mobile robotplatform for large-scale ad-hoc sensor networks," in Proceedings of the IEEE InternationalConference on Robotics and Automation, Vol. 2, 2002, pp. 1143-1148.

[11] J. P. Sheu, P. W. Cheng, and K. Y. Hsieh, "Design and implementation of a smartmobile robot," in Proceedings of the IEEE International Conference on Wireless andMobile Computing, Networking and Communications, Vol. 3, 2005, pp. 422-429.

[12] G. Wang, G. Cao, and T. LaPorta, "A bidding protocol for deploying mobile sensors,"in Proceedings of the 11th IEEE International Conference on Network Protocols,2003, pp. 315-324.

[13] M. A. Batalin and G. S. Sukhatme, "Efficient exploration without localization," in proceedings of the IEEE International Conference on Robotics and Automation, Vol.2, 2003, pp. 2714-2719.