iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 9 **Issue:** V **Month of publication:** May 2021

**DOI:** https://doi.org/10.22214/ijraset.2021.33891

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

# Application of Machine Learning in Routing and Wavelength Assignment: Literature Survey

Likhitha S[1], Sriraj Asnotikar[2], Surin Machaiah M T[3], Vinaykumar V G[4], Prof. Hamsaveni M[5]
[1, 2, 3, 4, 5]*Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysore, Karnataka, India.*

***Abstract: This paper presents the survey of the prominent issue in optical networks is Routing and Wavelength Assignment (RWA), due to the heavy traffic in network, there will be a need for wavelength assignment to the track in that network. The RWA problem is resolved using the approach of Genetic Algorithm and Ant Colony Optimization Algorithm, wherein up until now is noticed to be better than other optimization algorithms. By using Deep Q-Networks, a type of reinforcement learning, in optimizing the problem in selection of the routing path and wavelength assignment in an optical mesh network.***

## I. INTRODUCTION

Routing and Wavelength Assignment (RWA) is an optical networking problem with the goal of maximizing the number of connections in an optical network. In an optical network, each connection must be given a route and a wavelength that should be consistent for the entire path unless the wavelength is intentionally altered. Assigning these routes and wavelengths are done through various algorithms. This paper illustrates using one of the Reinforcement learning (RL) algorithms.

Reinforcement learning is an area of Machine Learning (ML) concerned with what action should be taken in an environment in order to maximize the cumulative reward. The algorithm considered here is Deep Q-Network (DQN).

The Deep Q-Network algorithm was developed by enhancing a classic reinforcement learning algorithm called Q-Learning with deep neural networks and a technique called experience replay. experience replay enables reinforcement learning agents to memorize and reuse past experiences, just as humans replay memories for the situation at hand.

### *A. Genetic Algorithm (GA)*

Genetic Algorithm (GA) are adaptive heuristic search algorithms that are based on the ideas of natural selection and genetics. They are used to generate high-quality solutions for optimization problems and search problems.

The main steps for executing GA are:

1) *Initialization:* A random set of individual chromosomes is initialized with random values. A unique cost metric is assigned to all, to identify their fitness. This function depends on the problem to be solved and also determines the fitness of each individual.
2) *Crossover:* Individual chromosomes are crossover with a specified probability, to produce the next generation individuals. After crossover, a new era of additional individuals is created in the total population. The cost of each is calculated again and the worst-performing ones are discarded. Usually, an upper bound in the number of individuals is maintained, and when this is reached all the rest are not considered for next-generation crossovers.
3) *Mutation:* In this step single gene is modified. The key idea is to insert random genes in offspring to maintain the diversity in the population to avoid premature convergence.

### *B. Ant Colony Optimization (ACO)*

The main objective of the network is to transfer the big amount of data in an easy manner to the receiver end. In optical communication there are 2 types of techniques that has been used such as time division multiplexing (TDM) and wavelength division multiplexing, In optical communication network all the source are send through a light path to different nodes for transferring of data, so we should assign wavelength to individual network, in this proposed method Ant Colony Optimization algorithm is used to solve this crisis. The basis of the Ant colony based RWA algorithm depends on launching ants as explorers, then it gone through the network to establish the connection request send by the user. Finally, through their network, ants put trails on the selected path to help the follower ants to optimally find paths of the network. Using these network trails become less attractive over time due to pheromone evaporation property which makes the trails disappear. Pheromone evaporation has a great role in avoiding the convergence and also helps to dynamic network exploration. Besides, evaporation helps to avoid stagnation which happens if all ants follow the same paths chosen by the first ants.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue V May 2021- Available at www.ijraset.com*

## II. REVIEW

### *A. Routing*

Routing is the process of selecting a path for traffic in a network. A network will have multiple nodes which are interconnected and for any source-destination node pair to communicate, there has to be a physical or virtual path. Routing determines this path using various algorithms and conditions that are not only general to all such networks but also specific to the current network. It is also not necessary that a source node and a destination node have one-to-one association. A message can also be delivered to all nodes in a network using one-to-all association or to only a few nodes using one-to-many and one-to-many-of-many associations.
Path selection involves applying a routing metric to multiple routes to select (or predict) the best route. Most routing algorithms use only one network path at a time.
In computer networking, the metric is computed by a routing algorithm, and can cover information such as bandwidth, network delay, hop count, path cost, load, maximum transmission unit, reliability, and communication cost.[1]

### *B. Wavelength Assignment*

Two of the most common methods for wavelength assignment are First Fit and Random Fit. First Fit chooses the available wavelength with the lowest index. Random Fit determines which wavelengths are available and then chooses randomly amongst them. The complexity of both algorithms is O($w$), where $w$ is the number of wavelengths. First Fit outperforms Random Fit.
There are several other wavelength assignment algorithms: Least Used, Most Used, Min Product, Least Loaded, Max Sum,[2] and Relative Capacity Loss.[3] Most Used outperforms Least Used use significantly, and slightly outperforms First Fit. Min Product, Least Loaded, Max Sum, and Relative Capacity Loss all try to choose a wavelength that minimizes the probability that future requests will be blocked.

## III. PRESENTATION OF DATA

A reinforcement learning task is to train an agent which interacts with its environment. The agent arrives at a state by performing some action. These actions result in rewards which could be positive or negative. The only purpose of the agent here is to maximize its total reward across an episode. An episode is anything and everything that happens between the Initial state and the terminal state within the environment. Agents are subjected to reinforcement learning to perform the best actions by experience. And hence this is the strategy or policy of the learning task.

### *A. Q Learning*

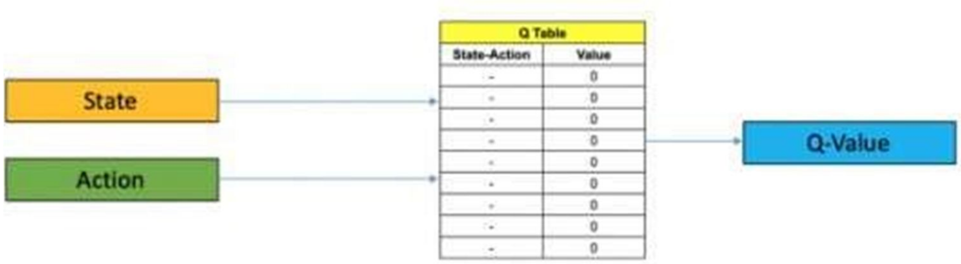



Fig 1: Q Learning

Assuming the reward for each action performed in every step is previously known, the agent will know what actions are to be performed in a particular state. The agent then performs the series of actions that in turn generates the maximum total reward. The total reward thus obtained is termed as the Q-value and hence formalize the strategy as:

$$Q(s,a) = r(s,a) + \gamma \max_{a} Q(s',a)$$

This equation states the Q-value obtained from the present state *s* and carrying out a certain action *a* yields the immediate reward r(s,a), in addition to the highest Q-value possible from the following state *s'*. $\gamma$ is known as the discount factor and it controls the future rewards that can be gained.

Q(s', a) once more depends on Q(s", a) and has a coefficient of gamma to the power 2. Thus, the current Q-value depends on Q-values of future states, giving rise to the equation:

$$Q(s,a) \rightarrow \gamma Q(s',a) + \gamma^2 Q(s'',a) \ldots\ldots\ldots \gamma^n Q(s''^{\ldots n},a)$$

Altering the value of gamma varies the contribution of future rewards. Due to the recursive nature of the equation, the arbitrary assumptions for all q-values can be done. With the gained experience, it converges to the optimal policy. In real world application, the equation can be updated and implemented as:

$$Q(S_t,A_t) \leftarrow Q(S_t,A_t) + \alpha\left[R_{t+1} + \gamma \max_a Q(S_{t+1},a) - Q(S_t,A_t)\right]$$

Here, α represents the learning rate or step size, which communicates to what extent old information is overridden by the newly acquired information.

*B. Deep Q-Networks*

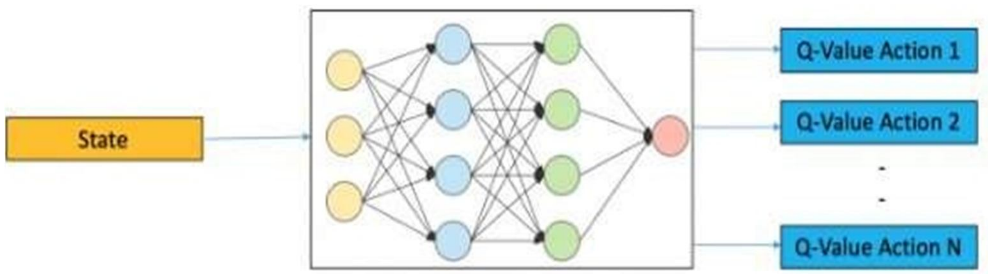



Fig 2: Deep Q-Network

In deep Q-learning, Q-value function is approximated using a neural network. The state is provided as the input and outputs the Q-value of all possible actions.

Following are the steps for deep Q-learning Networks:

1) User stores the experiences of past in the memory.
2) The subsequent action is set by the maximum output of the Q-network
3) The loss function here is mean squared error of the predicted Q-value and the target Q-value – Q*. This is a regression problem. Although, the target or actual value here is unsure, as reinforcement learning problem is dealt here. Visiting the updated Q-value equation which is derived from the Bellman equation:

$$Q(S_t,A_t) \leftarrow Q(S_t,A_t) + \alpha\left[\boxed{R_{t+1} + \gamma \max_a Q(S_{t+1},a)} - Q(S_t,A_t)\right]$$

The section in the box indicates the target. It can be debated that it predicts its own value, but R being the unbiased true reward, the gradient is updated by the network using backpropagation to converge at the end.

*C. RWA using Genetic Algorithms*

In this case, one of the k-shortest paths is represented by each gene in an individual chromosome, for a specific source-destination pair. Each chromosome consists of n genes, where n represents the number of all source-destination pairs, which inturn constitutes a result to problem in RWA.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue V May 2021- Available at www.ijraset.com*

### *D. Fitness Function*

A fitness function is an objective function type used to summarize as a single figure of merit. It is used in optimal design solutions, where it is helpful to guide the simulations. It is an application specific objective function used to evaluate relative effectiveness of the potential solutions. It can be implemented in which it is either mutable or constant.

### *E. Crossover*

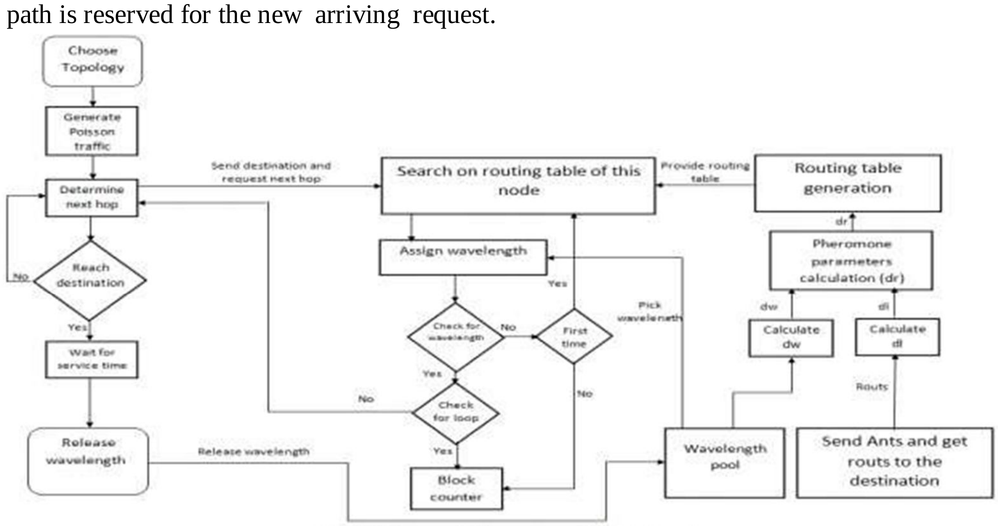
A n-point crossover is used for a chromosome, which is used when a path is assigned to a certain source-destination pair. The identity of the path is kept intact during the process. Through the conventional Roulette Wheel selection scheme, the individuals are chosen. A simple monotonic function is used in the process of mapping to the wheel, while intercalating between the best fit and the worst fit individual. After the process, an extra individual is generated and inserted in the population. If say, the population size is v, based on the fitness, the first v discrete members will be maintained and the rest will be rejected.

### *F. Mutation*

During mutation, regardless of the factor of fitness function, the chromosome created is replaced by itself. The worst-case scenario of the individual chromosome is selected for mutation. Hence in the population, a more fit individual is produced from a bad individual.

### *G. RWA using Ant Colony Optimization*

The network generated from a source traverses through all the nodes until it reaches a random destination according to ACO. At this phase, the P-route table is filled with adequate candidate paths. After that, the traffic is dynamically generated according to connection requests. When traffic is generated from source 's' to a random (target) destination 'd', then the algorithm searches in the P-route table of the source s into the list of all the routes. Given some of the possible destinations, the next path chosen is selected based on the highest probability for the next hop for each destination. The wavelength is randomly assigned among every available wavelength in the network if the chosen route has one or more unutilized wavelength. Then this procedure will be supported by each node in the target list for the origin node of each request to estimate a range of lengths for each nominee target. After the best route is selected for each connection demand in the first phase, the wavelength is selected in the second phase and finally, the selected path is reserved for the new arriving request.




Block diagram of the proposed ACO algorithm.

Fig 3: ACO Algorithm

## IV. CONCLUSION

### *A. RWA using DQN*

The nodes in the deep Q-networks structure represent the nodes of an optical fiber network and the connections between the DQN nodes are the possible paths that connect the nodes of the optical network. Assuming the state of the DQN as the transfer of data between source-destination node pair and the actions as the path that it takes, we obtain a list of Q-values for all the possible actions. Now observing the Q-values that are obtained, the route with the highest reward is selected. This will be the route for the optical network between this particular source- destination node pair.

### *B. RWA using Genetic Algorithms*

In this paper, a new approach for solving the RWA using GA in optical networks is presented. The proposed GA-RWA algorithm uses fitness functions to improve performance and further uses the maximum quantity of edge-disjoint paths to initialize the population set. The evaluation results have shown that this improves overall performance in terms of both computational time and the number of wavelengths.

### *C. RWA using Ant Colony Optimization*

The ACO algorithm is flexible in order to obtain better achievement. The simulation results indicate that an ant colony optimization algorithm consistently outperforms the other dynamic load-based routing algorithms in different network topologies. These interesting characteristics make the ant-based dynamic RWA algorithm very encouraging for the next models of WDM networks.

## REFERENCES

[1] Rainer Baumann, Simon Heimlicher, Mario Strasser, Andreas Weibel "A Survey on Routing Metrics", February 10, 2007, retrieved 2020-05-04.

[2] R. Barry and S. Subramaniam, "The MAX-SUM Wavelength Assignment Algorithm for WDM Ring Networks," Proceedings of Optical Fiber Conference, February 1997.

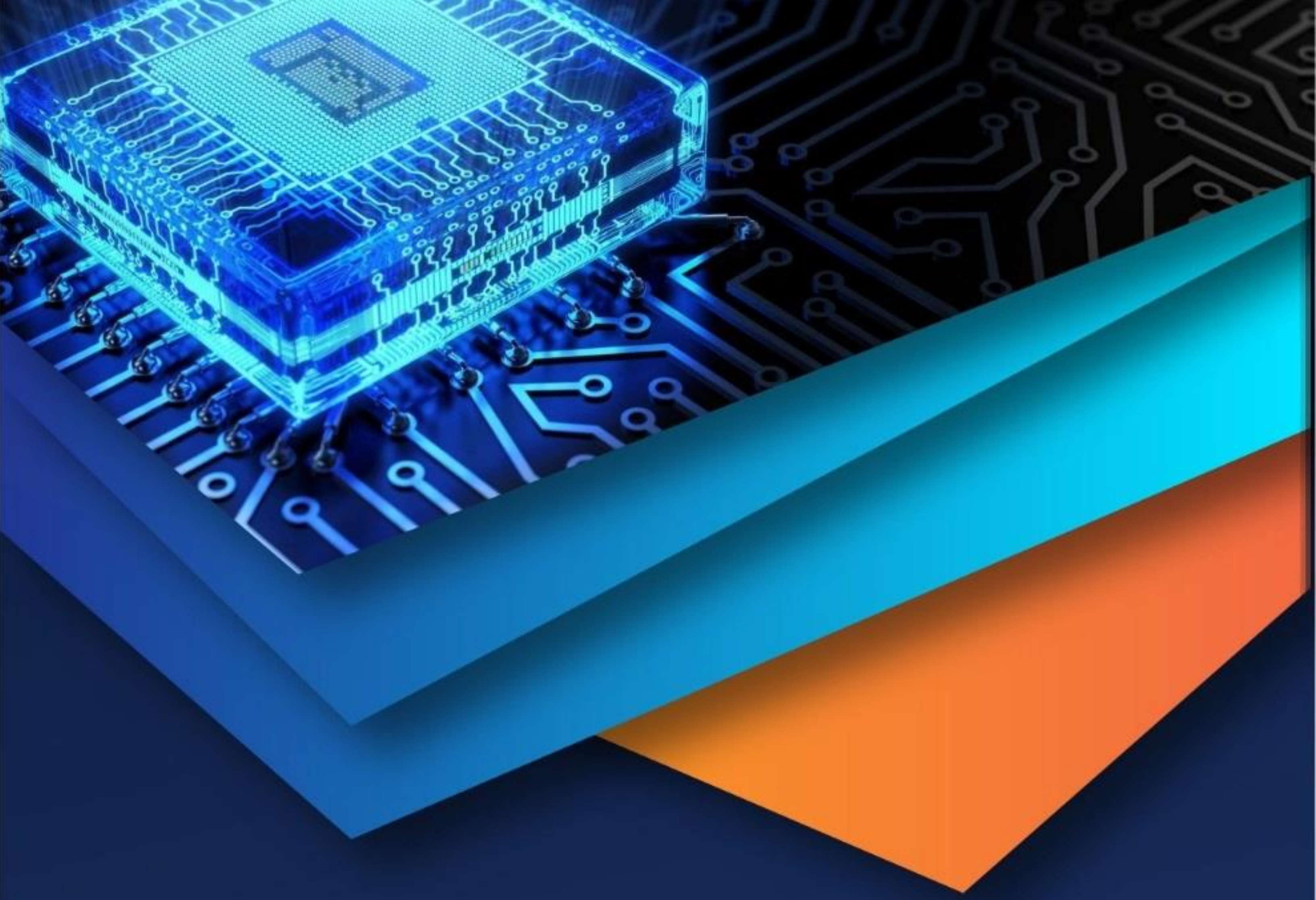
[3] X. Zhang and C. Qiao, "Wavelength Assignment for Dynamic Traffic in Multi- Fiber WDM Networks," Proceedings of International Conference on Communications, Vol 1, pp 406-410, June 1997.

[4] N. Banerjee, V. Mehta, and S. Pandey, "A genetic algorithm approach for solving the Routing and Wavelength Assignment Problem in WDM Networks", International Conference on Networks, ICN'04, Pointe-a-Pitre, Guadeloupe, French Caribbean (2004).

[5] Demetris Monoyios and Kyriakos Vlachos, "On the Use of Genetic Algorithms for Solving the RWA Problem Employing the Maximum Quantity of Edge Disjoint Paths".2008 10th Anniversary International Conference on Transparent Optical Networks.

[6] R. Hemalatha, G. Karthikeyan, R. Mahalakshmi (2018),' Implementation of Aco Algorithm in RWA'.

[7] Maysaa Ali Abdullah and Firas Mahmood Mustafa (2020),' Solving RWA Problem in All- Optical Networks based on Optimized Ant Colony Algorithm'.




10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY