Review of Neural Networks Application in UAV Routing Problems

Maksym Ogurtsov

V.M. Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine

2021

The problem’s review

The main attention in this work is paid to the analysis of the ways to use machine learning algorithms (ML) to solve combinatorial optimization (CO) problems, which arise when using unmanned aerial vehicles (UAVs), consisting of UAVs, their operators, etc., for which the distribution of initial data is unknown in advance. Mostly these problems may be formulated as VRP problems.

Novelty and general scientific significance

Although the use of NN for solving the CO problems is actively studied all over the world, its use for UAV-specific tasks has not been studied enough to date.

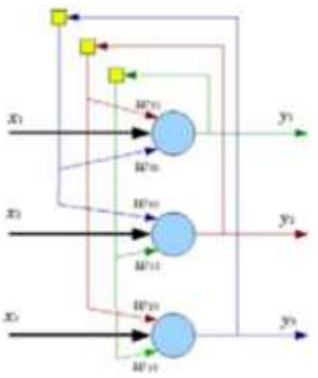
The use of NN to solve the UAV-specific CA problems can improve the quality of the resulting solutions, which will increase the efficiency of the UAVs use in general and potentially – to expand the UAVs scope.

Types of neural networks applicable to combinatorial optimization problems

- Hopfield Network
- Supervised learning
- Unsupervised learning
- Reinforcement learning
- Deep learning

Hopfield Network

The Hopfield neural network is a fully connected NN with a symmetric matrix of connections. If we modify the objective function of this NN so that it is equivalent to, for example, the objective function of the salesman problem for UAV, and use Lagrange multipliers to set penalties for violating the constraints of the problem, then such a network can be used to solve multiple CO problems. A limitation of Hopfield networks is that they are sensitive to hyperparameters and parameter initialization. Although such neural networks have several promising properties, their practical application has remained limited, in most cases only as research work.



Hopfield Network

Supervised learning

In supervised learning during the input, we have a set of input data (functions) / target pairs, the task is to find the function of converting input data into output, which for each set of input data provides output that meets the specified constraints and has the value of the objective function closest to the target. Finding such a function is called learning and may be completed by solving an optimization problem on a family of available functions.

Mathematically speaking, let *X* and *Y*, which have a common probability distribution *P*, be random variables that represent the input data and the objective function. Let ℓ be a minimization function of losses and let $\{f_\theta \mid \theta \in R^p\}$ is a family of machine learning models for optimization (in this case, parametric). The problem of controlled learning is defined as:

$$\min_{\theta \in R^p} E_{X,Y \sim P}\, \ell(Y, f_\theta(X)).$$

Unsupervised learning

In unsupervised learning the neural network does not have the goals of the problem it solves. Instead, it tries to get some characteristics of the general distribution of random variables observed during solving the problem. Because unattended learning has so far received little attention in CO context and its immediate use seems difficult, this method wasn’t considered in detail.

Reinforcement learning

Although the most successful methods of machine learning usually include families of controlled learning, where learning is performed based on educational input data, controlled learning is not applicable to most CO problems because it does not have access to optimal distribution of source data. However, you can compare the quality of the solution set using an evaluation data set for validation and determine the reward for the learning algorithm. Therefore, it is possible to adhere to the paradigm of neural reinforcement learning (RL) to solve the CO problems.

Defining the reward function is not always easy. Because RL is built on a dynamic process, it is naturally able to anticipate states / transitions that lead to future rewards. However, the above setting of rewards is difficult and not always effective, as it does not allow learning until the agent (randomly or through advanced approaches) solves the problem. In addition, when an approximate policy is applied (for example, using linear functions), learning does not guarantee finding the optimal solution and may fall to local lows.

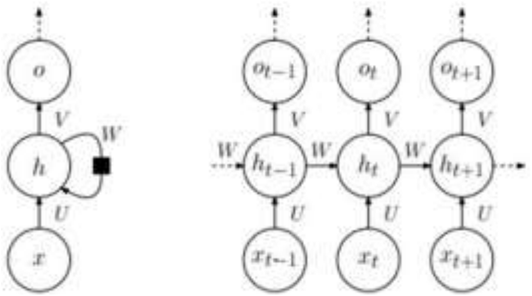
On a two-dimensional Euclidean graph containing up to 100 vertices, the RL for the travelling salesman task significantly exceeds the controlled approach to learning and allows to obtain results close to optimal, if enough time was used for learning.

Deep learning

Deep learning is a method of constructing NN, consisting of the large number of layers. Deep neural networks (DNNs) may be difficult to optimize, so several methods and approaches have been developed to make o better optimization, often by changing the architectural aspects of the network. Because DNNs have a large dimension amount, i.e., they can correspond to essentially any data set, they are prone to excessive content problem.

Recurrent neural networks

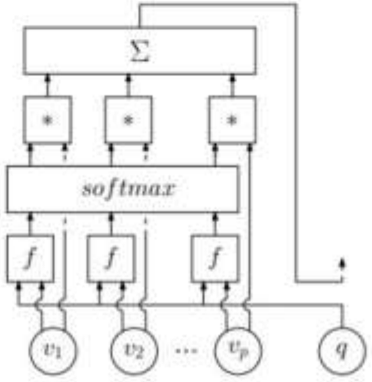
The first presented architectures are recurrent neural networks (RNN). RNNs can work on sequential data by exchanging parameters at different stages of the sequence. More precisely, the same NN block is consistently applied at each step of the sequence, i.e., with the same values of architecture and parameters at each step of time. The specificity of such a network is the presence of repeating layers: layers that take as input both the activation vector from the previous layer and the actual activation vector at the previous stage of the sequence (it is called the hidden state vector).



Recurrent neural networks architecture

Attention mechanism

Another important technique that makes the problem invariant to input data sizes is the attention mechanism. It can be used for data processing, where each point of the output data corresponds to a set of input data. In this context, parameter exchange is used to consider the fact that different sets of input do not have to be the same dimension. The attention mechanism is used to request information about all the elements in the set and combine this information for further processing in the NN.



NN attention mechanism

Approaches of using NN for solving UAV CO problems

Usage of RNN. Let us consider two approaches:

The first approach, called pre-training with reinforcement (PTR), uses a training kit to optimize a RNN, which performs parameterization of stochastic policy on test dataset, using the expected reward as a goal. During the tests, the reward is fixed, which provides conclusions about the quality of solutions based on the use of greedy sampling.

The second approach, called active search, does not involve prior RNN training. The search begins with a random policy and iteratively optimizes the RNN parameters on a single test instance, also using the expected reward as a goal, while maintaining the best solution selected during the search. The combination of PTR and active search works best in practice

Practical application

Considering the task of the traveling salesman problem, it is easy to make a greedy heuristic that builds a route, consistently selecting nodes among those that have not yet been visited, and therefore, building a permutation. The choice of the nearest node is an intuitive choice of criterion, but in most cases, it is far from optimal. Therefore, a greedy heuristic structure is built, where the node selection policy is studied using NN on a graph, a type of NN capable of processing at the input of a graph of any limited size using a message transmission mechanism. Thus, to select each node, the input of the NN getting a graphical image of the problem, supplemented by functions that indicate which of the nodes has already been visited. At the output of the NN issues the selected next node. This selected node is used to train the network through RL, and the partial length of the route is used as a reward. Of course, there are much better algorithms that allow you to get better solutions. However, it should be noted that given the above information, NN can be used to identify new, potentially more effective policies.

Conclusions

Thus, the paper reviews the approaches to the use of NN in the problems of CO. It is determined that the use of NN (including deep learning NN) is possible in the tasks of CO for UAVs in routing problems, so they can be used to solve problems that arise when planning the operation of UAVs. At the same time, recurrent NNs with nonparametric normalized exponential functions of controlled learning are successfully used to solve the problems of CO. It has been determined that the most effective to date is the use of RL training to solve such problems. This method is recommended to use.

Prospects for the use of research results: preparing mathematical model of the specific CA problem for UAVs, developing algorithms for solving this problem using the proposed NN architecture and comparing the results with the results of classical methods for solving such CA problem.

Thank you for attention!