Sofia University
“St. Kliment Ohridski”

Faculty of
Mathematics and Informatics

Neural Networks for Facility Location Problems

Vladislav Haralampiev

An abstract of a thesis presented for the partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Professional field: 4.6 Informatics and Computer Science
Doctoral program: “Computer Science” — Algorithms and Complexity

Scientific supervisor:
Assoc. Prof. PhD Minko Markov

Sofia, 2021

1 General description of the thesis

The main topic of the thesis is the use of neural networks for solving combinatorial optimization problems. The thesis proposes a new neural network approach to combinatorial optimization called **C**ompetition-**B**ased **N**eural **N**etworks (CBNN). The approach is designed for and tested on facility location problems but it can also be considered to be a general metaheuristic for combinatorial optimization.

The proposed CBNNs have several desirable properties. In the thesis, it is shown that the neural network approach asymptotically converges to an optimal solution of the modeled combinatorial optimization problem. It is also discussed how to estimate the speed of convergence. The theoretical guarantees provided by CBNNs are either analogous or are stronger than the guarantees of the other well-known metaheuristics for combinatorial optimization[1].

Asymptotic convergence is an interesting property showing that the method is not internally biased and indeed is a procedure for solving optimization problems. But, similarly to all other known algorithms for NP-hard problems, the runtime requirements for asymptotic convergence of CBNNs are often too bad to be useful in practice. What matters for a metaheuristic is to be able to find with a relatively small amount of effort a relatively good solution. In the thesis, it is intuitively justified why competition-based neural networks can be expected to quickly find a good solution. Additionally, for estimating the practical performance of CBNNs, the model is applied to six classical facility location problems: p-MiniSum, p-Hub, p-DefenseSum, Maximal Covering Location Problem, Flow Intercepting Facility Location, and Assignment problem. The results of the neural network approach are very promising. The method is often able to optimally solve the input instance. When the returned solution is not optimal, the difference, even in the worst case, is just several percent.

1.1 Motivation

Combinatorial optimization problems ask for an optimal object among a finite set of objects. The need to solve such problems arises very often. Crew scheduling, facility location, vehicle routing, assembly line balancing, and frequency assignment are just several examples of the wide range of combinatorial optimization problems that appear in practice. Sadly, most of these problems are NP-hard. This intuitively means that for finding an optimal configuration it is necessary to go through a very large number of candidate solutions. The input instances encountered in practice are relatively large and for such instances exhaustive search becomes too slow and expensive. A compromise strategy is needed that can provide a reasonable balance between the invested efforts (both computational and for development of the algorithm) and the quality of the returned solution. The metaheuristics for combinatorial optimization provide one such compromise. They target a sweet spot between returning an optimal result and being fast to execute and cheap to design. The metaheuristics are relatively quick, relatively easy to develop, and relatively reliably return a relatively good solution. This is often exactly what is needed in practice and metaheuristics have become a useful tool for solving optimization problems.

[1] C. Blum and A Roli. "Meta-heuristics in combinatorial optimisation: Overview and conceptual comparison". In: *ACM Computing Survey* 35(3) (2003).

A wide range of metaheuristics based on different phenomena have been developed. Among them, there is a class of neural network methods. A strong side of these methods is their parallelism and the possibility for a hardware implementation. The parallelism should not be underestimated because it allows to significantly speed up the algorithm in practice. It should be mentioned that the idea to use neural networks for solving combinatorial optimization problems is not new. Such methods were studied at least from the 1980s, but they never became a popular choice for solving optimization problems. The reasons for this are both objective and “emotional”.

The objective drawbacks of the existing neural networks for combinatorial optimization are divided into two groups: bad quality of the returned solution and very limited applicability. Two main types of neural networks for combinatorial optimization are known. One of them returns bad solutions and the other is applicable almost exclusively to the TRAVELLING SALESMAN PROBLEM. It is clear that with such drawbacks, the existing neural network approaches can not become popular.

The “emotional” reasons for the skepticism towards the neural networks for combinatorial optimization are largely historical. When Hopfield networks were proposed in the 1980s, this generated a lot of enthusiasm in the optimization community. Hopfield networks are a way of obtaining a meaningful digital result from analog computations and people started to ask themselves if neural networks are better suited for solving NP-hard problems than standard digital computers. Researchers then started to evaluate the potential of Hopfield networks and soon found out that the model has many problems. This, in turn, generated pessimistic views about the possibility to use neural networks for solving combinatorial optimization problems and blocked the development of the field.

The thesis expresses the opinion that the potential of neural networks for solving optimization problems is underestimated. The existing neural network approaches do have problems but their main idea to use massively parallel ensembles of neurons is not bad. The goal of the thesis is to provide an example of a neural network method that is applicable to a class of combinatorial optimization problems and that returns good solutions. This can serve as an argument for showing that the whole concept of using neural networks for solving combinatorial optimization problems is not flawed and that neural networks can successfully compete with the other metaheuristics.

Similarly to the methods from machine learning, the metaheuristics for combinatorial optimization make certain assumptions about the problem being solved. This is why it is not possible for a method to “solve everything”. From the large number of known combinatorial optimization problems, the thesis deals with facility location problems. Facility location offers a wide range of practically important tasks that are at the same time intuitive and easy to state but are hard to solve. The proposed idea of performing optimization through *competition* is also applicable to other problem classes. The thesis concentrates on facility location to be able to fully investigate the properties of the introduced neural network model.

1.2 Competition-Based Neural Networks

Neural networks are always systems consisting of independent computing units (neurons). These units operate locally. The goal is to design the system in such a way that from the combined local operations of the neurons a global meaningful behavior emerges. In the case of neural networks for combinatorial optimization, by global meaningful behavior it is meant a good solution of the modeled problem.

Neural networks can successfully model functions. This property is well-known and is used in machine learning. The difficulty with combinatorial optimization problems is that for them it is not enough to represent a function. Something in the neural network needs to make a choice. For example, assume there are two possible construction sites for a warehouse. The neurons need to somehow decide whether in the final solution the warehouse is going to be placed in the first site or in the second one.

In the thesis, it is suggested to use a mechanism of competition between the neurons as a way of making decisions. The neurons in the proposed model are divided into groups and compete inside their groups. The evolution of the whole system is guided by simple local rules such as "the strongest survives" or "neurons sometimes have luck". Each neuron only cares about itself and acts independently. In spite of this, in the thesis it is shown that if the system of neurons is left to evolve for long enough, then it reaches a configuration that represents a globally optimal solution of the modeled combinatorial optimization problem.

For formalizing the proposed optimization mechanism, the thesis defines a generalized facility location problem that is called the CBNN problem. The idea to solve optimization problems with systems of competing neurons is applied to the CBNN problem and the resulting algorithm is called a competition-based neural network. The algorithm is both described intuitively in human language and written in pseudocode. Chapter 4 investigates its theoretical properties. In Chapter 5, the algorithm is used for solving six practical facility location problems.

1.3 Results

The main contribution of the thesis is the introduction of a neural network that can compete with the established metaheuristics for combinatorial optimization. The neural network can compete both in theoretical guarantees and in empirical results on practical problems. This is an argument against the thesis that neural networks are not suitable for combinatorial optimization.

The theoretical guarantees of CBNNs can intuitively be summarized as follows: if the method is left to work for long enough, then it finds an optimal solution. The speed of convergence of the neural network can be estimated. Of course, it should be remembered that CBNNs solve NP-hard problems and are not a fast polynomial algorithm.

In practical situations, the guaranteed convergence of CBNNs takes too much time. In such cases, the thesis proposes to use a polynomial restriction of the algorithm and intuitively justifies why such a restriction can be expected to find a configuration that is close in quality to an optimal solution. The proposed method is applied to six classical facility location problems. The results are excellent: CBNNs always find solutions that are at most several percent worse than the optimal ones. In many cases, the neural network algorithm is able to find an optimal solution.

In spite of the very good results demonstrated by CBNNs, in the thesis it is suggested to perceive the method only as an initial step towards the creation of a high-quality neural metaheuristic for combinatorial optimization. The idea is expressed that the mechanism of competition between neurons is useful and probably it makes sense to consider it when developing more advanced metaheuristics. But what is the “best” way of using neural networks for solving optimization problems is an open question and still a lot of work needs to be done until a state is reached in which we have a good and reliable neural network method for combinatorial optimization.

2 Structure and content of the thesis

The thesis is 180 pages long. It consists of an exposition in five chapters, conclusion, two appendices, and bibliography. The thesis contains 17 figures, 6 pseudocodes, and 9 tables. The bibliography is on 10 pages and contains 107 sources.

2.1 Chapter 1 — Introduction

Chapter 1 briefly introduces combinatorial optimization problems and more specifically, the class of facility location problems. It continues with an intuitive presentation of six classical problems from this class: p-Hub, p-MiniSum, p-DefenseSum, Maximal Covering Location Problem, MaxiSum, and p-Center. The chapter ends with a discussion about the known methods for solving combinatorial optimization problems.

2.2 Chapter 2 — Metaheuristics for combinatorial optimization

Chapter 2 of the thesis concentrates on the metaheuristic approaches for combinatorial optimization. It starts with a description of the main ideas of eight established metaheuristics: Repeated local search, Simulated annealing, Tabu search, GRASP, VNS, Guided local search, Genetic algorithms, and Ant colony optimization. Conclusions are made about the properties that can be expected from good metaheuristics. These methods are randomized and use randomness as a way of making *hard* decisions. The metaheuristics find a local optimum of the target function and have a mechanism for escaping from local optima that are much worse that the globally optimal solution. Additionally, metaheuristics are in essence strategies for traversing parts of the solution space and aim to find a good balance between exploration of new regions and exploitation of the obtained information. During the search in the solution space, the metaheuristic approaches show *preference* towards good solution components. On one hand, this differentiates them from the “blind” search through random solutions and allows the metaheuristic methods to find a much better configuration. On the other hand, the preferences of every metaheuristic are reasonable only for a subset of all optimization problems and, in this way, every metaheuristic is suitable only for a certain subsets of all combinatorial optimization problems.

Chapter 2 continues with the introduction of the known neural networks for combinatorial optimization. It describes in detail the two main classes of neural networks for solving optimization problems: Hopfield networks[2] and self-organizing

[2] R. Rojas. *Neural Networks - A Systematic Introduction.* Springer-Verlag Berlin New-York, 1996.

approaches (more specifically, the elastic net method[3]). Additionally, Boltzmann machines that can be thought of as improved Hopfield networks are described. The chapter discusses the strengths and weaknesses of the presented methods. As a result of the investigation of the existing neural network approaches, in Chapter 2 it is concluded that the known methods have serious problems and for facility location problems they can not compete with the popular metaheuristics for combinatorial optimization.

2.3 Chapter 3 — Competition-Based Neural Networks

Chapter 3 introduces competition-based neural networks, the main topic of the thesis. The exposition starts with an intuitive description of the model that uses a "business" analogy: a comparison of the neural network to a set of competing companies that form the economy of an imaginary region. After this, a generalized facility location problem is defined that is called a CBNN problem. The idea to perform optimization through competition is applied to this problem and the resulting algorithm is called a competition-based neural network. The algorithm is written in pseudocode and is compared with the popular metaheuristics for combinatorial optimization. Additionally, several possible modifications of the algorithm are discussed.

2.4 Chapter 4 — Analysis of CBNNs

Chapter 4 of the thesis deals with the properties of competition-based neural networks and theoretically justifies why it can be expected for them to find a good solution.

CBNNs consist of independent neurons that are in essence binary variables with an attached update procedure. A configuration of the network is a concrete assignment of values, 0 or 1, to each of the neurons. During its operation, a CBNN creates a chain of configurations. By analyzing this chain, it is possible to make conclusions about the properties of the networks with competing neurons. The thesis proposes two views of the chain of configurations. From one side, it is possible to consider the exact value of each element of the chain (interpretation A). Such an interpretation is useful when analyzing the strategy of CBNNs for exploring the solution space. From another side, the sequence of configurations can be viewed as a Markov chain (interpretation B). In this case, for each element of the sequence, we are interested in the probability distribution of its possible values. This interpretation is useful for proving the asymptotic convergence of CBNNs to an optimal solution.

Chapter 4 consists of three main sections: empirical properties of CBNNs, proof of asymptotic convergence of the model to an optimal solution, and discussion about how to apply CBNNs in practical situations.

The first section empirically examines the behavior of the neural network when solving a concrete instance of a facility location problem. This allows to demonstrate characteristic properties of competition-based neural networks. It also allows to intuitively introduce concepts that are important for the proofs in the following sections.

[3] R. Durbin and D. Willshaw. "An analogue approach to the travelling salesman problem using an elastic net method". In: *Nature* 326 (1987).

In the start of the first section, interpretation A is used to investigate the strategy of CBNNs for balancing the exploration of new regions in the solution space and the exploitation of the obtained information (explore-exploit strategy). The operation of the neural network is clearly divided into three distinctive stages. In the beginning of the optimization, the neural method mainly explores the solution space and the value of the objective function almost does not decrease. In the second stage, a rapid improvement of the solution can be seen. Here, the exploration transforms into *exploitation* and *good components* of the final solution are formed. The third stage of the optimization is dominated by exploitation of the obtained information. The general structure of the final solution is already decided and only small improvements are made. The described shift from exploration to exploitation is controlled by a variable that is called *temperature* of the neural network. In Chapter 4, several other interesting observations are made about the behavior of CBNNs. For example, even when the temperature is very close to 0, the operation of the neural network is not completely equivalent to local search because of the independence of the neurons.

The second part of the first section uses interpretation B and the optimization is viewed as a Markov chain. The connection between solving a given problem and reaching a stationary distribution is discussed. It is empirically demonstrated that for the considered facility location problem and any temperature, the neural network reaches a stationary distribution. When the temperature is high, in this distribution every configuration is almost equally probable. When the temperature is low, in the stationary distribution only configurations that correspond to optimal solutions have significant probability. It is also demonstrated how to estimate the speed of convergence of CBNNs. As expected, convergence is fast for high temperatures and is slow for low temperatures.

The second section of Chapter 4 gives the proofs of the properties that were empirically noticed when investigating the behavior of CBNNs. It is shown that for any temperature, the Markov chain that corresponds to the operation of the neural network reaches a stationary distribution. In this distribution, the probabilities of the suboptimal solutions can be made arbitrary small by decreasing the temperature. This result is proven specifically for the p-MiniSum problem. The same result holds for all "reasonable" applications of CBNNs and the given proof can be used as a template. The combination of reaching a stationary distribution and the possibility to make arbitrary small the probability in this distribution of the suboptimal solutions shows the asymptotic convergence of competition-based neural networks to an optimal solution.

The final section of Chapter 4 discusses the application of CBNNs in practical situations in which we want to limit the runtime of the algorithm by a polynomial function of the size of the input instance. The mechanism of gradually lowering the temperature plays a central role for quickly finding a good solution. From the reasoning in the previous sections it can be seen that solving an optimization problem is equivalent to reaching a stationary distribution for a low temperature. During the operation of a CBNN, there are two extremes: high temperature for which the stationary distribution is reached quickly but corresponds to a poor solution, and low temperature which gives a good solution but requires a lot of time to reach the stationary distribution. The mechanism of gradually lowering the temperature combines these two extremes. Intuitively, for every temperature during the operation of the neural network, the mechanism maintains the network close to its stationary distribution. As a result, in the end of the optimization, when the temperature is the lowest, the neural network is close to its stationary distribution and finds a solution that is close in quality to an optimal one.

2.5 Chapter 5 — Applications of CBNNs

The first four chapters of the thesis are devoted to introducing competition-based neural networks and to proving their properties. In Chapter 5, CBNNs are applied to six facility location problems to demonstrate their good empirical performance. The exposition follows a common scheme. First, the facility location problem is presented together with its main variants. Then, the known methods for solving the problem are discussed. After this, it is described how to model the problem for solving with a CBNN and on what input data the resulting algorithm is evaluated. Finally, the results of the neural network on the input instances are given.

The six problems used in Chapter 5 are listed below.

- **p-MiniSum**
  This maybe is the first problem that comes to mind when talking about facility location. We have a number of clients and need to position p warehouses so as to minimize the transportation costs. The CBNN solver is evaluated on p-MiniSum instances from the Bulgarian roads data set. The data set was created specifically for the experiments in the thesis.

- **p-Hub**
  The goal here is to design a hub network. In this type of networks, there is a set of locations and hubs. The traffic between the locations is routed through the hubs. Postal delivery networks and airplane passenger networks are two examples of systems of this type. The goal of the p-Hub problem is to position the available hubs in such a way that the total cost of routing the traffic is minimized. Usually, for facility location problems locating the facilities is hard. But once the locations are selected, it is easy to assign the clients to them. The p-Hub problem is unusual because for it both the location step and the assignment step are hard. The CBNN solver is tested on the Australia Post data set[4]. This is a well-known data set for evaluating algorithms for the p-Hub problem.

- **p-Defense-Sum**
  In this problem, the goal is to position p facilities so as to maximize the sum of pairwise distances between the facilities. The p-Defense-Sum Problem is an example of an obnoxious facility location problem. In obnoxious facility location, the goal is to maximize distances instead of minimizing them. Another problem of this type is locating a garbage dump site: people want to push such facilities further away from their home. The p-Defense-Sum Problem is also an example of a facility location problem without clients. The CBNN solver is evaluated on p-Defense-Sum instances that were derived from the road network inside Bulgarian cities. The data set was specifically created for the experiments in Chapter 5.

[4] A. Ernst and M. Krishnamoorthy. "Efficient algorithms for the uncapacitated single allocation p-hub median problem". In: *Location Science* 4 (3 1996).

- **Maximal Covering Location Problem (MCLP)**
  MCLP is from the class of covering problems and is related to the classical Set Cover Problem (one of Karp's 21 NP-complete problems[5] that were shown to be NP-complete in 1972). We have a set of populated places and the goal is to position p cell phone towers so that they cover the maximal number of populated places. The CBNN solver for the MCLP Problem is evaluated on two sets of instances. The first one is derived from the map of Bulgaria and is created specifically for the experiments in the thesis. The second set of instances is created from Steiner Triple Systems. They are a well-known source of hard instances for set covering problems.

- **Flow Intercepting Facility Location (FIFL)**
  In the Flow Intercepting Facility Location Problem, we know the daily commute routes of people and want to position advertising boards so that the boards cover as many of the roads as possible. An obvious difference with the other facility location problems is that here the clients are not individual points but are roads. Apart from that, the problem is very similar to the other set covering problems. The CBNN solver is tested on instances that are derived from the road network of Sofia. The instances were specifically created for the experiments in the thesis.

- **Assignment Problem**
  In the Assignment Problem, there are n workers and n jobs. Every worker needs to be assigned to exactly one job so that no two workers are assigned to the same job. For every pair of worker i and job j there is a given profit for assigning i to j. The goal of the problem is to find a valid assignment that maximizes the sum of profits. The Assignment Problem is different from the rest of the problems in Chapter 5. It is not usually considered to be a facility location problem and a polynomial-time algorithm is known for it. The problem was selected for the experiments in Chapter 5 because its CBNN model illustrates one possible way of dealing with overlapping group constraints. The CBNN solver is evaluated on randomly generated instances of the Assignment Problem.

The six problems that are used for evaluating the empirical performance of CBNNs were selected so that they highlight different aspects of facility location. Overall, the neural network method is tested on more than 500 input instances of the problems. Large fraction of the input instances are medium in size so that the optimal solution for them can be computed with another algorithm and the quality of the solution returned by the neural network can be evaluated. Some input instances are relatively large. The number of variables for them is around 10000 and the maximum number of variables is more than 60000. For some inputs, the target function is a sum of more than $500 \cdot 10^9$ summands.

Competition-based neural networks demonstrate excellent performance on the test problems. They are often able to quickly find the optimal solution to the input instance. When the returned solution is not optimal, it is always at most several

[5] R. Karp. "Reducibility among Combinatorial Problems". In: *Complexity of Computer Computations. The IBM Research Symposia Series. Springer, Boston, MA.* (1972).

percent worse than the optimal one. This is not bad for a method that works out-of-the-box. The results of the CBNN solver on the test facility location problems are very competitive to the results of the other high-quality metaheuristics.

The main goal of the thesis is to show that neural networks can find good solutions to combinatorial optimization problems. This is why, when evaluating CBNNs on the test problems, in the center of attention is the quality of the returned solution and little attention is paid to the runtime of the algorithm. Moreover, the experiments use a simulation of the neural network on a sequential system. To get an idea about the real runtime of the neural network, a parallel implementation on an appropriate hardware needs to be used. It is interesting to note that even the sequential simulation of the neural network demonstrates reasonable speed. The resulting algorithm is not much slower than the other metaheuristics for combinatorial optimization and is faster than the integer programming library that is used for obtaining the optimal solutions for some of the problems.

2.6 Appendix A — Markov chains

Markov chains[6] are a well-studied model that is useful for analyzing competition-based neural networks. Appendix A introduces basic concepts from the theory of Markov chains and more specifically, the convergence to stationary distribution. It is demonstrated how to estimate the speed of convergence using the eigenvalues of the state transition probability matrix. It is also shown that small perturbations of the elements of the state transition probability matrix change the stationary distribution just slightly.

2.7 Appendix B — Datasets based on geographic data

Many facility location problems have a very natural interpretation in terms of road networks. For this reason, the road network of Bulgaria is often used in the thesis for generating realistic input instances of facility location problems. Appendix B describes the procedure that is used for processing geographic data. The project OpenStreetMap[7] was selected as a source of such data. The geographic information is extracted with Overpass XML queries. Appendix B shows an example query for getting the populated places and roads in a given area and describes one possible procedure for converting the data from OpenStreetMap to a graph representation of the road network.

[6] D. Isaacson and R. Madsen. *Markov Chains: Theory and Applications.* Wiley, 1976.

[7] OpenStreetMap contributors. *Planet dump [Data file from 27/12/2019]. Retrieved from `https://planet.openstreetmap.org`.* 2015.

3 Publications related to the topic of the thesis

V. Haralampiev. "Theoretical Justification of a Neural Network Approach to Combinatorial Optimization". In: *Proceedings of the 21st International Conference on Computer Systems and Technologies.* 2020.

V. Haralampiev. "Neural network approaches for a facility location problem". In: *International Scientific Journal Mathematical Modeling.* 2020.

V. Haralampiev. "Single facility location problems in k-trees". In: *58th Annual Science Conference of Ruse University and Union of Scientists - Ruse.* 2019.

V. Haralampiev. "Neural networks for facility location problems". In: *Annual of Sofia University "St. Kliment Ohridski".* 2019.

V. Haralampiev. "Dynamic facility location problems". In: *Young Researchers Conference Proceedings.* 2019.

4 Acknowledgements

A lot of people helped me during my study. I want to express my sincere thanks to Minko Markov, Georgi Georgiev, and Dobromir Kralchev for their advice and support throughout my PhD program. I am also very grateful to all the people from Sofia University for making my study a pleasant and hassle-free experience, both from scientific and administrative sides.

The Internet is an excellent source of information. It is not exactly clear who to thank for the existence of such a tool because it is the result of the collaborative effort of millions of people. But I was very lucky to do my research in a time when there is an easy and open access to information.

Finally, I want to thank my parents for their support!

5 Declaration

I declare that the thesis "Neural Networks for Facility Location Problems" is the result of my own original work that was done during my PhD study at Sofia University between years 2017 and 2021.