Leonard ROZENBERG[1], Robert RYCHCICKI[2]

West Pomeranian University of Technology (1), West Pomeranian University of Technology (2)

doi:10.15199/48.2015.02.15

Hybridization of the optimization process as an alternative way to find approximate solutions for NP-hard problems

Abstract*. This paper presents the results of calculations that demonstrate the possibility of using hybrid optimization method (with variable structure) for determining the approximate solutions to NP-hard problems. The Travelling Salesman Problem (TSP) is a classic combinatorial optimization subject, which has found widespread use in practice. Simple in definition, have remained hard to solve for many years. Not only an efficient solution would yield benefits in a substantial amount of routing problems, but it would also affect planning and logistics in a positive way.*

Streszczenie*. W prezentowanym artykule pokazano możliwość wykorzystania hybrydy optymalizacyjnej do uzyskania przybliżonego rozwiązania zadania NP-trudnego, czyli problemu obliczeniowego o ponad wykładniczym zapotrzebowaniu na moc obliczeniową. Do badań wybrano znany od wielu lat problem komiwojażera, którego od lat nie udało się ostatecznie rozwiązać. Wybór ten jednak umożliwił uzyskanie pokaźnego materiału porównawczego.* ***Wykorzystanie hybrydowej procedury optymalizacji jako alternatywy do określenia przybliżonych rozwiązań problemów NP - trudnych.***

Keywords: Monte Carlo method, hybryd algorithm, hybrid optimization, NP-hard problem
Słowa kluczowe: metoda Monte Carlo, algorytm hybrydowy, hybrydowa procedura optymalizacji, problemy NP.-trudne

Introduction

The concept of optimization is widely understood today as a set of procedures leading to the finding of the best solutions to all sorts of problems. Problems requiring optimization appear in our lives in a natural way since the dawn of history. Questions about the lowest cost or highest profit are as old as human activity. However, mathematical methods to help solve most of these problems in an effective manner arose relatively recently.

This involves, on the one hand, high demand created by economy's needs and, on the other hand, possibilities that arose thanks to the development of information technology. Technological advancement in recent years have caused changes in the meaning of the concept of applied mathematics. Discrete mathematics (combinatory and graph theory) has become an essential tool in studies dealing with mathematical foundations of computer science.

This material provides a summary of research conducted on the possibility of the use of hybrid optimization procedures in expert systems. Tasks involving analysis of large data sets in order to discover the laws that govern these sets, often come down to NP-hard problems. This state of affairs makes it necessary to use unconventional (approximate) optimization techniques. For instance, hybrid optimization algorithms, that combine several optimization methods into a hybrid structure.

Converting NP class problems using hybrid optimization

An important issue affecting whether or not we deem the optimization problem difficult or easy, is the amount of space, in which we have to make calculations. Each problem resolved on the computer must use a finite amount of this space, because what limits it is the size of machine memory. Even on a seemingly infinite space, the number of solutions provided by the algorithm is limited. Note, therefore, that any task solvable by a computer can be solved by enumerative exploration of all possible solutions' space.

We will focus on the issues of existence or non-existence of an algorithm with a polynomial complexity, that is, one whose solution can be calculated with the use of a deterministic computer. Problems, for which such a solution cannot be precisely stated, belong to NP-hard class optimization problems. One can convert any NP-complete problem to NP-hard one under a polynomial time constraint[1]. Amongst several popular NP-hard problems, one can specify:

- Traveling Salesman Problem – TSP,
- Minimum Graph Coloring,
- Solution problem of the minimal set of vertices of the graph of dominant (all other vertices neighbors),
- Search for a Minimum size of Reduct.

These difficult optimization problems can be only solved through the use of approximate techniques. The point is that if one cannot find the optimal solution precisely in an acceptable (often short) time, then one has to try to find a solution "close to" optimal. Close to, that is: of a determined distance from optimal solution. For most cases this "*close to*" solution is satisfactory[2].

A relatively simple solution to the problem can be one of simple random walk methods (Monte Carlo-method, or Vegas method). After some time, these methods are going to find a solution close to optimal with certain probability. Such algorithm can be made better in many ways, a very popular of which is to reduce the search space, if only we "suspect" that a solution is found to a certain area of space exploration. This increases the chances of obtaining a solution.

The solution presented in this case, involves more complicated approach, although it also uses an algorithm of a random search. The essence of the proposed approach to solve the problem is shown in Figure 1.

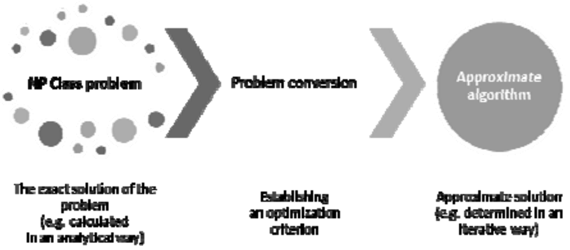

Fig. 1 Diagram of using the optimization algorithms to solving the NP class problems

Although optimization procedures require a mathematical (even if not direct) description of the problem,

[1] The concept of time complexity makes sense only in relation to the directly specified task "size", which is a property of the whole family of problems, and not of a single problem.
[2] Such a solution is called a VtR (Value to Reach) type solution.

one can convert most of the decision problems to such form through the use of mathematical modelling.

General Traveling Salesman Problem

General TSP is conceptually very simple: a traveling salesman must visit each city exactly once and return to the starting point. He should plan his route, taking into account the cost of travel between cities, so to minimize the total cost of the tour. The optimal solution is therefore to visit all the cities making the shortest route possible. It has been proven that the Traveling Salesman Problem is NP-hard, that is, one would not find an algorithm for determining an optimal solution which is of polynomial time complexity.

Based on the graph theory problem of a salesman can be represented as the task of finding the minimum Hamilton-cycle in a weighted graph. Individual cities are vertices of the graph, connected by weighted edges. The weight of the edge may correspond to the distance between places connected by the edge, travel time or cost. The task is therefore to find a permutation of locations (i1, i2, ..., in), for which the total cost is mini-mal. For the traveling salesman problem an effective algorithm (solving this problem in polynomial time) is not known. It belongs to NP-hard class of problems, therefore is it practically impossible to explore the complete solutions' space. Hence it is often chosen to verify the new ideas in the field of genetic algorithms. (Hrazdil, 2010).

The decision version of TSP problem ("Is there a Hamilton cycle in a full graph with a length, which is not greater than x?") is NP-complete, that is a NP-hard problem that belongs to the NP class as well. It should be noted that belonging to the NP class does not mean that one can verify the solution (that is, the word "Yes"/"No") in polynomial time, but means the ability to verify the certificate (proof) the correctness of the solution (in this case, such a certificate is a sequence of vertices; in the polynomial time we are able to check both whether the vertices create a Hamilton cycle, as well as whether the route length does not exceed x)" (Rychcicki, 2012). TSP, in the most general form, is not able to approximate, which is not a good basis for conducting comparative studies of hybrid algorithms. This section presents the general approach that allows you to convert NP-class problems to an approximable form.

In many real-world applications of the TSP, triangle inequality is met:

(1) $$\forall_{x,y,z\in G} c(x,y) \le c(y,z) + c(z,x)$$

where: G - set of verticals, $\forall_{x,y\in G}\ c(x,y)$ - weight of edge between vertices x and y.

In practice, this means that for any pair of cities it is always more effective (or at least "not less effective") to travel directly than indirectly through any other city. With the situation that we are dealing, the cities are placed on the plane, and travel cost measure is the Euclidean distance between them. (*translated quotation from*: Rychcicki, 2012).

To narrow the TSP even more, we can name a specific metric traveling salesman problem case, when the weight of the edge corresponds to the Euclidean distance between points in space represented by the vertices connected to the edge. This version of the problem is called Euclidean Traveling Salesman Problem - ETSP.

In spite of successive refinements of the problem, it remains a NP-hard problem. It is made possible, however, to implement certain TSP-solving algorithms more efficiently and the following properties begin to apply:

- Each pair of intersecting (with each other) edges can always be converted into a pair of non-intersecting ones (while maintaining the consistency of the route), whose total length is shorter.
- In the optimal route there are no intersecting edges.
- Among the points given in the input, these that define a convex hull, appear in the optimal route in the same order in which they appear on that convex hull. (Rychcicki, 2014)

In a later section, the results of the usage of algorithms for solving the ETSP in a heuristic and approximate manner, is shown.

Heuristic and approximate algorithms

The results of heuristic algorithms can theoretically differ by any distance from the optimal results. In practice, however, they often result in a pretty good solution, and do it in much less time than accurate algorithms. Among a set of heuristics these 3 were chosen, which also belong to the artificial intelligence algorithms:

- The Genetic Algorithm
- The Immunological Algorithm
- The Ant Algorithm

Some algorithms, such as approximated nearest neighbors algorithm, or the nearest insertions algorithm are often considered heuristics, however, for the metric traveling salesman problem (and thus ETSP) they are approximate and are therefore listed separately.

An important property of approximate algorithms (for TSP) is the existence of such restrictions that:

(2) $$C_{algorithm}^{approximated} \le \rho \cdot C_{solution}^{optimal}$$

for each problem dimension, where C means investigated solution

It has been proven that if P! = NP – for the general TSP does not exist any polynomial time approximate algorithm with constant relative constraint. The situation is different, however, in the case of the TSP with triangle inequality in effect. (*translated quotation from*: Rychcicki, 2014).

The set of studied approximate algorithms include:

- Minimum Spanning Tree algorithm (MST),
- Nearest Neighbor Algorithm,
- Nearest Insertion Algorithm.

For more information about these algorithms and their implementations and configurations for the TSP refer to (Rychcicki, 2014). A Hybrid Optimization algorithm (parallel HOP algorithm), with no special modification towards the TSP was also subject of a study.

Elimination of intersections

None of described algorithms warrant that edges of the resulting will have no intersections with each other. However, the Euclidean variant of the problem assures us that the optimal route will have such a feature (Rychcicki, 2012). Examination results endorse that applying intersection-elimination result in distinctive shortening of the length of the resulting route. It consequences the problem property that we can eliminate **every** intersection, by exchanging intersecting edges with non-intersecting ones, **making the total length of a route smaller**.

The comparative analysis

For comparison of selected algorithms, were implemented a special set of 20 tests, the number of vertices varying from 29 (cities of Western Sahara) to 10,639 (Finland cities).

Figure 2 contains chart set the length of routes in the TSP allows you to eliminate from further consideration immunologic and genetic algorithms, the results drastically differ from the other tested algorithms. Graph, which was created by the rejection of these two algorithms, includes Figure 3.

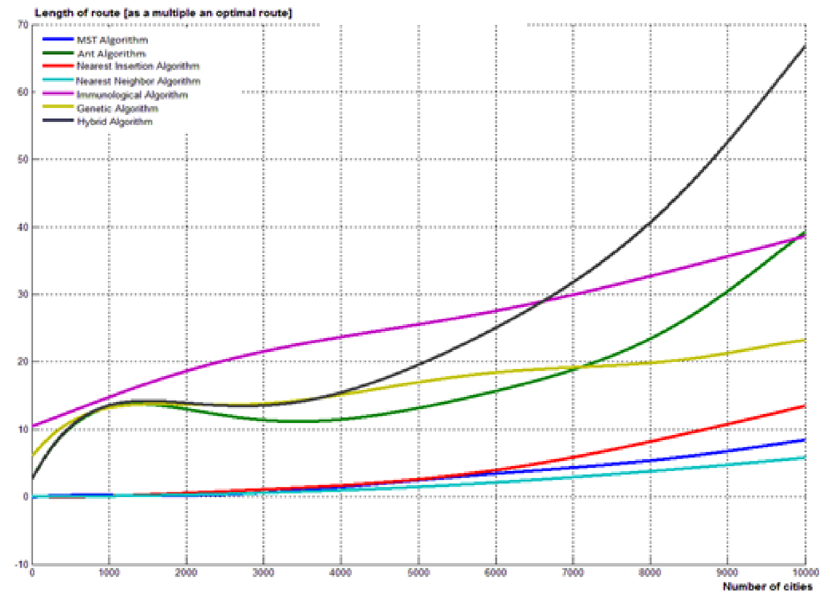

Fig.2. Comparison results for approximated algorithms solving TSP; Source: (Rychcicki, 2014)

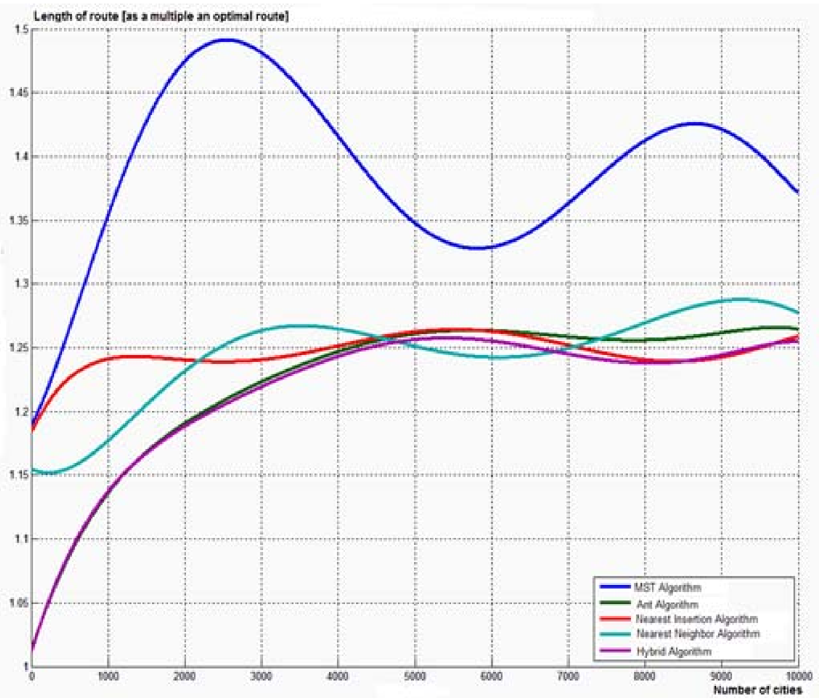

Fig.3. Comparison results for approximated algorithms solving TSP, not including genetic and immunological algorithms; Source: (Rychcicki, 2014)

Table 1. TSP-solving algorithms ranking Source: (Rychcicki, 2014)

Type of algorithm	Asymptotic algorithm complexity	Iterations count	Constraint (ρ) estimated during examination /theoretical value	Rank
Hybrid	$O(n^2 \log n)$	2500	1.2 / ∞	1
Ant	$O(n^2)$	2800	1.3 / ∞	2
Nearest Insertion	$O(n^2)$	1400	1.3 / 2.0	3
Nearest Neighbor	$O(n^2)$	500	1.5 / 7.5	4
MST	$O(n^2 \log n)$	900	1.6 / 2.0	5
Genetic	$O(n \log n)$	2200	150 / ∞	6
Immunologic	$O(n)$	2700	150 / ∞	7

The ranking of methods - short describing

Metadata generated during the tests, allow ranking optimization methods that solve ETSP. It is based on "quality" of solutions received by the given limitations with described number of iterations. It will be possible to define the criterion directly, as multiple of result in comparison with the optimal solution.

Conclusion and open problems

Analysis was shown in presented article for the possibility of using the hybrid optimization procedure to find out an approximate solution of NP-hard problem. The research shows that the hybrid optimization works well for obtaining an approximate solution, although the number of iterations is quite large. Due to the good quality of the obtained solutions and this procedure is straighter and easier in comparing to analytical methods it is interesting alternative. This idea is also easy to application.
Open problems are:

- Determining the optimal structure of the hybrid,
- Dynamic change in this structure,
- Using the parallel computing.

and many others.

Connecting optimization methods having different characteristics into the hybrid procedure results in obtaining a very universal operational research tool. Defined as such, the hybrid method allowed to optimize unimodal, as well as multimodal functions, effectively. It distinguishes the tested method from methods applied solely, even those more complicated.

Defined as such, the hybrid method allowed to optimize unimodal, as well as multimodal functions, effectively. It distinguishes the tested method from methods applied solely, even those more complicated.

It is not difficult to write the algorithm that finds the optimal solution, but the complexity of such an algorithm in practice exclude it from use. This algorithm is based on a complete review of the solutions; be-cause – in the case of the analyzed TSP – the number of possible ways of increasing the number of cities as n!, the complexity is O (n!). Similar problems occur in practical applications, so the number of „cities" can be very large, so the above algorithm cannot be suitable. Therefore approximate algorithms are being developed whose results somewhat „stand out" from the optimal performance. In practice, however, it does not matter whether the task such as controlling investment strategy will be extended by a few seconds ... (excluding High Frequency Trading market).

The main conclusion is the following: to solve NP-hard problems, use several methods, but so as to compare their results and choose the best for a particular case. For all NP-complete problems we are forced to use algorithms that generate only approximate solutions (VTR class). But the best way is to combine different techniques in special hybrid.

At the end of this material it is possible to point out that material presented here is for NP-hard problems, what numerical integration methods and mass-applying computer tools were for calculating integrals. Non-existence of universal solutions for many types of differential equations does not at all prevent from calculating approximate solutions (given set initialization conditions) which are appropriate for technical problems, as well as problems of optimal control (for example Riccati' equation). In many cases, the approximate algorithms yield satisfactory results so that they can be successfully used to solve real problems.

REFERENCES

[1] Buslenko N.P., Golenko D.I., Metoda Monte Carlo, PWN Warszawa 1967

[2] Białynicki-Birula I. Modelowanie rzeczywistości, Prószyński i S-ka, Warszawa, 2002

[3] De Jong, K. A., Genetic algorithms are NOT function optimizers. Foundations of genetic algorithms, 1993. Book 2, pp. 5-17

[4] Rychcicki, R., Szybka dyskretna transformata sinusowa. Studia i Materiały Informatyki Stosowanej, Szczecin, 2010 nr 2

[5] Rychcicki, R., Optymalizacja trasy w euklidesowym problemie komiwojażera. Praca inżynierska; Szczecin: 2012. Zachodniopomorski Uniwersytet Technologiczny,

[6] Rychcicki, R., Algorytmy heurystyczne i aproksymacyjne do optymalizacji trasy w problemie komiwojażera. Praca magisterska; Szczecin 2014; Zachodniopomorski Uniwersytet Technologiczny

***Authors**: dr hab. inż., prof. ZUT Leonard Rozenberg, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Wydział Informatyki, ul. Żołnierska 49, 71-210 Szczecin, E-mail: LRozenberg@zut.edu.pl; mgr inż. Robert Rychcicki, Instytut Elektrotechniki, Zakład maszyn Elektrycznych, ul. Pożaryskiego 28, 04-703 Szczecin, E-mail: RRychcicki@gmail.com*