

Research Article / Araştırma Makalesi

IMPROVEMENT OF TRANSPORTATION ROUTES IN MUNICIPALITIES BY DNA COMPUTATION METHOD

Asst. Prof. Salih Serkan KALELİ 

Ardahan University, Social Sciences VS, Ardahan, Turkey, (salihserkankaleli@ardahan.edu.tr)

Asst. Prof. Mehmet BAYĞIN 

Ardahan University, Ardahan, Turkey, (mehmetbaygin@ardahan.edu.tr)

Prof. Dr. Abdullah NARALAN 

Recep Tayyip Erdoğan University, Rize, Turkey, (abdullah.naralan@erdogan.edu.tr)

ABSTRACT

Today, municipalities generally provide faster and better quality and more regular transportation to the people living in those cities with the public transportation services provided in the metropolitan cities compared to the small cities. In addition, the public transport services offered to the people by the municipalities are constantly updated with the current road conditions and changing road routes in those cities and are developing accordingly. In this study, a bus route actively used by Erzurum Metropolitan Municipality was tried to be optimized and the results were shared. For this purpose, the actual data obtained were modeled using the traveling salesman problem, which is a type of optimization problem, and the current route used using the DNA Computation Algorithm was tried to be shortened.

Keywords: Optimization, Shortest Path, DNA Computing Algorithm.

BELEDİYELERDE ULAŞIM GÜZERGÂHLARININ DNA HESAPLAMA YÖNTEMİYLE İYİLEŞTİRİLMESİ

ÖZET

Günümüzde genellikle belediyeler küçük şehirlere nazaran büyükşehirlerde sağlamış olduğu toplu taşıma hizmetleri ile o şehirlerde yaşayan insanlara daha hızlı daha kaliteli ve daha düzenli bir ulaşım olanağı sağlamaktadır. Ayrıca yine belediyeler tarafından insanlara sunulan toplu taşıma hizmetleri o şehirlerdeki mevcut yol durumları ve değişen yol güzergâhları ile sürekli olarak güncellemekte ve buna bağlı olarak da gelişmektedir. Bu çalışmada ise Erzurum Büyükşehir Belediyesi tarafından aktif olarak kullanılan bir otobüs güzergâhı, optimize edilmeye çalışılmış ve sonuçlar paylaşılmıştır. Bu amaçla elde edilen gerçek veriler, optimizasyon problemi türü olan gezgin satıcı problemi kullanılarak modellenmiş ve DNA Hesaplama Algoritması kullanılarak kullanılan mevcut güzergâh kısaltılmaya çalışılmıştır.

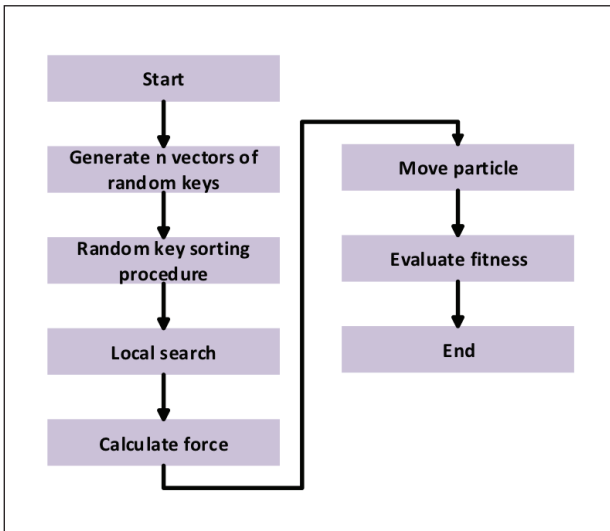
Anahtar Kelimeler: Optimizasyon, En Kısa Yol, DNA Hesaplama Algoritması.

1. Introduction

Today, optimization algorithms are actively used in many different areas and are also applied to many problems in daily life. One of the frequently encountered problems in the literature is the traveling salesman problem. The traveling salesman problem is basically guiding the principle that a traveler's goods are to be sold back to the cities in different cities and to the cities that have passed before. In this problem, which is very difficult to solve, no mathematical model is found completely and numerical calculations are insufficient. The solution space of this problem, which is frequently encountered in the literature, expands exponentially in parallel with the increase in the number of cities. Because the seller "n" the different city to travel without repeating, "n!" it brings different possibilities. In this way, optimization approaches are generally used for the problem types where the solution space is quite wide. Optimization methods for the solution of the traveling salesman problem are frequently used in the literature and there are quite different studies on this subject.

In one of these studies Ozkir & Topcu (2017) applied stochastic key based electromagnetism intuition to solve symmetric traveling salesman problems. The random key approach is adapted to electromagnetism intuition to solve vehicle routing problems. The flow diagram of the proposed algorithm is shown in Figure 1.

Figure 1: Examples from The Literature



Source: Özkar, V., & Topçu, B. (2018). Application of the random key based electromagnetism-like heuristic for solving travelling salesman problems. *Pamukkale University Journal of Engineering Sciences*, 24(1), 76–82.

In one of the studies related to the subject Cigdem & Karakose (2013) applied DNA computation algorithm to non-polynomial (NP) problems. NP problems take place in the difficult problem class, where the solution space is very wide, has no definite solution, and there is usually an approximate solution. For this purpose, the problem type of knapsack and

traveling salesman problem has been selected and the performance of the DNA computing algorithm developed in the study has been simulated on MATLAB. As a result of this study, it is seen that DNA computing algorithm gives good results especially in terms of cost and time when applied to these two types of problems.

Basically, DNA sequences are the basis of the DNA computing algorithm. With the use of these processes, many problems that take a long time can be solved in a short time. The DNA computing Algorithm was first developed by Adleman, and in the following years it has reached a very advanced and exciting point quickly (Adleman, 1994). In a study on DNA computing, Lipton (1995) focused on the solution of SAT (Satisfiability) problems consisting of logical equations. In this study, the whole solution was built with this set of DNA sequences and sequences of “1” and “0” or “True” and “False” has shown with values. In this study, biological steps such as synthesis, searching, polymerization and separation were followed and millions of biochemical reactions were carried out in the laboratory. Additionally, the DNA sequences coded with a set of possible solutions duplicated and placed at the same specific test tube and the solution obtained was subjected to parallel processing. Ouyang et al. (1997) developed a new method for the solution of maximum click problems using DNA molecules. In the study, the graphical problem of 7 knots and each knot connected with the edges was taken as an example. Here it is tried to obtain maximum cluster of nodes. In the study, each node is coded with a 6-bit binary number system. As the number of nodes in the selected example is low, there is not much problem, but the number of nodes has become more difficult to solve. Tomohiro et al. (1996) proposed a novel method of coding with DNA molecules and applied them in their studies. Maley (1998) conducted a study on the use of DNA computing in the light of computer programming and chemical reactions. Wood et al. (1999) In the solution of the Max 1 problems, the DNA computing algorithm is better than the genetic algorithm and studies have done. Chen et al. (1999) conducted a study showing that current molecular biology techniques will be used to complete DNA computing techniques. Forbes (2000) is a study that compares the digital computers with DNA computers, DNA processing has the ability to perform parallel processing, this feature has been said to be more effective than digital computers million times.

The main purpose of this study is to minimize the average travel time by optimizing the routes currently used by public transport systems. In this way, the average waiting and travel times of passengers and the amount of fuel consumed by public transportation vehicles can be minimized. There are various studies on route optimization in the literature. Compared to studies in which classical methods such as genetic algorithms are used, an optimization method which is quite new in the literature was preferred in this study. In this study, which is carried out by using the DNA computation algorithm, a different perspective is brought to the traveling salesman problems in the literature both theoretically and practically. The results obtained showed that this new optimization method used can be successful.

In this study, a traveling salesman problem based optimization application has been developed by using DNA computing algorithm. In this context, the bus routes which are actively used were examined by DNA computing method and new optimal routes were determined. The second part is the traveling salesman problem areas examined types of studies in the literature, the details of the proposed method is given in third section. In the fourth section, the simulation

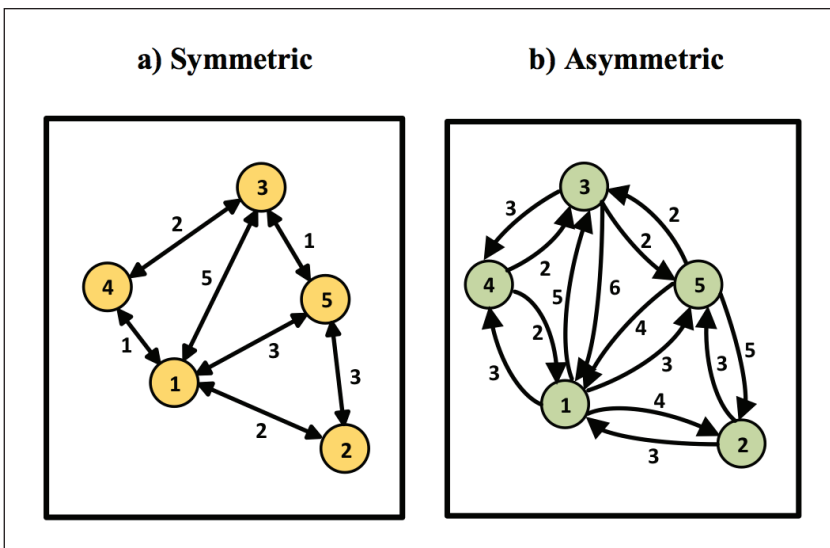
results obtained within the scope of the application are given and the results are shared in the fifth and last part.

2. Traveling Salesman Problem

The traveling salesman problem (TSP) is a type of problem that is frequently encountered in the literature and the solution cannot be calculated mathematically clearly (Naralan et al., 2017). In this study, this type of problem, which is frequently encountered in the literature, was tried to be solved by using DNA computing algorithm. The highway transport routes obtained from Erzurum Metropolitan Municipality Public Transportation Branch Directorate were examined and optimized. In the study, it was ensured that the route used by the bus in the selected line was shortened and accordingly the average travel time of the passengers was reduced.

The traveling salesman problem was tried to be solved on the basis of the proposed approach. There are many varieties of traveling salesman problem. These; the dynamic traveling salesman problem, the profitable traveling salesman problem, the multi-traveling salesman problem and the symmetric-asymmetric traveling salesman problem (Kara et al., 2011). The stops in the selected bus route were considered to be the cities that the traveling should face in the traveling salesman problem and the distances between the stops were considered as distances between the cities. Considering all these situations, it will be seen that this problem tried to be solved for real life will be asymmetric. Because the distance between the stops in the real life, such as the distance from the “A” stop to the “B” stop and the distance from the “B” stop to the “A” stop will not be the same. The most important reason for this situation is that some routes used are unidirectional and some of them are bi-directional. In this context, an image illustrating the symmetric and asymmetric traveling salesman problem is presented in Figure 2.

Figure 2: Traveling Salesman Problem Types

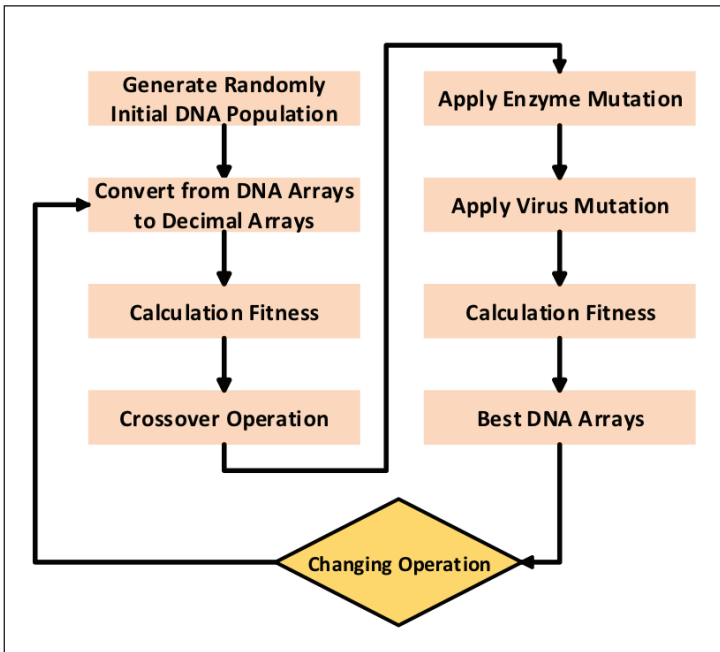


As can be seen from Figure 2, the distance of the traveler's departure and return paths is the same in the symmetric traveling salesman problem. In other words, it goes the same way and goes back in the same way. In the asymmetric problem type, there are different path routes where the traveling can choose between two points. The problem type which is tried to be solved within the scope of the study is also asymmetric problem class.

3. DNA Computing Algorithm for Asymmetric TSP

DNA computing algorithm is an optimization method based on the DNA structure in which the living things are kept (Karakose & Cigdem, 2013). The DNA computing algorithm is a set of processes that are used to solve a problem, especially in nonlinear problems as in other methods. DNA computing can be performed in two different environments: solution and electronics. Because the calculations performed in the solution environment are quite costly, calculation processes are carried out in electronic environment which is usually the second environment. The DNA computing algorithm has two different types and stages of mutation, unlike other algorithms. The first of these mutation processes is the enzyme mutation and the second one is the virus mutation (Muhammad et al., 2005). In the enzyme mutation process, the deletion of one or more DNA fragments is theoretically carried out from any DNA sequence. The virus mutation involves the addition of a new DNA fragment to replace this deleted DNA sequence. As a result of these processes, while the size of the DNA sequence does not change, new DNA sequences with different characteristics can be obtained. A flow diagram summarizing the working principle of the DNA computing algorithm is given in Figure 3 and details of the steps used in the algorithm are presented below.

Figure 3: DNA Computing Algorithm



Generation of Random DNA Sequences: The first step of the DNA computing algorithm is the generation of random DNA sequences. Adenine (A), Thymine (T), Guanine (G) and Cytosine (C) molecules are used in the production of DNA sequences in this structure based on DNA molecules in any living organism (Muhammad et al., 2005). In this paper, a route consisting of 18 stops was chosen. At this point, because it is a total of 4 molecules, all stops should be expressed using these molecules. In this context, since the molecules must be expressed in the quadratic base, the distribution of the stop numbers according to the molecules is given in Figures 4 and 5.

Figure 4: Symbol and Value Distribution of DNA Molecules

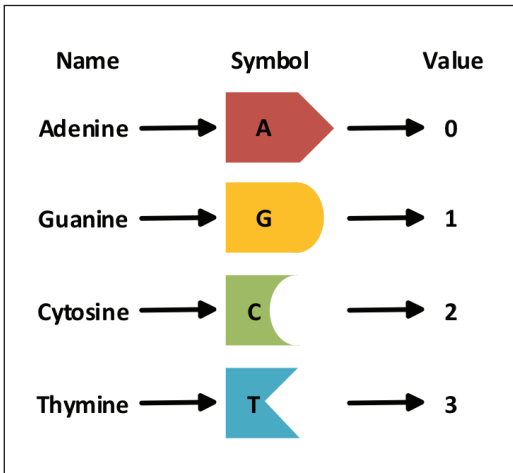
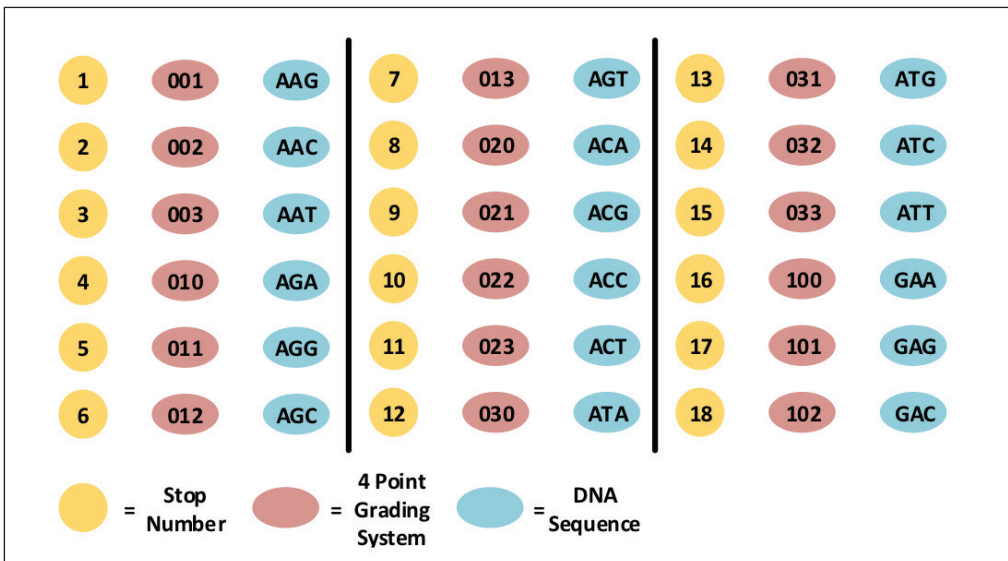


Figure 5: Distribution of DNA Molecules According to Stops



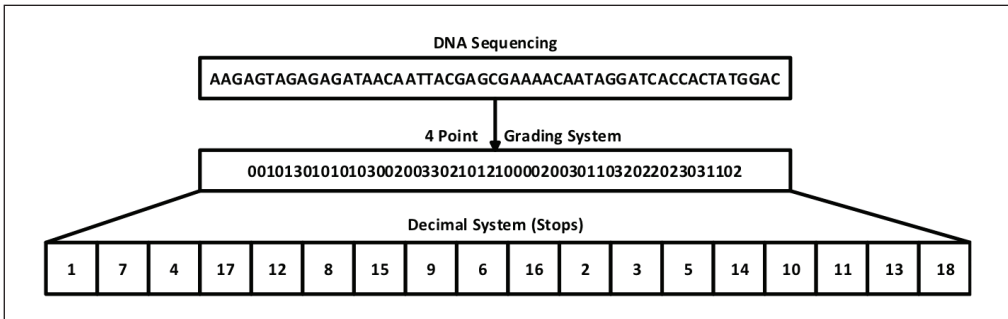
As can be seen from Figures 4 and 5, each molecule has a symbol and DNA sequences generated depending on this symbol. Since there are 4 symbols in total, the stops can be expressed in 4 levels and with these symbols, up to 64 stops (TTT) can be displayed. If there are more stops, the number of molecules should be increased. This situation is shown in the equation 1.

$$\begin{aligned}
 A &= 0, G = 1, C = 2, T = 3 \\
 AAA &= 0 \times 4^0 + 0 \times 4^1 + 0 \times 4^2 = 0 \\
 TTT &= 3 \times 4^0 + 3 \times 4^1 + 3 \times 4^2 = 0
 \end{aligned}
 \tag{1}$$

While random population is generated according to this equation, triple DNA sequences are joined side by side and a DNA helix of the length given in Equation 2 is obtained. In this context, a randomly generated sample DNA sequence for a route with 18 stops is given in Figure 6.

$$\text{DNA Sequence} = \text{DNA Array} \times \text{Stop Number}
 \tag{2}$$

Figure 6: Examples of a DNA Sequences



Convert DNA to Decimal: The DNA helix obtained in the first phase of the application is converted to the actual numerical values (decimal system) in the second stage of the application. The reason for this process is to provide ease of operation when doing conformity calculation.

Fitness Calculation: The fitness function has been obtained as a result of literature reviews and is based on the logic of taking the difference between the current stop and the next stop and repeating this process throughout all stops. The cumulative total is calculated for each difference and the total distance covered by the public transport vehicle for the new route is calculated. Equation 3 is used for fitness calculation.

$$\sum_{i=1}^{n-1} DT(P_i, P_{i+1})
 \tag{3}$$

n = Total number of stops,

P = Individual of population

DT = Distance Table

Crossover Operation: In the fourth step of the DNA computing algorithm, the selected DNA sequences are cross-over. Crossing is not performed on numerical values, but on DNA helix.

Enzyme and Virus Mutation Operation: In DNA computing algorithm, DNA sequences are subjected to a two-step mutation process. The first one is the enzyme mutation, which is the process of deleting any DNA fragments randomly selected from the DNA sequence. Virus mutation is the process of inserting a newly produced DNA fragment in place of the randomly deleted DNA sequence.

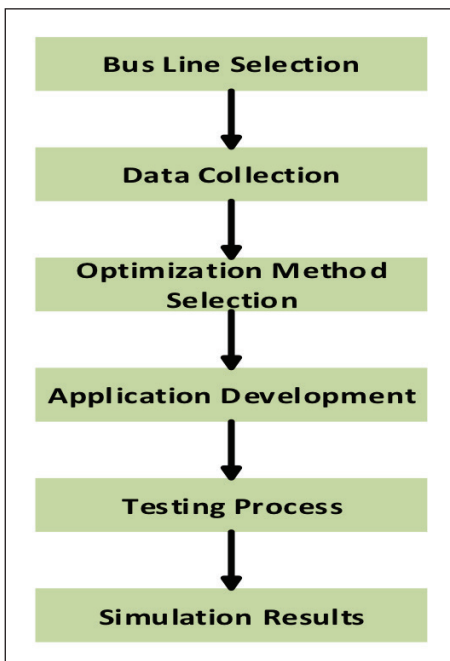
Fitness Calculation and Changing: At this stage, the mutated DNA sequences are converted back to numerical data and their fitness is calculated. If the newly calculated values are better than the fitness value of the DNA sequences initially produced, the displacement process is applied.

Termination: The DNA computing algorithm is terminated at a certain point by the termination process and the result is recorded as a solution.

4. Material and Methods

In this study, a simulation process was applied for route optimization and the real collected data were optimized by using the DNA computing algorithm. Basically, very good results were obtained in terms of distance in this study in which the average travel time was reduced. A block diagram summarizing the flow of the study is presented in Figure 7.

Figure 7: Flowchart of The Proposed Method



As can be seen from Figure 7, the road route to be optimized was selected in the first stage of the application. Later, data of this road route was collected and the optimization method to be applied was determined. Finally, the application development phase started and the data obtained were tested on the developed application. The data used in the study and the details of the method are presented in subsections.

4.1. Purpose of the Study

The main purpose of optimization methods is to find the best results quickly. In this study, a method that is relatively new in the literature was used. With this application in which DNA computing algorithm is used, it is ensured that the public transportation routes are optimized. Basically, thanks to this practice, where the average passenger waiting and travel time is reduced, the amount of fuel consumed by public transport vehicles is minimized. The main purpose of this developed application is that the bus starts from the starting point and stops at all stops and finally reaches the end point in the shortest way.

4.2. Data Collection

The main purpose of this study is to shorten the routes actively used by public transportation vehicles. In this way, shorter road routes are obtained and the average passenger travel time is reduced. For this purpose, firstly, Erzurum Metropolitan Municipality public transportation administration was contacted and GPS coordinates of the existing routes were provided. Later, these coordinates were visited and it was determined whether the stops were up-to-date. The reason for this situation is that it is not known whether the stops are up-to-date or not. After the coordinates of all the stops were determined, they were recorded and marked using the Google Maps application in digital environment. The data used within the scope of the study are updated data collected within the scope of the permission of Erzurum Metropolitan Municipality Public Transport Directorate.

4.3. Method

In order to process the data obtained within the scope of the study, a distance matrix file was primarily created. In this developed application, the problem is asymmetrical. In other words, the outbound route distance and return route distance are not the same. For this reason, a distance matrix file has been prepared and planned to be given as an introduction to the optimization algorithm. While preparing the distance file, Google Maps application was used and the distances between each stop were recorded. A small application has been developed for this process and the distances have been obtained using the Google Maps API. In the next stage of the system, the optimization method was chosen and the DNA computation algorithm, which is relatively new in the literature, was preferred. Finally, the DNA computing algorithm was developed on the MATLAB platform and tested using distance files. The new routes obtained with this simulation study were marked on the map and the results were observed. One of the main reasons for using the MATLAB platform while developing the application is that the results obtained can be easily visualized. In addition, the results obtained can be analyzed quickly thanks to the various calculation and statistical formulas it contains.

5. Findings

In this study, it is aimed to shorten a route which is actively used by using DNA computing algorithm. The DNA computing algorithm is coded in MATLAB environment and

all the steps mentioned in the previous sections for this algorithm were performed on this platform. No ready-made library was used at this point and the data used in the study are real time data. Some parameters of the DNA calculation algorithm used as an optimization method are presented in Table 1. Also, a part of the distance table for the data used in developing the application is presented in Table 2.

Table 1: DNA Computing Algorithm Parameters

Population Number	100
Number of Iteration	1000
Crossover Rate	1
Enzyme Mutation Rate	0.3
Virus Mutation Rate	0.3

Table 2: An Example Section of the Distance Table

Stop Number	1	2	3	4	5	6	7
1	0.0	1.1	1.1	1.8	2.1	1.0	1.6
2	2.0	0.0	0.3	1.4	1.7	2.2	2.2
3	1.7	0.9	0.0	1.1	1.4	1.9	2.8
4	1.9	2.1	2.0	0.0	0.3	2.2	3.1
5	1.6	1.8	1.7	1.1	0.0	1.8	2.8
6	1.7	3.1	2.7	2.4	1.8	0.0	1.2
7	1.5	2.3	2.6	2.3	2.6	0.9	0.0

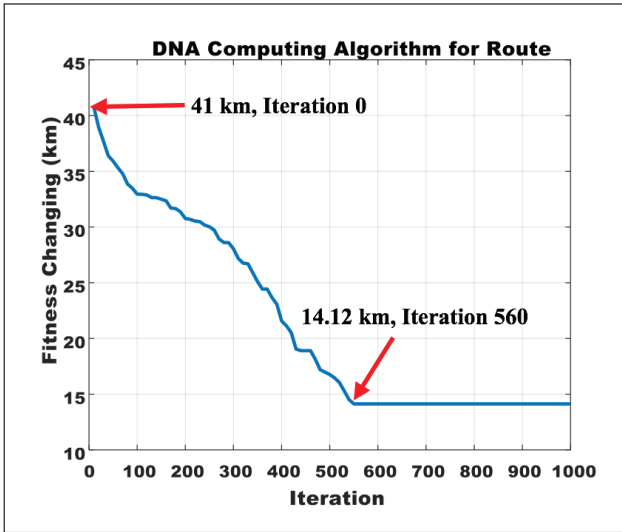
The route distance status given in Table 2 shows only a specific part for the selected route. Basically, the selected route has 18 stops. In the method, the starting stop is indicated by 1, and the end point is the stop 18. A GPS coordinate basically consists of latitude and longitude values. For this reason, the latitude and longitude values of each stop were determined and GPS coordinates were obtained. Detailed GPS coordinates of the stops used in this route are presented in Table 3.

Table 3: Stop Coordinate List

Stop No	Coordinate (Latitude, Longitude)	Stop No	Coordinate (Latitude, Longitude)
1	39.902638, 41.274524	10	39.911979, 41.266006
2	39.898729, 41.270129	11	39.908586, 41.265283
3	39.900073, 41.267221	12	39.901406, 41.266290
4	39.909412, 41.265402	13	39.892800, 41.248494
5	39.911734, 41.265663	14	39.898357, 41.262598
6	39.907539, 41.278788	15	39.899936, 41.266736
7	39.906147, 41.286567	16	39.898189, 41.270431
8	39.905089, 41.290898	17	39.897225, 41.272914
9	39.911879, 41.272602	18	39.898594, 41.275351

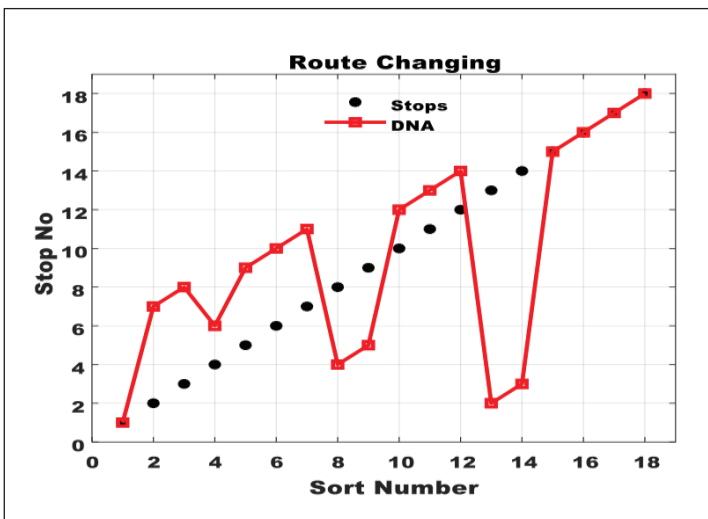
The coordinates used in the study are presented in Table 3. The distance matrix calculated according to these coordinate values is presented in Table 1. In this study, the results of fitness and route change for the selected bus line using DNA computing algorithm are shown in Figure 8 and 9, respectively.

Figure 8: Change of Fitness Value



As can be seen in Figure 8, the DNA computing algorithm calculates a distance of approximately 41 km in the first iteration. However, the system has improved with the process of increasing iteration steps and reduced the total distance to less than 15 km in the 560th stage.

Figure 9: Change of Route



In Figure 9, the curve indicated by the circle symbol shows the current stop sequence and the curve indicated by the square symbol shows the stop sequence obtained using the DNA computing algorithm. The DNA computing algorithm has improved results. When the new stop touring sequence is examined, it will be seen that an increase in efficiency of approximately 13% is achieved. This situation, which shows the new stop sequence, is as given in Table 4. The new stop layouts given in Figure 9 are presented in Table 4. In other words, public transportation vehicles will act according to the new numbered stop order obtained after the optimization process. An image that is marked on Google Maps of this optimization process using the DNA computing algorithm is given in Figure 10.

Table 4: Before and After Optimization Process

Before State	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
After State	1	7	8	6	9	10	11	4	5	2	3	12	13	14	15	16	17	18

Figure 10: Drawing The New Route On Google Maps



As can be seen from Figure 10, there are two different colors, red and blue on the route. While the red one of these colors shows the route before the optimization, the results of the route obtained after the optimization are presented with blue color. The values obtained as a result of the simulation results for 18 stops are given in Table 5.

Table 5: Performance Results of Optimization Process

Features	Values	Features	Values
Number of Stops	18	Difference	2.1440 km
Initial Distance	16.2710 km	Percentage	13%
Final Distance	14.1270 km		

As can be seen from Table 5, the route, which is normally 16 km, has been reduced to approximately 14 km after the optimization process. In this study, where the total number of stops is 18, approximately 2 km of gain has been achieved, and a total change of 13% has been achieved. A public transport vehicle makes many trips during the day. Considering this situation, the results obtained will be provided each time and the gain will be much higher.

6. Discussion and Conclusions

In this study, it is tried to optimize the problem of asymmetric traveling salesman which is an NP problem type by using DNA computing algorithm. The data used for this purpose is obtained from Erzurum Metropolitan Municipality Public Transportation Branch Office and used in real time. The new route was determined by taking into account the positions of the stations with the proposed approach and the performance of the stations and a performance of about 13% was achieved. In addition, the data obtained from the application are marked on Google Maps and the results are shown in detail.

The proposed approach provided a gain of approximately 2.1 km on the current route. Considering all voyage performed on this route during the day, it can be observed that the results obtained are extremely good. In addition, thanks to this performance, it is possible to reduce the fuel consumption and maintenance costs of the buses used in passenger transportation as well as to decrease the average passenger waiting and travel times.

It is planned to apply the method proposed within the scope of future studies to other routes. In addition to the DNA computing algorithm developed within the scope of this study, it is aimed to test other optimization methods frequently used in the literature and to compare the results. In addition, with this future study, it is aimed to update the bus stops of Erzurum Metropolitan Municipality, to record the stop coordinates and to display them in digital environment.

Acknowledgement

We thank to Erzurum Metropolitan Municipality and Public Transport Directorate for provided data in this study.

References

- Adleman, L. M. (1994). Molecular computation of solutions to combinatorial problems. *Science*, 266(4), 1021-1025.
- Baygin, M., & Karakose, M. (2013). Immunity-based optimal estimation approach for a new real time group elevator dynamic control application for energy and time saving. *The Scientific World Journal*, 1-12.
- Chen, J., Antipov, E., Lemieux, B., Cedeno, W., & Wood, D. H. (1999). DNA computing implementing genetic algorithms. *Proceeding of DIMACS Workshop on Evolution as Computation*, Princeton, NJ, USA, 39-49.
- Çiğdem, U., & Karaköse, M. (2013). Polinomal olmayan problemler için DNA hesaplama algoritması. *Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 29(1), 41-48.
- Forbes, N. (2000). Biological inspired computing. *Computing in Science & Engineering*, 2(6), 83-87.
- Kara, İ., Güden, H., & Koç, Ö. N. (2011). Genelleştirilmiş gezgin satıcı problemi için polinom büyüklükte yeni karar modelleri. *XI. Üretim Araştırmaları Sempozyumu*, 23-24 Haziran, İstanbul, 800-811.

- Karakose, M., & Cigdem, U. (2013). Qpso-based adaptive DNA computing algorithm. *The Scientific World Journal*, 1-8.
- Kuzu, S. (2014). Gezgin satıcı problemlerinin metasezgiseller ile çözümü. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 43(1), 1-27.
- Lipton, R. J. (1995). DNA solution of hard computational problems. *Science*, 268(5210), 542-545.
- Maley, C. C. (1998). DNA computation: Theory, practice, and prospects. *Evolutionary Computation*, 6(3), 201-229.
- Muhammad, M. S., Ibrahim, Z., Ono, O., & Khalid, M. (2005). Direct-proportional length-based DNA computing implementation for elevator scheduling problem. *IEEE Region 10 Annual International Conference, Proceedings/TENCON, 2007 Melbourne, Australia*.
- Naralan, A., Kaleli, S. S., & Baygin, M. (2017). Shortest path detection using clonal selection algorithm for Erzurum Metropolitan Municipality. *Muğla Journal of Science and Technology*, 3(2), 138-142.
- Ouyang, Q., Kaplan, P. D., Liu, S., & Libchaber, A. (1997). DNA solution of the maximal clique problem. *Science*, 278(5337), 446-449.
- Özkır, V., & Topçu, B. (2018). Application of the random key based electromagnetism-like heuristic for solving travelling salesman problems. *Pamukkale University Journal of Engineering Sciences*, 24(1), 76-82.
- Polynomial, N. E. W., Mathematical, S., For, M., Generalized, T. H. E., & Problem, S. (2006). Genelleştirilmiş Gezgin Satıcı Problemi için Polinom Büyüklükte Yeni Karar Modelleri. 800-811.
- Wood, D., Chen, J., Antipov, E., Lemieux, B., & Cedeño, W. (1999). A {DNA} Implementation of the Max 1s Problem. *GECCO-99: Proceedings of the Genetic and Evolutionary Computation Conference*, 2, 1835-1841.
- Yoshikawa, T., Furuhashi, T., & Uchikawa, Y. (1996). DNA coding method and a mechanism of development for acquisition of fuzzy control rules. *Proceedings of IEEE 5th International Fuzzy Systems*, September 8-11 New Orleans, LA, USA, 2194-2200.
- Yoshikawa, T., Furuhashi, T., & Uchikawa, Y. (n.d.). Tomohiro yoshikawa, takeshi furuhashi, yoshiki uchikawa. 2194-2200.