



Metaheuristic Methods for Water Distribution Network Considering Routing Decision

Ahmad Hakimi¹, Reza Mahdizadeh², Hossein Shokri Garjan³, Amir Khiabani⁴,
and Ajith Abraham^{5,6}(✉)

¹ Department of Industrial Engineering, University of Kurdistan, Sanandaj, Iran

² Department of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

³ Department of Industrial Engineering, Babol Noshirvani University of Technology, Babol, Iran
Hossen.shokri@nit.ac.ir

⁴ Department of Industrial Engineering, Islamic Azad University, Tehran, Iran
st_a_khiabani@azad.ac.ir

⁵ Machine Intelligence Research Labs (MIR Labs), Scientific Network for Innovation and
Research Excellence, Auburn, WA 98071, USA
ajith.abraham@ieee.org

⁶ Center for Artificial Intelligence, Innopolis University, Innopolis, Russia

Abstract. Today, the issue of monitoring water quality in distribution networks is of particular importance as the last step in the water production process. In this research, a change in the standard of vehicle routing is introduced, which is used in monitoring the quality of the water distribution network. In this study, a large number of teams visit candidate nodes in the water distribution network daily. Also in this research, mathematical modeling of the problem has been done so that there are a series of nodes (wells) that can not be visited directly. For this type of special node, the driver needs a key to open the well door and perform the required tests, which must visit the specific location where the well key is located. Two innovative algorithms based on genetic algorithm and neighborhood search, are proposed to solve this problem. Finally, some suggestions for developing the approach of this research are presented.

Keywords: Vehicle routing · Mathematical modeling · Heuristic model

1 Introduction

Providing safe drinking water is one of the important goals in human societies and achieving development and progress is possible in the shadow of the health of people in society and it is clear that the health of individuals depends on the supply of drinking water is desirable. Water distribution networks are vulnerable to accidental or deliberate contamination due to their vast geography and abundant access points such as reservoirs, wells and hundreds or thousands of kilometers of underground pipes. Determining pollution in drinking water distribution systems faces many challenges. An increasing number of people use the water of such a system every day and trust the safety and quality of water

in their life or work. In addition to the problems of contaminant detection, there are other reasons, such as the decay or growth of non-permanent components that occur during the transfer process, to reduce water quality in distribution networks. Effective and efficient water quality monitoring is one of the most important tools to ensure reliable drinking water supply to the consumer in a drinking water distribution system [1, 3].

Monitoring is a method of determining the quality of the aquatic environment and how the environment is affected by the spread of pollution and other human activities. A vital and effective way to monitor the distribution of the water network in real time is to use a water quality sensor. The sensors are placed at regular intervals in some parts of the network to monitor water quality. In other words, experiments are performed at predetermined times through located sensors and the results are monitored online. [2, 4] Due to various factors such as the high cost of online monitoring and the lack of such a system, it can be done in person. Candidate points in the water distribution network should be visited and the necessary tests should be performed by specialized manpower. Monitoring objectives are the most important factors that play a role in determining the direction of monitoring points. Depending on the purpose, finding a route for vehicles to visit different places has a special place.

The issue of vehicle routing in a supply chain is at the heart of distribution organization. Many companies involved in the delivery, collection and transportation of goods and passengers face this problem every day [6]. Because conditions vary from set to set, the goals and constraints of this issue are very diverse. Vehicle routing was first discussed by Danzig and Ramser [5] and evaluated as a hybrid optimization problem. An issue with vehicle routing is the effective use of a number of vehicles that must stop at stations to pick up a customer (product) or deliver a customer (product delivery). Customer refers to points of view on the network.

In this study, a large number of teams visit candidate nodes in the water distribution network daily and perform tests to control water quality. Also in this research, there are a series of nodes that can not be visited directly. For this type of special knot, the driver needs a key to open the well door and perform the required tests. Since the driver does not have the key, he must visit a specific place where the well key is located; Therefore, there is a need to visit a point in the path before visiting that well. Finally, the keys must be returned to their original location.

The ideas of this research are summarized as follows:

- There are two types of demand points (customer) that are visited differently. In the first type, the customer is visited directly, but in the second type, it is necessary to visit another point before and after visiting the customer. Also, visiting some of the first type of demand points requires more time.
- Using an innovative algorithm to find the transit routes of vehicles.

2 Literature Review

Since the vehicle routing problems face many goals and limitations due to the different and frequent changes of conditions from one situation to another, so most algorithmic studies and software development in this field are based on a number of basic models of problems. Vehicle routing is focused, and by creating flexibility in optimization

systems, this can be adapted to a variety of issues that may in fact be encountered. Routing problem-solving methods can be divided into three groups as follows: exact methods, innovative methods, super-innovative methods based on nature, combined methods (using an accurate or innovative method to find the minimum limit for solving the problem and using a metaheuristic method for obtaining the best possible answer is the exact methods including branch and boundary (B&B), branch and cut, dynamic programming (DP) and integer programming (IP), which has done a lot of research so far. Breakers et al. [7] illustrated that more than 93% of vehicle routing issues targeted routing costs, and about 40% of the articles surveyed focused on vehicle costs and driver rights.

Rabbani et al. [8] Presented a multi-objective model to minimize distribution costs, fuel consumption, and balancing drivers' workloads, and then used a multi-objective genetic algorithm to solve it. Islam et al. [9] proposed a hybrid algorithm combining particle swarm optimization and variable neighborhood search to solve the routing problem of clustered vehicles. Martinez et al. [10] explored a combination of workflow-based deployment and vehicle routing. The production stage is modeled as a deployment based on the workflow, and the second stage, the jobs produced must be delivered to the customer body [11]. Also, in order to minimize the last customer service time, a random variable neighborhood descent algorithm has been proposed in which various experimental factors, such as the use of alternative alternative solutions, solution demonstration and loading strategies, have been considered and analyzed [12]. An innovative large neighborhood search algorithm has been proposed by Chen et al. [13] to solve the problem of routing vehicles with time windows and delivery robots. The experiments, performance and effectiveness of the proposed solution algorithm provide insight into the use of automotive package delivery robots as an alternative service.

3 Model

The subject of vehicle routing is one of the familiar concepts in the field of operations research, which in the last two decades has been followed by great efforts and advances in this field. Vehicle routing problem refers to a problem in which a fleet consisting of several vehicles from one or more terminals or warehouses provides services to a set of customers located in different geographical locations, each of which has a specific demand and this in a way They do this to minimize the cost of doing so. During these routes, customers are met only once, and the requests of each of them are answered by only one vehicle.

In the water distribution network under study, which consists of many elements such as reservoirs, wells and pipes, sensors are placed in some parts of the network to check the water quality at regular intervals. Typically, many teams visit candidate nodes (clients) daily and perform the necessary tests. In this research, there are two types of demand points (customer):

- 1) The first type (households, shops and tanks): For this type of customers, the sampler can visit the selected points and perform the necessary tests on them. The difference is that for tanks, the sampler needs more time to perform the tests, which leads to more visiting time.

2) The second type (wells): For this type of customers, the sampler needs a key to open the well to perform tests, and because the key is not available, he performs visits according to a specific schedule; Therefore, before reaching the wells, he should go to the places where the key is kept, then continue working and finally return the key visit to the place.

Assumptions

- The number of customers is constant and equal to n .
- Service time is different for each type of customer.
- The communication network between customers and the warehouse is complete.
- It is assumed that there is a main network all customers and the warehouse.
- That is the device can move from one node to n other nodes.

Decision Variables

x_{ijk} : If the path i to j is visited by device k , it takes 1, otherwise it takes zero.

y_{ik} : If node i is visited by device k , it gets a value of 1, otherwise it gets a value of zero.

s_{ik} : Travel time to node i by k .

u : Variable added to the model for linearization of constraints.

z_1 : Variable added to the model to linearize the objective function.

3.1 Mathematical Model

The objective function of the problem is defined as the minimum sum of the moving times between nodes as well as the time of visiting each node and finally the maximum moving time between nodes, as follows:

$$\text{Min } z = \sum_{i \in v} \sum_{j \in v} \sum_{k \in k} x_{ijk} (c_{ij} + v_i) + z_1$$

S.t:

$$\sum_{k=1}^K y_{ik} = \begin{cases} K & i = 0 \\ 1 & i = 1, 2, \dots, n \end{cases} \quad (1)$$

$$\sum_{i=0}^n x_{ijk} = y_{jk} \quad j = 0, 1, \dots, n; \quad k = 1, 2, \dots, K \quad (2)$$

$$\sum_{j=0}^n x_{ijk} = y_{jk} \quad j = 0, 1, \dots, n; \quad k = 1, 2, \dots, K \quad (3)$$

$$\sum_{i=0}^n x_{iqk} - \sum_{i=0}^n x_{qjk} = 0 \quad j = 0, 1, \dots, n; \quad k = 1, 2, \dots, K \quad (4)$$

$$\sum_{i \in S} \sum_{i \in S} x_{ijk} \geq |S| - 1 \quad S \subset V \setminus \{0\}; |S| \geq 2; k = 1, \dots, K \quad (5)$$

$$S_{jk} \geq S_{ik} + C_{ij} - M(1 - x_{ijk}) \quad i = 0, \dots, n; i = 0, \dots, n; i \neq j; k = 0, \dots, K \quad (6)$$

$$S_{jk} \leq S_{ik} + C_{ij} + M(1 - x_{ijk}) \quad i = 0, \dots, n; i = 0, \dots, n; i \neq j; k = 0, \dots, K \quad (7)$$

$$S_{0k} = 0 \quad (8)$$

$$S_{p_i, i} + C_{p_i, i} \leq S_{ik} + M(1 - y_{ik}) \quad i \in V_2; k = 1, 2, \dots, K \quad (9)$$

$$u + C_{i, d_i} y_{ik} \leq S_{d_i, k} \quad i \in V_2; k = 1, 2, \dots, K \quad (10)$$

$$u \leq S_{ik} \quad i \in V_2; k = 1, 2, \dots, K \quad (11)$$

$$u \geq S_{ik} + M(1 - y_{ik}) \quad i \in V_2; k = 1, 2, \dots, K \quad (12)$$

$$u \leq M(1 - y_{ik}) \quad i \in V_2; k = 1, 2, \dots, K \quad (13)$$

$$y_{p_i, k} \geq y_{ik} \quad i \in V_2; k = 1, 2, \dots, K \quad (14)$$

$$y_{d_i, k} \geq y_{ik} \quad i \in V_2; k = 1, 2, \dots, K \quad (15)$$

$$z_1 \geq \sum_{i=0}^n \sum_{j=0}^n c_{ij} x_{ijk} \quad (16)$$

$$\sum_{i=0}^n \sum_{j=0}^n x_{ijk} (c_{ij} + v_i) \leq L \quad k = 1, 2, \dots, K \quad (17)$$

$$x_{ijk} \in (0, 1) \quad (18)$$

$$x_{ik} \in (0, 1) \quad (19)$$

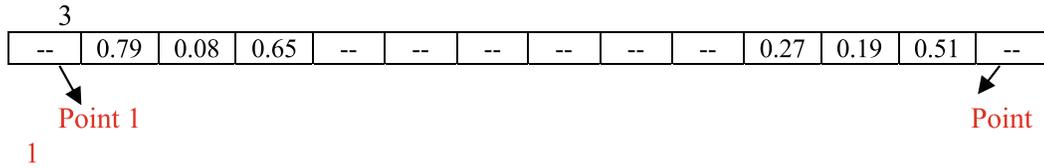
$$S_{ik} \geq 0 \quad (20)$$

$$u \geq 0 \quad (21)$$

$$Z_1 \geq 0 \quad (22)$$

The purpose of the problem is to minimize the sum of the time of movement between nodes and also the time of visit of each node and finally the maximum time of movement

To make each of the initial answers, we do this:



The second cells to the penultimate end of the chromosome must be filled in order. Candidate points must be identified for each step (to fill each cell). For this purpose, we consider 4 containers x_1, x_2, x_3 and x_4 .

x_1 represents the set of type 1 nodes that have not yet been selected. (Obviously in the second cell of chromosomes, this container contains all the points of type 1 problem).

x_2 Represents a set of type 2 nodes whose key has been removed. (Obviously in the second house of chromosomes, this dish is empty, because before the second house there was definitely no trip to the house of key points).

x_3 Indicates key points that have not been visited before.

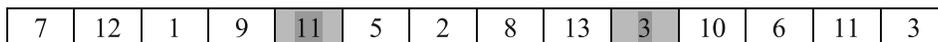
x_4 Indicates the nodes of the key point to which the trip was made, and all type 2 trips related to the key of this point have also been performed. (Ie key points we are allowed to return to).

Therefore, to fill each of the chromosome cells, a point from the X set must be selected.

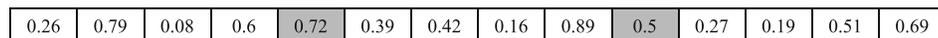
$x = [x_1, x_2, x_3, x_4]$ After selecting a node from the candidate X points and filling one cell of the chromosome, to fill the next cell of the chromosome, it is necessary to update the four containers x_1, x_2, x_3 and x_4 to determine the new candidate points to fill the next cell. This process continues to fill all the chromosome cells.

4.1 Definition of 3 Structures

1) Select the two key points that first appear on the chromosome (in this chromosome, points 3 and 11 are given twice on the chromosome under the assumptions of the problem. But the first choice of each is considered) and move those two points together. For example, if the arrangement of genes is as follows and points 3 and 13 are nodes of key points.



And the pseudo-code of random numbers related to this chromosome is as follows.



Then, after applying structure 1, the chromosome string is turned as follows, and then according to this string, the numbers of volunteer nodes are placed on the chromosomes.

0.26	0.79	0.08	0.65	0.56	0.39	0.42	0.16	0.89	0.72	0.27	0.19	0.51	0.69
------	------	------	------	------	------	------	------	------	------	------	------	------	------

2) Select a key point that first appears on the chromosome (on this chromosome point 8 is given twice on the chromosome under the assumptions of the problem. But the first choice is considered) and change the random number of that gene from the chromosome. For example, if the arrangement of genes is as follows and point 8 is a key point.

7	12	1	9	11	5	2	8	13	3	10	6	3	8
---	----	---	---	----	---	---	---	----	---	----	---	---	---

And the pseudo-code of random numbers related to this chromosome is as follows.

0.26	0.79	0.08	0.65	0.72	0.39	0.42	0.16	0.89	0.56	0.27	0.19	0.51	0.69
------	------	------	------	------	------	------	------	------	------	------	------	------	------

Then, after applying structure 2, the chromosome string is as follows, and then according to this string, the numbers of volunteer nodes are placed on the chromosomes.

0.26	0.79	0.08	0.65	0.56	0.39	0.42	0.97	0.89	0.72	0.27	0.19	0.51	0.69
------	------	------	------	------	------	------	------	------	------	------	------	------	------

3) Select a key point that appears on the chromosome a second time and change the random number of that gene from the chromosome. For example, if the arrangement of genes is as follows and point 6 is a key point.

7	12	1	9	6	5	2	8	13	3	10	6	3	11
---	----	---	---	---	---	---	---	----	---	----	---	---	----

Then, after applying structure 3, the chromosome string is as follows, and then according to this string, the numbers of volunteer nodes are placed in the chromosomes.

0.26	0.79	0.08	0.65	0.56	0.39	0.42	0.97	0.89	0.72	0.27	0.19	0.51	0.69
------	------	------	------	------	------	------	------	------	------	------	------	------	------

5 Solution Methodology

The response quality and convergence speed in the genetic algorithm are related to the algorithm parameters. These parameters are the number of generations that affect the quality of the response. Population size, which affects the amount of search space for the answer. The rate of intersection, which determines the probability of the intersection operator on a chromosome, and the rate of mutation, which indicates the probability of mutation on a chromosome. Variable Neighborhood Search (VNS) algorithm whose pseudocode is shown in Fig. 1. It was first proposed by Mladenovik and Hansen in 1997. The main idea of this algorithm is to systematically change the neighborhood structure during the search to avoid falling into the local optimal. The simplicity of

implementation and the quality of the results obtained from the VNS algorithm quickly made this algorithm a good way to solve optimization problems. The VNS algorithm begins by generating the initial answer and defining the neighborhood structures and using a neighborhood search method. The neighborhood structures of the algorithm are $l = \{1, 2, \dots, l_{\max}\}$ and N_l . Which N_l is l 's neighborhood structure. After determining the possible neighborhood structures, their order is determined. Two important points here are choosing the right neighborhood structures and determining the right order (for example, the order based on the magnitude of the neighborhood structures).

Generating quality initial solutions, defining neighborhood structures and determining their order and using the appropriate method for local search are the determining factors in the quality of the answers obtained from the algorithm. The algorithm starts using the generated initial answer (s_0) and the main loop of the algorithm is repeated until the termination criterion is established [14].

The main loop of the VNS algorithm consists of two main phases of shock and local search. In the shock phase, the algorithm moves from the current answer to the neighbor answer (s') using the l 's neighborhood structure. In the local search phase, the search is performed on the answer S' using local search methods to reach the local optimization (s'^*). Now in the move or not move section, if the local optimization obtained is better than the current answer S , it will be replaced and the search will return to N_l , otherwise from the neighborhood structure N_{l+1} to continue Search is used. This search continues as long as $l < l_{\max}$. Figure 1 shows the pseudocode of the variable neighborhood search algorithm.

```

Input: a set of neighborhood structures  $N_t, t= 1, 2, \dots, t_{\max}$ 
        S=generate initial solution ();
        Repeat
            t=1;
            While (t≤tmax)
            {
                S'=Shaking (S,  $N_t$ )
                S'*=Local search (S')
                if f(S'*)<f(S)
                    S ← S'*
                    t=1;
                else
                    t=t+1;
            }
        ;
    
```

Until stopping condition are met;
Output: The best solution;

Fig. 1. Pseudo-code variable neighborhood search algorithm.

6 Computational Experiments

In this step, the performance of the proposed algorithm is checked. For this purpose, data related to wells in Tehran are examined. The well network of this city has 87 nodes, 26 of which are nodes related to keyless wells and 53 nodes are key-type nodes. 8 nodes are considered as key locations. The beginning and end of the journey are done from a specific point. The goal is to minimize the total travel time between nodes and the time to visit each node, and ultimately the maximum time to move between nodes. The proposed algorithms for the mentioned problem are parameterized by comparing the results and the results are shown in Table 1.

Table 1. The result of GA-VNS, GA

	GA-VNS	GA
N_{max}	100	100
N_{pop}	100	100
P_c	0.7	0.80
P_m	0.25	0.35

The proposed model is first solved by Dijkstra method and based on the studied data and then by GA meta-heuristic algorithm and also GA-VNS heuristic algorithm and based on the structures defined for the model. The model code runs in Python on a computer system with 2.53 GHz Core i5 processor specifications and 8 GB of external memory. The results of model implementation are shown in Tables 2 and 3.

The exact solution of the model by Dijkstra method, in 97.09 s, has obtained the value 4392 for the objective function of the problem. In GA solution, the execution time of the model is significantly reduced to less than 8% of the exact solving time, but between the answer obtained and the exact answer, for three times of model execution, Gaps 7.93, 7.53 and 11.33 Has come. But the proposed hybrid algorithm, taking into account only the first neighborhood structure, and for three model runs, shows gaps of 4.84, 3.41 and 4.28. This Gap value is 1.84, 3.05 and 2.34 for the case where the first and second structures are considered, and 0.66, 0.39 and 0.57 for the case where all three defined neighborhood structures are considered. Therefore, according to the results of solving the model, it can be concluded that the proposed combined hybrid algorithm and also the neighborhood structures defined for it, has provided high accuracy for the model in providing the final answer to the routing problem. In a much shorter time and about 10% of the exact solution time, it provides a very close answer with a gap of less than 1%.

Table 2. The result of GA-VNS, GA, Dijkstra

Algorithm	Dijkstra	GA		
		1	2	3
Test number				
Execution time (s)	97.09	7.24	7.38	7.11
Total travel time (min)	4392	4723	4723	7890

Table 3. The result of GA-VNS

Algorithm	GA-VNS (k = 1)			GA-VNS (k = 2)			GA-VNS (k = 3)		
	1	2	3	1	2	3	1	2	3
Test number									
Execution time (s)	7.96	8.03	8.42	9.22	7.78	8.23	9.57	9.84	9.67
Total travel time (min)	4605	4542	4580	4473	4526	4495	4421	4409	4416

7 Conclusion and Future Works

In this research, the issue of vehicle routing in Mashhad water distribution network has been investigated. An issue whose primary purpose is to remove a key from a specific location and return it to that location. For this reason, the problem was examined in the form of a two-objective problem. The first goal is to minimize the total time of movement between nodes and the time of visiting each node and the second goal is to reduce the time of movement between nodes. A nonlinear mathematical model was linearized for the problem and then two methods of neighborhood search algorithm and genetic metaheuristic algorithm were used to solve it. In this research, the use of the desired algorithms and the reciprocal route intended for customers improved problem solving.

Considering the issues studied during the present study, suggestions for the development of this study have been presented:

- One of the determining parameters in the issue of vehicle routing is the capacity of the vehicles, which can be considered limited.
- Vehicles can go to the gas station if they run out of fuel on the roads.
- One of the things that can be used to bring the situation closer to the real world is to consider time windows. Whenever the service is less or more than the allowed service interval, a penalty is applied to the target function.
- Different drivers vehicles are considered.

Acknowledgement. This research has been financially supported by The Analytical Center for the Government of the Russian Federation (Agreement No. 70-2021-00143 dd. 01.11.2021, IGK 000000D730321P5Q0002).

References

1. O'Grady, J., Zhang, D., O'Connor, N., Regan, F.: A comprehensive review of catchment water quality monitoring using a tiered framework of integrated sensing technologies. *Sci. Total Environ.* **765**, 147266 (2020)
2. Chen, Y., Han, D.: Water quality monitoring in smart city: a pilot project. *Autom. Constr.* **89**, 307–316 (2018)
3. Rathi, S., Gupta, R.: Monitoring stations in water distribution systems to detect contamination events. *ISH J. Hydraul. Eng.* **20**(2), 142–150 (2014)
4. Hu, C., Li, M., Zeng, D., Guo, S.: A survey on sensor placement for contamination detection in water distribution systems. *Wireless Netw.* **24**(2), 647–661 (2016). <https://doi.org/10.1007/s11276-016-1358-0>
5. Dantzig, G.B., Ramser, J.H.: The truck dispatching problem. *Manage. Sci.* **6**(1), 80–91 (1959)
6. Fozooni, N., Daneshvari, H., Dastgoshade, S., Abraham, A.: Relationship between the monopoly of tobacco law and lung cancer using the theory of dynamic systems. In: International Conference on Innovations in Bio-inspired Computing and Applications, pp. 289–300, December 2020
7. Braekers, K., Ramaekers, K., Van Nieuwenhuysse, I.: The vehicle routing problem: state of the art classification and review. *Comput. Ind. Eng.* **99**, 300–313 (2016)
8. Rabbani, M., Navazi, F., Farrokhi-Asl, H., Balali, M.: A sustainable transportation-location-routing problem with soft time windows for distribution systems. *Uncertain Supply Chain Manage.* **6**(3), 229–254 (2018)
9. Islam, M.A., Gajpal, Y., Elmekawy, T.Y.: Hybrid particle swarm optimization algorithm for solving the clustered vehicle routing problem. *Appl. Soft Comput.* **110**, 1–20 (2021)
10. Thompson, M.G., et al.: Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers—eight US locations, December 2020–March 2021. *Morb. Mortal. Wkly Rep.* **70**(13), 495 (2021)
11. Repoussis, P.P., Tarantilis, C.D., Ioannou, G.: The open vehicle routing problem with time windows. *J. Oper. Res. Soc.* **58**(3), 355–367 (2007)
12. Goodarzian, F., Taleizadeh, A.A., Ghasemi, P., Abraham, A.: An integrated sustainable medical supply chain network during COVID-19. *Eng. Appl. Artif. Intell.* **100**, 104188 (2021)
13. Chen, C., Demir, E., Huang, Y.: An adaptive large neighborhood search heuristic for the vehicle routing problem with time windows and delivery robots. *Eur. J. Oper. Res.* **294**(3), 1164–1180 (2021)
14. Mladenović, N., Hansen, P.: Variable neighborhood search. *Comput. Oper. Res.* **24**(11), 1097–1100 (1997)