Accepted Manuscript

Title: Solving Vehicle Routing Problem by Using Improved K-Nearest Neighbor Algorithm for Best Solution

Authors: Mazin Abed Mohammed, Mohd Khanapi Abd Ghani, Raed Ibraheem Hamed, Salama A. Mostafa, Dheyaa Ahmed Ibrahim, Humam Khaled Jameel, Ahmed Hamed Alallah

PII:	S1877-7503(17)30426-X
DOI:	http://dx.doi.org/doi:10.1016/j.jocs.2017.04.012
Reference:	JOCS 661

To appear in:

Received date:	8-2-2017
Revised date:	30-3-2017
Accepted date:	16-4-2017

Please cite this article as: Mazin Abed Mohammed, Mohd Khanapi Abd Ghani, Raed Ibraheem Hamed, Salama A.Mostafa, Dheyaa Ahmed Ibrahim, Humam Khaled Jameel, Ahmed Hamed Alallah, Solving Vehicle Routing Problem by Using Improved K-Nearest Neighbor Algorithm for Best Solution, Journal of Computational Sciencehttp://dx.doi.org/10.1016/j.jocs.2017.04.012

Solving Vehicle Routing Problem by Using Improved K-Nearest Neighbor Algorithm for Best Solution

Mazin Abed Mohammed[1,2*], Mohd Khanapi Abd Ghani[1], Raed Ibraheem Hamed[3*], Salama A. Mostafa[4], Dheyaa Ahmed Ibrahim[5], Humam Khaled Jameel[5], Ahmed Hamed Alallah[5]

[1]Biomedical Computing and Engineering Technologies (BIOCORE) Applied Research Group, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia
[2]Planning and Follow Up Department, University Headquarter, University of Anbar, Anbar, Iraq
[3]Department of IT, College of Science and Technology, University of Human Development, Sulaymaniyah KRG, Iraq
[4]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn, Johor, Malaysia
[5]College of Computer Science and Information Technology, University of Anbar, Department of Information Systems
mazin_top_86@yahoo.com, khanapi@utem.edu.my, raed.alfalahy@uhd.edu.iq, semnah@yahoo.com, dheyaa.ibrahim88@gmail.com, kareskh91@gmail.com, hemedahmed24@gmail.com

Highlights

- (1) A comprehensive listed of active KNNACVRP.
- (2) Identified and established an evaluation criterion for KNNACVRP.
- (3) Highlight the methods, based on NNA operation, for selecting the best way.
- (4) KNNA finds a shorter distance for VRP routes.

Abstract

Context: Vehicle routing problem (VRP) is one of the many difficult issues that have no perfect solutions yet. Many researchers over the last few decades have established numerous researches and used many methods with different techniques to handle it. But, for all research, finding the lowest cost is very complex. However, they have managed to come up with approximate solutions that differ in efficiencies depending on the search space. **Problem**: In this study the problem is as follows: have a number of vehicles which are used for transporting applications to instance place. Each vehicle starts from a main location at different times every day. The vehicle picks up applications from start locations to the instance place in many different routes and return back to the start location in at specific times every day, starting from early morning until the end of official working hours, on the following conditions: (1) Every location will be visited once in each route, and (2) The capacity of each vehicle is enough for all applications included in each route. **Objectives**: Our paper attempt to find an optimal route result for VRP by using K-Nearest Neighbor Algorithm (KNNA). To achieve an optimal solution for VRP with the accompanying targets: (1) To reduce the distance and the time for all paths this leads to speedy the transportation of customers to their locations, (2) To implement the capacitated vehicle routing problem (CVRP) model for optimizing the solutions. **Approach**: The approach has been presented based on two phases: firstly, the algorithms have been adapted to solve the research problem, where its procedure is different than the common algorithm. The structure of the algorithm is designed so that the program does not require a large database to store the population, which speeds up the implementation of the program execution to obtain the solution; secondly, the algorithm has proven its success in solving the problem and finds a shortest route. For the purpose of testing the algorithm's capability and reliability, it was applied to solve the same problem online validated and it achieved success in finding a shorter route. **Finding**: The findings outcome from this study have shown that: (1) A universal listed of dynamic KNNACVRP; (2) Identified and built up an assessment measure for KNNACVRP; (3) Highlight the strategies, based KNNA operations, for choosing the most ideal way (4) KNNA finds a shorter route for VRP paths. The extent of lessening the distance for each route is generally short, but the savings in the distance becomes more noteworthy while figuring the aggregate distances traveled by all transports day by day or month to month. This applies likewise to the time calculate that has been decreased marginally in view of the rate of reduction in the distances of the paths.

***Keywords**: K-Nearest Neighbor Algorithm; Vehicle Routing Problem; Capacitated Vehicle Routing Problem*

1. Introduction

The vehicle routing problem becomes an area of research since it was studied by Dantzing and Ramser in1959. It has been investigated by many researchers for more than fifty years. The VRP has many applications in real life. It clarifies in a wide area of transportation and distribution such as transportation of individuals and items, conveyance service and garbage collection. All these problems have economic importance, particularly in developed countries. The economic factor in savings expenditures is a big motive for companies and researchers in an attempt to find the best way to resolve and improve transport efficiency [1]. The basic and traditional problem in routing, and which has been addressed by researchers for more than a hundred years, is the C Salesman Problem (TSP) who distributes goods to a group of cities and returns to his hometown. The condition is that the salesman visits each city once on his routing trip provided that he follows the shortest and the least costly route. In fact, the salesman problem is an arithmetic problem that can be easily represented by a graph showing a collection of nodes that represent the cities to be visited. Many of the researchers dealt with this problem and they used almost all kinds of algorithms to find the optimal solution to it, which to some extent succeeded in solving the problem when the number of cities was limited and failed when the number of cities was increased.

The VRP is summarized as follows: A fleet of vehicles (for distribution of goods) starts from one location and visits a group of scattered cities or customers and return to the same location with less distance and costs on the conditions [2]:

- Every city is visited by one vehicle only once within a single route.
- The capacity of each vehicle is enough for all cities included in the route.
- Routes begin and end at the same location.

The number of vehicles is supposed to be less than what can be proposed for routes as well as the number of routes is to be less than what can be provided to cover all cities. Static Vehicle Routing Problem (SVRP) means that all the information on the route is known in advance before starting and does not change after the route began. The problem will be dynamic if any restriction is imposed, like time, vehicle capacity or other variables. Exact algorithms, heuristic and metaheuristic were used to solve the VRP for optimality [2]. Many researchers have tried to solve the problem but failed to find the optimal solution. However, they have managed to come up with approximate solutions that differ in efficiency depending on the search space.

2. Research Problem

The problem of vehicle routing is one of the many problems that have no perfect solutions yet. Many researchers over the last few decades till now have established a lot of researches and used many methods with different techniques to handle it. But, for all research, finding the lowest cost is very complex and needing innovative methods to find an approximate best solution.

The problem statement is as follows: have a number of vehicles which are used for transporting applications to instance place. Each vehicle starts from a main location at different times every day. The vehicle picks up applications from start locations to the instance place in many different routes and return back to the start location in at specific times every day, starting from early morning until the end of official working hours, on the following conditions:

- Every location will be visited once in each route.
- The capacity of each vehicle is enough for all applications included in each route.

The paths must be created in a manner that every location is visited to just once by precisely one vehicle. All paths begin and end at the fundamental location, and the aggregate requests of all stations on one specific path should not surpass the limit of the vehicle. The distances among the different locations must be calculated to generate the population for the routes, and then the K-Nearest Neighbor Algorithm will be applied to find a best solution for the issue. We hope that the results of this research will help to reduce the transportation costs either by reducing the travel time or the distance for each route.

The major aim of this study is to improve the efficiency and accuracy of the VRP and improve the speed of preparing the solutions. The study investigates the most suitable parameters for the population based algorithm. By doing this, the study hopes to find the most feasible and efficient the vehicle routing problem. The aim of the study can be achieved by accomplishing the following objectives: (1) To reduce the distance and the time for all paths this leads to speedy the transportation of customers to their locations, (2) To reduce the travel transportation costs such as fuel utilization and also the vehicle upkeep costs, (3) To implement the algorithm can be used and applied for any problems like VRP, and (4) To implement the CVRP model for optimizing shuttle bus services.

3. Related Work

Vehicle Routing Problem becomes an area of research since it was studied by Dantzing and Ramser in1959. It has been investigated by many researchers for more than fifty years [1,2,3]. The VRP has many applications in real life. It clarifies in a wide area of transportation and distribution such as transportation of individuals and items, conveyance service and garbage collection. All these problems have economic importance, particularly in developed countries. The economic factor in savings expenditures is a big motive for companies and researchers in an attempt to find the best way to resolve and improve transport efficiency [1]. Concept of VRP can be described as the issues of designing shortest paths from one location to a group of geographically distributed locations (customers, cities, universities, warehouses, schools, stores, etc.) [1,2,3,4].

Shaw [7] used a large neighborhood search (LNS) with constraint programming technology for solving capacitated vehicle routing problems. The technique of LNS acts like a local search in making moves, but uses a tree-based search with constraint to evaluate the cost and validity of the move (remove and re-insertion of customer's visits). The research showed that the average solutions produced by LNS and its performance are close and competitive to the best published solutions of the best operations research meta-heuristic methods. Baldacci et al. [8] presented a new branch-and-cut algorithm to solve the CVRP based on a two commodity network flow formulation. They derived a new lower bound using Linear Programming (LP) and compare with the lower bound of different CVRP formulations. Their research showed that their algorithm is capable of solving the problem for large instance. Mohammed et al [1] consider the utilization of a GA in resolve CVRP in which a group of vehicles with limits on capability and moving time are accessible to benefit a group of customers and compelled by soonest and most recent time for serving. The outcomes demonstrate that GA s eligible to decide the best path for the vehicles while keeping up their requirements of limit and travel time. Nowadays, evolutionary algorithms like GA and ANN are the main interests of many researchers. GA may not solve the problem and find an optimal solution or may lead to the dead end, but, it can find an acceptable or feasible solution better than other algorithms in reasonable time.

According to Lopes et al. [10] an Ant Colony Algorithm is used to solve the CVRP. in their study proposed two stages of optimization: the first stage searches for demand paths and the second stage is enhancing every route as a TSP problem. their study outcomes are acceptable and good for small cases. Kuske et al. [11] showed a model for an Ant Colony Optimization (ACO) algorithm to solve the CVRP as a group of self-autonomous units. They displayed the self-sufficient conduct of each ant as a self-ruling unit which responds freely in a general domain while searching to objective. Their study outcomes demonstrated that the self-autonomous units improve Ant colony technique. Lyamine et al. [12] a Hybrid method to solve the problem by integrating an Ant Colony algorithm with a Savings algorithm. Their study outcomes demonstrated that this technique is competitive with the other approaches such as Tabu Search and Simulated Annealing. Vigo et al [13] clarified that heuristics are capable of solving the VRP for optimality because they can build a feasible solution and at the same time minimize the cost to as low as possible. He added that the approaches easily handle the variety of constraints that arise in real world and should take into consideration the simplicity and flexibility in addition to accuracy and speed. Yeun et al. [14] described the classic VRP and the types of dynamic problem. They also reviewed all the algorithms used to solve this problem for optimality. They stated that heuristics searches such as Genetic Algorithm, Tabu Search, Scatter Search and Simulated Annealing could obtain best solutions with different types of problem. Many researchers now are interested in using Genetic Algorithm, Fuzzy Logic (FL) and Neural Network to solve the search problem. Goel et al. [15] presented a new algorithm that is based on Large Neighbourhood Search to solve the dynamic vehicle routing problem. They used fast insertion methods for requests to achieve fast response. The first one is the sequential method in which all new requests are chosen and their possible insertions are determined. The other is the auction method in which all vehicles are sent out with a possibility and efficiency of insertion for new requests. Each request to be inserted is allocated to a vehicle with low cost to be inserted in its tour. They have proven that the algorithm performs well when the number of vehicles and transfer requests are high with less response time. Montemanni et al. [16] presented an algorithm for solving VRP using Ant Colony paradigm in which the DVRP is decomposed into a sequence of static VRPs. They proposed three major components. The first is an event manager, which records new orders. The second element is the Ant Colony algorithm which is based on the computational model inspired from real ant colonies function. The third is pheromone conservation in which a matrix including all the data for optimal solutions is used. If a static problem is similar to the next one, the information of good solution is passed on to the next problem. Their algorithm achieved good results compared with other heuristic techniques.

Csiszar et al. [17] presented route removal technique for VRP with Time Windows to minimize the number of routes and estimated cost. He presented two phase solutions: one for route elimination, and the other for the cost reduction. The outcomes demonstrate that his algorithm is competitive against the best algorithms. Azi et al. [18] used an exact algorithm for the first time to solve the VRP with multiple use of vehicle with which each one executes many routes according to a time window and customer demands. They used the column generation algorithm with a branch-and-price algorithm to facilitate the problem formulation. The column generation algorithm is used when customers are chosen according to profit to be gained from them, in case it is not possible to serve them all. The result shows that their method is limited by problem size which can be solved with 20 customers but many research problems have up to 40 customers. Xu et al. [19] used genetic algorithm to solve the VRP with Time Windows and Fuzzy Demand to reduce the total distance for vehicles routes and the delay times at the customers due to time violations. The GA is applied to solve this problem which is formulated in two stages; the first stage reduces the cost in both stages due to initial solutions and in the second stage, the failure for cost of route. Their research shows that their solutions are close to the best solutions. Xin et al. [20] showed the dynamic VRP with time windows which considers the time delay caused by traffic congestion during a day. They considered soft time windows which allows vehicle delay but with penalty. Their system is efficient and more flexible to resolve their problem compared with other heuristics. They considered routes planning and the accidents that appeared during the day. Based on travel time model, the algorithm calculated the speed that reflects the traffic congestion. They show three different methods to obtain three routes of different types in which the range from slow too high for average speeds are created. They aimed to reduce the time spent by routes and the customers time delay. Their research proves that the variable speed model is durable and achieves better results compared with the constant speed model. Nazif et al. [21] presented an optimized crossover operator in Genetic Algorithm application to solve the dynamic VRP with time window. Optimized crossover was applied by Aggarwal et al.'s genetic algorithm [22]. The crossover operator is the main factor to select individuals from a population for generating offspring. Instead of the traditional way to produce offspring, they used the swap node operator, which randomly selects and swaps two nodes from a parent and repeated on the second parent to create a second offspring. They test their results with other algorithms with good solutions and show that their results are competitive.

Tang et al. [23] studied the VRP with time windows (VRPTW) and proposed a mathematical model for it. He proceeded to solve the model by improved genetic algorithm. First, the initial population is computed then a pair of individuals is selected after computing the fitness values to make the crossover to generate two new individuals. He used a novel order crossover operator (NOX) which is better than other operators in generating a new child that is different from the parents. For mutation, he used a modified mutation operator which includes swapping mutation and inversion operator. In swapping, the genes are swapped on two selected positions. In inversion operator, the chromosome is found randomly. Using two cutting points, the operation to produce the child by inverting the substring between them. The research results show that the improved genetic algorithm is effective in achieving good solutions. Parragh et al. [24] presented adaptive variable neighborhood search (VNS) to solve MDVRPTW by introducing two parallel approaches of VNS using two cooperation schemes. The first scheme stores and manages the best found solutions and the most important search parameters. The second one reproduces the successful features of the successive VNS. Their research results show that in 11 cases out of 20, new best solutions are obtained. Carlsson et al. [25] presented two heuristics to minimize the maximal distance of a route. The first one is a linear programming-based approach with global improvement to assign customers to depots and generate routes for each vehicle. The second one is the partition heuristic which divides the service region into equal sub-regions that contain the same number of nodes. Then they proposed a fast approximation algorithm to generate good initial solutions. Their research techniques give good results than conventional local search methods. Cao et al. [26] presented a genetic algorithm to tackle the shortcomings of premature and slow convergence of classical genetic algorithm (GA). Their research shows that the results and performance of the proposed method is better than the traditional genetic applications. Martinovic et al. [27] proposed an adjustable simulated annealing with random initial solution algorithm to resolve Single-Commodity VRP with Pickup and Delivery Service. They show that their results are good for different number of instances in practical applications. Xuping et al. [28] presented a disruption management model and an improved GA to resolve the VRP with disruption events that may happen during the routes such as traffic accidents or vehicle breakdowns. The model is based on a series of solving simplified strategies based on the theory of disruption management. These strategies are used to simplify the solution of complicated optimization problem and simplify the solution space. When a vehicle on a scheduled task in a distribution system breaks down, the other transport vehicles or additional ones could compensate for that and complete the mission to deliver the goods of the disabled vehicle at the time, when serving the customer at the time is the most important goal for each solution. They improved the population and crossover operations in the genetic algorithm to improve the solution. They also improved the validity of strategies and algorithms by representing the disruption times and breakdown vehicles. Other studies on the VRP problem with different methods and techniques tried to get or find best solution are applied in this fields such as [29], [30], [31], [32].

4. Materials and Methods

K-NNA are non-deterministic, i.e. they are stochastic in decisions, which make them more robust. It is a key technology in random search to solve unclear and complex problems, which require large time space for best solution. It is an important technique in the search for the perfect choice of a solution set that is available for a particular design. KNNA is an excellent technique to solve huge problems with a high computational complexity, especially in computer science, where the problems need to enhance solutions. In this study, K-NNA is applied to solve and optimize the VRP. KNNA is a non-parametric approach utilized for classification & regression. In both cases, the input involves of the k nearest training cases in the element space [33,34,35]. The output relies on upon whether k-NN is utilized for classification or relapse:

- In k-NN classification, the output is a class membership. An object is classified by a dominant part vote of its neighbors, with the object being appointed to the class most normal among its k closest neighbors (k is a positive whole number, ordinarily small). If k = 1, then the object is essentially allocated to the class of that single nearest neighbor.
- In k-NN regression, the yield is the property estimation for the object. This esteem is the normal of the estimations of its k nearest neighbors.

In this study, the description of the optimization problem and its mathematical model is addressed. Since it was published in 1959, VRP receives a lot attention and becomes a focus of numerous researches, because it is very similar to the TS Problem, which is considered as the base of VRP. But there is a big difference between them, unlike the TS Problem; VRP is a multi-constrained optimization issue. Traditional VRP aims to determine the lowest cost of a route for a vehicle starting from a central depot to serve a set of locations and return back to the same main depot. The VRP has many constraints such as, vehicle capacity, time window, multi-depots and others, which gives it more complex solution space [1]. However, VRP is a general name given to an entire class of issues in which a group of paths for an armada of vehicles based at one or a few terminals must be fixed for various geologically scattered customers or stations. The goal of the VRP is to convey a group of customers with known requests on least cost vehicle paths beginning and ending at a stop. In the (Fig 1) provide a typical example of VRP problem and one of its conceivable solutions.

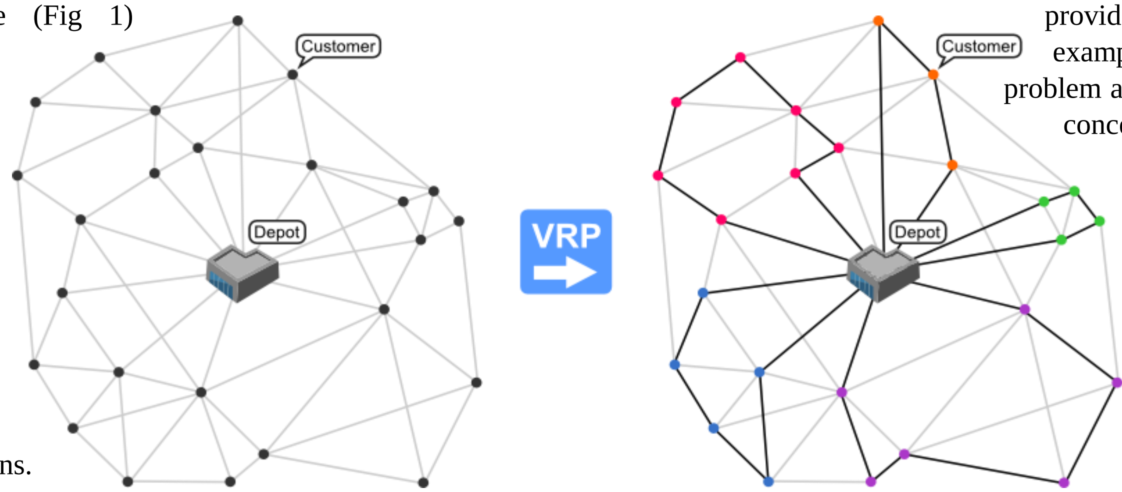

Fig 1 An example of a VRP (left) and its result (right)

To put the problem in perspective, we reproduce here the research problem for which the solutions to the VRP are sought. Simply and clearly stated, the VRP problem is the routing problem for cars to transport things within any country/city to many different locations. Cars start from the main (depot) location in different routes at specific times along the day. Each solution must include all locations, and each car must finish his route and return to the main location on the condition, that it does not visit each location more than once during each route. The transport costs its responsible company a lot of time and money yearly including fuel prices and the salaries of drivers and maintenance fees. The research aims to assist the companies in reducing these expenditures by using KNNA to find the shortest distance for vehicles routes.

4.1 Capacitated Vehicle Routing Problem (CVRP) Model

The CVRP used in this study is the general VRP, the methods for the VRP assist in solving the issue in situation of travelling applications. The goal of the VRP is to serve every one of the customers while reducing the aggregate travel separate under the capacity that the aggregate requests of the served users can't surpass the limit of the vehicle, where every vehicle starts from a similar depot, serves the users allocated to the warehouse, and comes back to the terminal with the area of the station and the users are given. In addition, each location (node) has its special demands, which will result to delaying the visited car at that node time [36,37]. The VRP can be described as follows:

- Number of vehicles must have enough one warehouse.
- Each vehicle begins from the main location, serves a group of stations and comes back to a similar stop. The aggregate request of locations served by the vehicle must not surpass the limit of the vehicle G.
- All N locations are served. Every user is served by precisely one vehicle once.
- The goal is to reduce the aggregate distance went by the set of the vehicles.

It is clear that VRP is a NP-difficult issue, in which finding the ideal result of VRP occurrence is hard and generally needs long computational time, thus, evolutionary algorithms have been utilized to locate a close ideal result in sensible measure of time [38,39]. The CVRP can be modeled as follows:

The VRP issue is given as a group of stations $N = \{1, 2, .., n\}$, staying at n different positions. Each two of positions(i, j), where $i, j \in N \ and \ i \neq j$, is connected with a distance traveled d_{ij} That is symmetrical $d_{ij} = d_{ji}$. The demand at point i denote by $q_i : i = \{1,2, \dots n\}$. The central depot is denoted by 0.

Let a graph $Gr = (S, D)$; $S = \{s_0, s_1, \dots, s_n\}$ in which Gr has $n + 1$ station and $D = \{(s_x, s_y): s_x, s_y \in S, x < y\}$ is the set of distances between the stations. Buses starting point is station 0 and the $S = S \backslash \{0\}$ related to n stations of bus stops. Let a cost variable be associated with each S of D, $c_{x,y}$; A set of buses, B with capacities of St students are located at the starting station 0; each of the B implements a rout of Gr in a straightforward cycle with the consideration of the St aggregate does not exceed the vehicle capacity.;. A $decision\,factor\,function,\ Df_{x,y,z}$ $(x, y = 0,1,2,3, \dots..n;\ z = 1,2,3, \dots, n;\ x \neq y)$ determines the rout options:

$$Df(x, y, z) \begin{cases} 1, & \text{true} \leftarrow \text{arc}\ (x, y) \\ 0, & otherwise \end{cases} \quad (1)$$

where z is the number of vehicles. The z is assumed 1 as the aim to reduce the route distance but not the number of vehicles. Since the cost of the route is a function of a route distance then; CVRP can be formulated as:

$$Minimize\ (\sum_{x=0, x\neq y}^{n} D_{x,y} Df_{x,y}) \quad (2)$$

Therefore, $x = 0,\ y = 1$ are subject to:

$$Df_{x,y} = 1, \forall\ x \in S$$

$$Df_{x,y} = 1, \forall\ y \in S$$

Since the slightest cost of every path in this issue relies on the distance, the distance should be reduced to raise chromosome fitness value. The fitness can be formulated as [1]:

$$f(\mathrm{x}) = 100 / \sum_{i=0, j=1}^{n} D_{ij} \quad (3)$$

4.2 K-Nearest Neighbor Algorithm

K-Nearest Neighbor Algorithm is effective in solving highly complex problems such as VRP. These are mainly optimization problems. The K-Nearest Neighbor technique generally using either the Euclidean distance or the cosine comparability between the training tuples and the test tuple be that as it may, goal of this paper, the Euclidean distance approach will be applied to implement our module. It is efficient to [40], [41]:

- Providing a faster and more accurate solution.
- Reducing error rate caused by accuracy in assumptions.
- In distributing a problem on parallel computers, which is difficult by linear programming, the distributed results can be easily compared.
- Stop the execution at any time, because there is a solution at any time, whatever it might be, better or worse.

KNNA are non-deterministic, i.e. they are stochastic in decisions, which make them more robust. It is a key technology in random search to solve unclear and complex problems, which require large time space for optimal solution. It is an important technique in the search for the perfect choice of a solution set that is available for a particular design. KNNA is an excellent technique to solve huge problems with a high computational complexity, especially in computer science, where the problems need to enhance solutions. KNNA can be applied in the following applications [42], [43]:

- Tour selection, the goal of tour optimization for travelling salesman problem and vehicle routing problem to reduce the costs and visit all locations with constraints.
- With Neural Networks, in which the operation of pattern selection or classification based on a previous input to select optimal pattern.
- Multi-constrained problems, such as routing of the network for the internet with different features.

- Many types of optimization problems such as Distributing packages, traffic optimization, puzzle optimization and others.

The following algorithm represents the implementation of the KNNA for VRP and the flowchart of nearest neighbor algorithm is showed in Fig 2.

Algorithm 1: The VRP model

1. begin;
2. generate the *InitialSolution* (group of routs) with applying a nearest rule between nodes i.e.:
   rout 1: 12 → 16 → 30 → 1 → 26 → 7 → 13 → 15,
   rout 2: 23 → 3 → 28 → 8 → 22,
   rout 3: 11 → 4 → 18 → 2 → 9 → 24,
   rout 4: 10 → 19 → 17 → 31 → 21 → 6 → 14 → 20,
   rout 5: 25 → 5 → 29 → 27;

//d denote to distance, q denote to demands, and s is the concerned solution.

3. calculate the fitness of *InitialSolution* as follow:
   $d_s = d_{Route_1} + d_{Route_2} + \cdots + d_{Route_N}$,
   $q_s = q_{Route_1} + q_{Route_2} + \cdots + q_{Route_N}$,
   $d_{Route_i} = d_{0,1} + d_{1,2} + \cdots + d_{n-1,n}$,
   $q_{Route_i} = q_1 + q_2 + \cdots + q_n$;

//Fitness_s is the fitness of solution

4. $Fitness_s = d_s + q_s$;
5. set best solution (*BestSol*) = *InitialSolution*;
6. loop for 10,000 iterations, in each iteration:
   6.1 generate new random solution (*NewSol*) with applying a nearest rule between nodes;
   6.2 calculate the fitness of a *NewSol*;
   6.3 if *NewSol* is better than saved *BestSol*, then *BestSol* = *NewSol*;
7. end of loop;
8. checking *BestSol* feasibility;
9. write out the path of each route stored inside the *BestSol*;
10. end;

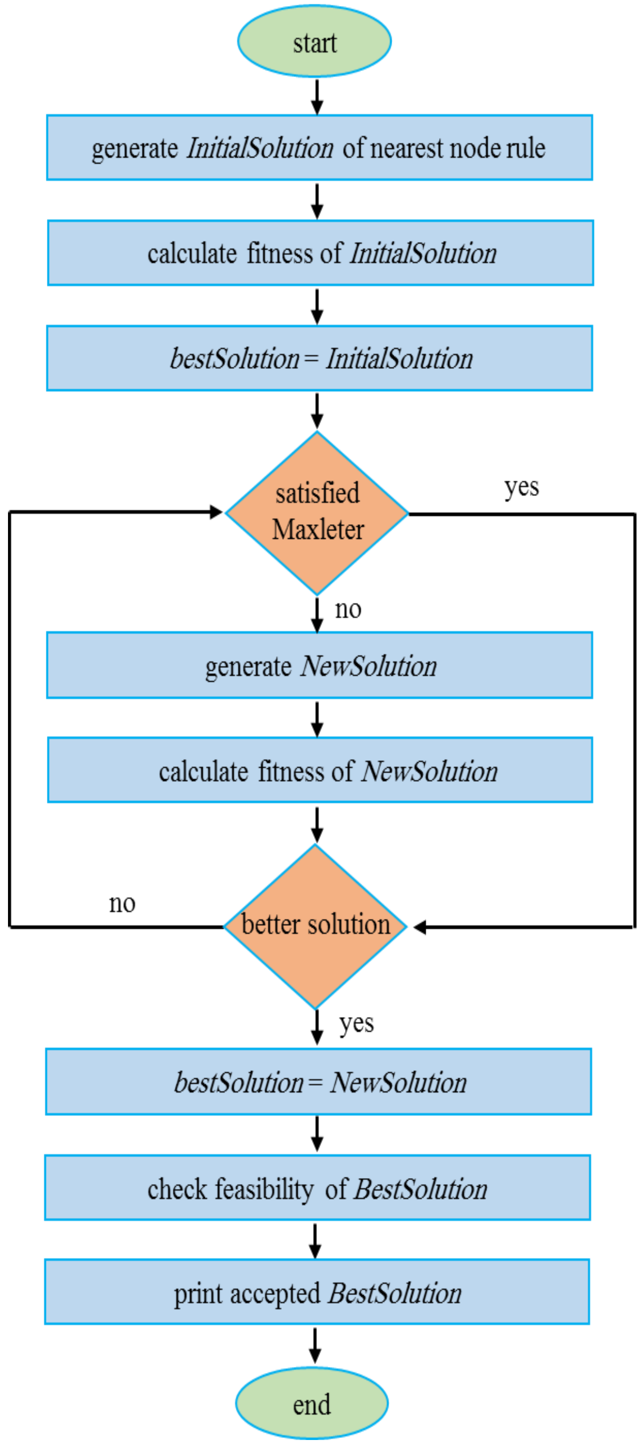

Fig 2 The nearest neighbor algorithm

4.3 Solution Generation

In this study, the solutions are generated respectively with several steps. At beginning, each vehicle starts from the depot (node number 0) and returns to the same depot to complete its route therefor, the first node of all vehicles is the depot. Then, the number of locations of each vehicle route will determine. After determining the number of nodes for each route of vehicle, the second node of each vehicle will be chosen randomly. The rest nodes of vehicle set will have determined by applying a nearest neighbor rules. The nearest neighbor rules are: the next node must be the nearest node among all the other nodes. The most important thing that must be taken into account is the individual node must be chosen once.

4.4 Fitness Evaluation

The "fitness function" is in control of assessment and giving back a positive whole number, or "fitness values", that reflects how ideal the result is: the lower the number, the better the result. The fitness values are then utilized as a part of a procedure of comparing between the generated *NewSol* and the stored *BestSol* for choosing which possible results will continue to another iteration, and which will ignore.

In this research, the distance of each route (d_{Route}) is calculated by collecting the whole distance from the depot (central node) to the last node for the individual vehicle. After that, the total distances of all vehicles are summed together and give the distance of the concerned solution (d_s). The second interested criterion of evaluation fitness is the demands of each node. As in the distance criterion, the demands of nodes in the individual route are collected to introduce this route demand (q_{Route}). The total demands of all vehicles inside the solution are summed together and give the demand value of the concerned solution (q_s). Finally, the total value of the solution fitness ($Fitness_s$) is given by collecting the total distances of vehicles (d_s) with the total demands of vehicles (q_s).

4.5 Feasibility Checking

The VRP problem is one of the problems that have no perfect solutions yet. Many researchers over the last few decades until now have established a lot of researches, and used many methods with different techniques to handle it. But, for all researches, finding the lowest cost is very complex and needing innovative methods to find an approximate optimal solution. Routing and scheduling problem is based on specific features of the rendered services, such as routing and scheduling objectives, number of vehicles, starting and ending location, and vehicles capacities. The objective of most routing and scheduling problems is to minimize the total cost, which is a function of vehicles number, capacity, route distance and vehicles operation and maintenance. Student's transportation is a service, provided by real study to facilitate the access of students to different locations, which represent the faculties and other departments in the main location. Therefore, the solution must schedule the problem to achieve minimum cost and on schedule.

The research problem is not simply, a set of locations have to be visited by a single vehicle, in a route contains all locations, on conditions, that the route should begin and end at same depot, and each location is to be visited only once. The locations may be visited in an order to achieve route feasibility with minimum cost, with or without restrictions, such as delivery time or vehicle capacity or others. The problem is very difficult to solve for optimality, but all researches try to determine a good minimum cost solution. VRP is multi objectives problem. That means, there are several issues that the result solution must meet it and in this work these issues are addressed. The first and most important one is, every location of the search space must be traveled precisely, once and every vehicle starts from depot and return to it. And at each location (node) the vehicle must wait exactly as the demand given for this location.

5. Results and Discussion

The VRP has many applications in real life. It appears in a wide area of transportation and distribution such as transportation of people and products, delivery service and garbage collection. All these problems have economic importance, particularly in developed countries. The economic factor in savings expenditures is a big motive for companies and researchers in an attempt to find the best way to resolve and improve transport efficiency. This is the traditional and basic vehicle routing problem in which all the information on the routes (customer location and demands) is known in advance before starting and will not be changed or updated after the routes began [42]. This means, no customers (cities) or new demands need to be inserted into one of the routes. In fact, if a new request has to be inserted, it will be very complex and need more planning process. If the capacities of all vehicles are identical and fixed, the problem is considered as the capacitated VRP. It's clear that the results showed to us after applying the algorithm of (KNNA) is very close to a global optimal we have existed in data set are taken for benchmarking. And from here we observe the results were good in terms of (time and fitness) as in an instance (23) as the example. We note that a global optimal is 333 and the value is 333 and this shows that the results are good and reliable, and this apply on another instances. K-nearest neighbor algorithm which used to solve Vehicle Routing Problem has been described. This studied problem has a several instances of standard benchmark with large search space with each instance. The applying of algorithm has proven its ability to solve these problems and find approximate solutions. The algorithms have been adapted to solve the research problem where its procedure is different than the common algorithm. The results show that the K-nearest neighbor algorithm successful in solving the transporting VRP. After applying the k-nearest neighbor algorithm to solve the VRP issue. And the results showed us as in the following table 1.

The constraints of the research problem, particularly the specific routes that must be followed by the buses, and the inability of paving new roads, result in limited search space, but at the same time, lead to access to good solutions. Although necessary, the reduction of expenses would not be significant, because the distances traveled by buses in their routes periodically every day, are short distances between internal locations within the main location. Therefore, the distance reduction rate for each route is small. This problem is unlike the traditional routing problem that deals with long distances and large spread of locations between cities or customers, which makes the issue of reducing costs very important, and has to be considered for measuring the ability of different algorithms for solving such problems. Although the distances between locations within the locations are short, and there is no ability to pave new roads, KNNA finds a shorter distance for routes. The proportion of reduction the distance for each route is relatively short, but the savings in the distance becomes greater when calculating the total distances traveled by all buses daily or monthly. This applies also to the time factor that has been reduced slightly based on the rate of reduction in the distances of the routes.

Any complex problem may have a very large number of solutions, no matter if some of these solutions are wrong and the others are correct. But always, there is a best solution, although in many cases, it is difficult to access it. The idea of KNNA lies in the generation of random solutions to such problem, examining these solutions, and comparing with certain criteria established by the algorithm designer. The best solutions are sought, and the less efficient ones are neglected. The efficient solutions are manipulated further to produce new solutions. The KNNA cannot find the optimal solution, but it is capable of finding an approximate solution. KNNA is a programming technique that imitates biological evolution as a problem-solving technique. In aligning with a specific problem at hand, KNN fitness function allows each candidate that needs to be solved to be quantitatively evaluated. The candidates may be the solution and by implementing KNNA makes the solutions better or improves the fitness level. The presence of KNNA is to evaluate each candidate at random according to the fitness function. Most of the time, the evaluation will not match/work, however, by chance a few may work even show weak or imperfect activity in solving the problem. These positive solution candidates are stimulated, reserved and permissible to reproduce. KNNA have been used widely in variety of fields that is difficult for human to handle. The solutions are more credible, efficient, faster, and of higher complexity than human engineers can solve. In many cases, KNNA produce solutions far better than the programmers who wrote the algorithm.

Table 1 Result for many instance

Instance	**Optimal**	**Our Method**	**Time(sec)**
A-n32-k5	784	797	0.85440
A-n33-k5	661	805	0.84362
A-n33-k6	742	869	0.85615
A-n37-k5	669	883	1.14100
A-n37-k6	949	1005	1.16998
A-n39-k6	831	977	1.32155
A-n44-k7	937	1137	1.7710
A-n45-k6	944	1129	1.85502
A-n45-k7	1146	1107	1.86844
A-n46-k7	914	1154	1.95561
A-n48-k7	1073	1203	2.28132

A-n53-k7	1010	1214	2.85678
A-n54-k7	1167	1219	2.90591
A-n55-k9	1167	1387	3.03500
A-n60-k9	1408	1436	3.7721
A-n62-k8	1290	1367	4.03635
A-n63-k10	1315	1511	4.37184
A-n65-k9	1177	1471	4.65246
A-n69-k9	1168	1552	5.6267
A-n80-k10	1764	1740	8.27966

KNNA is used to solve problems which cannot be solved theoretically without the need to find exact solutions with constraints in time and resources. The structure of the algorithm is designed so that the program does not require a large database to store the population, which speeds up the implementation of the program execution to obtain the solution. In this study, the execution time of the algorithm, obtaining the results, and drawing the route path is observed to be fast. The testing performance for KNNA method are shown in figure 3 and 4.

The evaluation has been made by comparing the different schemes and observing the performance and the efficiency of KNNA with each scheme. The results showed the ability to evaluate the performance of KNNA between the different operators, hence identifying the best results. Positive result has been obtained by many different configurations and settings as demonstrated by the KNNA scheme which produced a very good solution with best fitness value. The adaptive ratio demonstrated its ability as well to provide diversity in the population and as result, the KNNA was able to explore more variety of solutions and then better solution was found. The optimal settings and configurations found have demonstrated their accurateness and suitability to produce better quality route representing very good solution to the problem being addressed.

6. Conclusion

In this study, CVRP model is implemented for optimizing VRP services. K-Nearest Neighbor Algorithm (KNNA) is applied to solve this problem as it is eligible of solving numerous real life complex issues. The algorithms are written to perform all the KNNA operations. Since the least cost of the research issue based on minimizing the distance, the fitness value is calculated in a simple way as a function of distance. The structure of the KNNA is designed in a way that speeds up the search process. The algorithm achieves the research objectives through optimizing the distance of transportation routes in spite of the problem constraints. For testing the validity and reliability of the algorithm, it has been applied to the problem online, and has succeeded in solving the problem and finding the shortest route in a short time.

The presence of a single variable in the research problem (the distance), the small and limited number of station stops, makes it fairly easy to find a solution, and does not reveal the strength of the KNNA, which is known for its efficient in dealing with complex and ambiguous problems. There is no more than one physical road linking the different locations for collecting the students, which impedes the possibility of finding another physical roads and comparing between them to find a shortest distance route. It is suggested that another variable is added such as a heuristic function to represent the slope of the road or any other factors such as road smoothness or traffic jam or others. The slope will be used as a weight that will be multiplied by the distance between two consequent bus stops. This will help to determine the exact route time. The research aims to discover a best solution for VRP issue. The goal is to reduce the expenditure of transportation, which is a service provided free of charge to the problem. KNNA is used to solve the issue and it achieves the research objectives in a short execution time through in the epilogue, it has been demonstrated that the KNNA is effective in solving the VRP and finding approximate result as it is eligible numerous different issues. The strength of the KNNA comes from its capacity to be adjusted to solve any issue through consolidating with different methods, or by changing its methods as indicated by the issue as has been done in this study.

Acknowledgements
This research has been supported by Biomedical Computing and Engineering Technologies (BIOCORE) Applied Research Group, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia.

References

[1] Mohammed, M.A., Ahmad, M.S. and Mostafa, S.A., 2012, June. Using genetic algorithm in implementing capacitated vehicle routing problem. In Computer & Information Science (ICCIS), 2012 International Conference on (Vol. 1, pp. 257-262). IEEE.

[2] Mohammed, M.A., Obaid, O.I, Ahmad, M.S., Using Genetic Algorithm in Solving Vehicle Routing Problem for Optimal Solution, LAMBERT Academic Publishing in German, ISBN-13: 978-3-659-74963-6 Book Published on: 2015-08-12.

[3] Obaid, O.I, Mohammed, M.A, Ahmad, M.S., Solving Examination Timetabling Problem by Using Genetic Algorithm, LAMBERT Academic Publishing in German, ISBN-13: 978-3-659-76188-1, Book Published on: 2015-07-22.

[4] Ákos Kovács (2008). Solving the Vehicle Routing Problem with Genetic Algorithm and Simulated Annealing. Master Thesis. Högskolan Dalarna, Sweden.

[5] Obaid, O.I., Ahmad, M., Mostafa, S.A. and Mohammed, M.A., 2012. Comparing performance of genetic algorithm with varying crossover in solving examination timetabling problem. J. Emerg. Trends Comput. Inf. Sci, 3(10), pp.1427-1434.

[6] Shaw, P., 1998, October. Using constraint programming and local search methods to solve vehicle routing problems. In International Conference on Principles and Practice of Constraint Programming (pp. 417-431). Springer Berlin Heidelberg.

[7] Baldacci, R., Hadjiconstantinou, E. & Mingozzi, A. (2004). An exact algorithm for the capacitated vehicle routing problem based on a two-commodity network flow formulation. Operations Research, 723-738.

[8] Lopes, H.S., Dalle Molle, V.L. and Lima, C.R.E., 2005. An ant colony optimization system for the capacitated vehicle routing problem. In Proceedings of the XXVI Iberian Latin-America Congress on Computational Methods in Engineering CILAMCE 2005.

[9] Kuske, S., Luderer, M. & T◌nnieS, H, Autonomous units for solving the capacitated vehicle routing problem based on ant colony optimization. Electronic Communications of the EASST, 2010, 26.

[10] Lyamine, B, Amir , H .& Abder, K. A Hybrid Heuristic Approach to Solve the Capacitated Vehicle Routing Problem. Journal of Artificial Intelligence: Theory and Application, 2010,Vol.1,Iss.1.

[11] Vigo, D. Introduction to VRP. University of Bologna Dept. of Electronics, Computer Science and Systems (DEIS) and II Faculty of Engineering, 2007.

[12] Yeun, L. C., Ismail, W. A. N. R., Omar, K. & Zirour, M. Vehicle Routing Problem: Models and Solutions. Journal of Quality Measurement and Analysis JQMA, 4, 2008, 205-218.

[13] Goel, A. & Gruhn, V. Solving a dynamic real-life vehicle routing problem. Operations Research Proceedings, 2005, 367-372.

[14] Montemanni, R., Gambardella, L., Rizzoli, A. and Donati, A., 2003, April. A new algorithm for a dynamic vehicle routing problem based on ant colony system. In Second international workshop on freight transportation and logistics (Vol. 1, No. 1, pp. 27-30).

[15] Csiszár, S. Route Elimination Heuristic for Vehicle Routing Problem with Time Windows. Acta Polytechnica Hungarica, 2, 2005.

[16] Azi, N., Gendreau, M. & Potvin, J.-Y An exact algorithm for a vehicle routing problem with time windows and multiple use of vehicles. European Journal of Operational Research, 2010, 202, 756-763.

[17] Xu, J., Goncalves, G. and Hsu, T., 2008, June. Genetic algorithm for the vehicle routing problem with time windows and fuzzy demand. In 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence) (pp. 4125-4129). IEEE.

[18] Xin, Z., Goncalves, G. & Dupas, R. A genetic approach to solving the vehicle routing problem with time-dependent travel times. Control and Automation, 2008, 16th Mediterranean Conference.

[19] Nazif, H. & Lee, L. Optimized Crossover Genetic Algorithm for Vehicle Routing Problem with Time Windows. American Journal of Applied Sciences, 2010, 7, 95-101.
[20] Fonseca, C.M. and Fleming, P.J., 1995. An overview of evolutionary algorithms in multiobjective optimization. Evolutionary computation, 3(1), pp.1-16.
[21] Tang, K. S., Man, K. F., Kwong, S. & He, Q. Genetic algorithms and their applications. Signal Processing Magazine, IEEE, 1996, 13, 22-37.
[22] Parragh, S. N., Doerner, K. F. & Hartl, R. F. A survey on pickup and delivery problems. Journal für Betriebswirtschaft , 2008), 58, 81-117.
[23] Carlsson, J., Ge, D., Subramaniam, A., Wu, A. and Ye, Y., 2009. Solving min-max multi-depot vehicle routing problem. Lectures on global optimization, 55, pp.31-46.
[24] Cao, E. and Lai, M., 2007. An improved genetic algorithm for the Vehicle Routing Problem with Simultaneous Delivery and Pick-up Service. In Proceedings of the 6th Wuhan International Conference on E-Business (pp. 2100-2106).
[25] Martinovic, G., Aleksi, I. & Baumgartner, A. Single-commodity vehicle routing problem with pickup and delivery service. Mathematical Problems in Engineering, 2008, 1–17.
[26] Xuping, W., Xu, W., Zheng, W. & Xianpei, H. A Model and an Improved Genetic Algorithm for the Vehicle Routing Problem with Break-Down Vehicles. Innovative Computing, Information and Control (ICICIC), 2009, Fourth International Conference on.
[27] Kok A.L, E.W. Hans and J.M.J. Schutten Vehicle routing under time-dependent travel times: the impact of congestion avoidance. Operational Methods for Production and Logistics, University of Twente, Enschede, Netherland, 2010.
[28] M.A. Mohammed, M.K.A. Ghani, R.I. Hamed, M.K. Abdullah, D.A. Ibrahim, Automatic segmentation and automatic seed point selection of nasopharyngeal carcinoma from microscopy images using region growing based approach, journal of computational science (2017), http://dx.doi.org/10.1016/j.jocs.2017.03.009.
[29] Arunkumar, N., K. Ram Kumar, and V. Venkataraman. "Automatic Detection of Epileptic Seizures Using New Entropy Measures." Journal of Medical Imaging and Health Informatics 6.3 (2016): 724-730.
[30] Arunkumar, N., K. Ram Kumar, and V. Venkataraman. "Automatic Detection of Epileptic Seizures Using Permutation Entropy, Tsallis Entropy and Kolmogorov Complexity." Journal of Medical Imaging and Health Informatics 6.2 (2016): 526-531.
[31] Desrochers, M., Lenstra, J. K., & Savelsbergh, M. W. P. A classification scheme for vehicle routing and scheduling problems. European Journal of Operational Research, 1990, 46, 322–332.
[32] Barbucha, D. & Jedrzejowicz, P. (2007). An agent-based approach to vehicle routing problem. WASET, 26, 36–41.
[33] Singh, A., Deep, K. and Grover, P., 2017. A novel approach to accelerate calibration process of a k-nearest neighbours classifier using GPU. Journal of Parallel and Distributed Computing.
[34] Medina, J.L.V., Boqué, R. and Ferré, J., 2009. Bagged k-nearest neighbours classification with uncertainty in the variables. Analytica chimica acta, 646(1), pp.62-68.
[35] Mohammed, M.A., Al-Khateeb, B. and Ibrahim, D.A., 2016. Case based Reasoning Shell Framework as Decision Support Tool. Indian Journal of Science and Technology, 9(42).
[36] Eksioglu, B., Vural, A. V. & Reisman, A. (2009). The vehicle routing problem: A taxonomic review. Computers & Industrial Engineering, 57, 1472-1483.
[37] Mohammed, M.A., 2015. Design and Implementing an Efficient Expert Assistance System for car evaluation via fuzzy logic controller. International Journal of Computer Science and Software Engineering (IJCSSE), 4(3), pp.60-8.
[38] Suyanto, S., Hartati, S., Harjoko, A. and Van Compernolle, D., 2016. Indonesian syllabification using a pseudo nearest neighbour rule and phonotactic knowledge. Speech Communication, 85, pp.109-118.
[39] Peng, X., Cai, Y., Li, Q. and Wang, K., 2017. Control rod position reconstruction based on K-Nearest Neighbor Method. Annals of Nuclear Energy, 102, pp.231-235.
[40] Zhang, X., Li, Y., Kotagiri, R., Wu, L., Tari, Z. and Cheriet, M., 2017. KRNN: k Rare-class Nearest Neighbour classification. Pattern Recognition, 62, pp.33-44.
[41] Pan, Z., Wang, Y. and Ku, W., 2017. A new k-harmonic nearest neighbor classifier based on the multi-local means. Expert Systems with Applications, 67, pp.115-125.

[42] Xu, R., Morozov, K., Yang, Y., Zhou, J. and Takagi, T., 2016. Efficient outsourcing of secure k-nearest neighbour query over encrypted database. Computers & Security.
[43] Alfailakawi, M.G., Ahmad, I. and Hamdan, S., 2016. Harmony-search algorithm for 2D nearest neighbor quantum circuits realization. Expert Systems with Applications, 61, pp.16-27.

Author Biographies

Mazin Abed Mohammed

Email: mazin_top_86@yahoo.com

Phone: +60183801767

Office: Main Campus, FTMK Building, 2nd Floor

PhD Candidate at Biomedical Computing and Engineering Technologies (BIOCORE) Applied Research Group, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia**.** received his B.Sc. in Computer Science from College of Computer, University of Anbar, Iraq in 2008. He obtained his M.Sc. in Information Technology from College of Graduate Studies, Universiti Tenaga Nasional (UNITEN), Malaysia, in 2011.Currently PhD Student of Information & Communications Technology, University Teknikal Malaysia Melaka (UTeM), His current research interests include Artificial Intelligence, Biomedical Computing, Multimedia Applications and Optimization Methods.

Prof. Dr. Mohd. Khanapi Abd. Ghani

Email: khanapi@utem.edu.my

Phone: +606 3316677

Office: Main Campus, FTMK Building, 2nd Floor

PhD in Biomedical Computing from Coventry University, U.K. Masters in Software Engineering from Malaysia University of Technology (UTM), BSc (Hons) in Computer Science from Malaysia University of Technology (UTM) and Diploma in Computer Science from Mara Institute of Technology. His research areas of interest include electronic healthcare systems, telemedicine, healthcare knowledge management, system architecture and software reuse.

Assoc. Prof. Dr. Raed Ibraheem Hamed

Email: raed.alfalahy@uhd.edu.iq

Phone: +9647829377275

Office: Main Campus, University of Human Development, Sulaymaniyah KRG, Iraq

Dr. Raed Ibraheem Hamed is an associate professor (faculty staff member) Department of IT, College of Science and Technology, University of Human Development, Sulaymaniyah KRG, Iraq. He obtained his M.Sc. in Information Technology from University of Technology, Baghdad Iraq in 2004 and his Ph.D. in Computer Science from the University of Jamia Millia Islamia, New-Delhi-110025, India, in 2011. He carried out Doctoral Degree on Computer Science at the University of JMI. His personal research interests include Bioinformatics and computational technology, Petri Nets Modelling and Simulating, » Databases and Data Mining, Fuzzy Logic.

Dr. Salama A. Mostafa

Email: semnah@yahoo.com

Phone: +60 11-2323 1938

Office: Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn, Johor, Malaysia

Dr. Salama A. Mostafa obtained his BSc. in Computer Science from University of Mosul, Iraq in 2003. He obtained his MSc. in Information Technology from Universiti Tenaga Nasional (UNITEN), Malaysia in 2011. He obtained his PhD. in Information and

Communication Technology from the College of Graduate Studies, UNITEN, Malaysia in 2016. He has produced more than 35 articles in journals, conferences, and tutorials. His research interests are in the area of software engineering, artificial intelligence and their integration including software agents and intelligent autonomous systems.

Mr. Dheyaa Ahmed Ibrahim

Email: dheyaa.ibrahim88@gmail.com

Phone: +9647812883275

Office: Main Campus, University of Anbar, Iraq

received his B.Sc. in Computer Science from College of Computer, University of Anbar, Iraq in 2009. He obtained his M.Sc. in Computer Science from College of Computer, University of Anbar, in 2012. His current research interests include Artificial Intelligence, Biomedical Computing, medical image processing and Optimization Methods.

Mr. Humam Khaled Jameel

Email: kareskh91@gmail.com

Phone: +964781277457

Office: Main Campus, University of Anbar, Iraq

Mr. Humam Khaled Jameel, Bachelor of Information Systems in (Artificial Intelligence), College of computer science and information Technology, University of Anbar, Anbar, Iraq, 2015. His research areas of interest include electronic Framework, Artificial Intelligence, Intelligent systems.

Mr. Ahmed Hamed Alallah

Email: hemedahmed24@gmail.com

Phone: +9647701485796

Office: Main Campus, University of Anbar, Iraq

Mr. Ahmed Hamed Alallah, Bachelor of Information Systems in (Artificial Intelligence), College of computer science and information Technology, University of Anbar, Anbar, Iraq, 2015. His research areas of interest include Artificial Intelligence, Mobile Computing, HCI, Medical Image.

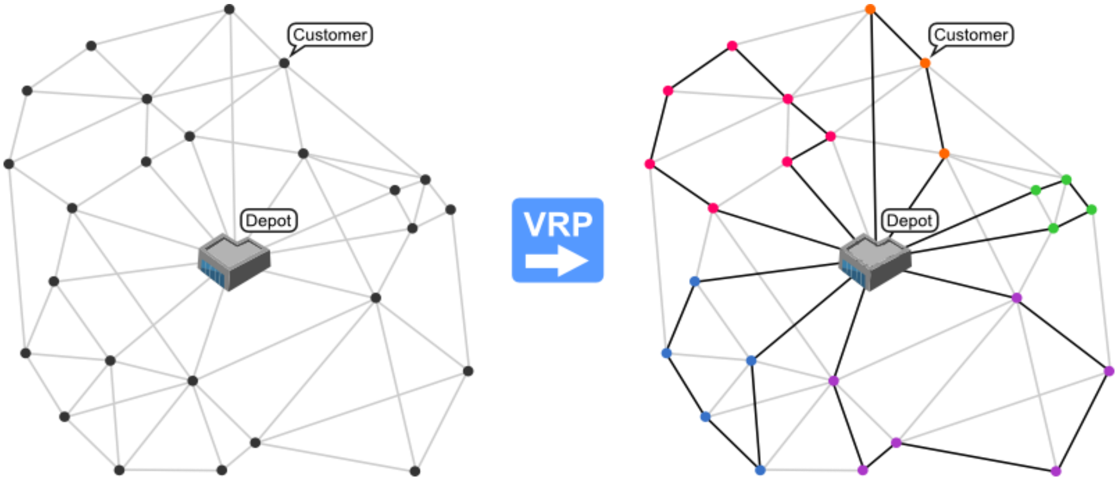

Fig 1 An example of a VRP (left) and its result (right)

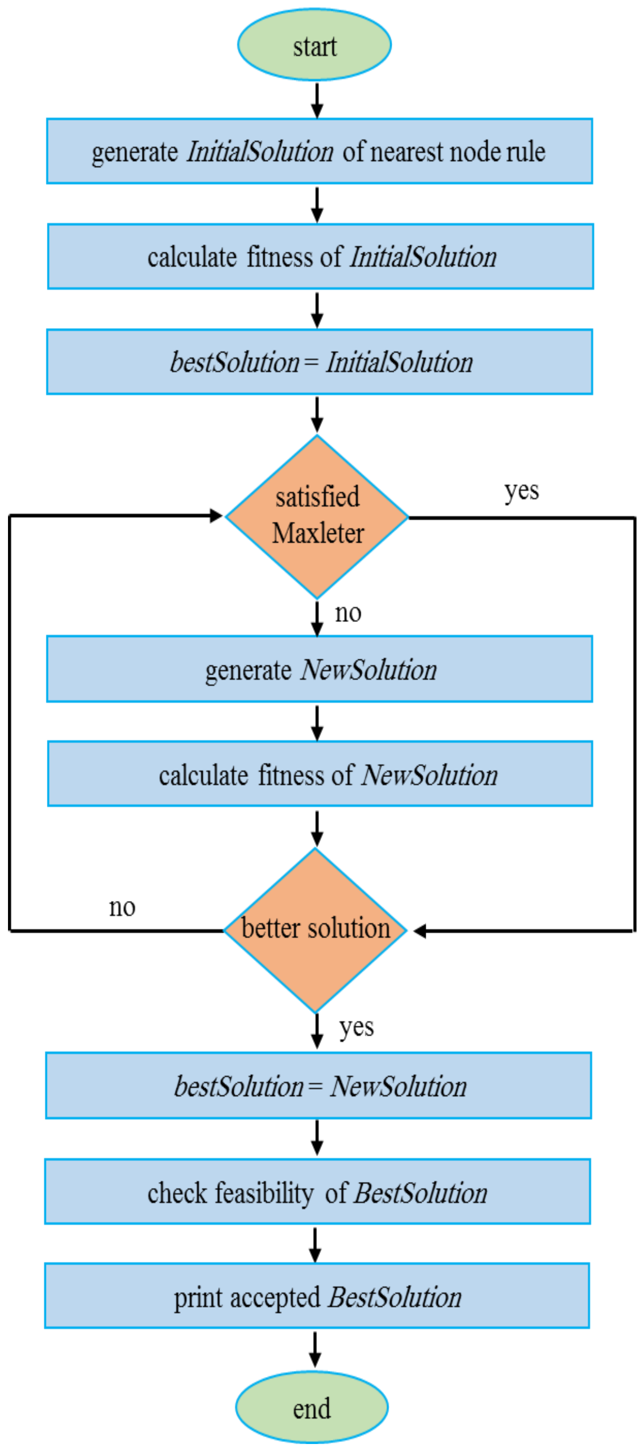

Fig 2 The nearest neighbor algorithm

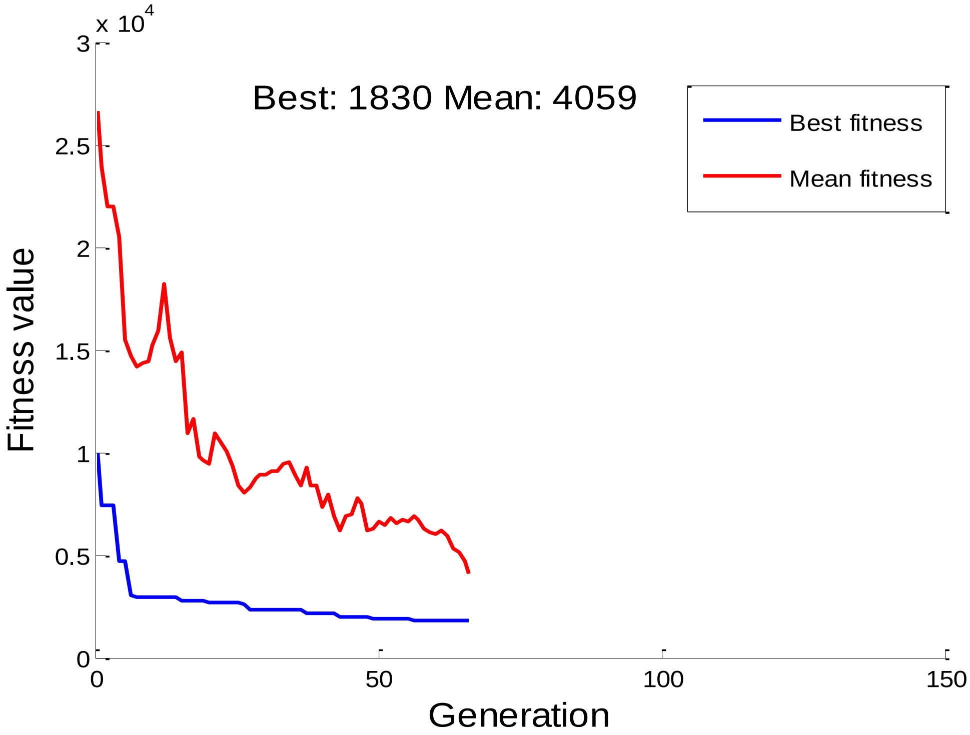

Figure 3: Test 1 Performance

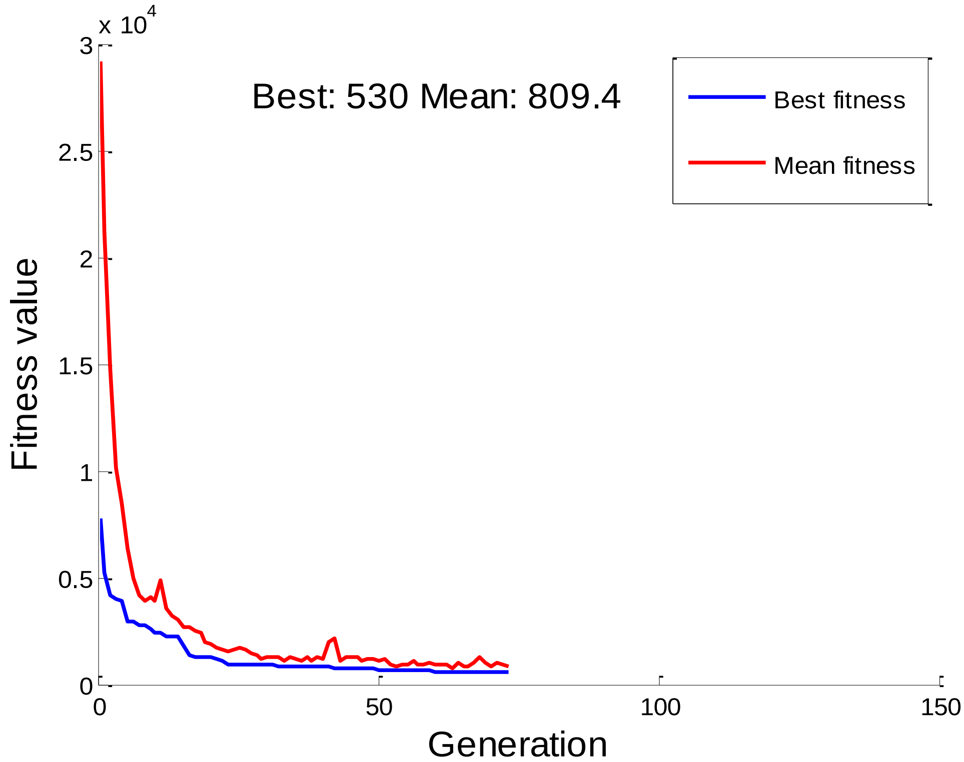

Figure 4**:** Test 2 Performance