Advanced Intelligent Technique of Real Genetic Algorithm for Traveling Salesman Problem Optimization

A.R. AWAD[1], I. VON POSER[2] , M.T. ABOUL-ELA[3]
[1]Department of Environmental Engineering, Tishreen University, Lattakia, Syria [2]Ingenierurtechnik,
[2]Merck KGaA., Darmstadt, Germany
[3]Civil Engineering Department, Minia University, Egypt

Abstract:-This work aims at solving the Traveling Salesman Problem (TSP) through developing an advanced intelligent technique based on real genetic algorithm (GA). The used GA comprises real-value coding with specific behavior taking each code as it is (whether binary, integer, or real), rank selection, and efficient uniform genetic operators. The results indicated, in comparison with the other applied optimization methods (linear, dynamic, Monte Carlo and heuristic search methods), that the real GA produces significantly the lowest distance (least cost tour) solution. It is concluded that real GA approach is robust and it represents an efficient search method and is easily applied to nonlinear and complex problems of the TSP in the field of solid waste routing system in the large cities.

Keywords:- Intelligent Technique, real genetic algorithms, optimization, traveling salesman problem, large-scale example.

1 Introduction

The problem of solid waste and its management is complex in small towns and critical in large metropolitan areas [1]. The routing is one of the main components of solid waste management in the cities where the collection takes 85% of the solid waste system cost and only 15% for disposal [2]. The traveling salesman problem (TSP) is one of the most widely studied and most often cited problems in operations research. For over fifty years the study of the TSP has led to improve solution methods for wide range of practical problems [3]. Those studies based mainly on several modeling approaches like linear, dynamic programming techniques and heuristic techniques [4, 5, 6]. Genetic algorithms have also been used successfully to solve this NP-problem. However, in many of such works, a little effort was made to handle the nonlinear optimization problem of routing solid waste collection.

In general, the TSP has attracted a great deal of attention because it is simple to state but difficult to solve [7]. The exhaustive algorithms for solving the TSP or node route problem are rarely very good in any sense [8]. They perform well for $n \leq 6$ and very badly for $n \geq 15$; it is time consuming and needs huge storage and memory time. It has been reported by Coney (1988) that a 21-location tour would require 77,100 years of computer time on a million operation per second (PC) [9]. Mathematical programming approaches have had rather limited success with this problem. According to Thieraut and Klekamp (1975) for a 20-node problem, integer linear programming requires 8000 variables and 440 constraints while dynamic programming is limited to 13-node problems [10].

The aim of this paper is to find the optimal routing for solid waste collecting in cities, taking Irbid City in Jordan as an example problem, through developing an advanced intelligent technique based on real genetic algorithm (GA). In addition, a comparison has been made with other optimization techniques such as linear, dynamic, Monte Carlo simulation, heuristic algorithms.

2 Encoding Schemes of GA

There are many variations of GAs but the following general description encompasses most of the important features. The analogy with nature requires creation within a computer of a set of solutions called a population. Each individual in a population is represented by a set of parameter values that completely propose a solution. These are encoded into chromosomes, which are originally sets of character strings, analogous to the chromosome found in DNA.

The GA search, sometimes with modification, has proved to perform efficiently in a large number of applications. This efficiency lies in the robustness of the search method that underlies the

GA approach and in the flexibility of the formulation itself. In contrast to traditional optimization methods that track only a single pathway to the optimal solution, genetic algorithms cover a whole population of possible solutions. The selection of an appropriate chromosome representation of candidate solutions to the problem at hand is the foundation for the application of genetic algorithms to a specific problem. With a sufficiently large population of chromosomes, adequate representation will be achieved. In real-value coding, variables that can take on continuous values are represented as a real (i.e. continuous) variable in the string. In this case, the string consists of a series of real values.

Real-value chromosomes have been used successfully in multireservoir system operation by various authors [11, 12], and in pipe networks optimization by the authors [13]. In this paper a new GA based methodology for optimal solving the TSP has been applied. This methodology takes by real representation with any code (binary, integer, or real) without any need to transfer from one to another, i.e. our GA, using chromosome as unit, works equally well with integer and non-integer decision variables. This is a specific behavior distinguishes our real-coding GA from other (binary, Gray, or integer) coding GAs. Thus the encoding scheme used herein proved to have the property of matching the effective problem search space towards the GA search space. But here in our example the nodes are taken as integer quantities working with one Chromosome which comprises all nodes encoded as A, B, C and so on (Table 1). Nevertheless, this choice does not limit our GA.

3 Example: A Real Problem of Solid Waste Routing

The use of GA to solve a traveling salesman problem will be illustrated by using real representation to solve the solid waste network studied initially in Reference [14].

Irbid is divided into six regions which are considered as separate solid waste generation areas. Each of these regions has its own department which regulates the solid waste services in the region. There is no specific routing basis for the vehicles being left to the driver's choice. Occasionally, one pick-up point may be missed. In regions which have two collection vehicles, they may meet at the same pick-up point several times. Once the solid waste is loaded into the vehicles, it is carried out of Irbid to the disposal site located far away.

The suggested procedure for solving the vehicle routing problem in the selected Region 2 of Irbid begins with a particular node closest to the garage or the previous region and ends with the nodes closest to the disposal site. This reduces the number of permutations considerably. Further detailed procedures can be found in Reference [15]. There are thirty-one pick-up points (nodes) in this Region 2 with about fifty containers distributed on it. Two major streets pass through the area of this region and divide the total area into three sectors, each having its own nodes and its own network, and each network has its distance matrix (Figure 1). Network I has 9 nodes (pick-up points), network II has 7 and finally network III has 15 nodes. The network III was used as a case study. A set of 15 nodes with 1 or 2 waste containers at each node were in service by vehicles. The collecting vehicles are equipped with compactors and have to collect the contents of about thirty full containers. The distances between nodes in Network III are shown in Figure 2. The routing problem in this part of Irbid is a node routing or traveling salesman problem (TSP).

Table 1: Coding of the network

A	1
B	2
C	3
D	4
E	5
F	6
G	7
H	8
I	9
J	10
K	11
L	12
M	13
N	14
O	15
P	16

Node No. 16 is added to indicate the ending point.

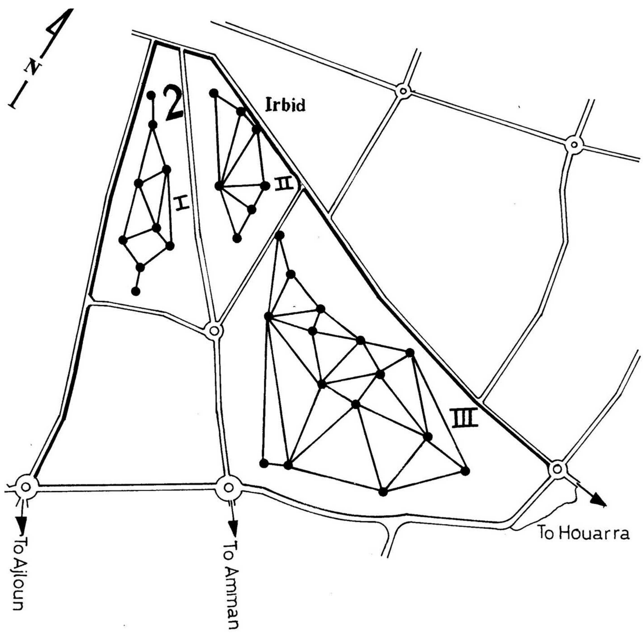

Figure 1: The constructed networks in region two

Our problem has been the construction of a tour through n points with the minimum distance, where the vehicle does not have to return to the starting point, i.e. the number of possible tours is (n-2)! . The application based on the fact that the shortest solid waste collection tour should begin with the first node No.1 (closest to the municipality garage), pass by all nodes once and only once, and end with the node No. 15 (closest to the disposal site out of the city).

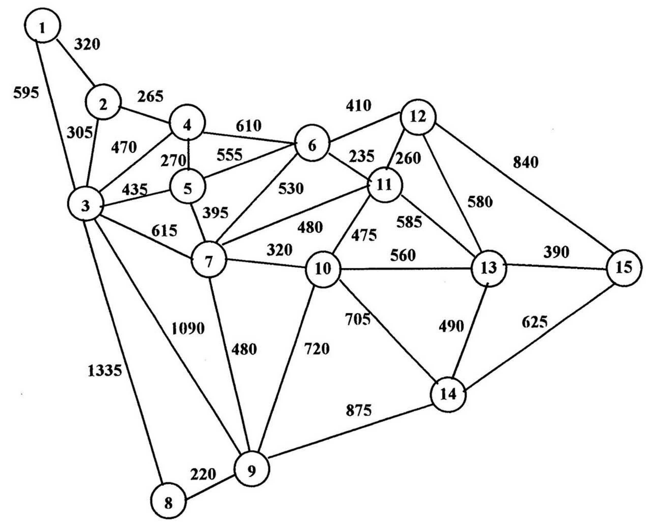

Figure2: Network (III) of region two

4 GA Optimization

The following procedures are required for the formulation of GA:

4.1 Coding

The genetic algorithm requires that the decision variables describing trial solutions to the solid waste routing problem be represented by a unique coded string of nodes (pick-up points). This coded string is similar to the structure of a chromosome of genetic code. As for the example, there are 15 decision variables to be made about the network. Each of these decision variables can take one node (pick-up point). A real string made of 15 substrings is used for representing the problem into a suitable form for use within a GA (Table 1). This string (chromosome) of 15 genes represents a route design for the network consisting of 15 nodes.

It is worthy to mention that the node (No. 16 or P) has not been coded but was used (as pseudo node) to indicate the ending point of the studied network.

4.2 Fitness

The fitness of a coded string representing a solution for the traveling salesman problem of solid waste routing is determined by the shortest cycle/cost provided that the collection vehicle passes through all nodes just for once on each.

The evaluation or objective function used is rather simple and determinant of the distance of a routing solution by summing the lengths of the nodes distances making up the network (Table2). The value of 10000m is a hypothetical value points to the node itself, but actually such a case never occurs. In the used coding scheme also infeasible solution can appear. In the distance matrix every infeasible route selection such as "C to J" will be penalized with a high distance of 5000m. An example for fitness evaluation is shown in Figure 3.

Table 2: Distance matrix among the nodes

	stri 0	inte 1	inte 2	inte 3	inte 4	inte 5	inte 6	inte 7	inte 8	inte 9	inte 10	inte 11	inte 12	inte 13	inte 14	inte 15	integer 16
stri		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	A	000	320	595	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	
2	B	320	10000	305	265	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	
3	C	595	305	10000	470	435	5000	615	1335	1090	5000	5000	5000	5000	5000	5000	
4	D	5000	265	470	10000	270	610	5000	5000	5000	5000	5000	5000	5000	5000	5000	
5	E	5000	5000	435	270	10000	555	395	5000	5000	5000	5000	5000	5000	5000	5000	
6	F	5000	5000	5000	610	555	10000	530	5000	5000	5000	235	410	5000	5000	5000	
7	G	5000	5000	615	5000	395	530	10000	5000	480	320	480	5000	5000	5000	5000	
8	H	5000	5000	1335	5000	5000	5000	5000	10000	220	5000	5000	5000	5000	5000	5000	
9	I	5000	5000	1090	5000	5000	5000	480	220	10000	720	5000	5000	5000	875	5000	
10	J	5000	5000	5000	5000	5000	5000	320	5000	720	10000	475	5000	560	705	5000	
11	K	5000	5000	5000	5000	5000	235	480	5000	5000	475	10000	260	585	5000	5000	
12	L	5000	5000	5000	5000	5000	410	5000	5000	5000	5000	260	10000	580	5000	5000	
13	M	5000	5000	5000	5000	5000	5000	5000	5000	5000	560	585	580	10000	490	5000	
14	N	5000	5000	5000	5000	5000	5000	5000	5000	875	705	5000	5000	490	10000	5000	
15	O	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	840	390	625	10000	
16	P	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	10	

Infeasible solutions, failed to meet the aforementioned node requirements, are not removed from the population. Instead, they are allowed to join the population and help guide the search, but for a certain price. Although this technique does not exclude infeasible solutions from the population, it reduces the probability of infeasible solutions remaining in future generations and permits convergence from all regions of the parameter space.

4.3 Selection, Crossover and Mutation

Detailed explanation and analysis of the selection, crossover and mutation operators for reproduction process using real representation with examples, features and processes can be seen in Reference [14]. The foregoing procedures of the GA formulations are shown in Figure 4. The used parameters for implementing our GA technique in the real example are the population size of 40 with probabilities of 0.8 for crossover and 0.4 for mutation. The optimization was carried out over 80 generations.

Furthermore, it is worthy to mention here, that an important aspect of crossover operator in application to a multivariate problem in real-representation comes from the fact that the gene comprises a single allele and it is itself the parameter value; while in the binary or Gray coding the crossover occurs only at the boundaries of the gene which consists of alleles or bits, thus exposing the gene to split. Therefore, what distinguishes our GA work is that the crossover comes only in between the genes, thus avoiding destroying any of them, so the coded information is not destroyed; contrary to what happens with most GAs based on binary coding and others.

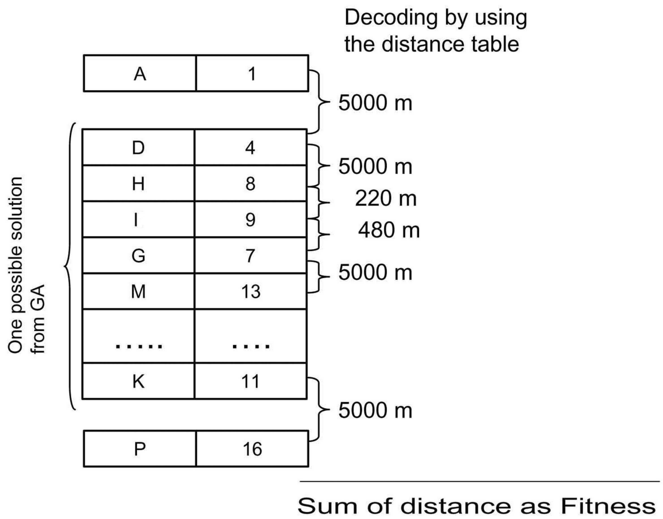

Figure 3: Example for fitness evaluation

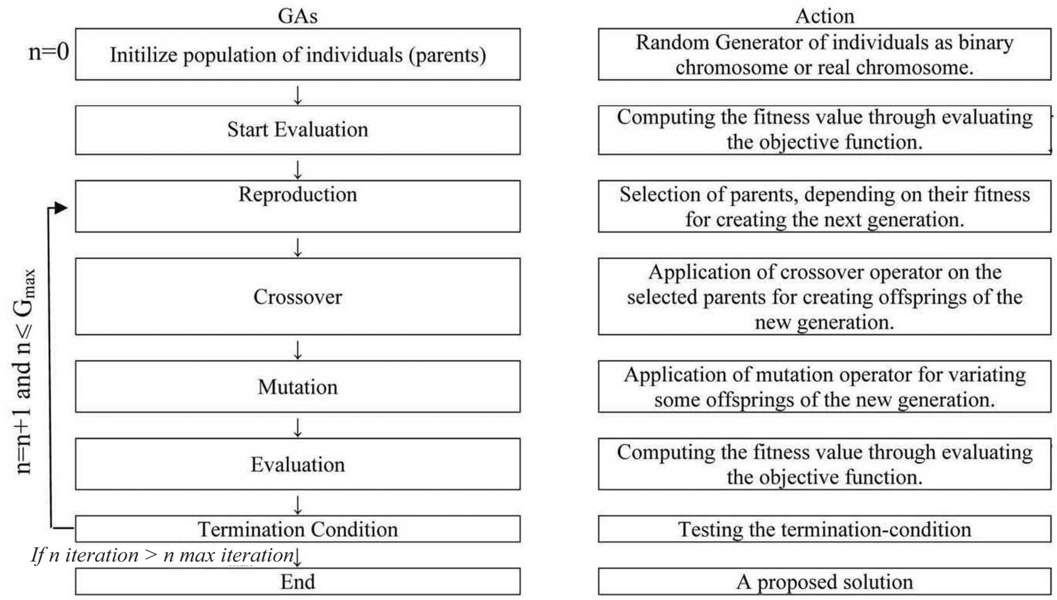

Figure 4: Flow-chart for the GA implementation together with the related actions

5 Results

The present work revealed by using linear and dynamic programming techniques, that the mathematical modeling has limited success with the TSP, with limited number of nodes and with the need to a large number of constraints; and that for a case like ours it takes much longer time. These results conform with the other reported results in the literature that mathematical programming approaches have had rather limited success with TSP.

By using GA, a computer program of three operators has been developed. The GA deals with a real-value chromosome comprising the 15 genes representing the nodes (route) network to be optimized. The fitness of the chromosome is computed through the objective or evaluation function, which determines the cost (distance) solution by summing the lengths of the node distances making up the network.

The GA searches for the minimum length of the solid waste routing, thus the objective function is supposed to be minimized. The total search space is 6.2×10^9 possible network routes. The reached results are presented in Figure 5, which shows a typical plot of the cost (distance) of the solution in each generation.

These results are presented as worst fitness, average fitness and best fitness to differentiate among them, and to show the optimality of the GA operators program. The GA found the best solution which led to the shortest – distance (lowest cost tour) of 6585m (6595-10) with 2060 simulations (evaluations) for collection vehicles routing in Irbid. Figure 6 shows the final route of this best solution. The computing time did not exceed 35 seconds CPU time on a PC (AMD 2.4 GHZ), while in the mathematical methods it takes much longer time and in the heuristic algorithms and their modification and Monte Carlo simulations it takes about 5 minutes.

The quality of the network routes solutions reached and the robustness of the method can be gauged by comparing with the other results from literature presented by the authors [15] who applied their own research work on the same example for modeling techniques of Monte Carlo Simulation and Heuristic Algorithms which resulted that the shortest tour was by Monte Carlo Simulation 6715 m (one million random trials) and by modified heuristic algorithm was 7945m.

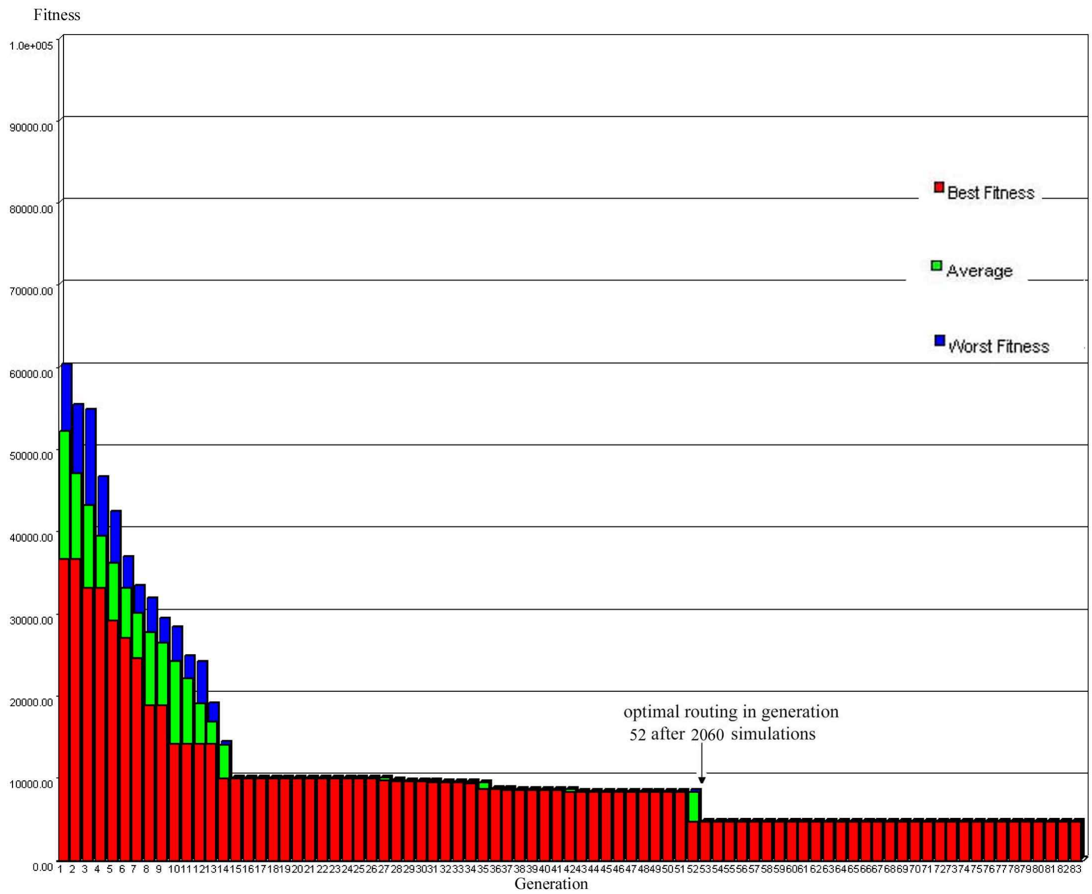

Figure 5: Worst, average and best of generation distances
for solid waste routing in the studied network of TSP

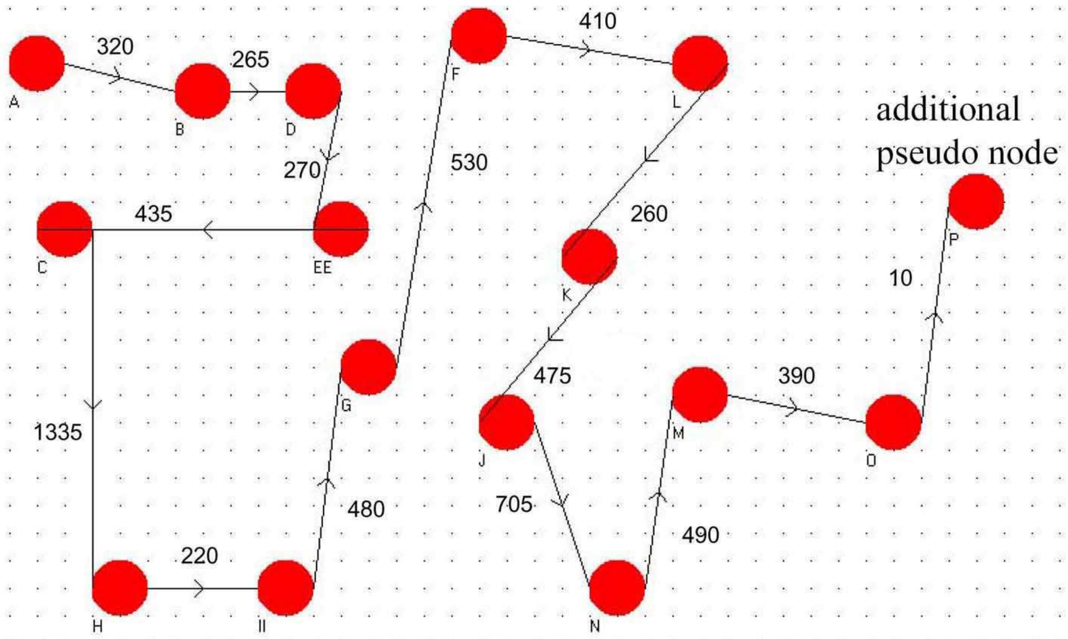

Figure 6: Shortest distance routing for solid waste collection
in the studied network of TSP

6 Conclusion

The results of our study, in comparison with the other applied optimization methods (linear, dynamic, Monte Carlo and heuristic search method), indicate that the real GA, through its specific behavior and through its efficient operators, presents significantly the lowest distance (cost tour) solution. Accordingly, it is concluded that real GA approach is robust and it represents an efficient search method and is easily applied to nonlinear and complex problems of the TSP in the field of solid waste routing system in the large cities. Furthermore, it has certain advantages over other applied optimization techniques such as linear, dynamic, Monte Carlo simulation and heuristic search methods in both cost and computer time.

Addition of other real constraints such as pick-up points (nodes) or extension of the studied waste collecting network for bigger serviced cities etc.., as decision variables can also be incorporated into our real GA by increase in the chromosome length by adding more genes for representing such additions in a route design for waste collecting problem in cities, i.e., more generations would be required to reach the optimum.

It is worthy to remark at the end that our (GA – TSP) program is not complicated to apply, since it doesn't need much mathematical sophistication for comprehending its mechanisms. So this program can be considered as a tool that gives the designer/decision maker a great number of alternative solutions for traveling salesman problem.

References

[1] Koushki AP, Hulsey JL, Bashaw EK. House hold solid waste traits and disposal site selection. J. Urban Planning & Development, ASCE 1997; 123(1): 1-9.

[2] Ludwig H, Black R. Report on the solid waste problem. J. of the Sanitary Engineering Division, ASCE 1968; 94, SA2: 355-370.

[3] Cook W. The traveling salesman problem and optimization on a grand scale. SIAM Annual Meeting, Montreal: June 18, 2003.

[4] Francis R, McGinnis L, White J. Location analysis. Eur. J. Operational Res. 1983; 12 (3): 220-252.

[5] Ramu VN, Kennedy JW. Heuristic algorithm to locate solid waste disposal site. J. Urban Planning & Development, ASCE 1994; 120(1): 14-21.

[6] Boyce ED, Lee H, Janson NB, Berka S. Dynamic route choice model of large-scale traffic network. J. Transportation Eng., ASCE 1997; 123 (4): 276-282.

[7] Rivett P, Ackoff RA. Manager's guide to operational research. New York: John Wiley and Sons Inc., 1963.

[8] Goodman S, Hedetniemi, S. Introduction to design and analysis of algorithms. New York: Mc Graw-Hill, 1985.

[9] Coney W. Traveling salesman problem solved with simulation techniques. Int. J. of Systems Science 1988; 19 (10): 2115-2122.

[10] Thieraut R, Klekamp R. Decision making through operations research. New York: John Wiley and Sons Inc., 1975.

[11] Oliveira R, Loucks DP. Operating rules for multireservoir systems. J. Water Resour. Res. 1997; 33 (4): 839 – 852.

[12] Wardlaw R. Sharif M. Evaluation of genetic algorithms for optimal reservoir system operation. J. Water Resour. Plng. and Mgmt., ASCE 1999;125(1): 25–33.

[13] Awad A, Von Poser I. Genetic algorithm optimization of water supply networks. Water Intelligence Online Journal (WIO), International Water Association (IWA) Publishing 2005; 8 (4): 1- 25.

[14] Abu-Hassan R, Routing of solid waste collection vehicles in Irbid, M.Sc. Thesis, Department of Civil Engineering, Jordan University of Science and Technology, 1988.

[15] Awad A, Aboul-Ela MT, Abu-Hassan R. Development of a simplified procedure for routing solid waste collection. International Journal for Science and Technology (Scientia Iranica) 2001; 8 (1): 71-75.