Optimal Routing for Solid Waste Collection in Cities by Using Real Genetic Algorithm

Ingo von Poser
Ingenierurtechnik, Merck KGaA.
Frankfurter Str. 250,
64293 Darmstadt, Germany
Email: Ingo.Poser@merck.de

Adel R Awad
Dept. of Environmental Engineering,
TishreenUniversity, Lattakia, Syria
00963 41 422201
Email: adel-a@scs-net.org

Abstract

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. The objective of this research paper is to find the best route for collecting solid waste in cities taking Irbid City in Jordan as an example problem. The routing problem in this example is a node routing or traveling salesman problem (TSP). This work has developed a methodology based on real genetic algorithm for effective solving the TSP. This GA comprises real-value coding with specific behavior taking each code as it is (whether binary, integer, real or name), 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, represents an efficient search method and is easily applied to dynamic and complex system of the well-known TSP in the field of solid waste routing system in the large cities.

1. Introduction

The increasing trend towards urbanization and population growth, combined with a growing environmental concern have created a critical situation for the management of house hold solid waste. The problem of solid waste and its management is complex in small towns and critical in large metropolitan areas [1]. The management of solid waste has become an important issue worldwide. Solid waste management is concerned with the control of generation, storage, collection, transportation, processing and disposal of waste according to the principles of public health, economic and other environmental consideration. 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 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 instruction per second computer [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].

Heuristic algorithms for solving the TSP are usually very fast, not complicated, and not limited to a definite number of nodes. They are based on hill-climbing theory. A work of Awad et al. (2001) has applied this method and also with its modification to find the minimum-cost tour for a 15-node problem of solid waste routing that gave good solution and in a very short programming time [11].

The objective of this research paper is to find the optimal routing for solid waste collecting in cities taking Irbid City in Jordan as an example problem, by using 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. Representation Schemes of GA

The theory behind GAs was proposed by Holland (1975) and further developed by Goldberg (1989) and

0-7803-9521-2/06/$20.00

others in the 1980s and 1990s [12,13]. 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 solution 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 [14].

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. Different representation schemes might lead to different performances depending on accuracy and computation time. As stated by Goldberg (1989) and Davis (1991), the best encoding is problem specific and may require some experimentation and modification of the crossover and mutation operators [13,15].

In a real-value coding, individual genes of a chromosome have initially allocated values randomly within the feasible limits of the variable represented. 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. The real-value coding operates significantly faster than binary coding and produces in general better results.

Real-value chromosomes have been used successfully in multireservoir system operation by various authors [16,17], and in pipe networks optimization by Awad and Von Poser (2005) [18].

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, real, name) whether individually or combined, without any need to change from one to another, i.e. our GA 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. In our example we work with one Chromosome which comprises all nodes coded as A, B, C and so on. (Table 1).

Table 1. Coding of the network

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

Node No. 1 is not coded being the starting point.

Node No. 16 is not coded but added to indicate the ending point.

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 by [19].

The collection operation in Irbid City begins when workmen with handcarts collect plastic bags containing residential solid waste. These bags are carried to the nearest pick-up point where there are steel containers of 1100 liter capacity. The containers are unloaded into special compact vehicles. 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. Every morning the vehicles are driven from the garage to the regions, where they begin to collect residential solid waste from the pick-up points. 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 node networks, description and procedures can be found in [11]. 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. Network I has 9 nodes (pick-up points), network II, 7 and finally network III, 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 1.

The routing problem in this part of Irbid is a node routing or traveling salesman problem (TSP). Our problem has been the construction of a tour through n points with a minimum distance.

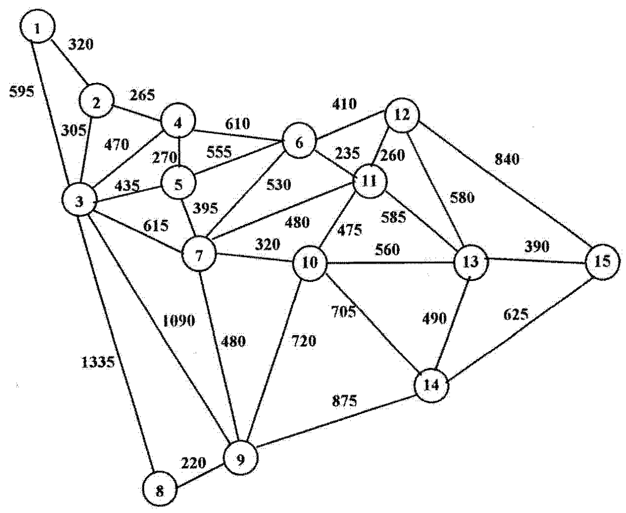

Figure 1. Network (III) of region two in Irbid city

4. Previous Studies

Previous to this case study the authors [11] applied their own research work for modeling techniques of Monte Carlo Simulation and Heuristic Algorithms which resulted that the shortest tour was by Monte Carlo Simulation 6715 m and by modified heuristic algorithm was 7945m. 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).

5. GA Optimization

The following procedures are required for the formulation of GA.

5.1. 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 14 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 14 substrings is used for representing the problem into a suitable form for use within a GA (Table 1). This string (chromosome) of 14 genes represent a route design for the network consisting of 15 nodes.

It is worthy to mention that the first node (No.1) has not been coded being the starting point. Also node (No. 16 or P) has not been coded but was used (as pseudo node) to indicate the ending point of the studied network.

5.2. 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 points to the node itself, but this case never occurs.

In the used coding scheme also infeasible solution can appear. In the distance matrix every infeasible route selection e.g. “C to J” will be penalized with a high distance of 5000m. An example for fitness evaluation is shown in Figure 2.

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.

5.3. Reproduction, Crossover and Mutation

Detailed explanation and analysis of the reproduction, crossover and mutation operators using real representation with examples, features and processes, and the foregoing procedures of the GA formulations as well can be seen in [18]. 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 herein, 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 should occur only at gene boundaries, because each gene consists of alleles, or bits, and crossover may split the genes. Regarding allele problem, in using real representation, the influence of how many bits have to be changed until the fitness get better is not touched by this representation. Therefore, what distinguishes our GA work, is

0-7803-9521-2/06/$20.00 ©2006 IEEE.

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	EE	F	G	H	II	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	EE	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	II	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	

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. Accordingly, it can be concluded that in real representation each gene keeps fully its structure, property and efficiency; while in binary or Gray representation the crossover operator works independent from the gene length.

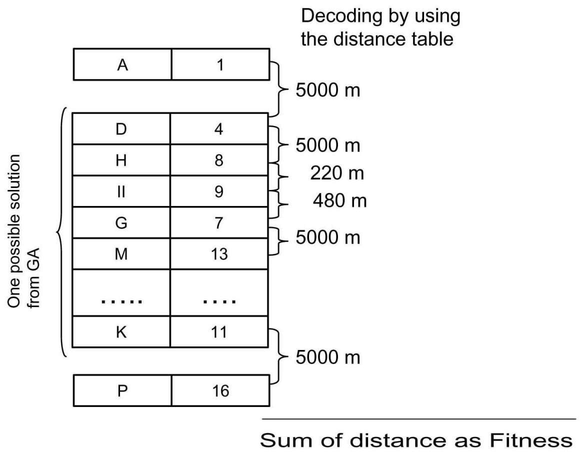

Figure 2. Example for fitness evaluation

6. 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 for a case like ours it takes ages. 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 is developed. The GA deals with a real-value chromosome comprising the 14 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 must be minimized. The total search space is 14^{14} or approx. 1.1×10^{16} possible network routes. The reached results are presented in Figure 3, 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 a 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 4 shows the final route of this best solution. The computing time did not exceeds 35 seconds CPU time on a PC (AMD 2.4 GHZ), while in the mathematical methods it takes ages and in the heuristic algorithms and their modification and MCarlo simulations it takes about 5 minutes or more. The quality of the network routes solutions reached and therefore the robustness of the method, can be gauged by comparing with the other results from literature presented in Awad et al. (2001) [11] 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. It is obvious that real coding GA has advantages over other applied optimization techniques such as linear, dynamic, Monte Carlo simulation and heuristic search methods in both cost and computer time.

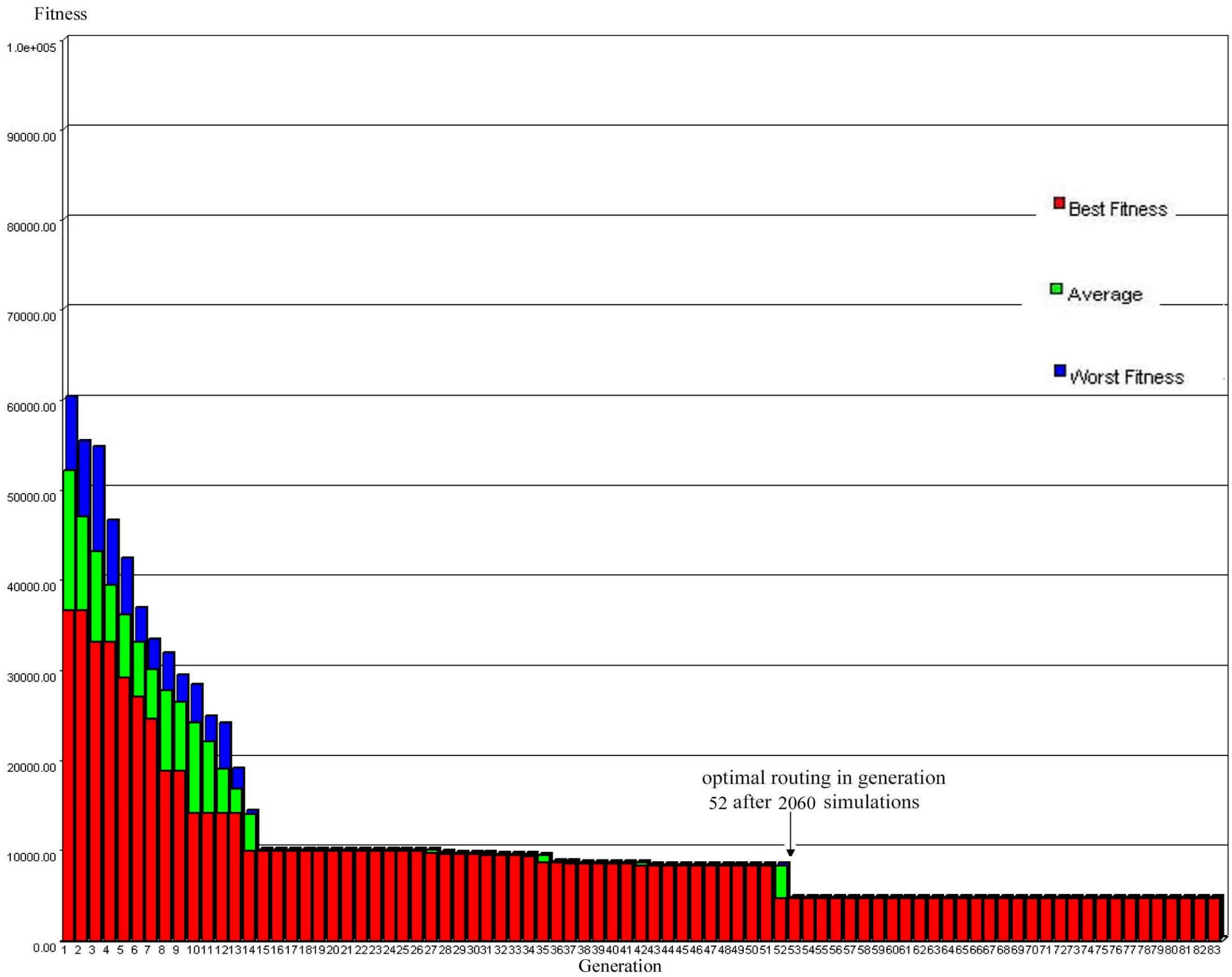

Figure 3. Worst, average and best of generation distances for solid waste routing in Irbid, Jordan

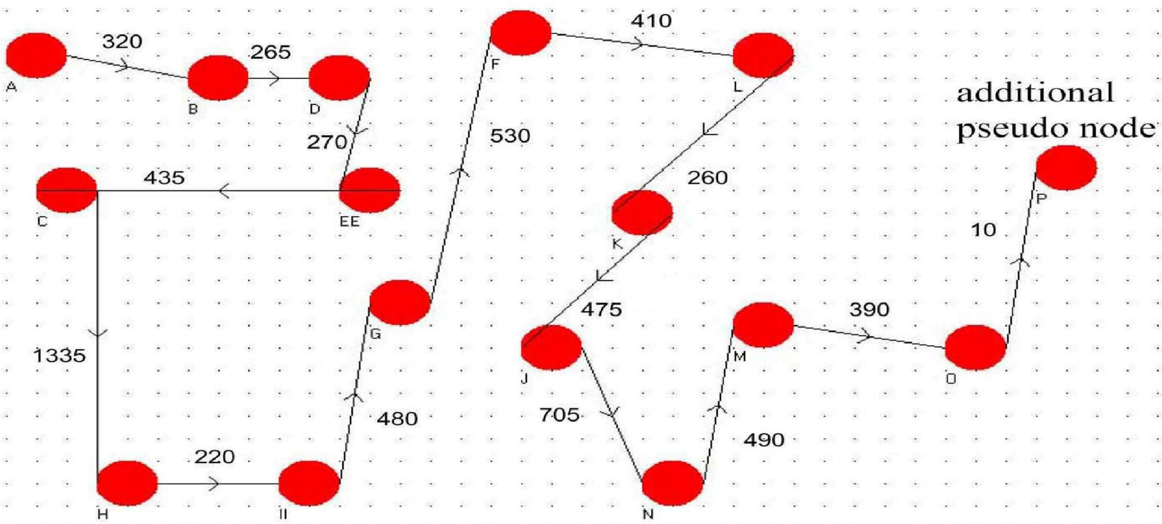

Figure 4. Shortest distance routing for collecting solid waste in the studied network of TSP

7. Conclusions

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, produces significantly the lowest distance (cost tour) solution. Accordingly, it is concluded that real GA approach is robust, represents an efficient search method and is easily applied to dynamic and complex system of the well-known TSP in the field of solid waste routing system in the large cities.

8. References

[1] A. P. Koushki, J. L. Hulsey, and E. K. Bashaw, "House Hold Solid Waste Traits and Disposal Site Selection", *J. Urban Plan.* & Devel., ASCE, 123(1), 1997, pp.1-9.

[2] H. Ludwig, and R. black, "Report on the Solid Waste Problem", *J. of the Sanitary Engineering Division*, ASCE, 94, SA2, 1968, pp. 355-370.

[3] W. Cook, "The Traveling Salesman Problem and Optimization on a Grand Scale", SIAM Annual Meeting, Montreal, June 18, 2003.

[4] R. Francis, L. McGinnis, and J. White, "Location Analysis", *Eur. J. Operational Res.*, 12 (3), 1983, pp. 220-252.

[5] V. N. Ramu, and J. W. Kennedy, "Heuristic Algorithm to Locate Solid Waste Disposal Site", *J. Urban Plan. & Devel.*, ASCE, 120(1), 1994, pp. 14-21.

[6] E. D. Boyce, H. Lee, N. B. Janson, and S. Berka, " Dynamic Route Choice Model of Large-Scale Traffic Network ", *J. Transportation Eng.*, ASCE, 123 (4), 1997, pp. 276-282.

[7] Rivett, P. and R. A. Ackoff, *Manager's Guide to Operational Research*, John Wiley and Sons Inc., New York, USA, 1963.

[8] Goodman, S. and S. Hedetniemi, *Introduction to Design and Analysis of Algorithms*, Mc Graw-Hill, NY, USA, 1985.

[9] W. Coney, "Traveling Salesman Problem Solved with Simulation Techniques", *Int. J. of Systems Science*, 19 (10), 1988, pp. 2115-2122.

[10] R. Thieraut, and R. Klekamp, *Decision Making Through Operations Research*, John Wiley and Sons Inc., NY, USA, 1975.

[11] A. Awad, M. T. Aboul-Ela, and R. Abu-Hassan, "Development of a Simplified Procedure for Routing Solid Waste Collection", *International Journal for Science and Technology (Scientia Iranica)*, 8 (1), 2001, pp. 71-75.

[12] H. J. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, Mich., 1975.

[13] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison – Wesley, Reading, Mass., 1989.

[14] D. A. Savic, and G. A. Walters, "Genetic Algorithms for Least – Cost Design of Water Distribution Networks", *J. Water Resour. Plng. and Mgmt.*, ASCE 123 (2), 1997, pp. 67 –77.

[15] L. Davis, *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York, 1991.

[16] R. Oliveira, and D. P. Loucks, "Operating Rules for Multireservoir Systems". *J. Water Resour. Res.,* 33 (4), 1997, pp. 839 – 852.

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

[18] A. Awad, and I. Von Poser, "Genetic Algorithm Optimization of Water Supply Networks", *Water Intelligence Online Journal (WIO)*, International Water Association (IWA) Publishing, 8 (4), 2005, pp.1- 25.

[19] R. Abu-Hassan, "Routing of Solid Waste Collection Vehicles in Irbid". MSc Thesis, Jordan University of Science and Technology, Jordan, 1988.