A Demand-Response Routing Model for Traffic Signal Maintenance Operations with Time-Dependent Travel Times

CHAO-HUA CHEN*, KAI-I LIAO
Department of Transportation Technology and Logistics Management
Chung Hua University
707, Section 2, WuFu Road, Hsinchu 300
TAIWAN
*erchen@chu.edu.tw

Abstract: - This study is aimed at planning traffic signal maintenance work using a demand-response routing model with time-dependent travel times. The routing model was able to produce timely maintenance routes for each maintenance vehicle compatible to the daily inspection plan and randomly revealed malfunctioning traffic signals. Three modules, the route-generating module, the decision-support module and the time-evolution module, constituted the model framework, and Genetic Algorithms (GAs) were used to solve for dynamic maintenance routes. Furthermore, different service strategies with different routing procedures based upon a case reconstructed from real world data were designed and tested to reflect the concerns for practical applications. Evaluation results show the routing model is flexible in structure and feasible enough for practical applications.

Key-Words: - Traffic Signal Maintenance Operations, Demand-Response Routing Model, Time-Dependent Travel Time, Genetic Algorithms

1 Introduction

Traffic signals play a critical role in both traffic system control and management. Even if a traffic signal system functions without breakdowns it may not necessary guarantee a smooth traffic flow. Furthermore malfunctioning traffic signals do have a significant effect on traffic flow and the derived impact tends to accumulate as time progresses, resulting in even more severe congestion if no proper response actions are made. Therefore, efficient traffic signal maintenance work is essential for preserving the smooth flow of a traffic signal system.

Traffic signal maintenance, which is complicated and dynamic in nature, has been categorized by Meyer [11] as including preventive maintenance, response maintenance and design modification. Preventive maintenance is performed on a routine basis; response maintenance is aimed at making a quick response to emergency situations as well as trouble shooting; and design modification deals with the need to monitor new equipment. Nonetheless, little attention has been put into the cost and procedure required for maintenance [11]. This can have serious consequences, especially because of the development and implementation of ever more advanced traffic signal systems. The problem becomes even more challenging in the face of limited resources and the growing number of traffic signals.

Routine inspections and repairs are usually scheduled manually based on the empirical experience of traffic operational officers. Such manually-planned schedules may be sufficient for a network with light traffic, but may not be good enough for a network with heavy traffic. Furthermore, emergency situations that occur during working hours can seriously disrupt daily maintenance plans, even preventing daily maintenance work if an immediate response must be made within a limited time frame. Attributes associated with each emergency situation are critical to maintenance operations. The travel time of the maintenance vehicle is another sensitive factor to be considered in routing plans.

Figures 1 and 2 demonstrate the frequency and ratio of different ranges of time required for the repair of malfunctioning traffic signals in Hsinchu city, Taiwan from April 2004 to March 2005. These two sets of data reveal only the aggregated monthly repair records for major traffic signals (around 200 in number) that were monitored by the city traffic control office. From the records we can see that more than half of the year there was a busy schedule of response maintenance work. If regularly scheduled preventive traffic signal maintenance works was incorporated, the schedule of signal maintenance and inspection would become even busier. The aim of this study is to develop a model that is able to resolve the aforementioned problems given the constraint of predetermined resources, such as the number of vehicles and technicians available.

The rest of the paper is organized as follows. In the next section, we define the core problem and describe the assumptions for the dynamic traffic signal maintenance problem (DTSMP). This is followed by a discussion of related works found in the literature. In the fourth section we describe the model framework and detailed routing procedures for different service strategies as well as measurement of the model objectives. In the fifth section, we discuss the evaluation results and suggestion for improving computational efficiency. The research findings and future research directions are summarized in the last section.

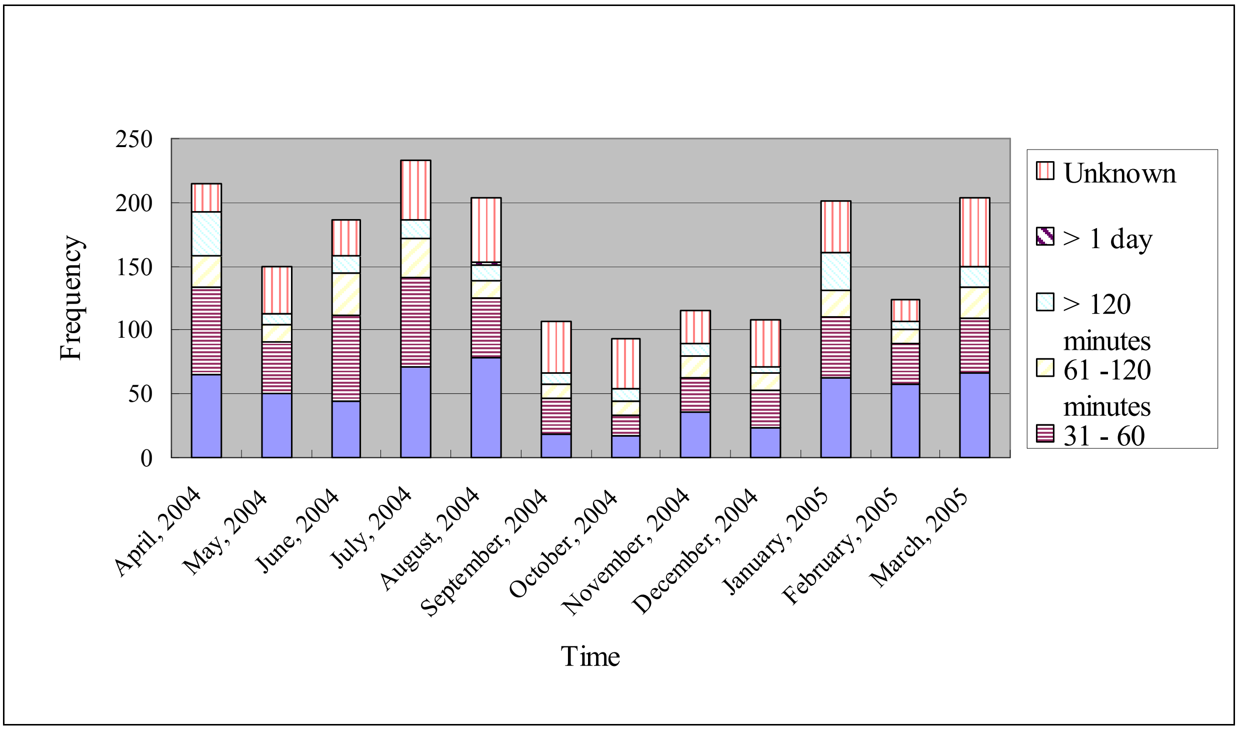

Figure 1 Frequency of Different Repair Times for Malfunctioning Traffic Signals

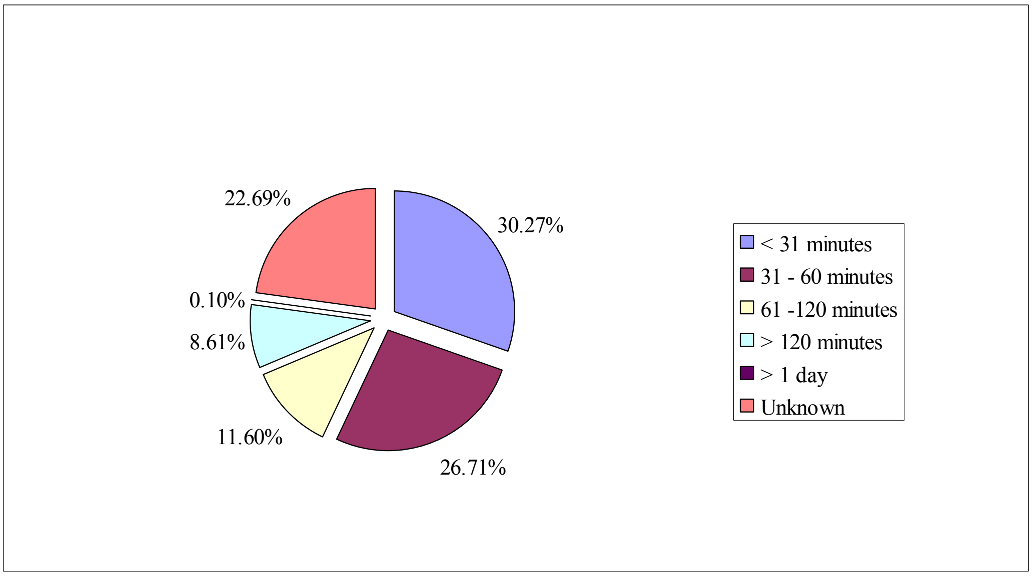

Figure 2 Ratios of Different Repair Times for Malfunctioning Traffic Signals

2 Problem Definition and Assumption

The DTSMP applied in this research is designed to treat traffic signal maintenance operations where technicians are carried in vehicles which traverse in the traffic network to carry out maintenance works. The dynamic nature of the DTSMP is a result of the nature of the service demands imbedded in the traffic network. The three aspects of this problem solution method are summarized below.

(1) Service demand: This consists of traffic signals that require inspection and/or repair.
(2) Supply: This represents the resources available to the traffic operations office, which include manpower, numbers of maintenance vehicles, quantity of spare parts, budget, working time period, service strategies, etc.
(3) Environment: This represents the background conditions such as the weather, traffic conditions, network characteristics, etc. The network characteristics include the sizes and types of traffic networks (symmetrical or asymmetrical) as is commonly considered in vehicle routing problems (VRPs).

Detailed assumptions for each aspect are summarized below. For service demand, it is assumed that (a) traffic signals are located at the nodes of a traffic network; (b) two types of traffic signals are considered, namely those that required regular maintenance and malfunctioning ones that require repairs. The latter can be revealed stochastically or authorized by higher authorities, and are termed emergency service demands; (c) any traffic signal that needs to be serviced within three working days. For supply of maintenance work, we assume that (a) there is one depot with multiple maintenance vehicles within the service area; (b) the maintenance vehicle has unlimited capacity and can traverse the traffic system as long as is required; (c) vehicles and technicians do not return to the depot until the end of an eight hour daily working time, or until all emergency service demands that have occurred in that day have been completed according to different service strategies; (d) each technician is equally capable and has the same work skills; (e) there are sufficient spare parts for traffic signal repair and they are always available for scheduled work; (f) the maintenance and repair times are pre-set and fixed based on historical records; and (g) the technician always complies with the order issued from the traffic operations office. In addition, we do not consider extreme weather conditions like snow storms, tropical storms or heavy rain. It is assumed that normal weather conditions do not influence the maintenance work. The traffic network taken into account in this study is asymmetric and composed of links and nodes, and vehicle travel time is time-dependent.

3 Literature Review

Given the dynamic nature of the attributes associated with real-time malfunctioning traffic signals such as the time-point at which the malfunction occurs, the quantity and locations of the DTSMP, the studied problem can be considered to be closer to a dynamic VRP (DVRP) than a static VRP. This view-point is confirmed by the research results of Psaraftis [13,14] and Ghiani et al [7]. Psaraftis [13] outlined twelve features with which to define a DVRP, namely (1) time dimension is essential; (2) problem may be open-ended; (3) future information may be imprecise or unknown; (4) near-term events are more important; (5) information updating mechanisms are essential; (6) re-sequencing and reassignment decisions may be warranted; (7) faster computation times are necessary; (8) indefinite deferment mechanisms are essential; (9) objective functions may be different; (10) time constraints may be different; (11) flexibility varies if vehicle fleet size is lower; and (12) queuing considerations may become important. According to Ghiani et al [7] the input data for a DVRP are explicitly time dependent. These results indicate that one requirement needed for a DVRP is dynamic (i.e. real-time or on-line) input information, which also happens to be the factor that complicates the DTSMP.

The DVRP is a new generation of the VRP, and has mainly been developed for planning vehicle routes in an on-going fashion, as vehicle locations, travel times and customer requests move over the planning horizon. The review results of Ghiani et al [7] indicated that the DVRP is the subject of on-going research and that more effort is required of researchers in the field before more efficient solution algorithms can be developed capable of fulfilling the real-time features of the DVRP. They also pointed out that for a weakly dynamic system, the focus is on minimizing routing cost, while for a strongly dynamic system, minimizing the expected response time becomes a key issue. This point echoes the twelfth feature of a DVRP as defined by Psaratftis [13]. Bertsimas and Van Ryzin [2,3] described a dynamic travelling repairman problem DTRP, which was classified by Psaraftis [14] as the simplest DVRP and reviewed by Ghiani et al [7]. This problem has an objective function that minimizes the average system time. In this current study, the system time represents the sum of the waiting time and service time for a demand. The review results show that the waiting time for each service demand has to be taken into account when planning dynamic routes in the DVRP. According to the special needs of the DTSMP, where preventing traffic disruption is a critical requirement, we must include a waiting time for each service demand in the objective function.

The DTSMP was previously studied by the authors [4,5]. In those studies, both fixed and stochastic travel times were considered with multiple objectives, the minimization of the total travel time and maximization of total utility gained from maintenance/repair works. Another concept, the “dynamic weight” was also proposed. This concept was designed to measure the potential traffic impact of a malfunctioning traffic signal. This study is and extension of the results of the previous aforementioned researcher, with the added consideration of time-dependent travel time and multiple vehicles as well as modified measurements of dynamic weight due to practical consideration.

Ichoua et al [10] made a complete review of work related to the issue of time-dependent travel time in the DTSMP, in which it is stated that it is step functions that have most often been used to compute the time-dependent travel time. In particular, they proposed a model based on time-dependent travel speeds, which satisfies the “first-in-first-out” property. Ichoua et al [10] also reported that time-dependent models offer a substantial improvement over fixed travel time models. This implies that a model with time-dependent travel speeds is closer to real-life conditions. This is also confirmed by the study of Hill et al [9]. In other words, strengthening the importance of time-dependent travel speeds increases the practical applicability of vehicle scheduling models. Hence, in this study, we adopt the concept of time-dependent travel speeds with some modifications when computing time-dependent vehicle travel time in the DTSMP.

4 The Demand-Response Routing Model

4.1 Model Framework

The conceptual framework of the demand-response routing model is presented in Figure 3. The structure consists of a time-evolution module, a decision-support module and a route-generating module. The time-evolution module controls the progress of the demand-response routing model by providing sequential execution time-points. These help by the decision-support module perform various functions such as receiving information, integration, communication, and feedback, as well as decision makings and commanding the execution of the route-generating module. The decision-support module processes and updates the information regarding daily maintenance demands, real-time repair demand-related data and timely road network data. Orders are received sequentially from the time-evolution module. The time-range for each sequential order is equal to the time needed for the repair or maintenance of a malfunctioning traffic signal. The time-evolution module is also responsible for maintaining the time progression, accumulation of time records needed for each work-task ensuring that the maintenance schedule follows the corresponding time sequence and that the daily working hours are reported to the decision-support module for making the decision when to stop the work day. The route-generating module receives service demands from the decision-support module to produce a timely route for each maintenance vehicle throughout the working process.

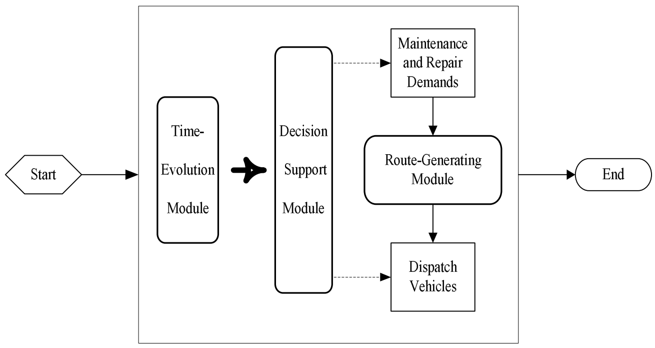

Figure 3 Dynamic Routing Framework

4.2 Routing Procedures under Different Service Strategies

The routing procedure is designed for handling multiple-vehicle conditions. The demand-response routing model uses the routing procedure to generate a dynamic maintenance route for each maintenance vehicle called on by the decision-support module. Due to the fact that there are too many possible service strategies that could be taken by the traffic operational offices in response to real-time emergency service demands, we simplify, so that only emergency service (ES) and general service (GS) strategies are explored here. The major difference between these two strategies lies in the vehicle dispatching rules used in responding to real-time emergency demands. For the ES strategy, an available vehicle is dispatched to serve the emergency demand immediately after the traffic operational office receives the orders. For the GS strategy, on the other hand, the emergency demand is assigned a priority or weight, and then this demand is pooled together with the remaining service demands and input into the route-generating model to produce a new route for that particular time. The stopping criterion for both service strategies is that all demands made in one day must be completed that is daily maintenance work does not stop until all emergency service demands occurring in that same day are completed, even if the eight-hour working time limit is violated. Additionally, it is assumed that in the real world, new demands are revealed randomly, either during the preparation period or during working stage of the routing procedure.

Figure 4 illustrates the specific steps of the routing procedure combining the ES and GS strategies. The procedure used for daily work begins with the integration of maintenance information prior to commencement of the work through the updating of real-time service demand data and the clock-time as well as decisions about vehicle dispatching during the work process, and ends with the completion of the eight-hour working time limit or the completion of the servicing of all emergency service demands. The double-criteria for the termination are designed to ensure that all emergency service demands revealed in one day are serviced without postponement. Each path finally traversed by individual maintenance vehicles is composed of a sequence of links assigned individually by the demand-response routing model. Furthermore, the dynamic route planned by the route-generating model will vary with the updating of the service demand data up to the time it is called.

According to the vehicle dispatching rule (for single or multiple vehicle conditions), maintenance vehicles are always dispatched one by one. In the multiple-vehicle case and under a GS strategy, each vehicle is selected randomly and sequentially dispatched to serve the demand, with the first priority being the maintenance route generated by the route-generating model during the preparation stage. During the working stage, vehicles are dispatched according to the sequence in which each vehicle completes its service, namely first finished-first dispatched. For multiple vehicles and under the ES strategy, vehicles are also selected randomly and dispatched to serve new service demands in the preparation stage, followed by the first come-first dispatched rule during the working stage.

4.3 Measurements of Model Objectives

Two objectives with equal weights are considered in the proposed routing model: maximum total utility and minimum total vehicle travel time. The measurements for these two objectives are discussed below.

(1) The Measurement of Utility

The term "utility" is a conceptual representation of the positive effect of the relief of traffic impact of malfunctioning traffic signals on road users. The level of potential effect is measured by the static and dynamic factors associated with each malfunctioning traffic signal. The static factors include the type of intersection and number of lanes in each direction at that intersection. For urban traffic networks, the types of roads include both main arteries and local streets. Obviously, the impact on traffic of a malfunctioning traffic signal located at the intersection of a 3-lane two-way arterial intersection will be considered more critical than that located at a 2-lane two-way local street. The dynamic factor is the time of the wait for maintenance or repair. In addition, if traffic signal repairs remain unaccomplished after two working days, these intersections are assigned the highest weights to assure their completion during the third working day.

We adopt the "dynamic weight" system developed by the authors [4,5] in this study. We combine the values of static and dynamic factors of a service demand where each factor's weight is computed using the entropy method. The dynamic weight of a malfunctioning traffic signal is calculated from the sum of its associated factors weights. The computation procedures for the dynamic weight and the impact utility are as follows:

Step 1: Computation of d_{ij}.

$$d_{ij} = X_{ij} / X_{ij}^*$$

where $X_{ij}^* = \underset{j}{Max}\, X_{ij}, i = 1, \ldots, m, j = 1, \ldots, n,$ and $0 \le d_{ij} \le 1.$

Step 2: Computation of P_{ij}.

$$P_{ij} = d_{ij} / D_i \quad \text{where } D_i = \sum_{j=1}^{n} d_{ij}, \text{ and } 0 \le P_{ij} \le 1.$$

Step 3: Computation of the entropy value e_{ij}.

$$e_i = -k \sum_{j=1}^{n} P_{ij} \ln P_{ij} \quad \text{where } k = 1/\ln n \text{ and } k > 0.$$

Step 4: Computation of the (dynamic) factor weight λ_i.

$$\lambda_i = \frac{1-e_i}{\sum_{i=1}^{n}(1-e_i)}$$

Step 5: Computation of the impact utility of the traffic signal P_j.

$$P_1 = \lambda_1 P_{11} + \lambda_2 P_{21} + \lambda_3 P_{31} + \cdots + \lambda_m P_{m1}$$
$$\vdots$$
$$P_n = \lambda_1 P_{1n} + \lambda_2 P_{2n} + \lambda_3 P_{3n} + \cdots + \lambda_m P_{mn}$$

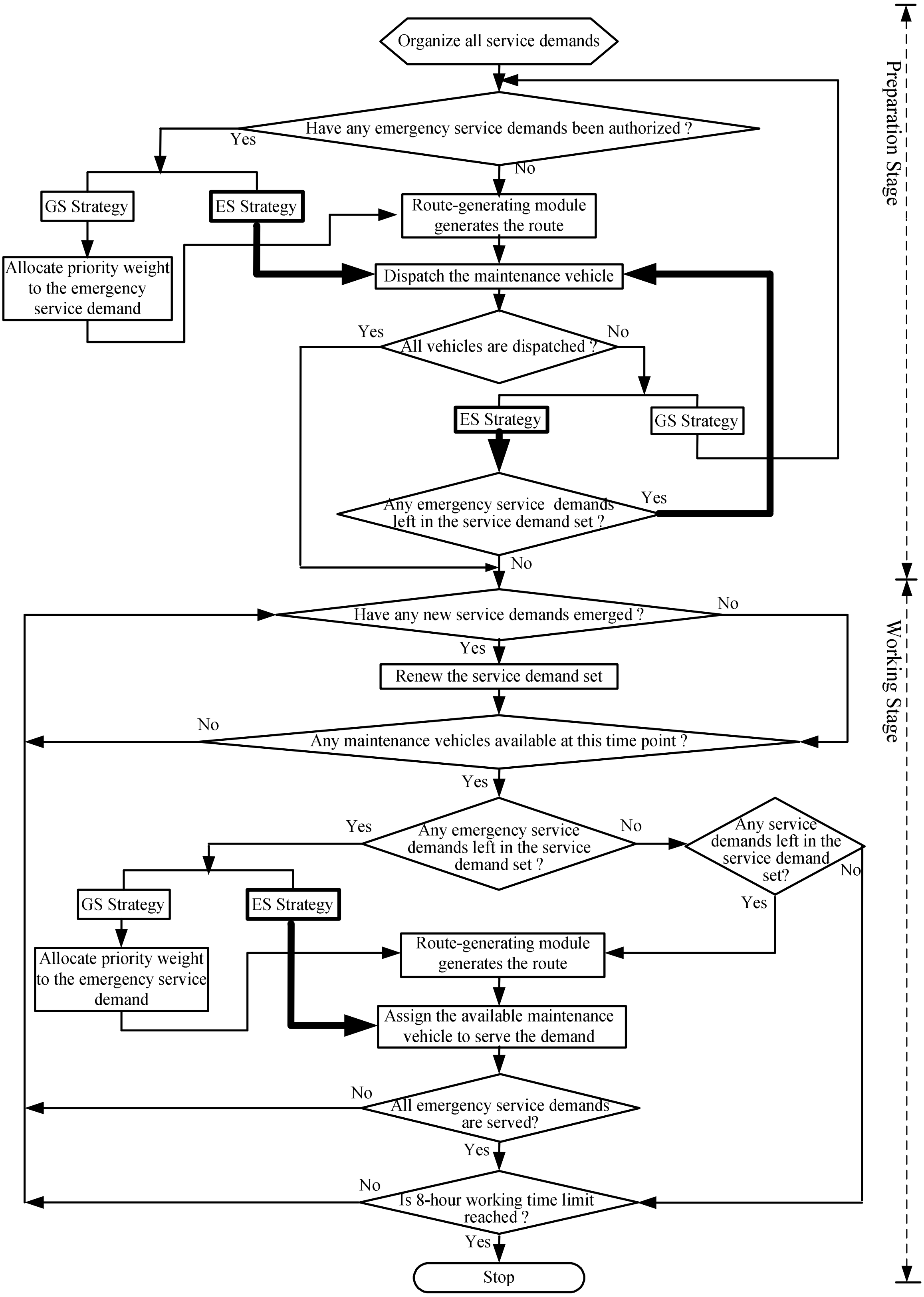

Figure 4 Dynamic Routing Procedure

(2) The Measurement of the Time-Dependent Travel Time

We extend the idea of Ichoua *et al.* [10] about using step-functions to compute vehicle travel speeds. We develop an approximate step-function for travel speed, as depicted in Figure 5, to estimate the travel speeds of the maintenance vehicle. Specifically, each maintenance vehicle's travel speed during working hours is assumed to follow a step-function with speed varying within a preset range in each time period, as specified by the user. For example, an 8-hour working time is divided into three time periods: morning (9:00 a.m. to 12:00 p.m.), noon (12:00 p.m. to 2:00 p.m.) and afternoon (2:00 p.m. to 5 p.m.). The travel speed in each time period is not constant but falls within a pre-defined range obtained from surveys or historical records.

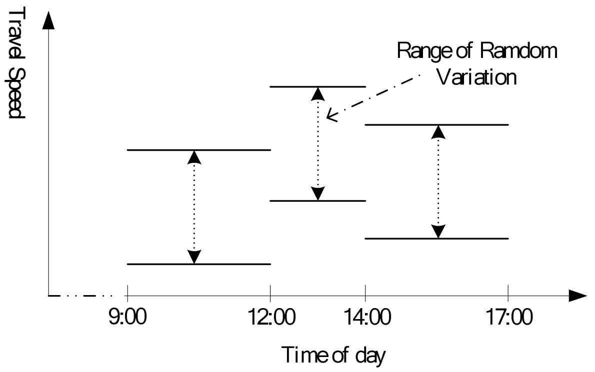

Figure 5 Illustration of the Approximation Speed-Function of Travel Speeds

The travel time of the maintenance vehicle between any two intersections can be computed using the steps detailed below, where the newly repaired traffic signal is called the "RO" and the next traffic signal to be repaired is called the "NO".

Step 1: Compute the time interval (denoted as Δt) between the time-point at which the maintenance vehicle leaves the RO and the end of the time period into which this time-point falls.

Step 2: Estimate the travel speed (denoted as s_1) for the maintenance vehicle from the approximate step-function of travel speeds based on the time-point at which it leaves the RO.

Step 3: Use s_1 and Δt to compute the travel distance and denotes it as l ($l = s_1 \times \Delta t$).

(1) If $l > l'$ (l' represents the distance between the RO and NO), use s_1 to compute the travel time as l'/s_1.

(2) If $l \le l'$ (this means the maintenance vehicle will traverses two time periods), compute the travel time using the following formula:

$$Travel\ time = \left(\frac{l' - l}{s_2}\right) \times 60 + \Delta t$$

, where s_2 is the estimated speed of the maintenance vehicle in the time period next to the time period to which s_1 belongs.

4.4 Model Form

The demand-response routing model proposed to solve the DTSMP is a TDTSP-type model since vehicles are dispatched one by one and it is assumed that they cannot return to the depot until the end of working time or all daily emergency service demands are completed. Furthermore, the assumption of unlimited vehicle capacity makes the TDVRP a TDTSP. The mathematical formulation of the static TSP model used at each time point to generate a dynamic maintenance route for each "available" maintenance vehicle is presented below.

Objective function:

$$Z = \left\{ Max. \sum_{j \in N} P_j Y_j \ \ \& \ \ Min. \sum_{i \in N} \sum_{j \in N} \tau_{ij} x_{ij} \ , \quad i \ne j \right\} \tag{1}$$

Constraints:

$$\sum_{i \in N} x_{ij} = 1 \qquad (j = 1,2,3,...,n) \ , \tag{2}$$

$$\sum_{j \in N} x_{ij} = 1 \qquad (i = 1,2,3,...,n) \ , \tag{3}$$

$$X = (x_{ij}) \in S \ , \tag{4}$$

$$x_{ij} = 0 \text{ or } 1 \qquad (i, j = 1,2,3,...n) \ , \tag{5}$$

Where

P_j = the dynamic weight of the j^{th} malfunctioning traffic signal,

$$Y_j = \begin{cases} 1 & \text{if malfunctioning signal } j \in N \text{ is repaired or maintained,} \\ 0 & \text{otherwise.} \end{cases}$$

τ_{ij} = the dynamic travel time between malfunctioning traffic signals i and j,

$$x_{ij} = \begin{cases} 1 & \text{if } arc(i.j) \in A \text{ is travelled,} \\ 0 & \text{otherwise.} \end{cases}$$

N is a node set and $N = \{1,2,3,..., n\}$,

A is an arc set and $A = \{(i, j) | i, j = 1,2,3,...,n\}$.

S is the sub-tour breaking constraint.

4.5 Application of Genetic Algorithms

Genetic Algorithms (GAs), first proposed by John Holland in 1975, are applied here to find a near-optimal solution for the routing model. The GA is simply a computational method which simulates nature's evolutionary method to solve optimization problems. GAs can be applied for either optimization or classification. Their advantage is that they are better able to find global optimum in highly nonlinear spaces such as is the case with the TDTSP [8]. The core mechanisms of GAs are natural evolution and natural selection. A population of chromosomes is repeatedly modified through the application of genetic operators. A gene is the basic structure of a chromosome. It can be either binary coded or order coded depending on the problem types. A chromosome is a component of a population and is usually used to encode a solution for the task at hand. The GA process usually starts with a population of *n* randomly generated chromosomes and continues through a number of generations until some predetermined criterion is satisfied. The process is continued using three operators, reproduction, crossover and mutation. The chromosome with the optimum, or highest, fitness value of the last population is selected. The fitness value is the information used in the GA to determine survival or death during the evolutionary process. The design of the fitness value depends on the type of problem without fixed types of representations.

The encoding of GA for the traffic signal maintenance problem is illustrated in Table 1 where each gene is represented by the number of a specific traffic signal that requires maintenance, which is a node number in the traffic network. Each chromosome is represented by a routing path, which is constituted of a sequence of traffic signals requiring service. The population is a set of routing paths.

Table 1 GA Encoding for the Traffic Signal Maintenance Problem

Genetic Algorithms	**The Studied GA Model**
Gene	Number of a traffic signals requiring service
Chromosome	Routing path
Population	Set of routing paths

The detailed GA procedure designed for this study is presented in Figure 6. The key settings are discussed below.

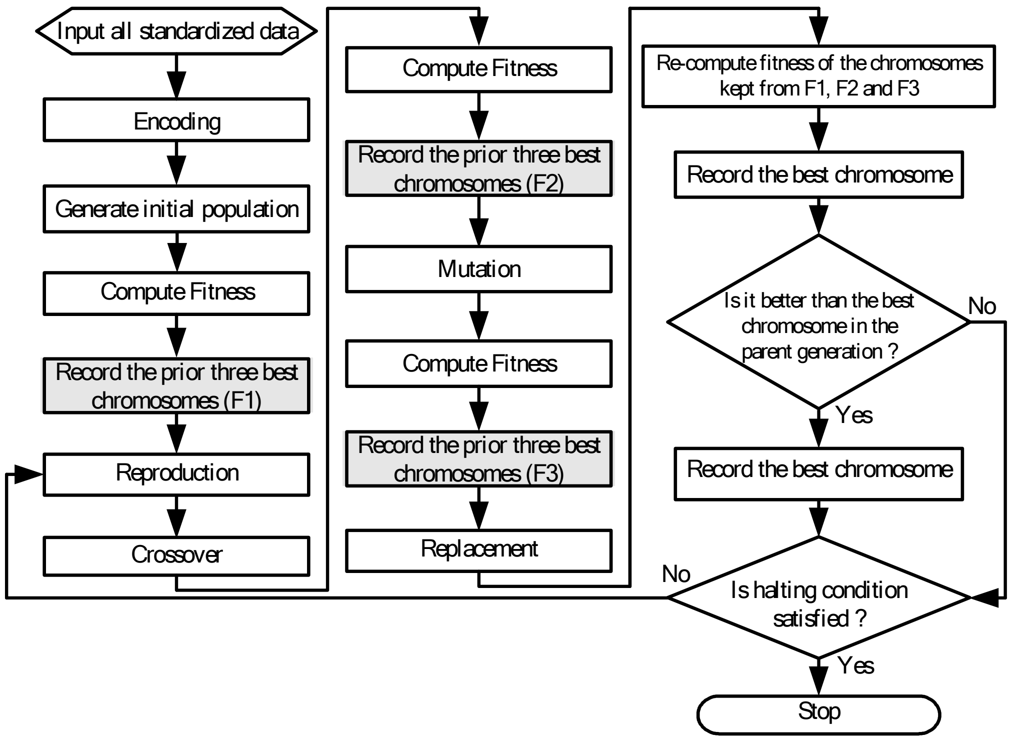

Figure 6 Flowchart of the Designed GA

(1) Population: An initial population is generated randomly. The initial population size is determined with simultaneously consideration of the number of generations, crossover rate and mutation rate.

(2) Fitness: Fitness is the basis for evaluating a chromosome. Each chromosome's fitness is

measured by the values of two objectives of the routing model. The utility of a repaired traffic signal is measured by a weight and travel times. The steps of the fitness calculation are described below.

Step 1: All chromosomes in the population are sequenced according to the value and order of the associated weight of each gene inside a chromosome. Then ranked value of each chromosome indicates its total utility.

Step 2: All chromosomes in the population are sequenced according to the value of the total travel time of each chromosome. Then ranked value of each chromosome indicates its total travel time.

Step 3: These two ranked values obtained from Steps 1 and 2 are summed up for chromosome and sequenced one more time. The new ranked value of each chromosome indicates its final fitness.

Note that some chromosomes may have the same final fitness values as others although their contents are quite different. Therefore, all are retained without prejudice for further operations.

(3) Reproduction: Reproduction is a process whereby fitter individuals are selected from the population to create offspring. Since the fitness of a chromosome in the problem at hand is a ranked value, the Tournament Selection is adopted to reproduce offspring. As suggested by Abrams [1] a group of 5 individuals is randomly chosen from the population, and the fittest individual from that group is chosen as the parent.
(4) Crossover: A Partial-Mapped Crossover (PMX) operator [6] is used to produce offspring.
(5) Mutation: A Reciprocal Exchange operator [6] is adopted to produce offspring.
(6) Replacement: The offspring generated by the operations of reproduction, crossover and mutation replace all parent chromosomes.
(7) Re-computation of the Fitness: The best chromosome is selected from those chromosomes with the prior three best fitness values obtained from each previous step of the fitness computation. The fittest chromosome is then compared with the best chromosomes in the parent generation to determine the fittest chromosome of this generation.
(8) Halting Condition: In this study the stopping condition is a pre-set number of generations determined based on simultaneous consideration of the population, crossover rate and mutation rate.

5 Model Testing and Evaluation

The dynamic, stochastic and complicated nature of the DTSMP means that many factors such as the time-dependent travel time, quantity of malfunctioning traffic signals and those scheduled for maintenance, quantity of real-time as well as new emergency service demands and the different time-points at which they are known may individually or jointly affect the model's solution result, regardless of which heuristic method is used. There are two factors focused upon when evaluating the method investigated in this study, the first is to evaluate the GA model's capability of fulfilling real-life applications and the second one is aimed at improving the model's computational efficiency. The first part include three aspects, which are (a) the model's general performance under different service strategies; (b) the model's capability in servicing real-time service demands given the revealing of different time-points and different levels of significance (measured by weights); and (c) the model's performance in servicing randomly revealed emergency service demands under different service strategies.

5.1 The Test Problem

The test problem is modified from a real world case with a traffic network constituted of 91 signal-controlled intersections in Tainan County, Taiwan. The experimental network is constructed from this real network and its associated geometric data, but contains only 50 randomly selected traffic signals. Moreover, maintenance and repair records for malfunctioning traffic signals are also referenced from Tainan County's traffic operational office so as to set the necessary traffic signal maintenance information and data. A summary of data for the test problem follows.

(1) Service strategy: the proposed general or emergency service strategy.
(2) Environmental related data: (a) the depot is located at traffic signal number 50; (b) traffic signals are numbered from 1 to 49; (c) ranges of travel speeds for the maintenance vehicle in different time periods are illustrated in Table 2; and (d) the daily working-time limit is 8 hours.
(3) Supply related data: One or two maintenance vehicles
(4) Demand related data: (a) traffic signals scheduled for service prior to the start of work are numbers 1, 4, 5, 7, 10, 12, 14, 16, 21, 24, 26, 27, 31, 35, 36, 40, 41, 43, and 46; (b) known malfunctioning traffic signals are illustrated in Table 3; (c) randomly revealed malfunctioning

traffic signals are numbers 3, 9, 23, 39, and 49; (d) emergency service demands are traffic signal numbers 18 and 47; (e) different repair times required for each traffic signal are illustrated in Table 4; (f) the maximum wait for service of a malfunctioning traffic signal is set as 3 days; and (g) road attributes associated with each traffic signal are illustrated in Table 5.

Table 2 Ranges of Vehicle Travel Speeds in Different Time Periods

Range of Random Speed (km/hr)		
Morning (9:00 - 12:00)	Noon (12:00 - 14:00)	Afternoon (14:00 - 17:00)
17.00 - 40.00	40.00 - 66.05	30.00 - 45.00

Table 3 Associated Wait for Servicing of the Malfunctioning Traffic Signals

Wait for Service (day)	Traffic Signal Number
1	6, 19, 22, 28, 32, 37, 38, 42, 45, 48
2	2, 8, 11, 17, 20, 30, 33, 34, 44
3	13, 15, 25, 29

Table 4 Required Repair Time for the Traffic Signal

Required Repair Time (minutes)	Traffic Signal Number
5	8, 16, 23, 27, 28, 42, 45
10	1, 2, 10, 11, 17, 21, 22, 31, 32, 34, 43, 49
15	4, 24, 35, 41, 46
20	5, 6, 25, 29, 48
25	19, 33, 44
30	12, 15, 30, 37, 38, 40
40	3, 9, 13, 18
50	20, 39
60	7, 26, 47
80	14, 36

Table 5 Road Attributes Associated with each Traffic Signal

Road Attributes		Traffic Signal Number
Road Level	State	4, 8, 10, 13, 17, 23, 27, 31, 36, 41, 42, 44, 45, 48
	County	1, 3, 6, 9, 11, 14, 18 19, 20, 22, 24, 25, 28, 30, 34, 35, 38, 39, 43, 46, 49
	City	2, 5, 7, 12, 15, 16, 21, 26, 29, 32, 33, 37, 40, 47
Number of Lane	One	2, 6, 7, 11, 12, 14, 18, 22, 26, 28, 35, 41, 44, 46, 47, 49
	Two	1, 3, 4, 5, 8, 9, 13, 15, 16, 19, 21, 24, 25, 27, 30, 31, 32, 34, 37, 38, 40, 42, 43, 45, 48
	Three	10, 17, 20 ,23, 29, 33, 36, 39

5.2 Parameter Analysis of Genetic Algorithms

In this study we utilize the same range of values for each parameter commonly used by other researchers. These include (a) population size (P): 20 to 100; (b) crossover rate (P_c): 0.5 to 1.0; (c) mutation rate (P_m): 0.001 to 0.05; and (c) evolution generations (I): 100 to 1000. In total there were 72 combinations performed in the simultaneous search for better values from the four parameters. The ranges are illustrated in Table 6. Each combination was tested ten times to counter the random effect of the initial population. The results are shown in Figure 7. The best combination of parameters is (P, I, P_c, P_m) = (100, 1000, 0.9, 0.05).

Table 6 Parameter Analysis Data for the GA Model

Parameters	Test Range	Incremental Interval of the Test Range
Cross Rate	0.5 - 1.0	0.1
Mutation Rate	0.01 - 0.05	0.02
Evaluation generations	500, 1000	
Population Size	50, 100	

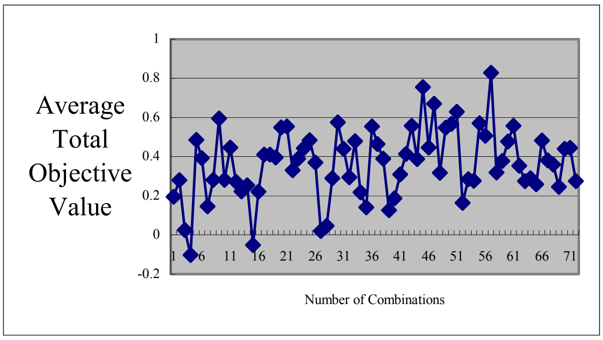

Figure 7 Results of Parameter Analysis of the GA Model

In our evaluation of the best combination of parameters, we modify a TSP example, eil51, from a reference TSP test bank (http://www.tsp.gatech.edu/concorde/bench99.html) to validate it in a static sense. The network size of eil51 is equivalent to the test problem. Its optimal solution value (total travel distance) is 426. Each node contained in the optimal solution of eil51 is first designated a weight according to its order in the path before the data are fed into the GA model for validation. The validation results are shown in Table 7. The average value is computed from an average solution of 30 runs in which two are equivalent to the optimal solution.

Table 7 Validation Results of the Best Combination of Parameter Values

Average Value	Maximum Value	Minimum Value	Numbers of Optimal Solution Reached
456	510	426	2

5.3 Numerical Results

The outcomes presented below are based on the first day's maintenance results. In the following sections, the evaluation results indicating the model's capability for practical applications are discussed. This is followed by a discussion of the consequence of the improvement of computational efficiency.

5.3.1 Model's Capability for Practical Application

Evaluation results of the model's performance under different service strategies with different numbers of maintenance vehicle are illustrated in Tables 8 and 9. Table 8 shows each vehicle's service sequence. Detailed information about the results is further illustrated in Table 9. It can be seen that all emergency demands are repaired within the working time limit and more malfunctioning traffic signals are repaired than those in the scheduled maintenance demands. This holds true for both service strategies. The result of more scheduled maintenance demands left behind is because their dynamic weights are often less comparing to those of malfunctioning traffic signals. In addition, please note that unequal numbers of real-time service demands for different cases is primary due to their occurrences being randomly generated by the computer program.

Table 8 First Day's Maintenance Results

Service Sequence under GS Strategy			Service Sequence under ES Strategy		
Two-vehicle Case		One-vehicle Case	Two-vehicle Case		One-vehicle Case
First Vehicle	Second Vehicle		First Vehicle	Second Vehicle	
13, 2, 8, 3, 34, 17, 29, 15, 6, 46, 47	5, 30, 18, 45, 23, 44, 33, 19, 32	5, 13, 17, 25, 18, 30, 33, 23	13, 18, 17, 20, 2, 39, 42, 15, 48	11, 44, 8, 33, 5, 29, 6, 47	5, 13, 18, 17, 25, 30, 33

Table 9 Detailed First Day Maintenance Results

	GS Strategy				ES Strategy			
	Two-vehicle Case		One-vehicle Case		Two-vehicle Case		One-vehicle Case	
	Traffic Signal Number	Serviced Demand (%)	Traffic Signal Number	Serviced Demand (%)	Traffic Signal Number	Serviced demand (%)	Traffic Signal Number	Serviced Demand (%)
Emergency Service Demands	18, 47	18, 47 (100)	18	18 (100)	18, 47	18, 47 (100)	18	18 (100)
Scheduled Maintenance Demands	1, 4, 5, 7, 10, 12, 14, 16, 21, 24, 26, 27, 31, 35, 36, 40, 41, 43, 46	5, 46 (10.05)	1, 4, 5, 7, 10, 12, 14, 16, 21, 24, 26, 27, 31, 35, 36, 40, 41, 43, 46	5 (5.00)	1, 4, 5, 7, 10, 12, 14, 16, 21, 24, 26, 27, 31, 35, 36, 40, 41, 43, 46	5 (5.00)	1, 4, 5, 7, 10, 12, 14, 16, 21, 24, 26, 27, 31, 35, 36, 40, 41, 43, 46	5 (5.00)

Malfunctioning Traffic Signal	Known Before Work	2, 6, 8, 11, 13, 15, 17, 19, 20, 22, 25, 28, 29, 30, 32, 33, 34, 37, 38, 42, 44, 45, 48	13, 30, 45, 2, 8, 44, 34 17, 33, 29,19, 15, 6,32 (60.86)	2, 6, 8, 11, 13, 15, 17, 19, 20, 22, 25, 28, 29, 30, 32, 33, 34, 37, 38, 42, 44, 45, 48	13, 17, 25, 30, 33 (21.74)	2, 6, 8, 11, 13, 15, 17, 19, 20, 22, 25, 28, 29, 30, 32, 33, 34, 37, 38, 42, 44, 45, 48	13, 11, 44, 17, 8, 20, 33, 2, 29, 42, 6, 15, 48 (56.62)	2, 6, 8, 11, 13, 15, 17, 19, 20, 22, 25, 28, 29, 30, 32, 33, 34, 37, 38, 42, 44, 45, 48	13, 17, 25, 30, 33 (21.74)
	Real-time Revealed	3, 23	3, 23 (40.00)	49, 3, 23	23 (33.3)	3, 9, 23, 39, 49	39 (20.00)	23, 39, 3	- (0.00)

The model's capability for servicing real-time service demands can be evaluated by investigating how well the model answers the following two questions: "Can the model handle a randomly emerging demand immediately after it appears?" and "How long does it take to respond to a randomly emerging demand?" For this evaluation, the weight of the random demand is defined as having one of three different levels, high, middle or low, prior to analysis. The other control factors include (1) one maintenance vehicle for the general service strategy, (2) one real-time revealed service demand whose number is 3; (3) the service time for each malfunctioning traffic signal is the same constant; and (4) all other service demand conditions remain unchanged.

Regarding the first question, the results after an average of 30 runs are demonstrated in Figure 8. Note that random demand with high-level weights can almost be serviced, except when it shows up near the end of the working time limit. The demands with high and median level weights also have a higher tendency to be repaired earlier. However, the demands with low-level weights cannot be repaired for any one condition. In addition, please note that the time periods shown in the y-axis of Figure 8 are segmented according to the time-point at which each traffic signal maintenance work is completed. The evaluation results regarding to the second question are presented in Table 10. The results are also an average of the results of 30 runs. The results indicate that there is no wait for the ES strategy and an average of four service demands have to be completed before the real-time demand is repaired under the GS strategy. This outcome is also as expected, since ES strategy was designed to carry out immediately response service.

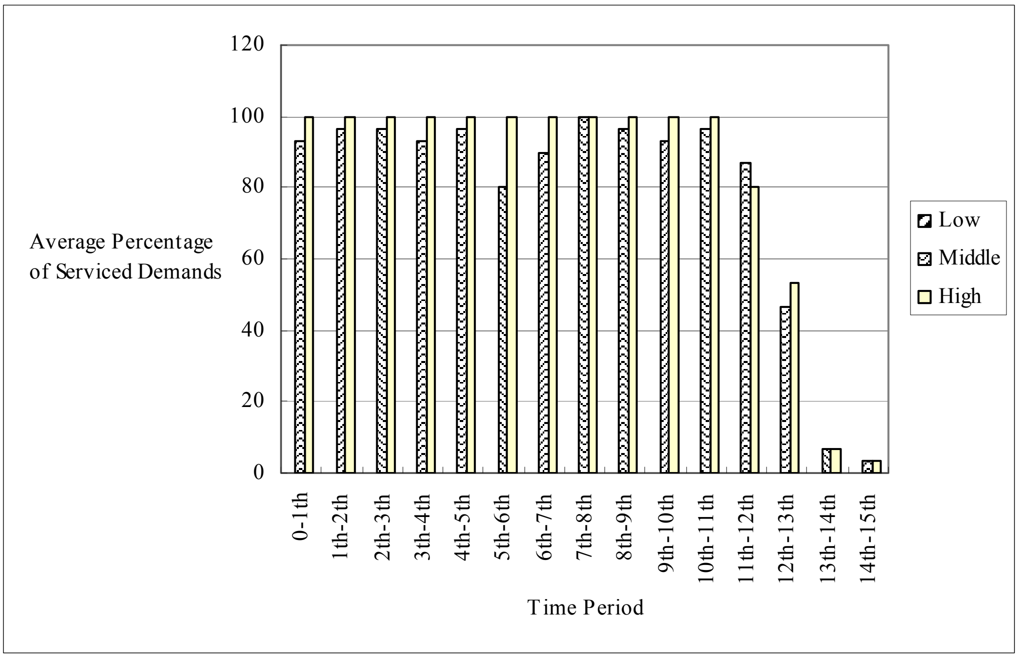

Figure 8 Results of the Demand-Response Model for the First Question

Table 10 Model Performance for Servicing a Randomly Revealed Service Demand

GS Strategy		ES Strategy	
Percentage of Accomplished Service Demand	Average Wait	Percentage of Accomplished Service Demand	Average Wait
100%	4 Service Demands	100%	None

Although the aforementioned evaluation results for the proposed demand-response routing model reveal promise for practical application there is still room for discussion when implementation of this model in real world is intended. For example, the one-depot assumption made in this study is feasible for small and mid-size networks but not large networks. The determination of how big a small or mid-size network is depends upon the nature of the problem and this requires further study for definition. When the proposed routing model is applied in large networks, it is always possible to divide the network into smaller feasible sub-networks; however whether this is an optimal approach or not still requires further investigation. In addition, the assumption of normal weather made in this study assumes that maintenance work could always be postponed if bad weather conditions are encountered. In other words, bad weather conditions are handled in the model as a time period for postponement and the accumulation of service demands that occurred during that period can be incorporated into the work schedules after work resumes given some necessary information renewal work done for the service demand set.

5.3.2 Computational Issues

Given the special features of the fitness calculation for the proposed GA model, it is found that the ranking selection procedure consumes the largest amount of time, about 78% of the total computation time. Hence, a parallel computing technique was applied to reduce the total computation time. The parallel computing set consisted of (a) Software: Window2000/XP with JDK (Java Develop Toolkit), version 1.4 and (b) Hardware: Intel Pentium 4 with double CPU, 256MB RAM and 20G HD space. The results of the average of 30 runs are shown in Table 11. It can be seen that there was a 58% saving of time after the adoption of this parallel computation structure.

6 Conclusions and Future Research

Traffic signal maintenance and inspection scheduling is a challenging research issue because of its dynamic nature. The maintenance routes are limited by the resources of the traffic operation offices, and are affected by the different attributes of the service demands. A demand-response routing model with time-dependent travel time, which can account for traffic impact derived from malfunctioning traffic signals and the wait for service is proposed. The goal is to plan timely maintenance routes according to daily inspection plans and randomly revealed malfunctioning traffic signals. The maintenance routes were solved by using GA and evaluated under different service strategies based on real world data. The results show that the demand-response model (which balances weights obtained for the maximization of total utility and minimization of total travel time) proposed in this study is able to (a) account for both scheduled and randomly occurring service demands; and (b) react reasonably to real-time service demands and respond immediately after the occurrence of an emergency service demand under the emergency service strategy. Moreover, the utilization of a parallel computing technique can significantly improve the computational efficiency of the proposed GA model, increasing its promise for future real-life applications.

In view of the complexity of the DTSMP, further investigation in theory and exploration for real-life applications are necessary. The future research includes the following directions: (a) to explore the effect of different combinations of weights of the two objectives to the solution results; (b) to investigate the significant effect of incorporating supply cost into model's objective function to the solution results; (c) to study the effect of network size to the solution results; (d) to compare the performance of different solution methods; and (e) to develop a computer-human interface to construct the demand-response model to be an interactive decision support system for real world applications.

Table 11 Comparison of System Processing Time

Average System Processing Time (mini second)	
Sequential Computing Technique	Parallel Computing Technique
11849.67	4976.86

Acknowledgements

This work was partially supported by the National Science Council of Taiwan under Grant NSC922211E216011.

References:

[1] J.P. Abrams, *A Hierarchical Genetic Algorithm for the Travelling Salesman Problem*, Thesis, School of Computer Science, Carleton University, 2003.

[2] D.J. Bertsimas and G. Van Ryzin, A Stochastic and Dynamic Vehicle Routing Problem in the Euclidean Plane, *Operations Research*, Vol.39, No.4, 1991, pp.601-615.

[3] D.J. Bertsimas and G. Van Ryzin, Stochastic and Dynamic Vehicle Routing Problem in the Euclidean Plane with Multiple Capacitated Vehicles, *Operations Research*, Vol.41, 1993, pp.60-76.

[4] C.-H. Chen and K. Liao, A Dynamic Traffic Signal Maintenance Problem: Framework and Preliminary Results, *Proceedings of the Fifth Asia Pacific Industrial Engineering and Management Systems Conference*, Gold Coast, Australia, 2004, pp.29.1.1-29.1.16.

[5] C.-H. Chen and M. Wang, 2002. A Vehicle Routing Model for Traffic Signal Maintenance Works, *Proceedings of the 7th ROC Transportation Network Conference*, Hsinchu, Taiwan, 2002, pp.363-374.

[6] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Design*, John Wiley & Sons Inc., New York, 1997.

[7] G. Ghiani, F. Guerriero, G. Larporte, and R. Musmanno, Real-Time Vehicle Routing: Solution Concepts, Algorithms and Parallel Computing Strategies, *European Journal of Operational Research*, Vol.151, 2003, pp.1-11.

[8] D.E. Glodberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Boston, 1989.

[9] A.V. Hill, and W.C. Benton, Modelling Intra-City Time-Dependent Travel Speeds for Vehicle Scheduling Problems, *Journal of the Operational Research Society*, Vol.43, No.4, 1992, pp.343-351.

[10] S. Ichoua, M. Gendreau and J.-Y. Potvin, Vehicle Dispatching With Time-Dependent Travel Times, *European Journal of Operational Research*, Vol.144, 2003, pp.379-396.

[11] M.D. Meyer, *A Toolbox for Alleviating Traffic Congestion and Enhancing Mobility*, Institute of Transportation Engineers. Washington D.C., 1997.

[12] S. Michalewicz, *Genetic algorithms + Data Structures = Evolution Programs*, Springer-Verlag, London, 1996.

[13] H.N. Psaraftis, Dynamic vehicle routing problems, *Vehicle Routing: Method and Studies*, Elsevier Science Publishers B.V., North-Holland, 1988.

[14] H.N. Psaraftis, Dynamic Vehicle Routing: Status and Prospects, *Annals of Operations Research*, Vol.61, 1995, pp.143-164.