Computer-Aided Civil and Infrastructure Engineering **24** (2009) 459–469

Optimal Long-Term Infrastructure Maintenance Planning Accounting for Traffic Dynamics

ManWo Ng*, Dung Ying Lin & S. Travis Waller

Department of Civil, Architectural, and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA

Abstract: *Periodic infrastructure maintenance is crucial for a safe and efficient transportation system. Numerous decision models for the maintenance planning problem have been proposed in the literature. However, to the best of our knowledge, no model exists that simultaneously accounts for traffic dynamics and is intended for long-term planning purposes. This article addresses this gap in the literature. A mixed-integer bi-level program is introduced that minimizes the long-term maintenance cost as well as the total system travel time. For the solution approach we utilize a genetic algorithm in conjunction with mesoscopic traffic simulation. The model is illustrated via a numerical example.*

1 INTRODUCTION

Periodic infrastructure maintenance is crucial for a safe and efficient transportation system. Maintenance planning problems can be categorized as either long-term or short-term. (We want to note that different classification schemes are possible, for instance, based on the way infrastructure deterioration is modeled. However, to avoid repetition, we have subdivided our discussion using the long-term/short-term criterion only.) In long-term planning problems, the typical trade-off is the timely allocation of scarce resources over time such that infrastructure elements are maintained above critical service levels. Further, it is important to account for various realities regarding maintenance cost (extended periods of neglect typically result in higher maintenance costs), budget constraints, and infrastructure deterioration. On the other hand, short-term maintenance planning problems require as input a list of facilities to be maintained. The optimal operational schedule is to be determined to minimize user delay within budget constraints. For an introduction to the road deterioration and management problem, we refer to Paterson (1987), Jiang and Adeli (2003), Karim and Adeli (2003), and Jiang and Adeli (2004).

In long-term planning models, the selection of appropriate deterioration models for infrastructure elements is of crucial importance. In the literature, two major approaches can be found. The first approach uses deterministic empirical deterioration curves. Tsunokawa and Schofer (1994) utilize exponential decay in the modeling of pavement deterioration. In Miyamoto et al. (2000) bridge deterioration is represented by fourth-order polynomials; Maji and Jha (2007) use parabolic equations, whereas Frangopol et al. (2001) use piecewise linear representations. Kong and Frangopol (2003) obtain the reliability index of a bridge by the superposition of its reliability function and maintenance-induced reliability profiles. Although the aforementioned models provide critical insights, they may not fully characterize the stochastic nature of material decay. The second approach focuses on the explicit representation of the stochastic nature of deterioration (Pasupathy et al., 2007). Most notable are the formulations based on Markov decision processes (MDPs), where transition probabilities are obtained either via statistical estimation techniques (e.g., see Golabi et al., 1982; Carnahan, 1988; Madanat and Wan Ibrahim, 1995; Mishalani and Madanat, 2002), or via expert opinions. More recently, researchers have examined the use of state-space specifications of time series models to estimate infrastructure performance models (e.g., see Chu and Durango-Cohen, 2007; Chu and Durango-Cohen, 2008a,b).

Numerous long-term planning models have been formulated for the infrastructure maintenance problem

*To whom correspondence should be addressed. E-mail: *mwng@mail.utexas.edu*.

DOI: 10.1111/j.1467-8667.2009.00606.x

based on both paradigms. Although many critical contributions exist, a small sample is discussed in this article to place the research fully in context. For the deterministic paradigm, Ouyang and Madanat (2004) provide a mixed-integer nonlinear programming formulation. Maji and Jha (2007) develop a nonlinear nonconvex formulation for the maintenance optimization problem. A genetic algorithm is utilized for the solution procedure. Other mathematical programming models were proposed by Al-Subhi et al. (1990) and Jacobs (1992). Under the stochastic paradigm, Guignier and Madanat (1999) present an integer programming model for the joint optimization of maintenance and improvement activities. The model is formulated under steady state assumptions. Smilowitz and Madanat (2000) extend the latent MDP formulation to the network level. Latent MDPs were introduced as models for non-error-free infrastructure inspections. They give rise to randomized policies. Models are typically solved via linear programming algorithms. Guillaumot et al. (2003) present an adaptive model in which transition probabilities of the deterioration model were updated via Bayes' theorem after inspections were made on the infrastructure. Although stochastic models capture the stochastic nature of deterioration, the solution of large-scale problems tends to be computationally challenging because of the curse of dimensionality (Bellman, 1957), despite various approximation techniques that have been proposed to counteract scalability issues (e.g., see Bertsekas, 2001).

The above models are classified as long-term planning models. Planning horizons can potentially lie several years into the future. The objective of these models tends to be maintenance cost minimization. If user cost (in this article, the term user cost refers to the total delay in the road network caused by maintenance activities, unless stated otherwise) is considered, user cost parameters are introduced to quantify this cost (e.g., see Guignier and Madanat, 1999). Recent research quantifies user cost by solving some traffic assignment problems. Uchida and Kagaya (2006) are the first to introduce traffic equilibrium constraints to the long-term pavement maintenance problem. A model to minimize user and maintenance costs under the probit-based user equilibrium model is formulated and solved. In their model, user cost is defined as travel time and driving cost. Ouyang (2007) proposes a dynamic program that simultaneously minimizes user cost and maintenance cost, under deterministic user equilibrium. The model is solved via approximate policy iteration. To the best of our knowledge, all equilibrium models considered in the long-term maintenance planning problem are some form of static equilibrium. This article introduces dynamic traffic equilibrium to the long-term infrastructure maintenance planning problem. Apart from the more realistic representation of traffic flow compared to static equilibrium models (static models are synonymous with steady state, time-invariant conditions, which is obviously erroneous, especially during the peak periods), the proposed model alleviates the following concern about static traffic assignment-based formulations: suppose that we have a road segment consisting of three consecutive streets. The decision to do maintenance on the street in the middle will only yield increased travel time for the street itself under static traffic assignment models. The upstream street will be spared. However, reality tells us that not only will there be an increase in travel time for this specific link, but also upstream streets will likely be affected by this local decrease in capacity. Therefore, static models tend to underestimate user delays when congestion is present. Dynamic traffic assignment (DTA) removes this deficiency and hence accounts for a more accurate representation of user delay.

Short-term maintenance planning models consider user delay as the most important entity to be minimized. Unlike the long-term planning models, in short-term planning problems the road segments to be maintained are assumed to be given. The challenge then becomes finding the optimal maintenance schedule such that user delay is minimized (e.g., see Fwa et al., 1998). Short-term planning models typically encompass some form of traffic equilibrium. For example, Chang et al. (2001) utilize a tabu search approach for the work zone scheduling in a small network, assuming dynamic traffic equilibrium. In Cheu et al. (2004) a genetic algorithm-based microsimulation method was used to obtain optimal maintenance schedules. They illustrated their method on a network with 74 nodes, 160 links, 13 origin–destination pairs, and 10 prespecified maintenance jobs. A computational time of nearly 7 days was recorded. Ma et al. (2004) proposed several techniques to reduce the computational time of the aforementioned approach. For a network with 397 nodes, 986 links, 22 origin–destination pairs, and 20 work zones, they reported a computational time of 28,800 minutes.

To the best of our knowledge, no maintenance planning model exists that simultaneously accounts for traffic dynamics and is intended for long-term planning purposes. This article addresses this gap in the literature. We formulate a long-term planning model for the infrastructure maintenance problem that accounts for dynamic traffic conditions. Furthermore, we utilize mesoscopic traffic simulation to evaluate the impact of maintenance decisions on road users. We choose mesoscopic simulation because of the following reasons (Burghout et al., 2005). First, microsimulation models tend to be highly sensitive to input data, especially in congested conditions. Mesoscopic models are less sensitive in this respect. Because we consider long-term planning problems that are typically characterized by various sources of uncertainty, it is more natural to choose mesoscopic modeling. Second, the calibration of a microsimulation model is not trivial because of the large number of parameters involved. Mesoscopic models typically require less parameters to be calibrated. However, we have to note that the calibration of origin–destination parameters remains a challenge, for micro- as well as mesoscopic simulation. Finally, the use of microsimulation tends to be limited to smaller sized networks because of its higher computational requirements. Consequently, it may suffer from boundary effects. The proposed model is based on the Cell Transmission Model (CTM) introduced by Daganzo (1994, 1995). Moreover, we follow the paradigm of deterministic deterioration. In the mathematical formulation of the model, we limit our attention to a single-destination problem (it should be noted that no formal mathematical statement of planning models exists in the current literature that account for traffic dynamics, even for single-destination problems, e.g., Cheu et al., 2004), because no general CTM-based formulation for the multiple-destination problem exists (Li et al., 1999). However, the solution procedure adopted is also appropriate for multiple-destination problems, as we demonstrate in our numerical example.

The remainder of this article is organized as follows. In Section 2, we first briefly discuss the two main components of our model, that is, CTM and the modeling of the deterioration curve. Then a mixed-integer bi-level program is formulated to determine optimal maintenance schedules. The simulation-based genetic algorithm solution procedure for the bi-level program is discussed in Section 3. Section 4 presents a numerical example to illustrate the use of our model. Finally, we summarize our main contributions and findings in Section 5.

2 MODEL FORMULATION

2.1 The cell transmission model

CTM was first introduced by Daganzo (1994, 1995). CTM is a discrete approximation of the hydrodynamic traffic flow model of Lighthill and Whitham (1955a,b) and Richards (1956). It is based on a piecewise linear relationship between traffic flow and density. The result is a set of simple difference equations. In this model, roads are divided into cells of appropriate length such that free-flowing vehicles travel at most one cell per unit time. The cells are connected by cell connectors that do not represent physical entities in a road network. They are merely used to represent flow between consecutive cells. Cells can be classified into different categories. For our purpose, it is sufficient to introduce the following: source cells (sink cells) are the cells where traffic enters (exits) the network; sink cell connectors are cell connectors incident to sink cells. (For a more detailed discussion, the reader is referred to Daganzo, 1995, or Ziliaskopoulos, 2000.) By the inclusion of parameters such as jam densities, cell capacities, and backward wave speeds, real-world traffic phenomena such as shockwaves and queueing can be accurately accounted for. However, CTM as introduced by Daganzo (1994, 1995) cannot be easily incorporated into mathematical programming models. The transformation into linear equations suitable for linear programming (LP) purposes was first proposed by Ziliaskopoulos (2000) in which an LP was presented for the single-destination system optimum dynamic traffic assignment problem. The user optimal formulation was later introduced by Ukkusuri (2002).

In the current model, cells can be thought of as (part of) some infrastructure element (e.g., roads, bridges, and guardrails) that needs to be maintained. Maintenance gives rise to capacity reductions on the road, for example, in the form of reduced speeds and detours. The idea behind the proposed model is to propagate traffic from the source cells to the sink cell when travelers simultaneously attempt to reach the destination in the shortest possible time, while minimizing the total system travel time and the monetary cost. In this process, it is guaranteed that the state of the infrastructure is above a predetermined critical service level. The total system travel time can be used as a proxy for user delay. Henceforth, we shall use these two terms interchangeably.

2.2 Deterioration curves

In this article, maintenance can be interpreted in a very broad sense. Every activity related to maintenance, repair, and rehabilitation (MR&R) can be captured in our model provided that it gives rise to capacity reductions on the road. This requires the specification of maintenance activity durations, associated costs, and induced reliability index profiles.

Reliability index profiles are discussed in relation to bridge maintenance by Kong and Frangopol (2003). However, they can easily be adopted for more general use. To illustrate reliability index profiles, consider a piece of infrastructure whose reliability index deteriorates according to some empirically determined function $r(t)$, where t denotes time. In the upper part of Figure 1, we have depicted a typical quadratic reliability index function. Suppose that there is a certain type of maintenance that yields instantaneous improvement

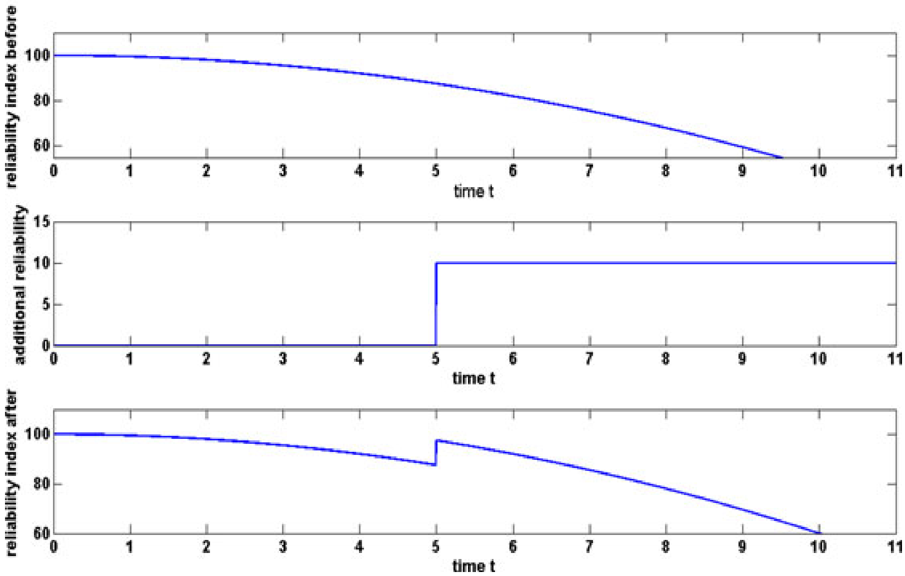

Fig. 1. Reliability index profile before maintenance (*upper*); additional reliability profile induced by maintenance (*middle*); the result of the superposition of the original reliability function with the induced profile (*lower*).

in the reliability index of 10 index points (on a scale of 100) of the infrastructure element under consideration (other possible induced reliability profiles are discussed by Kong and Frangopol, 2003) and assume that this maintenance is done at time $t = 5$ (see the middle part of Figure 1). The state of the infrastructure element under this maintenance action is obtained by the superposition of the induced reliability profile with the original reliability function $r^0(t)$ (see the lower part of Figure 1).

Maintenance cost can be determined in a similar fashion to account for the fact that timely maintenance decreases maintenance costs in the future (Maji and Jha, 2007). The exact mathematical expressions for the maintenance costs are given in the next section in which we present the proposed bi-level maintenance planning program.

2.3 A mixed-integer bi-level program

The proposed maintenance planning model is a bi-level program (Bard, 1998). In the upper level, maintenance agencies decide on the maintenance locations and the type of maintenance to be performed. They anticipate all possible reactions in terms of route choice of the road users and try to minimize the maintenance cost as well as user delay. In the lower level, individual road users choose their routes such that their travel times are minimized. CTM ensures an accurate representation of traffic flow.

To date, no formal model statement for the maintenance planning problem can be found in the literature when traffic dynamics is accounted for, be it single- or multiple-destination problems. In the following, we present a formal description of the single-destination formulation of the planning model. The solution procedure easily allows the solution of multiple-destination problems.

Sets

$C =$ set of cells
$C_S =$ sink cell
$C_R =$ set of source cells
$T =$ set of disjoint time intervals
$M =$ set of time intervals at which maintenance is performed. Note that M is a subset of T
$E =$ set of cell connectors
$E_S =$ set of sink cell connectors
$FS(i) =$ cell connectors emanating from cell i
$RS(i) =$ cell connectors incident to cell i
$A_i =$ set of maintenance activities for cell i

Parameters

$U_t =$ cost at time t that yields user-optimal flows
$\lambda =$ weight parameter chosen by the decision maker
$B_m =$ available budget at time interval m
$\delta_i^t =$ ratio of link free-flow speed and backward propagation speed for cell i at time t
$\zeta_i =$ initial number of vehicles in cell i
$d_i^t =$ vehicle inflow/demand at cell i at time t
$N_i^t =$ maximum possible number of vehicles in cell i at time t

$Q_i^t =$ maximum number of vehicles that can flow into or out of cell i at time t
$\chi_{ij} =$ change in N_i^t if maintenance of type j is done on cell i
$\phi_{ij} =$ change in Q_i^t if maintenance of type j is done on cell i
$r_i^{crit} =$ the critical reliability level of cell i
$r_{it}^0 =$ the initial reliability function of cell i (as a function of t)
$\Delta r_{mij} =$ additional reliability induced by maintenance of type j on cell i at time m
$\gamma_{mij}^0 =$ cost of maintenance of type j on cell i at time m given no prior maintenance
$\Delta\gamma_{mij} =$ reduction in maintenance cost for cell i given that maintenance of type j is done on cell i at time m
$l_j =$ length of duration of maintenance type j
$H =$ a large number

Decision variables

$x_i^t =$ number of vehicles in cell i at time t
$y_{ij}^t =$ number of vehicles moving from cell i to cell j at time t
$z_{mij} =$ indicator variable. It equals 1 if maintenance of type j is done on cell i at time m, 0 otherwise

Model formulation

$$\min_{x,y,z} \sum_{t\in T}\sum_{(i,j)\in E_s} t\cdot y_{ij}^t + \lambda \sum_{m\in M}\sum_{i\in C\backslash C_S}\sum_{j\in A_i} z_{mij} \times \left(\sum_{j\in A_i}\gamma_{mij}^0 z_{mij} + \sum_{k<m}\sum_{j\in A_i}(\Delta\gamma_{kij}z_{kij})\right) \tag{1}$$

Subject to

$$r_{it}^0 + \sum_{j\in A_i}\sum_{m\le t}\Delta r_{t-m,ij} z_{mij} \ge r_i^{crit} \quad \forall i \in C\backslash C_S, \forall t\in T \tag{2}$$

$$\sum_{j\in A_i}\sum_{i\in C\backslash C_S}\gamma_{mij}^0 z_{mij} + \sum_{i\in C\backslash C_S}\sum_{k<m}\sum_{j\in A_i}(\Delta\gamma_{kij}z_{kij}) - B_m \le H(1-z_{mij}) \quad \forall i\in C\backslash C_S, \forall m\in M, \forall j\in A_i \tag{3}$$

$$z_{mij}\in\{0,1\} \quad \forall i\in C\backslash C_S, \forall m\in M, \forall j\in A_i \tag{4}$$

$$\min_{x,y} \sum_{(i,j)\in E_S}\sum_{t\in T}\left(U_t\cdot y_{ij}^t\right) \tag{5}$$

Subject to

$$N_i^t = N_i^0 + \sum_{m\le t}\sum_{j\in A_i}\chi_{ij} z_{mij} I_{[m\le t\le m+l_j]} \quad \forall i\in C\backslash C_S, \forall t\in T \tag{6}$$

$$Q_i^t = Q_i^0 + \sum_{m\le t}\sum_{j\in A_i}\phi_{ij} z_{mij} I_{[m\le t\le m+l_j]} \quad \forall i\in C\backslash C_S, \forall t\in T \tag{7}$$

$$x_i^t - x_i^{t-1} + \sum_{(i,j)\in FS(i)} y_{ij}^{t-1} - \sum_{(j,i)\in RS(i)} y_{ji}^{t-1} = d_i^t \quad \forall i\in C\backslash C_S, \forall t\in T \tag{8}$$

$$\sum_{(i,j)\in FS(i)} y_{ij}^t - x_i^t \le 0 \quad \forall i\in C\backslash C_S, \forall t\in T \tag{9}$$

$$\sum_{(j,i)\in RS(i)} y_{ji}^t \le \delta_i^t\left(N_i^t - x_i^t\right) \quad \forall i\in C\backslash(C_R\cup C_S), \forall t\in T \tag{10}$$

$$\sum_{(j,i)\in RS(i)} y_{ji}^t \le Q_i^t \quad \forall i\in C\backslash(C_R\cup C_S), \forall t\in T \tag{11}$$

$$\sum_{(i,j)\in FS(i)} y_{ij}^t \le Q_i^t \quad \forall i\in C\backslash C_S, \forall t\in T \tag{12}$$

$$x_i^0 = \zeta_i \quad \forall i\in C\backslash C_S \tag{13}$$

$$y_{ij}^0 = 0 \quad \forall (i,j)\in E \tag{14}$$

$$x_i^{|T|} = 0 \quad \forall i\in C\backslash C_S \tag{15}$$

$$x_i^t \ge 0 \quad \forall i\in C\backslash C_S, \forall t\in T \tag{16}$$

$$y_{ij}^t \ge 0 \quad \forall (i,j)\in E, \forall t\in T \tag{17}$$

Equation (1) denotes the (maintenance agency's) objectives of minimizing the total system travel time over the entire planning horizon (first term) and the maintenance cost (second term). The weight $\lambda \ge 0$ can be used by the decision maker to indicate the relative importance of each objective. Equations (2)–(4) are maintenance-related constraints. More specifically, Equation (2) states that the reliability level of each cell must be maintained above some critical service level, accounting for its maintenance history. Equation (3) expresses the constraint that the budget is not to be exceeded at any time. In this constraint we acknowledge that maintenance done in the past reduces the present maintenance cost. Note that the constant H is chosen in such a way (i.e., sufficiently large) that the constraint is trivially satisfied in case no maintenance is performed at time interval m. In Equation (4) we state the binary nature of the maintenance decision variable. Equation (5) is the objective function that yields user optimal (UO) flows (Ukkusuri, 2002; Ukkusuri and Waller, 2008), provided that U_t satisfies

$$U_t - U_{t-1} > (U_{|T|} - U_t) \sum_{t,i} d_i^t$$

and where $|T|$ denotes the cardinality of the set T. (In the literature, certain publications have made the distinction between UO and user equilibrium (UE), although others have considered them to be equivalent. Here, we are essentially employing the UE definition, but prefer the term UO because for a CTM-based model true equilibrium may not exist (i.e., all used paths might not be equal in travel time due to the noncontinuous response of flows in CTM), but the concept of users being unable to unilaterally improve their paths is maintained (i.e., anticipatory user optimality is achievable although true equilibrium is not present). For a more detailed discussion, we refer to Waller and Ziliaskopoulos, 2006.) The remaining equations impose traffic flow restrictions that are governed by CTM. Because maintenance decisions will reduce cell capacities, we need to keep track of these, which is accomplished via Equation (6). Here $I_{[t \geq a]}$ denotes the indicator function that equals one if and only if $t \geq a$; it equals zero otherwise. For example, the capacity of a cell at time t will be affected by a previous maintenance effort (of type j performed at time $m < t$), as long as $m \leq t \leq m + l_j$, where l_j denotes the duration of the maintenance. Equation (7) represents a similar relation for vehicle flows. Equation (8) imposes conservation of flow at the cell level, where d_i^t denotes the exogenous traffic inflow (demand) at time t into cell i. Further, Equation (9) limits the number of outgoing vehicles from cell i to the number it contains. Equation (10) expresses the same limitation, except that it considers ingoing vehicles, while incorporating possible capacity reductions. Similar relations for cell connectors are given in Equations (11)–(12). In Equations (13)–(15), we enforce initial/final conditions on the cell network. The last two constraints, Equations (16) and (17), acknowledge the nonnegative nature of the decision variables in question.

As a final remark, we want to note that in case the physical dimension of some infrastructure (e.g., a bridge) spans over several cells, and if the facility has to be maintained in its entirety, we can simply add an additional constraint to ensure that consecutive cells are maintained at the same time.

3 SOLUTION PROCEDURE: A GENETIC ALGORITHM APPROACH

In Section 2 we presented a mixed-integer bi-level program for the long-term infrastructure maintenance planning problem. Bi-level problems are known to be NP-hard (Bard, 1998). We will base our solution strategy on genetic algorithms (GAs). GAs have a demonstrated record of solving challenging optimization problems (e.g., see Adeli and Cheng, 1993; Adeli and Cheng, 1994a,b; Chan et al. 1994; Fwa et al., 1994; Adeli and Kumar, 1995a,b; Kim and Adeli, 2001; Sarma and Adeli, 2001; and more recently, Ukkusuri et al., 2007; Dridi et al., 2008). Further, GAs allow for a parallel implementation to further reduce the computational time for problems involving a large number of decision variables (Cantú-Paz, 2000). This strategy is, however, outside the scope of the current article. (Adeli and Cheng, 1994b, and Adeli and Kumar, 1995b, were the first to apply parallel GAs to civil engineering problems.) Next we present a concise review of GAs and their particular use in this article. For a more extensive discussion we refer to Goldberg (1989).

GAs were first developed by Holland (1975). A GA is an iterative procedure aimed at finding the global optimum in complex optimization problems. At each iteration, it maintains a population of candidate solutions (also known as chromosomes) that evolve according to principles from natural selection (survival of the fittest) and genetics (crossover and mutation).

The way we represent candidate maintenance schedules in terms of chromosomes is depicted in Figure 2. The planning horizon consists of T years and the total number of cells in the transportation network is equal to I. Furthermore, in each year we can select from J_i types of maintenance activities for cell i. The first TJ_1 entries of the chromosome represent the maintenance activities for cell 1 over the entire planning horizon. A value of one indicates that the specific maintenance type will be performed for cell 1 in the year in question; otherwise the entry will contain a zero value. The next TJ_2 entries encode the same information, but for cell 2 and so on.

The operation of mutation can now be defined as the random replacement of 1s with 0s (or vice versa), where each entry in the chromosome is subjected to this replacement operation with some small mutation probability and independent of other entries. Crossover, on the other hand, is a pairwise operation. That is, chromosomes are (randomly) ordered in pairs and a random crossover point in the chromosomes is determined. The partial maintenance schedules on one side of the crossover point are then interchanged, resulting in two new chromosomes.

The fitness of each chromosome is determined by the total system travel time it induces over the planning period and its associated monetary cost. The total system travel time is determined via mesoscopic traffic simulation. For this purpose, we use the visual interactive

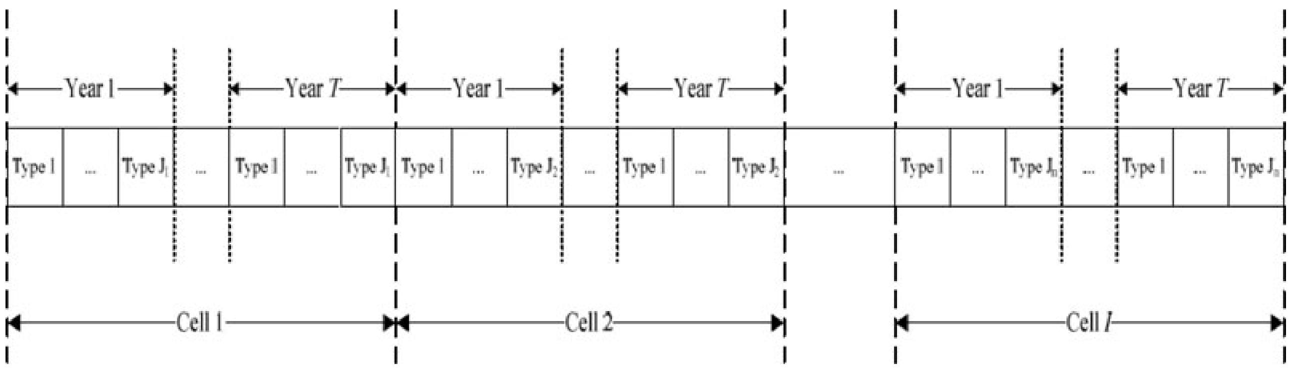

Fig. 2. Chromosome representation of maintenance schedules.

system for transport algorithms (VISTA). VISTA is an innovative transportation modeling framework that integrates spatiotemporal data and models for a whole spectrum of transportation applications (Ziliaskopoulos and Waller, 2000). Its traffic simulator is RouteSim, which is a mesoscopic simulator based on CTM (Ziliaskopoulos and Lee, 1996).

4. NUMERICAL EXAMPLE

The numerical example in this section is based on the hypothetical network shown in Figure 3. Chang et al. (2001) used a similar-sized network for demonstration purposes for the scheduling of short-term maintenance activities. The numerical experiment is primarily intended to demonstrate the workings of the proposed model.

The test network consists of 13 nodes, 4 of which represent centroids (nodes 100, 101, 102, and 103). Centroids form the origins and destinations for all traffic. Furthermore, there are 24 two-lane links (numbered 1 to 24 in the figure). They are chosen in such a way that each link corresponds to one cell. The links connecting to the centroids are not considered for maintenance. The origin–destination (OD) table for year 1 is presented in Table 1, and the hourly distribution of the OD demand can be found in Table 2. For instance, 5% of the OD demand departs in the time interval from 0 to 300 seconds and 10% of the OD demand departs between 900 and 1,200 seconds. The OD demand in subsequent years is obtained randomly based on the base year data. It should be noted that the test network has multiple destinations.

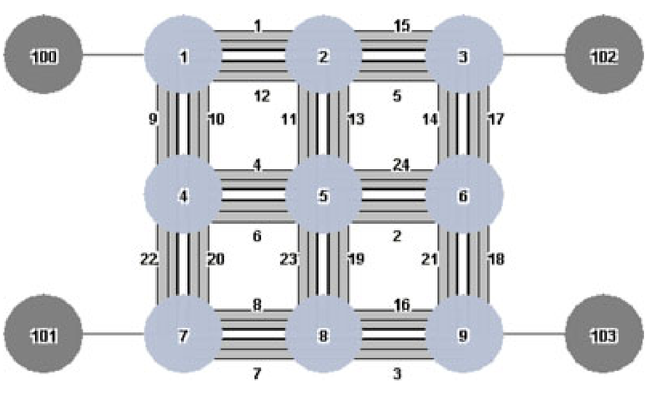

Fig. 3. Test network.

We consider a planning period of 3 years. For simplicity, we assume that maintenance only takes place at the beginning of a year and that it completes at the end of that same year, that is, $l_j = 1$ year for all j. Furthermore, maintenance reduces the capacity of a cell with 50% during an entire year (note that for real-life applications, this assumption might be relaxed to account for time-dependent capacity reductions), and saturation flow rates on the different links were arbitrarily chosen, having values between 500 and 3,800 vph. Based on this assumption, the capacity of the road network changes at most once a year. To determine the total system travel time induced by a given maintenance plan, we perform traffic simulation. We simulated 3 hours (using a simulation time step of 6 seconds)—each for 1 year of the planning horizon—as a representative peak period evaluation. One typical peak hour simulation may not be representative of the total system travel time

Table 1
Base year OD table

Origin	*Destination*	*Demand*	*Origin*	*Destination*	*Demand*
100	101	100	102	100	100
100	102	100	102	101	100
100	103	100	102	103	100
101	100	100	103	100	100
101	102	100	103	101	100
101	103	100	103	102	100

Table 2
Hourly OD distribution

Start time (second)	*Duration (seconds)*	*Percentage (%)*
0	300	5
300	300	5
600	300	5
900	300	10
1,200	300	10
1,500	300	15
1,800	300	15
2,100	300	10
2,400	300	10
2,700	300	5
3,000	300	5
3,600	300	5

over the entire project duration. However, our contention is that a model that examines a sample of the total project duration while maintaining detailed traffic flow representation (via dynamic traffic modeling) is superior to a model that treats traffic through static conditions (thereby omitting the impact of bottlenecks and traffic dynamics). A similar strategy was used in Chang et al. (2001) and Cheu et al. (2004) to reduce the computational time in the context of short-term maintenance planning. Further, we assumed an annual maximum budget of 650 units. Without loss of generality, we set $\lambda = 0$ (i.e., we only consider system travel time as our objective) and we allow for one type of maintenance only. Each time maintenance is performed, the reliability level of a cell (on a scale of 100) increases by 5 index points. We use the following functions to model deterioration and cost $r_{it}^{0} = r_i^{initial} - t^2$, $\gamma_{mij}^{0} = 5t^2$ and $\Delta\gamma_{mij} = -10$, where $r_i^{initial}$ denotes the reliability index of cell i at the beginning of the planning horizon. Values for $r_i^{initial}$ and r_i^{crit} are summarized in Table 3.

Table 3
$r_i^{initial}$ and r_i^{crit}

Cell	$r_i^{initial}$	r_i^{crit}	*Cell*	$r_i^{initial}$	r_i^{crit}
1	80	75	13	78	70
2	82	80	14	90	75
3	90	90	15	85	81
4	80	76	16	90	72
5	80	75	17	85	80
6	90	70	18	80	80
7	90	85	19	81	77
8	80	75	20	90	73
9	70	70	21	86	71
10	85	85	22	83	80
11	90	90	23	90	75
12	89	73	24	80	80

Table 4
Maintenance schedule

Cell	*Year 1*	*Year 2*	*Year 3*	*Cell*	*Year 1*	*Year 2*	*Year 3*
1	–	X	X	13	–	X	X
2	X	X	X	14	–	X	–
3	X	X	X	15	X	X	X
4	X	X	–	16	X	X	X
5	X	–	X	17	X	X	–
6	X	X	X	18	X	X	X
7	–	X	X	19	X	X	–
8	–	X	X	20	–	X	X
9	X	X	X	21	X	–	–
10	X	X	X	22	X	X	X
11	X	X	X	23	–	X	–
12	–	X	X	24	X	X	X

We employed a slight variation of the GAs developed at the Kanpur Genetic Algorithms Laboratory (http://www.iitk.ac.in/kangal/index.shtml). In these genetic algorithms, the penalty method was adopted to handle constraint violations: whenever a constraint violation is detected in the evaluation of the chromosome, the objective value incurs a large penalty that reduces the chance that the chromosome be considered in the next iteration. Our numerical results have demonstrated that this approach is very effective in obtaining feasible solutions. The number of generations used is 50. The population size equals 50 as well, whereas the crossover and mutation probabilities were set to be equal to 0.6 and 0.001, respectively. Karoonsoontawong and Waller (2006) have shown through extensive numerical testing that the above parameters perform very well for CTM-based network design problems. At each iteration, the DTA component evaluates the chromosomes by running traffic simulation. We used the Linux operating system with an Intel® 3.00GHz Xeon(tm) CPU, 32-GB memory and the gcc compiler (version 4.1.1). The computer code was implemented in C++. For the 50 GA iterations, we recorded a CPU time of 2,858 minutes and 12.55 seconds (about 2 days; cf. Cheu et al., 2004, and Ma et al., 2004). The budget spent in years 1, 2, and 3 are 80, 300, and 520 units, respectively. The maintenance schedule obtained is outlined in Table 4, where "X" represents the performing of maintenance. The resulting total system travel time is 6,479,595 seconds (about 75 days) and the cell reliability indices at the end of each year are given in Table 5.

In year 1, only 16 cells are maintained, compared to 22 and 18 cells in years 2 and 3, respectively. At the end of year 1, the sum of the differences between the cell reliabilities and their critical values is lowest in the entire planning period. If no maintenance were done in year 2, various cells would not have met their thresholds. This

Table 5
Cell reliability indices after maintenance

Cell	*Year 1*	*Year 2*	*Year 3*	r_i^{crit}	*Cell*	*Year 1*	*Year 2*	*Year 3*	r_i^{crit}
1	79	81	81	75	13	77	79	79	70
2	86	88	88	80	14	89	91	86	75
3	94	96	96	90	15	89	91	91	81
4	84	86	81	76	16	94	96	96	72
5	84	81	81	75	17	89	91	86	80
6	94	96	96	70	18	84	86	86	80
7	89	91	91	85	19	85	87	82	77
8	79	81	81	75	20	89	91	91	73
9	74	76	76	70	21	90	87	82	71
10	89	91	91	85	22	87	89	89	80
11	94	96	96	90	23	89	91	86	75
12	88	90	90	73	24	84	86	86	80

explains the relatively large number of maintenance activities in year 2. Finally, because year 3 is the last period in the planning horizon (i.e., the state of the infrastructure beyond year 3 is irrelevant to the model), relatively little maintenance is done to keep the total system travel time low. This suggests the use of this model on a rolling horizon basis.

5 CONCLUSIONS

In this article we have addressed a gap in the current literature by providing a decision support model that is intended for long-term infrastructure maintenance planning in which we minimize user delay as well as maintenance cost. Existing maintenance planning models can be categorized into two groups:

1. *Long-term and static*: In these models maintenance locations are determined dynamically over time. They do not consider traffic dynamics to quantify user cost. State-of-the-art planning models in this category account for user delay via static traffic equilibrium. We introduced a model based on dynamic traffic equilibrium, which is inherently more realistic than static models.
2. *Short-term and dynamic*: Maintenance locations are assumed to be given in these planning models. Typically, traffic dynamics is accounted for by the use of microsimulation. We employed mesoscopic simulation to circumvent complicating issues associated with microscopic traffic simulation. Furthermore, mesoscopic modeling allows for a formal mathematical problem statement, albeit currently only for the single-destination problem.

The bi-level model presented in this article can be seen as a generalization of the short-term and dynamic planning models because it relaxes the assumption of the presence of prespecified maintenance locations. Consequently, the solution of the presented model is computationally much more challenging. We presented a numerical example that illustrated the model and our hybrid solution strategy (GA in conjunction with mesoscopic traffic simulation). Due to the computational burden imposed by traffic dynamics (Cheu et al., 2004; Ma et al., 2004), we employed a modest grid network and analysis. However, the analysis provides a clear demonstration of the approach and is consistent with fundamental research in this area, particularly considering the novelty of modeling network traffic dynamics (Peeta and Ziliaskopoulos, 2001). Further refining the solution method parameters and devising increasingly efficient optimization techniques (e.g., parallel genetic algorithms) will be a long-term research effort, which will proceed from the basis provided here.

REFERENCES

Adeli, H. & Cheng, N.-T. (1993), Integrated genetic algorithm for optimization of space structures, *Journal of Aerospace Engineering, ASCE*, **6**(4), 315–28.

Adeli, H. & Cheng, N.-T. (1994a), Augmented Lagrangian genetic algorithm for structural optimization, *Journal of Aerospace Engineering, ASCE*, **7**(1), 104–18.

Adeli, H. & Cheng, N.-T. (1994b), Concurrent genetic algorithms for optimization of large structures, *Journal of Aerospace Engineering, ASCE*, **7**(3) 276–96.

Adeli, H. & Kumar, S. (1995a), Concurrent structural optimization on a massively parallel supercomputer, *Journal of Structural Engineering, ASCE*, **121**(11), 1588–97.

Adeli, H. & Kumar, S. (1995b), Distributed genetic algorithms for structural optimization, *Journal of Aerospace Engineering*, **8**(3), 156–63.

Al-Subhi, K. M., Johnston, D. W. & Farid, F. (1990), A resource constrained capital budgeting model for bridge maintenance, rehabilitation and replacement, *Transportation Research Record*, **1268**, 110–17.

Bard, J. F. (1998), *Practical Bi-level Optimization Algorithms and Applications*, Kluwer Academic Publishers, Dordrecht, the Netherlands.

Bellman, R. E. (1957), *Dynamic Programming*, Princeton University Press, Princeton, NJ.

Bertsekas, D. P. (2001), *Dynamic Programming and Optimal Control*, vol. 2, 2nd edn., Athena Scientific, Belmont, MA.

Burghout, W., Koutsopoulos, H. N. & Andréasson, I. (2005), Hybrid mesoscopic-microscopic traffic simulation, *Transportation Research Record*, **1934**, 218–55.

Cantú-Paz, E. (2000), *Efficient and Accurate Parallel Genetic Algorithms*, Kluwer Academic, Boston.

Carnahan, J. V. (1988), Analytical framework for optimizing pavement maintenance, *Journal of Transportation Engineering*, **114**, 307–22.

Chan, W. T., Fwa, T. F. & Tan, C. Y. (1994), Road maintenance planning using genetic algorithms I: formulation, *Journal of Transportation Engineering*, **120**(5), 693–709.

Chang, Y. Y., Sawaya, O. B. & Ziliaskopoulos, A. K. (2001), A tabu search based approach for work zone scheduling. *Proc. 80th Annual Meeting, Transportation Research Board*, Washington, DC.

Cheu, R. L., Wang, Y. & Fwa, T. F. (2004), Genetic algorithm-simulation methodology for pavement maintenance scheduling, *Computer-Aided Civil and Infrastructure Engineering*, **19**(6), 446–55.

Chu, C.-Y. & Durango-Cohen, P. L. (2007), Estimation of infrastructure performance models using state-space specifications of time series models, *Transportation Research Part C*, **15**(1), 17–32.

Chu, C.-Y. & Durango-Cohen, P. L. (2008a), Incorporating maintenance effectiveness in the estimation of dynamic infrastructure performance models, *Computer-Aided Civil and Infrastructure Engineering*, **23**(3), 174–88.

Chu, C.-Y. & Durango-Cohen, P. L. (2008b), Estimation of dynamic performance models for transportation infrastructure using panel data, *Transportation Research Part B*, **42**(1), 57–81.

Daganzo, C. F. (1994), The cell transmission model: a simple dynamic representation of highway traffic consistent with the hydrodynamic theory, *Transportation Research Part B*, **28**(4), 269–87.

Daganzo, C. F. (1995), The cell transmission model, part II: network traffic, *Transportation Research Part B*, **29**(2), 79–93.

Dridi, L., Parizeau, M., Mailhot, A. & Villeneuve, J. -P. (2008), Using evolutionary optimisation techniques for scheduling water pipe renewal considering a short planning horizon, *Computer-Aided Civil and Infrastructure Engineering*, **23**(8), 625–35.

Frangopol, D. M., Kong, J. S. & Gharaibeh, E. S. (2001), Reliability-based life-cycle management of highway bridges, *Journal of Computing in Civil Engineering*, *ASCE*, **15**(1), 27–34.

Fwa, T. F., Cheu, R. L. & Muntasir, A. (1998), Scheduling of pavement maintenance to minimize traffic delays, *Transportation Research Record*, **1650**, 28–35.

Fwa, T. F., Tan, C. Y. & Chan, W. T. (1994), Road maintenance planning using genetic algorithms II: analysis, *Journal of Transportation Engineering*, *ASCE*, **120**(5), 710–22.

Golabi, K., Kulkarni, R. & Way, G. (1982), A statewide pavement management system, *Interfaces*, **12**, 5–21.

Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA.

Guignier, F. & Madanat, S. (1999), Optimization of infrastructure systems maintenance and improvement policies, *Journal of Infrastructure Systems*, *ASCE*, **5**(4).

Guillaumot, V. M., Durango-Cohen, P. L. & Madanat, S. (2003), Adaptive optimization of infrastructure maintenance and inspection decisions under performance model uncertainty, *Journal of Infrastructure Systems*, **9**(4), 133–39.

Holland, J. H. (1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI.

Jacobs, T. L. (1992), Optimal long-term scheduling of bridge deck replacement and rehabilitation, *Journal of Transportation Engineering*, *ASCE*, **118**(2), 312–22.

Jiang, X. & Adeli, H. (2003), Freeway work zone traffic delay and cost optimization model, *Journal of Transportation Engineering*, *ASCE*, **129**(3), 230–41.

Jiang, X. & Adeli, H. (2004), Object-oriented model for freeway work zone capacity and queue delay estimation, *Computer-Aided Civil and Infrastructure Engineering*, **19**(2), 144–56.

Karim, A. & Adeli, H. (2003), Radial basis function neural network for work zone capacity and queue estimation, *Journal of Transportation Engineering*, *ASCE*, **129**(5), 494–503.

Karoonsoontawong, A. & Waller, S. T. (2006), Dynamic continuous network design problem: linear bi-level programming and metaheuristic approaches, *Journal of the Transportation Research Board*, **1964**, 104–17.

Kim, H. & Adeli, H.(2001), Discrete cost optimization of composite floors using a floating point genetic algorithm, *Engineering Optimization*, **33**(4), 485–501.

Kong, J. S. & Frangopol, D. M. (2003), Life-cyle reliability-based maintenance cost optimization of deteriorating structure with emphasis on bridges, *Journal of Structural Engineering*, *ASCE*, **129**(6), 818–28.

Li, Y., Ziliaskopoulos, A. K. & Waller, S. T. (1999), Linear programming formulations for system optimum dynamic traffic assignment with arrival time-based and departure time-based demands, *Transportation Research Record*, **1832**, 121–42.

Lighthill, M. J. & Whitham, J. B. (1955a), On kinematic waves I: flood movement in long rivers, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **229**(1178), 281–316.

Lighthill, M. J. & Whitham, J. B. (1955b), On kinematic waves II: a theory of traffic flow on long crowded roads, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **229**(1178), 317–45.

Ma, W., Cheu, R. L & Lee, D.-H. (2004), Scheduling of lane closures using genetic algorithms with traffic assignments and distributed simulations, *Journal of Transportation Engineering*, **130**(3), 322–29.

Madanat, S. & Wan Ibrahim, W. H. (1995), Poisson regression models of infrastructure transition probabilities, *Journal of Transporation Engineering*, *ASCE*, **21**(3), 267–72.

Maji, A. & Jha, M. K. (2007), Modeling highway infrastructure maintenance schedule with budget constraint, *Transportation Research Record*, **1991**, 19–26.

Mishalani, R. & Madanat, S. (2002), Computation of infrastructure transition probabilities using stochastic duration models, *Journal of Infrastructure Systems*, *ASCE*, **8**(4), 139–48.

Miyamoto, A., Kawamura, K. & Nakamura, H. (2000), Bridge management system and maintenance optimization for existing bridges, *Computer-Aided Civil and Infrastructure Engineering*, **15**, 45–55.

Ouyang, Y. (2007), Pavement resurfacing planning for highway networks: parametric policy iteration approach, *Journal of Infrastructure Systems*, **13**(1), 65–71.

Ouyang, Y. & Madanat, S. (2004), Optimal scheduling of rehabilitation activities for multiple pavement facilities: exact and approximate solutions, *Transportation Research Part A*, **38**(5), 347–65.

Pasupathy, R., Labi, S. & Sinha, K. C. (2007), Optimal reconstruction periods for stochastically deteriorating infrastructures, *Computer-Aided Civil and Infrastructure Engineering*, **22**(6), 389–99.

Paterson, W. D. O. (1987), *Road Deterioration and Maintenance Effects: Models for Planning and Management*, The Johns Hopkins University Press, Baltimore, MD.

Peeta, S. & Ziliaskopoulos, A. K. (2001), Foundations of dynamic traffic assignment: the past, the present and the future, *Networks and Spatial Economics*, **1**(3–4), 233–66.

Richards, P. I. (1956), Shockwaves on the highway, *Operations Research*, **4**, 42–51.

Sarma, K. C. & Adeli, H. (2001), Bi-level parallel genetic algorithms for optimization of large steel structures, *Computer-Aided Civil and Infrastructure Engineering*, **16**(5), 295–304.

Smilowitz, K. & Madanat, S. (2000), Optimal inspection and maintenance policies for infrastructure networks, *Computer-Aided Civil and Infrastructure Engineering*, **15**, 5–13.

Tsunokawa, K. & Schofer, J. L. (1994), Trend curve optimal control model for highway pavement maintenance: case study and evaluation, *Transportation Research Part A*, **28**(2), 151–66.

Uchida, K. & Kagaya, S. (2006), Development of life cycle cost evaluation model for pavements considering drivers' route choices, *Transportation Research Record*, **1985**, 115–24.

Ukkusuri, S. V. S. K. (2002), Linear programs for the user-optimal dynamic traffic assignment problem. Master thesis, University of Illinois at Urbana-Champaign.

Ukkusuri, S. V. S. K., Tom, V. M. & Waller, S. T. (2007), Robust transportation network design under demand uncertainty, *Computer-Aided Civil and Infrastructure Engineering*, **22**(1), 6–18.

Ukkusuri, S. V. S. K. & Waller, S. T. (2008), Linear programming models for the user and system optimal dynamic network design problem: formulations, comparisons and extensions, *Networks and Spatial Economics*, **8**(4), 383–406.

Waller, S. T. & Ziliaskopoulos, A. K. (2006), A combinatorial user optimal dynamic traffic assignment algorithm, *Annals of Operations Research*, **144**(1), 249–61.

Ziliaskopoulos, A. K. (2000), A linear programming model for the single destination system optimum dynamic traffic assignment problem, *Transportation Science*, **34**(1), 37–49.

Ziliaskopoulos, A. K. & Lee, S. (1996), A cell transmission based assignment-simulation model for integrated freeway/surface street systems, *Transportation Research Record*, **1701**, 12–23.

Ziliaskopoulos, A. K. & Waller, S. T. (2000), An internet-based geographic information system that integrates data, models and users for transportation applications, *Transportation Research Part C*, **8**, 427–44.