Associative Memory Techniques for Quick Estimation of Network Performance Measures

R. Palavadi Naga and Yue Yue Fan

Many important decision-making processes in transportation planning and engineering involve repetitive computation of network performance measured by total network delay, throughput, network efficiency, and other measures. The computational complexity imposed by repetitive evaluation of these measures, especially under user equilibrium conditions, is a serious obstacle to timely decision making in network-related problems. This study applies associative memory techniques, which are conceptually and computationally simple, to quick estimation of these performance measures. The results of the numerical experiments are encouraging, and the relative error on average was found to be less than 2%. Furthermore, the applicability of this approximation method to bilevel network problems, a class of important but complex problems, is explored through a study of the network recovery problem, which seeks a quick and effective repair strategy for disturbed networks following natural or human-induced disasters.

Transportation engineering and planning studies usually require that individual transportation facilities in a region be studied as a whole system. Because of its spatial feature, a transportation system is often modeled as a network consisting of nodes and various links connecting these nodes. Standard network performance measures used in the field of transportation include total travel time, network efficiency, total throughput, and other measures. Almost every transportation system study involves evaluation of network performance criteria. To name a few, examples include cost-effectiveness analysis of various network improvement policies, loss assessment of natural and human-induced disasters (*1*), network resilience studies (*2*), and as part of the solution to a large number of important bilevel problems such as the network recovery problem (NRP) (*3, 4*).

However, resources required to compute these performance measures for networks of realistic size may be quite extensive, in spite of significant developments in theory and computation (*5–7*), especially if traffic is assumed to be in a user equilibrium (UE) condition. The Franke-Wolfe algorithm, for instance, required about 150 s of CPU time on a personal computer to solve one UE traffic assignment problem for highway network of the San Francisco Bay area in California, with about 1,120 zones and 10,647 nodes. The computational burden may not seem to matter if only a single computation is needed, but it becomes a serious drawback in the context of bilevel problems, in which the solution procedure involves repeated computation of performance criteria. Nevertheless, many times a quick and effective approximation of network performance measures may be acceptable for practical reasons, especially when a timely decision is needed.

In this paper, an alternative approach is introduced that uses the concept of associative memory (AM) to predict various network performance measures simultaneously. This approximation approach trades off a little of the accuracy of conventional methods for significant savings in computational time, thus potentially leading to a reduction in the decision-making time. AM techniques are inspired by the human memory and have broad applications in science and engineering, particularly in system identification and pattern recognition (*8*). The research presented here demonstrates good potential for application of AM techniques to transportation network problems. First, the construction and evaluation of AM models that predict the performance for a given network configuration are introduced. The usefulness of AM models in solving bilevel problems will be explored further in the context of the NRP, which seeks an optimal strategy for repairing damaged network components following a disturbance to the network.

ESTIMATION OF NETWORK PERFORMANCE CRITERIA WITH AM MODELS

Introduction of AM Techniques

AM models belong to the class of neural computing techniques and are basically mappings between ordered sets of input and output signals (Figure 1). The inspiration for constructing these models stems from the associative nature of the human memory, which connects items that are similar or contrary or if they occur simultaneously or in close succession (*8*). There are two different categories of AM models: autoassociative memories and heteroassociative memories. Autoassociative memories are capable of recollection of the complete version of a given incomplete or noisy pattern but cannot map a totally different key pattern that was not memorized earlier. Heteroassociative memories, however, are mappings between two different sets of patterns and thus are capable of producing an associated output pattern for any of the new input patterns. Heteroassociative memories are commonly used in situations involving parameter estimation.

The three main examples of such models are simple associative memory (SAM), recurrent associative memory (RAM), and multicriteria associative memory (MAM). The model used for estimation of network performance measures in this study is the SAM model. In simpler terms, the whole of the AM model in this study may be

Department of Civil and Environmental Engineering, Institute of Transportation Studies, University of California, Davis, CA 95616. Corresponding author: Y. Y. Fan, yyfan@ucdavis.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2039, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 75–82.
DOI: 10.3141/2039-09

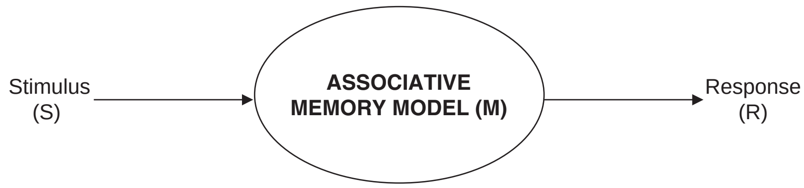

FIGURE 1 Representation of AM models.

thought of as a black box that transforms a given set of inputs into the set of required outputs. The details of stimulus and response vectors for this study are discussed later.

AM techniques have been applied in a number of situations in which unknown system parameter estimation was involved, including nonlinear system identification, passive ranging, remote sensing, and image processing (*9, 10*). Kalaba et al. (*9*) further showed the application of AM techniques even when the data available were noisy. Neural computing techniques, such as neural networks, have been used widely to solve highly complex and nonlinear transportation problems (*11*), particularly in areas regarding parameter estimation, driver behavior, traffic pattern analysis, and others. However, applications of the AM techniques to transportation are yet sparse in number, even though, as pointed out by Ferris et al. (*6*), more comprehensive comparative studies that incorporate new paradigms in neural computing techniques are needed.

The major advantage of AM models over the other neural computing methods is the resulting savings in time. The compact representation of AM models (e.g., polynomial) enables easier integration of these models into bilevel problems. However, the AM approach being a single-layered process, it can only at best be as good as the neural network approach, which is multilayered in most situations. Potential applications of SAM, RAM, and MAM to the problem of traffic assignment were investigated by Kim (*12*) to determine highway flows, and fairly good results were achieved. The major improvement of the AM model in the current work is the capability of simultaneously capturing several optimization problems and approximating multiple network performance measures at once.

Definitions of Network Performance Measures

Two network performance measures are chosen to demonstrate the capability of AM models in simultaneously estimating multiple performance measures: total travel time and global efficiency.

Total Travel Time

Total travel time is one of the most commonly used system performance measures and constitutes the objective function of a number of bilevel problems such as the network design–capacity expansion problem and the optimal pricing problem. The expression for computing this measure is

$$\text{total travel time} = \text{TT} = \sum_a V_a \times t_a\left(V_a\right)$$

where V_a is the volume on link *a*, and $t_a(V_a)$ is the time needed to traverse link *a* with volume V_a.

The link volumes here are usually determined under a UE assumption, given the physical representation of the network and the origin–destination (O-D) demand matrix. Commonly adopted functions describing the relationship between link travel times and flows are usually nonlinear, convex, and monotone increasing. In spite of considerable progress in developing solution algorithms for the UE problem, the computational resources needed to arrive at the volumes are not insignificant, especially for large transportation networks.

Global Efficiency

Global efficiency was introduced by Latora and Marchiori (*13*) for assessing the efficiency of transporting information (or people) in networks. The global efficiency of a network may be used as a measure for quantifying the ease with which people can move around in the network. This measure is widely used in network vulnerability and resilience studies (*2*). The expression for computing the global efficiency given the physical configuration of the network is

$$\text{global efficiency} = \frac{1}{N\left(N-1\right)}\sum_{i,j}^{N}\frac{1}{t_{ij}}$$

where N is the total number of nodes in the network and t_{ij} is the shortest path from node i to j.

Computation of global efficiency requires finding the shortest path between every node pair in the network. However, the shortest paths are based on the physical distances between the nodes. Thus this measure is a property of the network configuration alone and does not account for congestion effects.

Construction of AM Models

An AM model, as described earlier, establishes a mapping between a stimulus (input) vector and a response (output) vector. The AM models in the context of this study are required to predict the network performance measures for different realizations of network configurations. The stimulus vector thus needs to represent the network configuration in some manner, whereas the response vector would be constituted by the various performance measures. Simple AM models operate on the given stimulus vector to generate the response vector in the following manner:

$$R = MS$$

where R is the response vector, S is the stimulus vector, and M is the characteristic matrix associated with the AM model.

The crux of the AM approach thus lies in the construction of this transformation matrix M. The construction of AM matrices primarily involves two steps: training of the matrices and testing of the trained matrices. Before the training of matrix M is discussed, the functional form of the stimulus vector and the number of training cases to be used must be determined.

A natural way of constructing the stimulus vector would be to stack all the link capacities to form a column vector. This structure for the stimulus vector is referred to as the linear stimulus vector and the resultant mapping as linear associative mapping. However, the relationships between the stimulus and response vectors in most problems, especially in transportation, are highly nonlinear in nature. Inspired by Kalaba et al. (*9*), the authors investigated the performance of second-order polynomial (quadratic) AM models in which the stimulus vectors are constructed as follows:

$$S = \left[C_1, C_2, \ldots, C_{76}, C_1^2, C_2^2, \ldots, C_{76}^2, C_1C_2, \ldots, C_1C_{24}, \ldots, C_2C_3, \ldots, C_{75}C_{76}\right]^T$$

where C_i is the capacity of the ith link of the network.

This choice of stimulus vector has led to improved approximation. The downside of using such a stimulus vector in transportation network problems is the increased size of the stimulus vector. Thus, a structure with only partial terms of the quadratic AM model was also investigated as a part of this study, in which the stimulus vector is constructed as follows:

$$S = [C_1, C_2, \ldots, C_{76}, C_1^2, C_2^2, \ldots, C_{76}^2]^T$$

This functional form will be referred to as a partial quadratic structure.

Apart from the functional form of the stimulus, the number of training cases to be used is another important parameter that needs to be specified beforehand. There are no governing rules about the number of training cases that ought to be used in different problems. Numerical experiments were thus conducted with different numbers of training cases in the construction of matrix *M*. The results of these experiments will be discussed in the next section.

Once the functional form and the number of training cases are decided on, the AM models are constructed in the following manner. For the sake of simplicity, the construction procedure for linear associative memories is described first. At the initial step, different network configurations with up to 10% of the total number of links damaged were randomly selected. Damaged links are assumed to have only one-third of their original capacity. The UE model was used to compute the total travel time. Similarly, the lengths of damaged links were set to a high value and the global efficiency associated with each of the network configurations was computed. The two performance measures thus computed were then used to form the following response vector:

$$R = [\mathrm{TT}, \mathrm{GE}]^T$$

where TT represents the total travel time and GE is the global efficiency.

For each of the training cases, a stimulus vector (column vector of capacities) and a response vector were generated. The stimulus matrix ($\check{S}$) and the response matrix ($\check{R}$) were then constructed by aggregating all the stimulus and response vectors, respectively, with each column in $\check{S}$ and $\check{R}$ corresponding to each chosen training case. The mapping matrix *M* can be solved by

$$M = \check{R}.\check{S}^+$$

where $\check{S}^+$ is the Moore–Penrose generalized inverse of the stimulus matrix $\check{S}$.

The procedure for constructing nonlinear AM models is similar except that each individual stimulus vector is cast into the required structure by means of appropriate operations. The matrices so built were then tested with a different set of network configurations and the approximated measures were compared with the true measures. Detailed results from model validation are reported in the next section.

An implicit assumption when the total travel time is estimated by using the AM approach is that the demand for the transportation network is a constant across various networks. This static demand assumption, as will be shown later, can be relaxed.

Numerical Experiments and Results

A highly aggregated version of the network in Sioux Falls, South Dakota (24 nodes, 76 links), was used in this study to construct and evaluate the AM models. It is assumed that incidents affecting the Sioux Falls network can damage up to seven links of this network. AM models with the three different stimulus structures were constructed to quickly estimate the network performance at various damage levels. The training of the AM models was carried out by using up to 6,000 randomly chosen network configurations. The testing, however, was performed on a different set of 1,000 networks.

The following two error measures were used to evaluate the performance of the AM models:

- Average relative error. This statistic quantifies the error between a predicted and observed value for a single network performance criterion in the following manner:

$$\mathrm{ARE} = \frac{\sum\left(\left|r_{\mathrm{act}} - r_{\mathrm{obs}}\right| / r_{\mathrm{act}}\right)}{N}$$

where

N = total number of predictions,
r_{act} = actual value of network performance criterion, and
r_{obs} = predicted value of network performance criterion.

The foregoing statistic measures the goodness of the model with respect to each individual network performance criterion.

- Root-mean-square error. The root mean square associated with an AM model is defined in the following manner:

$$\mathrm{RMSE} = \sqrt{\frac{\sum\left(\mathrm{norm}\left(R_{\mathrm{act}} - R_{\mathrm{obs}}\right)\right)^2}{N}}$$

In addition, the standard deviations of the errors across all the test networks were computed. The performance of the three AM models is summarized in Table 1.

The testing of the AM matrices revealed that the errors in the predictions made by the AM models are reasonably low. The AM models seem to perform exceedingly well in the context of global efficiency (ARE = 0.27%). The ARE corresponding to total travel time was fairly low and was around 1.69%. The standard deviations of the errors were also found to be low. The higher values for errors in the total travel time might be due to the degree of variation in this measure across networks when compared with global efficiency. The models constructed with the quadratic stimulus

TABLE 1 Evaluation of AM Models

	Stimulus Structure		
	Linear	Partial Quadratic	Quadratic
Total travel time			
ARE	6.91%	2.86%	1.69%
Standard deviation of RE	5.45%	2.67%	1.53%
RMSE	8.37	4.32	2.32
Global efficiency			
ARE	0.64%	0.54%	0.27%
Standard deviation of RE	0.68%	0.64%	0.33%
RMSE	0.004	0.0037	0.0019

structure outperformed the other models in the prediction of both performance measures.

The consistency of the foregoing predictions was then checked by constructing 10 different AM matrices and subsequently testing them with 1,000 test cases each. The average ARE and RMSE across the 10 different models were about 1.71% and 2.66, suggesting that the differences across matrices might be minor. There might, however, be significant variation in the performance of models that are built with relatively fewer training cases since there is a higher possibility of underrepresentation of certain types of networks. This conjecture about variation in performance of AM matrices trained with fewer cases has yet to be supported by rigorous numerical experiments.

The results shown in Table 1 indicate that there is a marked increase in the performance of the models on progressing from linear to quadratic structures. There is a decrease of almost 5.22 percentage points in the ARE (total travel time) corresponding to the quadratic form when compared with the linear structure. A similar trend observed in RMSEs suggests that the relationship between the stimulus and response vectors is nonlinear. There is also a noticeable improvement from partial quadratic to quadratic AM models. This finding seems to indicate that the combination terms $C_i.C_j$ have a significant impact on the predictive power of the AM models. These terms may capture the effect of interactions between different links and thus contribute significantly to the performance of the models.

As mentioned in the previous section, the performance of various models that were constructed with different numbers of training cases was evaluated. As expected, a significant variation in performance of the AM models was observed on varying the number of training cases. It was observed that the AM models performed satisfactorily even when relatively fewer training cases are used. For instance, the total travel time ARE was found to be 4% to 6% for models that were constructed with even fewer than 1,000 training cases. The testing here was conducted by computing the average of the errors across 10 models with 1,000 cases for varying numbers of training cases.

Another interesting observation made across all testing cases is the large spike in the error values as the stimulus matrix approaches a square matrix. It is thought that this sudden extreme error is caused by the limited capability of MATLAB in handling the inversion of a nearly singular matrix when the matrix approaches a square and becomes highly ill-conditioned. A more rigorous investigation of this issue is still in progress. Once the stimulus matrix moves away from the square structure, the average prediction error decreases and seems to become stable after a certain number of training cases (about 4,000 cases in this study). It should be noted that even though training may require a significant amount of computational resources, this computation is done offline. Once the AM model is trained, it can be used in a real-time decision process to support quick computation.

The other performance measures that were approximated by using the AM techniques include the average volume-to-capacity ratio under UE conditions and the total throughput (maximal multicommodity flow). Fairly good results were obtained from numerical experiments with these measures (ARE for mean *V/C* ratio ~ 0.3%, ARE for total throughput ~ 8.5%).

To test for the nature of potential dependence of AM model effectiveness on network structures, the following experiments were carried out. Four different networks that were constituted by 25 nodes and varying numbers of links were considered for attempts at extending the AM models. The four networks here possess a progressively increasing degree of redundancy for the links. For instance, damage to even a single link in Network *a* (Figure 2) will lead to disconnection in the network. However, every node pair in Network *d* has several alternate paths, thus lessening the chances of disconnection.

AM models that predicted global efficiency were constructed for each of the four networks with up to 10% of the total number of links damaged in each case. The results of the model evaluation for each of these networks are summarized in Table 2. It can be observed that the ARE is less than 1% for most of the cases. The results of these preliminary tests suggested that the AM techniques for predicting global efficiency are very much applicable to other networks, both trans-

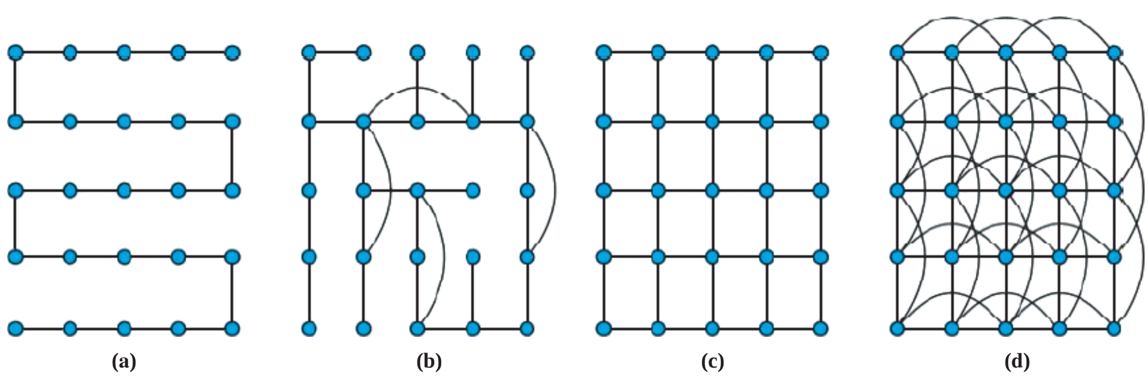

FIGURE 2 Four networks with different structures. (Courtesy of L. Dueñas-Osorio, Rice University.)

TABLE 2 Evaluation of AM Model for Total Travel Time

	Networks			
	(a)	(b)	(c)	(d)
Number of links	48	58	80	140
Number of damaged links	5	6	8	14
Average relative error (RE)	0.013	0.009	0.005	0.001
Std deviation of RE	0.012	0.01	0.005	0.001

portation networks and other types. Similar experiments for total travel time based on the UE model could not be carried out because of the lack of demand data for these networks.

Overall, the performance of the AM models coupled with the simplistic nature of the models make this approach appealing. The entire UE problem, as will be shown in the following section, can be replaced with a simple matrix multiplication in a number of situations in which a repeated solution to the UE problem is involved. A highly nonlinear problem can easily be transformed into a simple linear or polynomial one. Another major advantage of the AM approach is its ability to approximate a number of optimizations in one shot, thus producing estimates in a single operation. In addition, AM models when constructed by using data directly from the field can be particularly useful since they can map the stimulus to the response vectors without any prior knowledge of the relationship between stimuli and responses. Moreover, AM models have also been shown to be capable of handling noise in the input data by training the system with noisy data.

APPLICATION OF AM TO NRP

A large number of the problems being faced by transportation planners and system engineers can be formulated as bilevel optimization problems. The lower level of such problems typically involves an optimization problem, such as a traffic assignment problem, for computing the values of various network performance criteria. The higher level, in contrast, involves the solution to another optimization problem that maximizes or minimizes certain systemwide parameters that are of social importance. Bilevel problems have long been recognized as one of the most difficult problems to solve in the field of transportation (*4*). The general form of a bilevel program may be represented as follows:

$$\min_y F(x,y)$$

where y is the upper-level decision variable and x is the optimal solution of $\text{Min}_x\, f(x,z)$ such that $g(x,z) \leq b$. (F, f, and g represent general functions, z represents variables in the lower-level problem, and b is a constant.)

A typical example of a bilevel problem is the NRP, wherein it is necessary to determine an optimal repair strategy for a given damaged network. Consider a damaged network following a large-scale disaster such as an earthquake. If the resources available for repairing the network were unlimited, a straightforward decision would be to repair all damaged links. However, in reality, resources are often scarce, thus necessitating that repair jobs be ranked by priority. The challenge in planning an optimal recovery plan would then be to identify the set of damaged links that, when repaired, will yield the maximum societal benefit under certain resource (budget) constraints. Assuming the total travel time to be an appropriate system performance measure, the lower-level problem here would be to solve the UE problem for various network configurations resulting from feasible repair decisions. The upper level, however, involves determining the optimal repair decision such that the savings in total travel time (computed in the lower level) are maximized (given the total repair budget).

A conventional formulation of the aforementioned NRP is as follows:

maximize

$$\text{benefit} = \text{TT}(\Phi) - \text{TT}(X)$$

such that

$$\text{cost}(X) \leq \text{budget}$$

where

X = repair strategy (binary vector indicating links to be repaired),
TT (X) = total travel time computed from UE model when strategy X is implemented,
cost (X) = cost of implementing $X = C_0 \times N$,
C_0 = cost of repairing one damaged link (it should be noted that the cost of repairing different damaged links should be different in reality, but it is assumed identical in this study for the sake of data simplicity), and
Φ = null vector representing base scenario in which no repair is done.

It is assumed that recovery of a link is a binary decision; that is, partial recovery is not considered. Total time saving is set to be the objective function as an illustration; the choice of other network performance criteria or multiple objectives is permissive.

As mentioned earlier, repeated estimation of the UE flows at every step in the lower level of the NRP problem makes such a problem computationally burdensome. Apart from this difficulty, the nonconvex nature of the objective function in the upper level further complicates search algorithms for global optima (*4, 14*). However, the pervasive and critical nature of this class of problems, given the current threat level of natural or human-induced disasters, demands development of solution algorithms that can effectively produce approximate solutions.

The complexity of the foregoing problem is vastly reduced by incorporation of the AM models in this problem. The objective function here is transformed into a linear or quadratic function (depending on the stimulus structure), thus converting the problem into a standard integer programming problem. The steps involved in the solution algorithm of the NRP by the AM approach are outlined as follows:

1. Following the procedure described in the section on numerical experiments and results, a simple AM model was first constructed by using 6,000 training cases to establish a relationship between the network configuration and the total travel time under UE flow conditions. A quadratic structure was adopted for the stimulus vector, which was constituted by the link capacities of a network configuration. The response vector (R), in contrast, is a singleton set—the total travel time under UE flow conditions for the damaged network. The constructed AM model was then validated by using a different set

TABLE 3 Evaluation of AM Approach to NRP

Number of Damaged Links	Error in Objective Value						
	0%	0%–10%	10%–20%	20%–30%	30%–40%	40%–50%	>50%
2 links	96.00%	3.00%	0.00%	1.00%	0.00%	0.00%	0.00%
3 links	91.00%	5.50%	0.00%	1.00%	1.00%	0.00%	1.50%
4 links	87.00%	8.00%	3.33%	0.67%	0.67%	0.33%	0.00%
5 links	78.50%	15.75%	3.00%	1.75%	0.50%	0.25%	0.25%
6 links	75.00%	20.40%	2.40%	1.60%	0.20%	0.00%	0.40%
7 links	68.00%	26.50%	3.50%	1.67%	0.33%	0.00%	0.00%
All cases (2–7 links)	77.90%	17.24%	2.62%	1.43%	0.43%	0.10%	0.29%

of 1,000 randomly generated testing cases. The approximation errors expressed in terms of ARE and RMSE were 1.47% (Std = 1.37%) and 2.07, respectively.

2. Once the AM model was validated, it was used as a part of the solution to the NRP by simply replacing the function TT with the product of matrix *M* and vector *S* at every step.

On average, the proposed approximation method predicts the exact repair solution for over 75% of the networks that were used for testing. Table 3 shows the details of how the solutions with the AM approach to the NRP compare with the exact solutions at various network damage levels.

Table 3 was obtained by computing the solutions to the NRP under different network configurations and budgetary constraints. Given the number of damaged links *n,* 100 different configurations were randomly chosen for each of the possible budget limits, ranging from allowing repair of one to $n - 1$ damaged links (repair of all *n* links is not considered because of its triviality). For instance, given that three links are damaged, 100 NRPs with a budget limit of C_0 and another 100 NRPs with a budget of $2C_0$ were solved by using the exact and AM methods. Table 3 compares the extent to which the objective values of the NRP differ when the problem is solved by using the two approaches. For instance, the AM approach in the networks where three links are damaged predicts the exact solution over 91% of the times, and for another 5.5% of the cases, the solution from the AM approach is within 10% of the actual solution.

On the whole, the AM approach thus seems to perform reasonably well by predicting approximate solutions whose errors are less than 10% for almost 90% of the situations irrespective of the number of decisions involved. Another measure that can be used to compare the AM method with the exact solution is the average Hamming distance. The Hamming distance between two binary vectors is defined as the number of positions in which the corresponding numbers are different (*8*). For instance, the Hamming distance between 00110 and 10010 is 2. The average Hamming distance between the approximated repair strategy and the true optimal solution is 0.46.

Lack of a convenient computational tool deters wide use of NRP models by practitioners (*15*). The usual rule of thumb adopted in practice is to give links that carry more traffic higher priority (*16*). Table 4 shows the extent to which this strategy differs from the true optimal solution.

It can be seen from Table 4 that the performance of this rule drastically degenerates with an increase in the number of decision variables involved. For instance, the link volume approach matches with the true optimal solution 73% of times when only two links were damaged. This percentage decreases to as low as 18.33% when 10% (~7 links) of the links in the network are damaged. The ineffectiveness of this practice reinforces the need to adopt other methods such as the AM method, which fares far better in terms of its nearness to the optimal solution with reasonable amounts of computational requirements.

As noted earlier, the construction of the aforementioned AM models is based on the assumption that the demand for the transportation network at the end of the recovery period remains unchanged. However, in reality the demand may fluctuate for various reasons. Fluctuations in demand may cause a significant variation in total travel time. This variation could potentially lead to discrepancies in the calculation of benefits that are accrued by adopting different repair

TABLE 4 Link Volume Approach to NRP

Number of Damaged Links	Error in Objective Value						
	0%	0%–10%	10%–20%	20%–30%	30%–40%	40%–50%	>50%
2 links	73.00%	2.00%	4.00%	1.00%	2.00%	2.00%	16.00%
3 links	51.50%	10.00%	7.50%	8.00%	4.00%	3.00%	16.00%
4 links	40.33%	17.67%	11.33%	7.00%	4.33%	5.33%	14.00%
5 links	33.50%	22.75%	14.75%	8.50%	6.75%	2.25%	11.50%
6 links	23.00%	27.20%	19.00%	9.80%	5.80%	3.40%	11.80%
7 links	18.33%	34.50%	19.33%	9.33%	3.83%	4.17%	10.50%
All cases (2–7 links)	31.24%	24.24%	15.38%	8.43%	4.86%	3.57%	12.29%

strategies. The numerical experiment here showed that there can be as much as a 10% increase in total travel time when a noise of 25% is added to each of the demand matrix's elements. Given the trends in demand fluctuations, it would thus be preferable to construct models that can predict the network performance indicators under any of the demand conditions that are likely to be encountered. This aspect is especially critical in the context of planning a recovery following an unforeseen incident, where the timing of the disturbance may not be known beforehand. These perturbations in the demand matrix can be accommodated in the AM approach through slight modification of the model structure.

In the variable demand context, the AM models are required to predict the network performance under varying network structures as well as demand. The stimulus vector hence needs to contain terms related to both of these variables. The response vector remains the total travel time under UE conditions. The adopted functional form of the stimulus vector is

$$S = \left[C_1, C_2, \ldots, C_{76}, C_1^2, C_2^2, \ldots, C_{76}^2, C_1^3, C_2^3, \ldots, C_{76}^3, D_{1.2}, D_{1.3}, \ldots, D_{23.24}, D_{1.2}^2, D_{1.3}^2, \ldots, D_{23.24}^2, D_{1.2}^3, D_{1.3}^3, \ldots, D_{23.24}^3\right]^T$$

where C_i is the capacity of the ith link, and $D_{i,j}$ is the demand from node i to node j.

The training step here needs to be preceded by the determination of the nature of the demand fluctuations. The assumption in this regard was that the elements in the demand matrix can vary between a fixed percentage of the base value. With the assumption that this fixed percentage is 25%, 6,000 training cases were generated by varying the network configuration and by adding a noise ranging from –25% to +25% for each element in the demand matrix. The mapping matrix M was then constructed by inverting the stimulus matrix and taking its product with the corresponding response matrix. The matrix thus built was tested by using an additional 1,000 realizations. The stimulus elements were link capacities and OD demand, and the response vector was total travel time. The evaluation of the AM model with demand perturbations is summarized as follows:

ARE	3.9%
SD of relative error	3.2%
RMSE	5.44

A fair amount of consistency in the performance of AM models was observed on construction of 10 such models.

The AM models built with the expanded stimulus vector in association with the demand forecasting models can then be used to solve the NRP at any point in time when demand may differ from what was predicted. The AM models can be built with the demand matrix elements in the stimulus vector when fluctuations in demand are anticipated. Otherwise, if the demand matrix is expected to vary to a small extent, link capacities alone need to be included in the stimulus vector.

The approximate solution to the NRP with perturbed demand by using an AM approach was compared with the true optimal solution. No significant drop in the performance of the AM approach was observed when it was compared with the performance of the AM approach to the static demand NRP. The average Hamming distance between the approximated and the true solutions in the perturbed demand scenarios was 0.46. The gap between the objective values caused by the true optimal solution and the approximated strategy was still within 10% for almost 90% of the scenarios that were tested.

Application of the AM models for solving the NRP, on the whole, seems to be promising, with the benefit of a fairly competent performance for low amounts of computational resources. In addition to the computational advantage, the AM approach to the NRP is very flexible. For instance, it is a relatively simple task to solve the NRP with varied objective functions that involve computation of multiple performance criteria. Furthermore, this method can easily be extended to any of the other bilevel problems that deal with improving network performance. In view of these results, the approximate solution from the AM models can be used as an excellent starting point in the search for global optima, thus cutting down the time expended in reaching the neighborhood of the optimal solution.

DISCUSSION OF RESULTS

The estimation of various network performance indicators can hardly be overemphasized given the numerous decision-making processes in which they are needed. The presented approximation approach based on AM techniques was found to have reasonably low errors in prediction. The authors' experience in quickly solving NRP problems using AM models motivates more future investigation incorporating AM techniques into the general class of bilevel problems.

The choice of this method may be appropriate only in situations where it would be worthwhile to compromise on accuracy in exchange for computational time savings. For instance, in long-term planning in which online decision time is not a critical issue, it may be more beneficial to spend more effort seeking the true optimal strategy. The approximate method, however, seems to perform far better than the rules of thumb used in highway retrofit and recovery practice. This approach can also be highly useful in real-time decision making. Examples of such a situation include signal timing control and dynamic congestion pricing, which are also typical bilevel problems. Other suitable applications of the AM approach include sensitivity analysis of network performance against some uncertain network parameters such as demand.

However, the AM approach, because of its heuristic nature, has some disadvantages too. Similar to other neural computing methods (e.g., neural networks), this approach represents a sort of black box and provides little understanding of the underlying principles. The other potential disadvantage can be the high setup time associated with preparation of the many training cases. Future work includes identifying the functional form of the stimulus vector and finding a good composition for training cases that would perform well for a particular problem. This aspect of training is more likely to be dependent on the nature of the problem itself. At this stage of the research, it is not possible to provide any general rules for these choices.

Through the paper such terms as "true" optimal solution and the "approximated" solution are often used. It should, however, be emphasized that no matter what modeling or solution methods one chooses to use, these are approximation of the reality. Similarly, only the reality is the best test bed for judging the effectiveness of a decision.

REFERENCES

1. Okuyama, Y., and S. E. Chang, eds. *Modeling Spatial Economic Impacts of Disasters.* Springer-Verlag, 2004.
2. Albert, R., H. Jeong, and A.-L. Barabasi. Attack and Error Tolerance of Complex Networks. *Nature,* Vol. 406, 2000, pp. 378–382.

3. Magnanti, T. L., and R. T. Wong. Network Design and Transportation Planning: Models and Algorithms. *Transportation Science,* Vol. 18, 1984, pp. 1–55.
4. Yang, H., and M. G. H. Bell. Transport Bilevel Programming Problems: Recent Methodological Advances. *Transportation Research,* Vol. 35B, 2001, pp. 1–4.
5. LeBlanc, L. J., E. K. Morlok, and W. P. Pierskalla. An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem. *Transportation Research,* Vol. 9, 1975, pp. 309–318.
6. Ferris, M. C., A. Meeraus, and T. F. Rutherford. Computing Wardropian Equilibrium in a Complementarity Framework. *Optimization Methods and Software,* Vol. 10, 1999, pp. 669–685.
7. Dafermos, S. Traffic Equilibrium and Variational Inequalities. *Transportation Science,* Vol. 14, No. 1, 1980, pp. 42–54.
8. Kohonen, T. *Self-Organization and Associative Memory.* Springer-Verlag, 1987.
9. Kalaba, R., Z. Lichtenstein, T. Simchony, and L. Tesfatsion. Linear and Nonlinear Associative Memories for Parameter Estimation. *Information Sciences,* Vol. 61, 1992, pp. 45–66.
10. Kagiwada, H. H., J. K. Kagiwada, and S. Ueno. Kalaba's Associative Memories for System Identification. *Applied Mathematics and Computation,* Vol. 45, 1991, pp. 135–142.
11. Dougherty, M. A Review of Neural Networks Applied to Transport. *Transportation Research,* Vol. 3, 1995, pp. 247–260.
12. Kim, G. *Rapid, Artificial Intelligence Estimates of Los Angeles Highway Flows: Applications to Seismic Risk Analysis and Congestion Pricing.* PhD dissertation. University of Southern California, Los Angeles, 1997.
13. Latora, V., and M. Marchiori. Efficient Behavior of Small World Networks. *Physical Review Letters,* Vol. 87, 2001, pp. 198701-1 to -4.
14. Yin, Y., Genetic-Algorithms-Based Approach for Bilevel Programming Models. *Journal of Transportation Engineering,* Vol. 126, No. 2, 2000, pp. 115–120.
15. Werner, S. Current Developments and Future Directions for Seismic Risk Analysis of Highway Systems. *Proc., Technical Council on Lifeline Earthquake Engineering (TCLEE) Sixth U.S. Conference and Workshop on Lifeline Earthquake Engineering,* ASCE, Aug. 2003, pp. 849–858.
16. Werner, S. D., C. E. Taylor, J. Moore, and J. S. Walton. *Seismic Retrofitting Manuals for Highway Systems.* Vol. I. *Seismic Risk Analysis of Highway Systems, and Technical Report for Volume I,* Multidisciplinary Center for Earthquake Engineering Research, Buffalo, N.Y., 1999.

The Transportation Network Modeling Committee sponsored publication of this paper.