Jurnal Ekonomi Malaysia 41 (2007) 17 - 41

Exploring an Evolutionary Traffic Flow Landscape with a Fitness Approach

Kingsley E. Haynes
Rajendra G. Kulkarni
Roger R. Stough

ABSTRACT

The road patterns of major metropolitan areas and constituent jurisdictions evolve slowly through a complex set of independent and interdependent decisions producing a transportation network. The resulting network must be used for variety of commuting and spatial interaction activity. A typical trip taker spends considerable time on the road to reach the workplace and other destinations. Adding more links to existing road networks and/or increasing traffic capacity by adding lanes does not necessarily decrease travel times. However, a dense redundant network of roads provides a trip taker with alternate routes when traffic congestion occurs. Such issues raise the question of, how to evaluate the flow characteristics of the entire road network of a jurisdiction or its larger region? We explore a methodology to evaluate fitness criteria for road networks based on Kauffman's evolutionary complexity or NK model (1993) and develop an information theoretic measure of the order or organization in transportation networks.

ABSTRAK

Bentuk jalanraya bagi kawasan metropolitan dan kawasan sekitar berubah perlahan-lahan melalui rangkaian pengangkutan yang kompleks sama ada secara bebas atau bergantungan antara satu sama lain. Rangkaian ini perlu digunakan untuk pelbagai aktiviti perhubungan dan interaksi ruang. Seorang pengguna jalanraya mengambil masa yang agak lama untuk sampai ke tempat kerja dan destinasi lain yang dituju. Dengan menambahkan lebih banyak hubungan kepada jalanraya sedia ada tidak semestinya mempercepatkan masa perjalanan. Walau bagaimanapun, rangkaian jalanraya yang padat menyediakan pengguna jalanraya dengan jalan-jalan alternatif apabila berlakunya kesesakan. Isu ini menimbulkan persoalan seperti bagaimana menilai ciri-ciri aliran dalam *rangkaian jalan di pelbagai wilayah? Kita mencuba satu kaedah untuk menilai kriteria kesesuaian bagi rangkaian jalan raya berdasarkan kepada kekompleksan evolusi Kauffman atau model NK (1993) dan mengembangkan satu ukuran maklumat secara teori bagi susunan atau organisasi yang berkait dengan rangkaian pengangkutan.*

INTRODUCTION

Urban road networks are characterized by traffic congestion, incidents and accidents (Lave 1985), resulting in travel delays for commuters and other trip takers. The interaction costs of such congestion in a regional economy are enormous and factoring in work time lost to business and commuters makes the sums astronomical (Arnott & Small 1994). Increasing capacity of existing freeways by adding more lanes is not always possible or environmentally desirable and does not always ease the delays. The much studied Braess's paradox tells us that congestion may increase rather than decrease as capacity is increased (Murchland 1970). However, the costs of incidents and accidents could be reduced if the trip taker is provided with timely warnings of such events. Intelligent Transportation Systems (ITS) traveler management systems hold the promise of providing information on traffic conditions. However, providing data on traffic conditions alone may be of little help if there is no underlying processing framework to evaluate and disseminate the processed information.

Figures 1 and 2 show the schematics of the existing traffic management arrangement framework in many metropolitan settings. An urban traffic flow network is divided among a number of zones and has one or more traffic management center (TMC) in each zone. The data collected by non-intrusive surveillance equipment (Figure 2) in each zone is processed by a traffic management center (TMC) in that zone and disseminated to users. In this paper we develop an analytical model of the TMC data processing unit and suggest some underlying considerations that need to be assessed in developing a decision framework for traffic guidance and management. The analytical model is based on concepts borrowed from evolutionary biology, especially the concept of fitness landscapes (Kauffman 1993) and information theory (Appelbaum 1996; Suhir 1997) to describe the organization or order of traffic networks.

So far, with the exception of a few, much of the modeling efforts to describe network traffic flows are based on physical analogs such as fluid

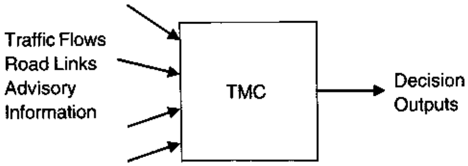

FIGURE 1. Input data on traffic flow on road links

dynamics (Herman & Prigogne 1979), percolation (Kulkarni & Stough 2000) and spin glass physics (Kulkarni et al. 1996). A survey of the literature shows many attempts to model the dynamics as well as the equilibrium/disequilibrium network flow conditions that exist on urban road networks. Both analytical and simulation/experimental studies have been done (Friesz et al. 1996; Friesz et al. 1993; Friesz et al. 1994; Mahmassani 1995; Mahmassani et al. 1990; Mahmassani et al. 1992, Mahmassani and Peeta 1992; Koutsopoulos 1995). While some modelers have addressed the stochasticity of traffic flows by trying to reduce the randomness - following the so-called micro-simulations approach (for pioneering work see Mahmassani and Herman (1987), Mahmassani et al. (1990), and the TRANSIMS (1995) model developed by the transportation

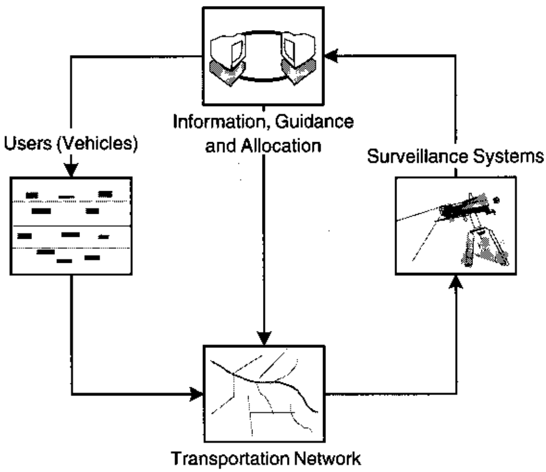

FIGURE 2. A Traffic management center (TMC) flow chart
(Schematic inspired by Mahmassani 1995.)

group at Los Alamos National Laboratory (Barrett et al. 1995, Smith et al. 1995).

MOTIVATION

In various degrees and shades biological models have been adapted by fields as different as cosmology (evolutionary universe; Linde 1994; Coleman et al. 1991), economics (evolutionary economics; Tu 1992; Krugman 1994, 1995; Arthur 1989), and sociobiology (Wilson 1995). Although many workers in other fields view evolution as a concept to describe gradual changes, as opposed to revolutionary changes in system behavior (Fabian 1998), this perspective is not necessarily consistent with the evidence of punctuated evolution. Whatever the points of view, all evolutionary systems consist of a large number of agents whose interactions give rise to complex system-wide behavior: micro-level actions giving rise to macro-level patterns of behavior (Shelling 1978). Network traffic flow is also an evolutionary system.

Consider for example, an urban road traffic network consisting of a large number of road segments. Workday traffic on a segment of a highway typically has the profile of the morning and evening rush hour peaks with the intervening troughs for the rest of the day. However, it is quite unlikely that the traffic pattern profile on a given day matches exactly with those of previous or following workdays. Indeed, the stochasticity of traffic patterns arises in part as a consequence of "non-collaborative" trips taken by commuters. Commuters are "aware" of other commuters' plans to travel only to the extent that they are going to share the limited resources of time and road space with other unknown commuters. The commuters do not collaborate or inform each other of their intended trips and schedules and plan accordingly for their journeys. Commuters mostly follow a loose schedule that they create from their day to day experiences of trips on the roads. Thus, traffic is an aggregate of the multitudes of decisions executed by commuters in a non-collaborative manner, giving rise to traffic patterns despite master plans intended to regulate traffic on the roads. Can information theory models help explain the traffic patterns resulting from actions of distributed agents? Do biological and information theory models help in understanding concepts such as organizational order and adaptation? Is such a metaphor for the life-like processes a good explanation of traffic flow networks?

The appeal of biological models is that although they are somewhat *imprecise and fuzzy,* they provide a powerful explanatory framework for dealing with systems consisting of independent but interacting agents that change behavior to suit the dynamics of the environment. A model based on these concepts could provide us with some insight into network traffic flows and assist in the development of tools to assist ITS systems and improve the traffic flows in the network. For example, suppose that we know about a specific property that is beneficial to a population of agents. Then, at least in theory, an abstract landscape defined in terms of this specific property can be constructed and the dynamics of this property observed as the population adapts to and modifies the changing landscape. In evolutionary biology such landscapes are defined by survival fitness. In the analytical model described in this paper, we use the *ease of flow of traffic on the network* as a fitness property to construct abstract traffic flow landscapes, similar to the N-K model fitness landscapes proposed by Kauffman (1993).

Some important differences between the analytical model presented here and the classic N-K model are important to note. The N-K model describes interactions among agents in terms of boolean functions and relies on autonomous boolean networks for emergence of organization. Here we used an information theoretic approach to compute the ability of networks to display organization and to show how such a method can be utilized by ITS to adapt to changes in network flows. Additionally, we explain how ITS related technologies can help maintain an overall good fitness of traffic networks for the benefit of both the users and the traffic managers.

This should not be seen as a fully developed alternative to classic approaches to traffic forecasting in the traditional planning literature but rather a complexity approach supplement that has proved helpful in other settings (economics, institutional analysis, demography, etc.) and may prove of interest in the systems approach in transportation modeling.

BACKGROUND

An urban region's road network consists of many types of roads — highways, major and minor roads, arterials and connecting roads. For a traffic fitness landscape only those roads and links that are referred to as primary and secondary roads/links as described in the TIGER/Line™files Census Feature Class Codes (1992) are included. Links are segments on highways, major roads and arterials. Segments are characterized by the levels of service (LOS) (Highway Capacity Manual 1985) which vary depending on the traffic conditions on these segments. According to the Special Report 209 of the Highway Capacity Manual, "the concept of levels of service is defined as a qualitative measure describing operational conditions within a traffic stream, and their perception by motorists and/or passengers. A level-of-service (LOS) definition generally describes these conditions in terms of such factors as speed and travel time, freedom of maneuver, traffic interruptions, comfort and convenience, and safety" (Transportation Research Board 1985: 1-3).

FITNESS CRITERIA FOR TRAFFIC FLOWS

If we assign a numerical value to each LOS, then in theory, we could use these values as a fitness measure of each segment. Unlike the concept of fitness in biology associated with survivability of a species, the fitness in the traffic flow model refers to the idea of relative ease (high fitness) or difficulty (lower fitness) of traffic on segments of roads in transportation network. One can use the LOS concept to represent traffic flow conditions at any time over an entire road transport network. To quantify the qualitative concept of levels of service we suggest a very simple method in Appendix A based on fractals.

NETWORK TRAFFIC FLOW LANDSCAPES

The word "landscapes" has topological connotations. Even though the word evokes different images to different people, certain properties are common to many landscapes. For instance, landscapes have morphology such as multiple peaks (either sharp or gentle) and troughs and connecting ridges. The topology of a region makes it clear that to reach point '*P*' on one of the peaks from point '*Q*' on another peak involves finding the best possible route between these two points, avoiding regions of valleys. Alternately, one may want to avoid the peaks and move between valley regions. The landscape image suggests more than one peak that may satisfy a given set of criteria and that not all peaks and valleys are reachable easily. Thus, one may envision assigning a fitness value to criterion or criteria set and then creating a visual image with peaks for good fitness values and valleys for bad fitness. How does one use the idea of fitness to create traffic network landscape?

A network of '*N*' segments with '*L*' number of LOS has L^N possible configurations. With six LOS (*A* through *F*), we have 6^N possible configurations for N segment network. Each configuration is just one LOS different than its neighboring configurations, $D = N^*(L-1)$ configurations, where D is the number of neighbors. For example, for $N = 3$ segments and $L = 2$ LOS, then number of configurations is $L^N = 8$ or 000,001,010, 100,101,111. This can be represented as an eight vertices cube where each vertex has $N^*(L-1) = 3$ neighbors and each vertex represents one of the eight distinct configurations.

We associate each LOS of a segment with a computed fitness value. Note that the LOS of one segment may affect the LOS of its neighboring segments. We suggest an analytical expression to take these interactions into account. The number of such interacting segments is the interaction parameter, represented by an integer '*K*'. For simplicity assume that the fitness values of each segment are additive. Then, a configuration where all segments of the network have LOS '*A*', has the highest fitness value (free flow traffic) of 'Alpha'. At the other end is a configuration where every segment has a LOS of '*F*' with the lowest fitness value (traffic jams on all segments) of 'Omega.' Usually, the network with '*N*' segments will be in one of the L^N configurations with a value of fitness that is less than 'Alpha' and more than 'Omega'. But, the overall fitness of the network is not a simple additive process, since the traffic on each segment is affected by other segments, and more by near segments (nearest neighbor or contiguous) than far segments. Hence, each segment's fitness is a function of the fitness of neighboring segments. If there are '*K*' such neighbors that affect the fitness of any segment, the value of '*K*' can be between zero and '$N-1$'. When $K = 0$, each segment has traffic flows that are independent of all other segments. On the other hand, when $K = N - 1$, each segment's traffic flow is affected by the traffic flows on all other segments.

Description of traffic networks in terms of L^N configuration is analogous to a system with L^N states. A truly random system would show ergodic behavior such that the probability of such a system being in any one of these states is the same. But are traffic flows truly random? Probably they are not. Since traffic systems show patterns at the macro level, we need to assess the degree to which these patterns are constant or the degree to which they evolve and organize into new patterns that are mostly beneficial to participating agents. We can illustrate the development of the analytical model by a simple example, almost a cartoon of a real life network.

ANALYTICAL MODEL OF TRAFFIC FLOW LANDSCAPE

Consider an urban road network of N segments. The segments can be either sections between mile stones or distances between consecutive traffic signals or any other consistently defined measure across the network. The traffic flow levels on each segment are determined in terms of the LOS. Thus each segment can have 'A' through 'F' levels of service with LOS of 'A' for free flow (+1.0) and LOS of 'F' for no flow (0.0). At any instant a set of segments with specific flow levels constitutes a configuration of the entire road network among all possible configurations. For 'N' segments with 'L' flow levels, the number of possible configurations is L^N. For two types of flows (identified by A and F) the total number of possible configurations of segment flows is 2^N. Let 'K' refer to number of segments interacting with each other. Next we explain the road configurations for different values of 'K'.

1. $K = 0$ Case

When 'K' is assumed to be zero, each segment contributes to the overall fitness independent of all other segments. The entire network configuration may be represented as a combination of +1.0 and 0.0s. Thus, two extreme states that one can find for the N segment network are trivial. One of these states has all the segments blocked thus the network segments are represented as the following vector:

$$\left(F_1, F_2, F_3, \mathrm{K}\ , F_N\right) \tag{1}$$

and the total fitness M_g, is given by:

$$M_g = \sum_{i=1}^{N} f\left(F_i\right) = 0.0 \tag{2}$$

The other state has all the segments in the free flow condition and may be represented as the following vector:

$$\left(A_1, A_2, A_3, \ldots, A_N\right) \tag{3}$$

and the total fitness M_g, for this state is given by:

$$M_g = \sum_{i=1}^{N} f\left(A_i\right) = +N \tag{4}$$

Every other state has a total fitness contribution that is between 0.0 and N.

A physical network of N segments, each of which can have any one of L levels of service has L^N distinct configurations. Since each of these configurations are distinct, they can be represented by a hypercube whose dimensions are determined by $N^*(L-1)$. Again as shown in the example of $N = 3$ and $L = 2$ we have 8 distinct configurations and they can be represented as vertices if a $N^*(L-1) = 3$ dimensional cube. For $N = 4$ segments and $L = 2$, the total distinct configurations are $L^N = 2^4 = 16$, which can be represented as 16 vertices on a $N^*(L-1) = 4$ dimensional cube. One can prove with induction that in general, for N segments and L levels of service, there L^N distinct configurations which can be represented as vertices of a $N^*(L-1)$ dimensional cube.

Table 1 shows each configuration and its total fitness contribution. It is clear from the fitness values that there are multiple configurations with the same fitness value. Configurations 2, 3, 5 and 9 all have fitness of 1/4, while configurations 8, 12, 14 and 15 have fitness of 3/4. Six configurations have a fitness of 1/2 (Configurations 4, 6, 7, 10, 11 and 13). If we assign probabilities to each of these fitness values, then it is clear that the configuration with fitness value of 1/2 has a higher probability of occurrence. Intuitively, a configuration with all blocked segments also has a very small probability as does the configuration with all segments

TABLE 1. Configuration vs. fitness

Configuration	Fitness value	Avg. Fitness $K = 0$
1 = F,F,F,F	0,0,0,0	0
2 = F,F,F,A	0,0,0,1	¼
3 = F,F,A,F	0,0,1,0	¼
4 = F,F,A,A	0,0,1,1	½
5 = F,A,F,F	0,1,0,0	¼
6 = F,A,F,A	0,1,0,1	½
7 = F,A,A,F	0,1,1,0	½
8 = F,A,A,A	0,1,1,1	¾
9 = A,F,F,F	1,0,0,0	¼
10 = A,F,F,A	1,0,0,1	½
11 = A,F,A,F	1,0,1,0	½
12 = A,F,A,A	1,0,1,1	¾
13 = A,A,F,F	1,1,0,0	½
14 = A,A,F,A	1,1,0,1	¾
15 = A,A,A,F	1,1,1,0	¾
16 = A,A,A,A	1,1,1,1	1

in a free flow situation. A very simple coding scheme gives quite a bit of information on the condition of a traffic network. This schema can be extended to a greater number of LOS to achieve a more realistic model of traffic flow conditions.

2. $K = N - 1$

Consider the case when each segment interacts with all the other segments of the network. Since, we do not know yet how each segment affects the other segments, we assume that the complex interactions are multiplicative in nature, that is if segments have high fitness values, and then the result of interaction would be a high fitness contribution value. On the other hand if one or more segments have lower fitness values, then the result of the interactions accordingly reflect a low fitness contribution value.

Another way to represent the interactions is to use a modified Tanner function (Tanner 1961; Paelinck and Klassen 1979),

$$\pi\left(f_i'\right) = \sum_{j=1}^{K+1} \pi\left(f_j\right) * \left(\left(d_{ij} - 1\right) \times \exp\left(-\alpha \times d_{ij}\right)\right) \tag{5}$$

where, 'α' is a proportionality constant, 'd_{ij}' is the distance (steps) between segment 'i' and segment 'j', $\pi(f_j)$ is the fitness of segment 'j', $\pi(f_i)$ is the fitness of segment 'i' and $\pi(f_j')$ is the computed fitness of segment 'i' as a result of the 'k+1' interactions between segment 'i' and other segments.

Next, consider a set of configurations, each with 4 segments as before. But now, we represent the individual fitness values as random numbers between 0 and +1. Then the total fitness of a configuration M_g is

$$M_g = \sum_{i=1}^{N} \pi\left(f_i'\right) \tag{6}$$

where 'g' is a configuration, $\pi(f_j')$ is the fitness potential of a segment 'i' calculated using equation (5)

Once more, the fitness landscape is constructed as an N*(L – 1) dimensional abstract hypercube where each vertex represents a configuration with a fitness value that is a contribution from the segments of that configuration. Although the fitness landscape is bounded from above and below by 0 and N, it is now much more rugged and has multiple peaks interspersed with deep valleys. For a simple network of $N = 4$ and $L = 2$ (near free flow and near total blockage) we obtain a hypercube with $2^4 = 16$ configurations represented by its vertices. Each configuration on

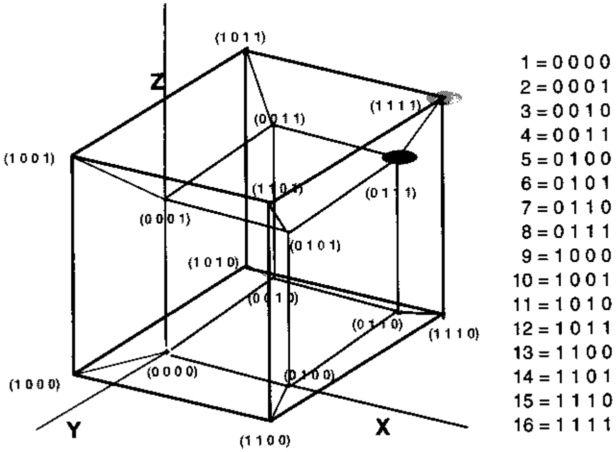

FIGURE 3. Hypercube representation of genotypes for genes N

a vertex has a fitness value that is either a peak or a valley depending on the result of interactions among the four segments of each configuration. Every vertex has a configuration that has 'D' other neighboring vertices with their own configurations and all of these differ from each other in one LOS of a segment. Accordingly, the total fitness also differs for each vertex (see figure 3).

The two extreme cases of $K = 0$ and $K = N - 1$ show how the fitness landscape can change from a one maximum and one minimum fitness landscape to a multiple peak rugged landscape (Figure 4.) Note that the landscape for the $K = N - 1$ case was generated using random values between 0.0 and 1.0 for fitness of each segment, and the interactions among the segments were thought to be of a multiplicative type. The configurations 4, 5 and 8 have a higher fitness than the rest. The overall fitness level becomes smaller as the number of interactions increases from $K = 0$ to $K = 3$. This is due to the conflicting nature of interactions among the segments and results in reduced fitness maxima.

As noted earlier, fitness level is expressed between 0 and 1 (real number). When there is interaction between neighboring segments represented as multiplicative in nature, the result will be less. Consider two neighboring segments. If fitness of one segment is 0.5 and fitness of its neighbor is also 0.5 and if they are independent then total fitness would be $0.5 + 0.5 = 1$, however, if the interaction is, say multiplicative in nature then, total fitness will be $0.5 \times 0.5 = 0.25$. Thus when the entities are not independent of each other, fitness values fall as K increases. Similar results would be obtained using Tanner function or radial function or any other interaction function where the influence of one entity over the other decreases with distance.

In general, as the value of K changes the fitness landscape also changes – from a single maximum/single minimum fitness landscape to a rugged multiple maximum/multiple minimum fitness landscape whose maxima and minima have reduced fitness values. To create a fitness landscape for the entire network, we use a random fitness function or some other function that reflects the general traffic conditions. In the later case the function could be a weighted combination of a number of traffic properties, such as peak flow times, numbers, density of the traffic, speed limits on the segments or any other relevant property of the traffic on the road network.

The specification of the traffic flow fitness landscape model is complete when the assignment of the fitness vectors for all the segments is done. The traffic flow landscape is represented as a hypercube in an $N^*(L-1)$ dimensional space for 'L' LOS traffic flow condition. Each of the configurations at the vertex is one LOS different from its 'D' neighbors and accordingly, its fitness is slightly different than the rest of its 'D' neighbors.

Now it is possible to estimate the overall fitness of the road network for all configurations using Equation (6) as follows:

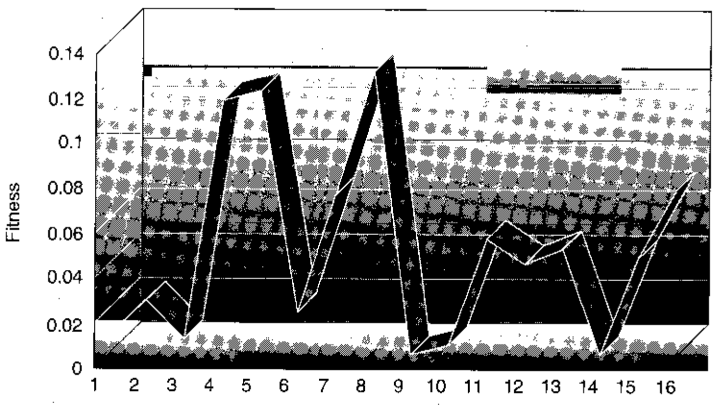

FIGURE 4. Fitness landscape, $N = 4$, $K = 3$, $L = 2$

$$\Gamma = \frac{1}{L^N} \sum_{g=1}^{L^N} M_g \tag{7}$$

where M_g is the fitness of configuration 'g'. Equation (7) serves as the general fitness index of road networks.

EXPLORING THE TRAFFIC FLOW LANDSCAPE

Since the traffic flows on roads change dynamically, they do not lend themselves easily to modeling. One cannot associate a single equilibrium point at which the traffic flows settle down into a regular pattern. Instead, the traffic flows follow multi-equilibria metastable behavior, jumping from one configuration *gx* to the next configuration *gy* on the fitness landscape hypercube. The new configuration *gy* may or may not be in the immediate neighborhood of *gx*. The process of moving from this configuration to the next continues as the flow dynamics change on various segments. The movement over the fitness landscapes may not always result in a better fitness configuration. Consider again an $N = 3$ segment, $L = 2$ LOS network, there are total of $L^N = 8$ configurations, each distinct configuration when represented as a vertex of a 3 dimensional cube, has $D = L^*(N - 1) = 3$ neighbors. Thus for a vertex and its 3 neighbors, the probability that its fitness is better than the three neighbors is $1/(D + 1) = ¼$. Hence, the probability that g_x has a better fitness value than its D neighbors is given by:

$$p(g_x) = \frac{1}{D+1} \tag{8}$$

The higher fitness value of configuration g_x makes it a locally optimal configuration among its neighbors. For a landscape consisting of L^N configurations, the total number of such local optima is given by:

$$E \approx \frac{L^N}{D+1} \tag{9}$$

A large number of locally optimal configurations for a traffic flow landscape exist. The local optima are the multiple equilibria that are scattered all over the traffic fitness landscape.

ATTRACTORS AND ATTRACTOR BASIN

Next, we consider a four line sub-network of road segments that are adjacent to each other such that one segment's flows go to the next segment $N = 4$ segments, ('a', 'b', 'c', 'd') and $L = 2$ LOS and $0 <= K <= 3$ interactions. We rank the 16 possible configurations according to their fitness values. If we assume that these segments are such that, during a time period '*t*', the traffic on segment 'a' moves on to segment 'b' and so on, we can represent the current traffic conditions on the four segments with '0101.' Then, as the traffic moves in one time period, the flow conditions on the successor configuration could be such that the resulting configurations are one or more, but less than *N*, LOS different from the neighbor configurations.

For a network consisting of a large number of segments, the successor configuration could vary from being one LOS different from its neighbor to *N* LOS different. If the successor configuration is the same from one instant to another or if the successor configurations change back into the original configuration then the original configuration becomes the *attractor*. In other words, if a set of different configurations corresponding to small scale perturbations in the flows on segments have the same successor configuration, then the members of the set form the so-called *attractor basin* and the successor configuration may be designated as the *attractor configuration* or *meta stable equilibrium configuration*. On the other hand, if the successor configurations are all wildly different it is an indication that the traffic flow patterns are changing chaotically and the network has become unstable.

SEARCH FOR LOCAL OPTIMA

If we could construct a traffic flow landscape at a given instant, then, in theory, it would be possible to estimate the time needed to reach an optimal solution. Suppose that a traffic landscape has been constructed and currently the entire network is represented as a configuration g_x in this landscape. For a $D = N^*(L - 1)$ dimensional hypercube, a configuration at a vertex is at least one LOS different than its '*D*' other neighbors. If we start at a worst fitness configuration, then moving to any of its neighboring '*D*' vertices would lead us to a configuration that has better fitness than the previous one. Since the total number of configurations is L^N, the rank order of the new configuration is between 2 and L^N. If we continue to move randomly to a vertex that has better fitness than the previous one, then every such move makes the new configuration halfway closer to the remaining configurations. As the improvement continues, the process slows down such that for every such move the time to search for a fitter neighbor doubles. For example, using the fitness values mentioned in Table 1, if we start at the worst fitness configuration (0000), then there is a 69% chance that the new configuration will have fitness of ½ or more and only 25% chance that the new configuration has fitness of ¼. In the later case, for the next step, there are 11 configurations that have fitness values better than the current one. Thus chance of fitness value of ½ is 6 in 11, while that being ¾ or more is 5 in 11. Suppose, now the fitness value is ½. Then for the next step, only 5 configurations have fitness values that are more than ½. And thus, chance of fitness value being ¾ is 80%. If the fitness value becomes ¾ then in the next step, the fitness value of 1 is achieved. Thus for 16 configurations with fitness levels between 0 and 1, starting at the configurations with worst fitness (0), it takes 4 steps to reach the fitness level of 1.

Another way to look at the ranked fitness landscape starting from the worst fitness vertex g_w is similar to travelling down a tree whose root is the current vertex and at each level of the tree, each node (the new configuration) branches to '*D*' other configurations which are 1 LOS different and accordingly have a different fitness level (See Figure 5). Remember, the tree should read as ordered tree and then mapping from ordered tree to binary tree is a standard operation carried out in computer science. Typically, for an ordered tree, for a node *N*, left child becomes

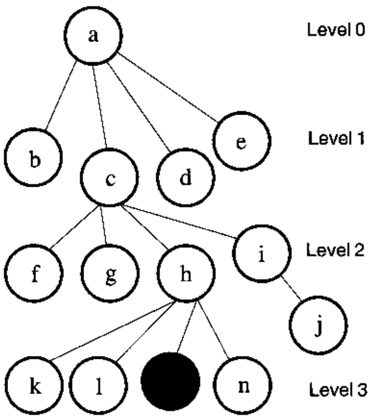

FIGURE 5. Partial fitness tree

left branch in the binary tree and the right child becomes right branch and then on to next level.

At the start of a given time period, we can construct a configuration tree that has the root configuration corresponding to the current traffic flows on the segments. Next we search randomly for a better fitness configuration on one of the '*D*' branches. When a configuration is found, we are at the new configuration and repeat the process until we reach a local optimum. The total time or the number of levels becomes a measure of the time needed to reach a locally optimal configuration. Alternately, the configuration tree in Figure 6 can be changed into a binary tree (Horowitz & Sahni 1985; Wilson 1988), such that at each node or configuration there are just two branches (see Figure 6.) Every time we reach a node, we take the branch that leads to a better fitness configuration until we reach a node that corresponds to a locally optimal configuration.

Similarly, we could use the above techniques to analyze the impact on the basis of increasing (decreasing) the total number of segments '*N*' or changing the number of interactions per segment '*K*' to generate a configuration tree and searching for a locally optimal configuration

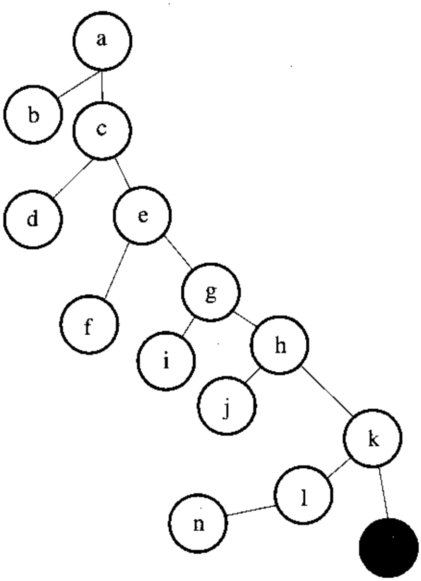

FIGURE 6. Partial fitness binary

that has an equal or better fitness than the configurations in its neighborhood.

ORGANIZATION AND TRAFFIC FLOWS

Consider a system characterized by a large number of configurations. If we have little or no information about each of these configurations, then the only meaningful thing we can express about these configurations is that each configuration occurs with probability 'p_i' where 'i' ranges over all possible configurations and that all such probabilities are equal. The amount of information 'I' (Applebaum 1996; Suhir 1997) obtained from this system in configuration 'i' is given by

$$I = -\log_2 (p_1) \tag{10}$$

The negative number is because the information is measured as a log of probability. Since log of a number between 0 and 1 is –ve, to make the sign of the information right (+ve), by definition one puts –ve sign in the front of the log operation.

The expression for information I in equation (10) can be converted from logarithm of base 2 to logarithm of base 10 as follows:

$$\log_{10} (I) = \frac{\log_2 (I)}{\log_2 (10)} = 0.30103 \times \log_2 (I) \equiv \Theta \log_2 (I). \tag{11}$$

Since Θ is a constant, it will be ignored in the following discussion as the log operation is assumed to be to the base 10.

Let us define a fitness function 'V'. The fitness function computes fitness value M_g (Equation 6) for each configuration, where 'g' ranges over [1, L^N]. Note that 'V' is a one-to-many function, i.e. many configurations have the same value of fitness. Let R_j, (where $j <= g$), denote the number of configurations with the same fitness values. Since the values that R_j can take a priori are unknown, R_j can be expressed as a random variable.

$$R_j = V(M_g) \tag{12}$$

where 'j' ranges over an interval [1, r] and as before 'g' ranges over interval [1, L^N]. Let P_j denote the probability distribution associated with R_j. Then, the expectation 'E' of such a distribution is defined as the amount of uncertainty or information entropy 'S' in the system and is given by:

$$S = -\sum_{j=1}^{r} P_j \times \log(P_j). \tag{13}$$

If there is little or no information about the configurations and hence, the associated fitness values, all that can be said about such a system is that the number of configurations with similar values of fitness (R_j) occur with equal probability or R_j has a uniform probability distribution. Such a system is said to have maximum uncertainty or maximum entropy S_{max} (Jaynes 1979) and is characterized by disorder or disorganization. On the other hand, a decrease in entropy of the system indicates increasing order or organization in the system. How does increasing order or decreasing disorder of a system occur?

Entropy 'S' of a traffic network changes according to probability distributions of a random variable 'R', which is determined by the fitness 'f' of each configuration, which in turn depend on the type of flows on various segments. Hence, at any instant, the difference in 'S_{max}' and 'S' indicates the organization or degree of order in the network. Thus we define an order parameter as

$$O = \kappa\left(S_{max} - S\right) \tag{14}$$

where, 'k' is a proportionality constant and 'O' is the inherent organizing capacity of the network.

The discussion of entropy and order follows from the treatment of uncertainty in information theory. An information theoretic definition of entropy 'S' of a system is a measure of uncertainty in a system (Applebaum 1996; Wilson 1969, 1973; Haynes et al. 1980; Haynes and Phillips 1981; Haynes and Storbeck 1978). If we can reduce uncertainty in the information content of a system then we will have reduced the entropy of the system and accordingly its internal fitness will have improved (see Equation 13), indicating organized traffic flow.

Consider a tiny network of $N = 21$ segments and $L = 6$ LOS; this network has a staggeringly high number, 6^{21} configurations. With no prior knowledge of flow conditions, one must assume that each of these configurations is equi-probable. But as soon as we gather traffic flow information on even a small number of segments, the total number of possible configurations decreases. Further, by defining a discrete random variable that takes on values according to a fitness criteria (see Equation 12), we can find the probability distribution of such a random variable. From this the entropy, order can be computed using Equations (10) through (13). If a specific level of service on these segments is maintained, then every time we make an observation we are certain to find a specific level of service. The probability of such segments is 1 and the information content is zero, since log(1) = 0. As flow of traffic on a definite number of segments becomes certain, the overall entropy of the network decreases, indicating an increase in organization of the network. The surveillance equipment to monitor traffic flows (Figure 2) on segments of a network can provide information that would reduce uncertainty in traffic flows and increase its fitness. Additionally, if TMC is able to maintain higher flows on different segments, it modifies the probability distribution of the flows, which in turn will modify fitness landscapes that are favorable for efficient flows.

Alternately, measurement of flows on all segments gives the current state of the network. In terms of the traffic flow landscape we now know the vertex representing the state of the network. If this vertex corresponds to a good fitness, then the TMC can maintain flows in the network corresponding to that fitness level. Conversely, if a traffic network is on a bad fitness vertex, then the TMC can adjust to improve the fitness of the network and possibly evolve towards a region of better fitness on the fitness landscape. The information on network flows could be used as input for ITS (Intelligent Traffic System) technologies such as ATMS (Advanced Traffic Management System) and ATIS (Advanced Traveller Information Service). In theory TMCs can be distributed across a traffic flow network, each TMC monitors and manages a subset of segments of the network and helps to maintain efficient network flows.

CONCLUSION AND FUTURE DIRECTIONS

Defining fitness vectors for traffic flows on a road network creates a rugged fitness landscape. For large values of N (the number of segments in a network) with K segments ($K < N$) influencing each of N segments, a very complex traffic flow fitness landscape is generated. For $K = 0$, a simple traffic flow landscape with a single global optimum is obtained. Such a traffic network has all it's segments independent of each other and incidents affect only local flows.

As the value of K increases, more complex traffic flow landscapes evolve. The other extreme occurs when $K = N - 1$ and each segment influences the flow of traffic on all other segments. This type of landscape has an infinite number of local optima and is very rugged. A fitness landscape with $K = N - 1$ indicates a network in which an incident on any one of the segments affects flows across the entire network. It is a highly unstable network.

The traffic flow landscape model depends primarily on the values of '*N*' and '*K*'. A TMC-like system can influence flows on different segments of a part of a network such that overall traffic flows across that part of the network can be improved by reducing the level of uncertainty (entropy) in the flows – or increasing organization or order in the network.

Note that we do not address issues of interfacing all these TMCs. In fact, each TMC is considered to be in operation independent of all the others. How to coordinate TMCs will be addressed in reports on future research. One of the features of the fitness landscape is its relative independence of factors such as the fitness values and variations in parameter '*L*' (Kauffman 1993). Since the traffic landscapes are mainly dependent on values of '*N*' and '*K*', each TMC should be able to develop processes to maintain a level of service on segments of a network that would always give a better fitness configuration.

Appendix A:

Consider a road segment of length L and width W. Then the total area A of the road segment is given by:

$A = L^D * W^D$ or if we express w as fraction of l then

$A = L^D * (\gamma * L^D)$ or alternately it may be expressed as follows:

$$A = \gamma * L^{2D} \tag{A1}$$

Equation (A1) can be re-written as:

$$A \alpha L^{2D} \tag{A2}$$

where $D = 1$, the dimension of the road segment. Now consider a stream of vehicles traveling on the segment of the road. Thus at any instant there are a finite number of vehicles occupying a finite amount of space on a section of the road.

Since, the vehicles on a road are discrete objects and occupy finite and discrete amount of space, we can express the total area occupied by the vehicles as follows:

$$a = n * l^d * w^d = n * l^d * \left(\delta * l^d\right), \tag{A3}$$

where a is the average area occupied by a vehicle of average length l and average width w and δ is a fractional measure for converting w into l. Equation (A3) may be expressed as:

$$a = n^* \delta^* l^{2d} \quad \text{or} \quad a\alpha\left(n^* l^{2d}\right) \tag{A4}$$

Let us express the average value of vehicle length l in terms of the length of section of the road, then equation (A4) can be written as:

$$a = n^* \delta^* l^{2d} \quad \text{or} \quad a\alpha\left(n^* (\varepsilon^* L)^{2d}\right) \tag{A5}$$

From equations (A2) and (A4) the density of vehicle occupancy ρ may be expressed as:

$$\rho\alpha\frac{n^* (\varepsilon^* L)^{2d}}{L^{2D}} \quad \text{or} \quad \rho\alpha\frac{n^* \varepsilon^{2d} * L^{2D}}{L^{2D}}. \tag{A6}$$

We can express the density function ρ by introducing a proportionality constant β in equation (A6) and get the following equation:

$$\rho = \frac{\beta^* n^* \varepsilon^{2d} * L^{2d}}{L^{2D}} = (conts.)^* L^{2(d-D)}. \tag{A7}$$

Taking logarithm on both sides of equation (A7) gives us the following equation:

$$\log(\rho) = \log(conts.) + 2(d - D)^* \log L. \tag{A8}$$

Since, $d = 1$ we can get the following equation:

$$d = \frac{\log(\rho) = \log(conts.) + 2\log(L)}{2^*\log(L)}, \tag{A9}$$

From equation (A8) we can get an expression for d as follows

$$d = 1 + \frac{\log(\rho) = \log(conts.)}{2^*\log(L)}. \tag{A10}$$

The value of d varies between a minimum of zero (free flow) and maximum = 1 (blocked segment) (see equation A10). This computed value of d can be used as a measure of the level of service for assigning fitness values to sections of roads.

Appendix B

KAUFFMAN'S *NK* MODEL

The *NK* model describes emergence of order in biological systems as a result of a multitude of complex, random, epistatic (non-reciprocating and inhibitory) binary interactions among the most fundamental agents of self-organization, the genes. A population of genes (genotypes) evolves over a fitness landscape (a type of hill-climbing) as it adapts to changes in the environment. To get a better understanding of the *NK* model, given below are definitions of the biological terms. Gene is the basic unit of inheritance. Genotype is a possible configuration or arrangement of genes. Allele is a variation of a gene. Fitness is any "well defined property" and the fitness landscape is a distribution of this property across an ensemble (Kauffman 1993).

The *NK* model of the evolutionary biologist Kauffman explains how a variety of genotypes is able to adapt to so-called rugged fitness landscapes of the environment in which these genotypes evolve. The '*N*' stands for number of genes and '*K*' stands for number of interactions any single gene has with other genes. Each gene may have '*L*' alleles. Alleles are the variations in each gene that give rise to a physical trait such as eye color. Each gene contributes to the overall fitness of a genotype. At the same time each is influenced by '*K*' genes that are either nearest neighbors or are spatially separated from the gene. Thus the result of all the interactions between '*N*' genes and '*K*' influencing genes is a fitness landscape with multiple peaks and valleys. The peaks are associated with fitness values. Depending on the value of '*K*', the landscape varies from a simple profile ($K = 0$) to one with a very complex profile ($K = N - 1$). The former ($K = 0$) refers to an environment in which each gene is independent of all its neighbors and the latter refers to a situation when each gene is influenced by all the genes ($K = N - 1$) in a genotype.

REFERENCES

Applebaum, D. 1996. *Probability and Information: an Integrated Approach.* Cambridge University Press. Cambridge. U.K.

Arnott, R. & K. Small. 1994. The Economics of Traffic Congestion. *American Scientist* 82: 446-455.

Arthur, Brian. 1989. Competing Technologies, Increasing Returns, and Lock-in by Historical Events. *Econ. J.* 99: 116-131.

Barrett, C., K. Berkbigler, L Smith, V. Loose, R. Beckman, J. Davis, D. Roberts & M. Williams. 1995. *An operational description of TRANSIMS*, LA-UR-95-2393, Los Alamos National Laboratory, Los Alamos, New Mexico.

Coleman, S., J. B. Hartle, T. Piran & S. Weinberg. 1991. Quantum Cosmology and Baby Universes. Singapore: World Scientific.

Fabian, A. C. 1998. *Evolution: Society, Science and the Universe*. Darwin College, Cambridge, U.K.: Cambridge University Press.

Friesz, T. L., D. H. Bernstein & R. R. Stough. 1996. Dynamic Systems, Variational Inequalities and Control Theoretic Models for Predicting Time-Varying Urban Network Flows. *Transp. Science* 30: 1.

Friesz, T. L., D. H. Bernstein, N. J. Mehta, R. L. Tobin & S. Ganjalizadeh. 1994. Day-to-day Dynamic Network Disequilibrium and Idealized Driver Information Systems. *Operations Research* 42(6): 1120-1131

Friesz, T. L., D. H. Bernstein, R. L. Tobin & B. W. Wie. 1993. A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem. *Operations Research* 37(6): 893-901.

Haynes, K. E. and J. Storbeck. 1978. The Entropy Paradox and the Distribution of Urban Population. *Socio-Economic Planning Science* 12: 1-6.

Haynes, K. E., F. Y. Phillips & J. W. Mohrfeld. 1980. The Entropies: Some Roots of Ambiguity. *Socio-Economic Planning Sciences* 14: 137-145.

Haynes, K. E. & F. Y. Phillips. 1981. Minimum Information Models and Individual Choice: Application to Linear and Loglinear Regression and Competitive Spatial Interaction Models. Ed. G. E. Lasker, in *Applied Systems and Cybernetics* II: 733-740.

Herman, R. & I. Prigogine. 1979. A Two-fluid Approach to Town Traffic. *Science* 204: 148-151.

Holland, J. H. 1995. *Hidden Order:How Adaptation Builds Complexity*. Reading MA.: Addison-Welsley.

Horowitz, E. & S. Sahni. 1985. *Fundamentals of Computer Algorithms*. Rockville: Computer Science Press Inc.

Jaynes, E. T. 1979. Where Do We Stand in Maximum Entropy? Eds. R. D. Levine and M. Tribus, in *The Maximum Entropy Formalism*. The MIT Press, Cambridge.

Kauffman, S. A. 1993. *Origins of Order*, Oxford University Press, New York.

Kauffman, S. A. & S. Levin. 1987. Towards a General Theory of Adaptive Walks on Random Landscapes. *J. Theoretical Biology* 128: 11.

Koutsopoulos, H. N. 1995. Traffic Simulation Models: Overview and Basic Elements. MIT Summer Professional Program 1.10s, *Modeling and Simulation for Dynamic Transportation Management Systems*.

Krugman, P. 1994. Complex Landscapes in Economic Geography. *The American Economic Review* 84(2): 412-417.

________________. 1995. *Development, Geography and Economic Theory*. Cambridge. MA.: MIT Press.

Kulkarni, R. G., R. R. Stough, K. E. Haynes 1996. Spin Glass and he Interactions of Congestion and Emissions. *Transportation Reach Board C* 4(6): 407-476.

____________. 2000. Toward a Percolation Model of Accessibility. *Computers, Environment and Urban Systems* 24: 421-434.

Lave C. A. 1985. Speeding, Coordination, & the 55 MPH Limit. *American Economic Review* 1159-1164.

Linde, A. 1994. *The Self-Reproducing Inflationary Universe.* Scientific American.

Mahmassani, H. 1995. Dynamic Traffic Assignment for ATIS/ATMS Operation and Evaluation: Issues, Formulations. Algorithms, Implementation and Numerical Experiments, *MIT Summer Professional Program 1.10s, Modeling and Simulation for Dynamic Transportation Management Systems.*

Mahmassani, H. & R. Herman. 1987. Interaction of Trip Decisions and Traffic Systems Dynamics. *European Journal Operations Research* 30: 304-317.

____________. 1990. Interactive Experiments for the Study of Tripmaker Behavior Dynamics in Congested Commuting Systems. In *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*, P. Jones (ed.). Aldershot, England: Avesbury, Gower Publishing. 272-298.

Mahmassani, H. R., Jayakrishnan & R. Herman. 1990. Network Traffic Flow Theory: Microscopic Simulation Experiments on Supercomputers. *Transaction Research A* 24A: 149-162.

Mahamassani, H. S., T. Hu & R. Jayakrishnan. 1992. Dynamic Traffic Assignment and Simulation for Advanced Network Informatics. Proceedings of the Second International Capri Seminar on Urban Traffic Networks.

Mahmassani, H. & S. Peeta. 1992. Dynamic Assignment for Route Guidance Decisions. Paper presented at the 39th North American Meetings of the Regional Science Association.

Murchland, J. D. 1970. Braess' Paradox of Traffic Flow. *Transp. Research* 4: 391-394.

Paelinck, J. H. P. & L. H. Klaassen. 1979. Spatial Economics. Hants., England: Saxon House, Teakfield Ltd.

Schelling, T. C. 1978. *Micromotives and Macrobehavior,* New York: W.W. Norton.

Smith, L., R. Beckman, K. Baggerly, D. Anson & M. Williams. 1995. *Overview of TRANSIMS, the TRansportation ANalysis and SIMulaton System Technical Report, LA-UR-95-2393.* Los Alamos National Laboratory, Los Alamos, New Mexico.

Suhir, E. 1997. Applied Probability for Engineers and Scientists. New York: McGraw-Hill.

Tanner, J. C. 1961. Factors Affecting the Amount of Travel. *Road Research Laboratory,* Technical Paper # 51, HMSO, London.

TRANSIMS. 1995. Sandia National Laboratories, Santa Fe, NM, USA.

TIGER/Line™Files. 1991. Census Feature Class Codes, Appendix E, TIGER/Line Census Files, 1990 prepared by the Bureau of the Census, Wash. D.C., The Bureau.

Transportation Research Board. 1985. Special Report No. 209, Highway Capacity Manual, National Research Council, Wash., D.C.

Tu Pierre, N. V. 1992. *Dynamical Systems: an Introduction with Applications in Economics and Biology*. New York: Springer Verlag. Berlin.

Wilson, A. G. 1969. The Use of Entropy maximizing models in the theory of trip distribution, mode split and route split. *Journal of Transp. Econ. Policy* 3:108-126.

______________. 1973. Further Developments of Entropy Maximizing Transport Models. *Transportation Planning Technology* 1: 283-306.

Wilson, E. O. 1995. *On Human Nature*. London: Penguin.

Wilson, R. 1988. *An Intro. to Dynamic Data Structures*. England: McGraw-Hill Books (UK) Ltd.

The School of Public Policy
George Mason University
Fairfax, VA 22030, USA
e-mail: khaynes@gmu.edu