Transaction A: Civil Engineering
Vol. 16, No. 3, pp. 226–237

Estimating Crash Risk Using a Microscopic Traffic Model

S.M. Sadat Hoseini[1,*] and M. Vaziri[1]

Abstract. *In this research, a microscopic model is developed that combines car following and lane changing models, describing driver behavior as a utility maximization process of drivers for reducing crash risk and increasing speed. This model is simulated by a cellular automata simulator and compared with the real data. It is shown that there is no reason to consider the model invalid for driver behavior in basic segments of the freeways in Iran, under non-congested conditions. Considering that the uncertain position of vehicles is caused by their acceleration or deceleration, a probability function is calibrated for calculating the presence probability of vehicles in their feasible cells. By multiplying the presence probability and impact of a crash, the crash risk of cells is calculated. An idea for estimating the crash risk of vehicles is introduced, named total risk. Total risk is the sum of risks on the path of the considered vehicles. It is shown that, when the difference between vehicle characteristics such as brake deceleration increases, crash risk also increases, and vice versa.*

Keywords: *Microscopic traffic simulation; Utility maximization; Cellular automata; Crash risk.*

INTRODUCTION

Microscopic and macroscopic approaches can be used for traffic flow analysis. In the microscopic approach, traffic behavior is modeled based on the movement of vehicles individually, while in the macroscopic approach traffic parameters such as density, volume and the average speed of vehicles are considered [1].

Most traffic system simulation applications today are based on the simulation of vehicle-vehicle interactions and are microscopic in nature. Traffic flow analysis is one of the few areas where macroscopic simulation has also been used. Most of the well-known macroscopic applications in this area originated in the late 1960's or the early 1970's. The British TRANSYT-program is an example of the macroscopic simulation of urban arterial signal control coordination, and the American FREQ- and FREFLO-programs plus the corresponding German analysis tool are related to motorway applications [2].

Research on macroscopic models has been conducted extensively to study the aggregate behavior of traffic flow. Many studies have also been implemented considering the movement behavior of drivers in a microscopic approach. In most of these studies, vehicle movements are defined in two models: Car following and lane changing [3-6]. In these models, car following and lane changing models are presented separately. Usually in car following models, it is assumed that drivers tend to reach their desired speed in their current lane as they prevent any collision with the front vehicles. When a driver is not able to reach his desired speed, lane changing to the neighboring lanes is considered [7-10]. Most lane changing models are based on the assumption that drivers evaluate the current and adjacent lanes and choose the lane with a higher average speed. Even in the few models that consider the untidiness of traffic flow and the lack of driver observance of driving lanes, the drivers' behavior is modeled according to their position in the driving lanes of the freeway [11].

In this research, a model has been implemented that joins the car following and lane changing models as a utility maximization process for decreasing crash risk and increasing speed. The proposed model is suitable for simulating driver behavior in basic freeway segments, where driver destination does not affect the movements.

A time-based simulation software is provided that

1. *Department of Traffic, Police University, Tehran, Iran.*
*. *Corresponding author. E-mail: Sadathoseini1@yahoo.com*

Received 31 May 2006; received in revised form 17 January 2007; accepted 7 February 2007

uses cellular automata to simulate traffic behavior. In cellular automata, a basic freeway segment is divided into cells. Individual vehicles in each simulation time step occupy single cells and their movements in these cells describe the movement of vehicles on the freeway [12-14]. Cellular automata are used for modeling pedestrian walking behavior as a discrete choice model [15] where there are no regulations and no lane to be considered. It has been shown that driver behavior in many countries cannot be modeled by traffic regulations and drivers do not consider driving lanes [16,17]. The proposed model is based on the realities of driver behavior on the freeways of such countries.

There are many studies that use image processing techniques for detecting the position of vehicles to be used for calibrating and validating the microscopic traffic models [18]. Considering the fact that many drivers do not pay attention to driving within the driving lanes in Iran, the use of ordinary image processing software will encounter many difficulties. This fact necessitates the development of new image processing techniques that can detect vehicle positions without considering driving lanes. In this research, an image processing computer program has been invented for determining the position of vehicles in the images taken from the motorway. This data is used for calibrating and evaluating the proposed model.

A new concept for measuring the crash risk of vehicles has been introduced, named total risk. Total risk is the sum of risks on the path of the considered vehicles. It has been shown that when the difference between vehicle characteristics like brake deceleration increases, the crash risk also increases, and vice versa.

There are four major phases in this research: Formulation of the proposed model, preparation of an image processing system for data collection of the real positions of vehicles, preparation of a micro simulation program for simulating the proposed model, evaluation, and measuring of crash risk using the proposed model.

MODEL DESCRIPTION

In the proposed model, every driver tries to increase his utility at the next time step. At first, feasible cells for each vehicle are determined and the utility of the feasible cells is calculated. Then, the cell with a higher utility is selected by a logit model. The same process is done for every vehicle at every time step until the end of the simulation duration. In other words, the following problem is simplified and solved for every vehicle ID at every time step, t:

$$\text{Max}\, U(X, Y, ID, t + \Delta t), \tag{1}$$

s.t.:

$$X(ID,t) + V_x(ID,t).\Delta t \leq X(ID, t+\Delta t) \leq X(ID,t) + V_x(ID,t).\Delta t,$$

$$Y(ID,t) + V_y(ID,t).\Delta t \leq Y(ID, t+\Delta t) \leq Y(ID,t) + V_y(ID,t).\Delta t.$$

where:

$X(ID,t)$:	longitudinal position of vehicle ID at time t,
$Y(ID,t)$:	lateral position of vehicle ID at time t,
$V_x(ID,t)$:	longitudinal speed of vehicle ID at time t,
$V_y(ID,t)$:	lateral speed of vehicle ID at time t,
$U(X,Y,ID,t+\Delta t)$:	utility of vehicle ID to be in the cell (X,Y) at time $t+\Delta t$.

The utility of a cell for a vehicle is determined by the maximum safe speed and risk value of that cell for the considered vehicle as will be explained in the following sections. The constraints of the above mentioned optimization problem are used to determine the feasible cells. In this way, the proposed model combines lane changing and car following models as an optimization process for increasing speed and decreasing crash risk. The algorithm of the proposed model is shown in Figure 1.

As can be seen in Figure 1, drivers calculate the crash risk and maximum safe speed of their feasible cells. Then, the rank value of feasible cells is calculated, using the maximum safe speed and crash risk. Drivers consider the rank of their feasible cells and randomly choose a cell, with respect to its rank, to go to the next time step. The probability of choosing a cell with higher rank is more than choosing a cell with a lower rank, but it is not deterministic and the cell with the highest rank is not necessarily chosen. After choosing the cell with a higher rank, the driver must adjust his speed to the maximum safe speed of the chosen cell with respect to the deceleration or acceleration of his vehicle. If the current speed of vehicle is less than the maximum safe speed, the driver will increase his speed with respect to the acceleration rate of his vehicle. If the current speed of the vehicle is more than the maximum safe speed the driver will decrease his speed with respect to the deceleration rate of his vehicle.

In the next sections, the calculation of maximum safe speed, crash risk, utility and rank value of cells will be described.

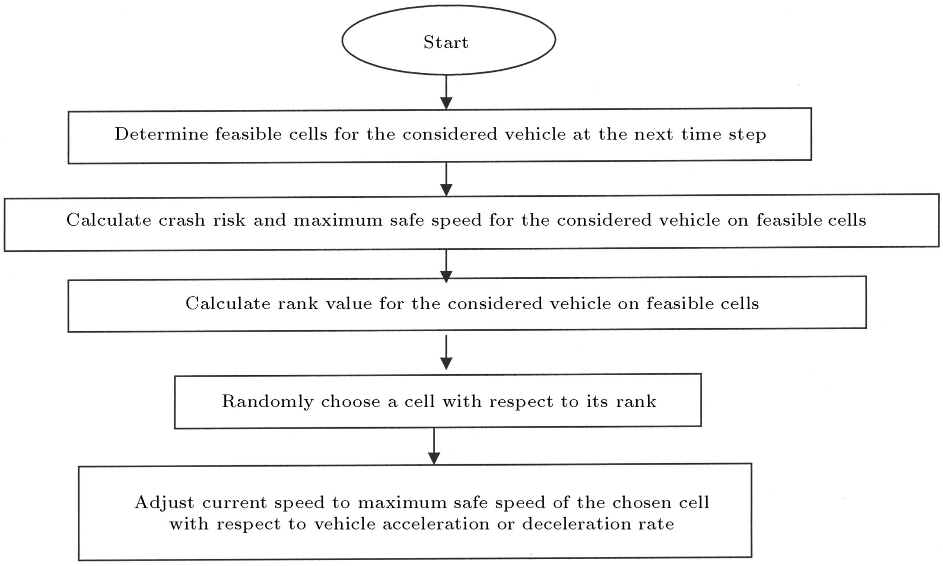

Figure 1. Applied algorithm in the proposed model.

Maximum Safe Speed

The maximum safe speed, V_{MSS}, of a cell is the maximum speed that a vehicle can reach on that cell, avoiding collision with the front vehicle. The maximum safe speed of a cell is determined by the position and speed of its front vehicles. A function must be defined for the V_{MSS} value of each cell.

Suppose vehicle 2 is following vehicle 1, their speeds are V_2 and V_1, respectively, the distance between vehicles is d meters and L is the length of the front vehicle. It is necessary to calculate a safe distance for reducing the probability of collision between vehicles.

If the front vehicle sees an obstacle and brakes at t_0, vehicle 2 will brake after reaction time τ at $t_0 + \tau$. In order to prevent a collision between vehicles, it is important that there exists a safe distance between vehicles according to their speed and brake deceleration.

Considering the movement equation of vehicles for avoiding a collision, the following equation must be applied where a_1 and a_2 are the brake deceleration of vehicles 1 and 2, respectively:

$$L + V_2\tau + \frac{V_2^2}{2a_2} = \frac{V_1^2}{2a_1} + d. \tag{2}$$

Equation 3 shows the maximum safe speed of vehicle 2 to avoid an accident.

$$V_{\mathrm{MSS}}(X,Y,ID,t) = V_2 =$$

$$-\,a_2\tau + \sqrt{a_2^2\tau^2 - 2a_2\left(L - d - \frac{V_1^2}{2a_1}\right)}, \tag{3}$$

where $V_{\mathrm{MSS}}(X,Y,ID,t)$ is the maximum safe speed of cell (X,Y) for vehicle ID at time t. In this way, the maximum safe speed of a cell can be calculated, using the distance between that cell and its front vehicle, d, speed of front vehicle, V_1, and brake deceleration of two vehicles, a_1 and a_2.

Crash Risk

The idea of risk is complex, but conventionally its definition has been consistent. Risk can be defined as a measure of the probability and impact of adverse effects [19]. Crash risk is calculated by multiplying the impact and probability of a crash. In this research, the crash risk of a cell for vehicle ID is the probability of collision between vehicle ID and any other vehicle on that cell, multiplied by the impact of collision between them. Impact is set to 1 for all crashes. As there may be more than one vehicle that can cause a collision on a cell, the crash risk of vehicle ID on cell (X,Y) at time t is the maximum risk created by the other vehicles and can be calculated by the following equation:

$$\mathrm{Risk}\,(X,Y,ID,t) =$$

$$\mathrm{Max}\,\{\mathrm{Impact}(X,Y,ID,t).\ P(X,Y,i,t)\}$$

$$\text{for all} \quad i \neq ID, \tag{4}$$

where:

$\mathrm{Risk}(X,Y,ID,t)$:	is the crash risk for vehicle ID on the cell (X,Y) at time t,
$\mathrm{Impact}\,(X,Y,ID,t)$:	collision impact of vehicle ID on the cell (X,Y) at t which is set to 1,

$P(X,Y,i,t)$: presence probability of vehicle i to be in cell (X,Y) at t.

An accident occurs when two vehicles occur on the same cell. Therefore, collision probability is the presence probability of two vehicles on the same cell. Consider vehicle i to be on cell $(X_i(t),Y_i(t))$ at t and on cell $(X_i(t+\Delta t),Y_i(t+\Delta t))$ at $t+\Delta t$. The longitudinal and lateral position of vehicle i at $t+\Delta t$ can be calculated by the following equations:

$$X(i,t+\Delta t)=X(i,t)+V_x(i,t).\Delta t, \tag{5}$$

$$Y(i,t+\Delta t)=Y(i,t)+V_y(i,t).\Delta t. \tag{6}$$

where is $V_x(i,t)$ is the longitudinal speed of vehicle i at t, and $V_y(i,t)$ is the lateral speed of vehicle i at t.

Considering the independency of the lateral and longitudinal positions of vehicles, the presence probability of vehicle i on cell (X,Y) can be calculated by the following equation:

$$P(i,X,Y,t+\Delta t)=P(i,X,t+\Delta t).P(i,Y,t+\Delta t), \tag{7}$$

where:

$P(i,X,Y,t+\Delta t)$ probability of vehicle i to be on cell (X,Y) at $t+\Delta t$,
$P(i,X,t+\Delta t)$ probability of vehicle i to be on longitudinal position X at $t+\Delta t$,
$P(i,Y,t+\Delta t)$ probability of vehicle i to be on lateral position Y at $t+\Delta t$.

The longitudinal position of vehicle i at $t+\Delta t$ is calculated by the following equation:

$$X(i,t+\Delta t)=X(i,t)+V_x(i,t).\Delta t+\frac{1}{2}a_x(i,t).\Delta t^2, \tag{8}$$

where $a_x(i,t)$ is the longitudinal acceleration rate of vehicle i at t that can be calculated by the following equation:

$$a_x(i,t)=\frac{V_x(i,t+\Delta t)-V_x(i,t)}{\Delta t}. \tag{9}$$

For calculating the longitudinal speed, the longitudinal position of vehicle i at time t, $X(i,t)$, is determined and compared with the longitudinal position of the same vehicle at $t+\Delta t, X(i,t+\Delta t)$. The longitudinal speed of vehicle i is calculated by the following equation:

$$V_x(i,t)=\frac{X(i,t+\Delta t)-X(i,t)}{\Delta t}. \tag{10}$$

For estimating the longitudinal positions of vehicles, it is important to notice that drivers can calculate the speed of other vehicles in the last time steps, but they do not know the change of speed at the next time step. Considering that acceleration is a change in speed, the uncertainty of longitudinal speed is caused by acceleration of vehicles. Supposing that acceleration is distributed as a normal distribution with average μ_{ax} and standard deviation σ_{ax}; the longitudinal position of vehicle i at $t+\Delta t$ can be estimated by the following distribution:

$$X(i,t+\Delta t)\approx \quad \mathrm{Normal}_x\left(X(i,t)+V_x(i,t).\Delta t+\frac{1}{2}\mu_{ax}.\Delta t^2,\frac{1}{2}\sigma_{ax}.\Delta t^2\right). \tag{11}$$

Considering Equation 11, the presence probability of vehicle i to be on longitudinal position X can be calculated by the following equation:

$$P(i,X,t+\Delta t)=\int_{X-\frac{l_v}{2}}^{X+\frac{l_v}{2}}\mathrm{Normal}_x(x;x_i(t)+V_x(i,t).\Delta t+\frac{1}{2}\mu_{ax}.\Delta t^2,\frac{1}{2}\sigma_{ax}.\Delta t^2).dx. \tag{12}$$

The position of each cell is determined by the position of the center of that cell. The length of all cells is equal, shown by l_v. Therefore, the lower and upper boundaries of the integral are determined as $X-l_v/2$ to $X+l_v/2$.

In the same way, as mentioned regarding longitudinal position, the lateral position can be calculated by the following equation:

$$Y(i,t+\Delta t)=Y(i,t)+V_y(i,t).\Delta t+\frac{1}{2}a_y(i,t).\Delta t^2, \tag{13}$$

where $a_y(i,t)$ is the longitudinal acceleration of vehicle i at t and can be calculated by the following equation:

$$a_y(i,t)=\frac{V_y(i,t+\Delta t)-V_y(i,t)}{\Delta t}. \tag{14}$$

For calculating the lateral speed, the lateral position of vehicle i at time t, $Y(i,t)$, is determined and compared with the lateral position of the same vehicle at $t+\Delta t, Y(i,t+\Delta t)$. The lateral speed of vehicle i is calculated by the following equation:

$$V_y(i,t)=\frac{Y(i,t+\Delta t)-Y(i,t)}{\Delta t}. \tag{15}$$

For estimating the lateral position of vehicles, it is important to notice that drivers distinguish the lateral position of other vehicles in the last time steps, but they do not know the change of lateral position at the next time step. Considering that lateral acceleration is the change in lateral speed; the uncertainty of lateral speed is caused by the lateral acceleration of vehicles. Supposing that acceleration is distributed as a normal distribution, with mean μ_{ay} and standard deviation σ_{ay}; the lateral position of vehicle i at $t+\Delta t$ can be estimated by the following distribution:

$$Y(i,t+\Delta t) \approx \mathrm{Normal}_y(Y(i,t)+V_y(i,t).\Delta t + \frac{1}{2}\mu_{ay}.\Delta t^2, \frac{1}{2}\sigma_{ay}.\Delta t^2). \tag{16}$$

Considering Equation 16, the presence probability of vehicle i to be in lateral position Y can be calculated by the following equation:

$$P(i,Y,t+\Delta t) = \int_{Y-\frac{W_v}{2}}^{Y+\frac{W_v}{2}} \mathrm{Normal}_y(y; y_i(t)+V_y(i,t).\Delta t + \frac{1}{2}\mu_{ay}.\Delta t^2, \frac{1}{2}\sigma_{ay}.\Delta t^2).dy. \tag{17}$$

Positions of cells are determined by the position of their centers. The width of all cells is equal and is shown by Wv. Therefore, the lower and upper boundaries of the integral are determined as $Y - W_v/2$ to $Y + W_v/2$.

Rank Value

After calculation of the risk value and maximum safe speed of each cell for the considered vehicle, feasible cells are checked to find the cell which provides the maximum rank. The rank of a cell is calculated by a logit model, shown in the following equation:

$$\mathrm{Rank}\,(X,Y,ID,t+\Delta t) = \frac{e^{U(X,Y,ID,t+\Delta t)}}{\sum e^{U(i,j,ID,t+\Delta t)}}$$

$$\text{for all feasible cells} \quad (i,j), \tag{18}$$

where:

Rank $(X,Y,ID,t+\Delta t)$:	rank of cell (X,Y) for vehicle ID at time $t+\Delta t$,
$U(X,Y,ID,t+\Delta t)$:	utility of cell (X,Y) for vehicle ID.

It is assumed that the utility function is in a nonlinear form and can be shown in Equation 19.

$$U(X,Y,ID,t+\Delta t) = \frac{V_{\mathrm{MSS}}\,(X,Y,ID,t+\Delta t)^{\alpha}.\rho^{D}}{\mathrm{Risk}\,(X,Y,ID,t+\Delta t)^{\beta}}. \tag{19}$$

As the drivers prefer to move straight forward, the cell that is in front of the vehicle is chosen more than the others. Therefore, a dummy variable, D, is defined to consider the effect of being in the same lateral position. This variable is defined in the following equation.

$$D = \begin{cases} l & \text{for the cells in the same lateral position} \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Equation 19 can be written in linear form as below:

$$\ln(U(X,Y,ID,t+\Delta t)) = \alpha.\ln(V_{\mathrm{MSS}}(X,Y,ID,t+\Delta t)) - \beta.\ln(\mathrm{Risk}(X,Y,ID,t+\Delta t)) + D\ln\rho. \tag{21}$$

In Equation 18, it is assumed that drivers consider only three cells in front of their vehicles as feasible cells at the next time step. In this way, three equations for these three cells can be derived from Equation 18:

$$\mathrm{Rank}\,(X+1,Y,ID,t+\Delta t) = \frac{\exp(U(X+1,Y,ID,t+\Delta t)}{\sum_{Y'=Y-1}^{Y+1}\exp(U(X+1,Y',ID,t+\Delta t))}, \tag{22}$$

$$\mathrm{Rank}\,(X+1,Y-1,ID,t+\Delta t)) = \frac{\exp(U(X+1,Y-1,ID,t+\Delta t))}{\sum_{Y'=Y-1}^{Y+1}\exp(U(X+1,Y',ID,t+\Delta t))}, \tag{23}$$

$$\mathrm{Rank}\,(X+1,Y+1,ID,t+\Delta t) = \frac{\exp(U(X+1,Y+1,ID,t+\Delta t)}{\sum_{Y'=Y-1}^{Y+1}\exp(U(X+1,Y',ID,t+\Delta t))}. \tag{24}$$

After calculating the rank of the three feasible cells for each vehicle, the driver chooses a cell, with respect to its rank, for moving to the next time step.

Vehicle Acceleration and Deceleration

When a vehicle chooses a cell to move to the next time step, it must adjust its speed to the maximum safe speed on that cell. This cannot be done immediately and must be done according to the vehicles' acceleration or deceleration.

If the current speed of vehicle ID is less than the maximum safe speed, it can increase its speed with respect to its acceleration. In this way, the speed of the vehicle at the next time step can be determined, using the following equation:

$$V(ID,t+\Delta t) = V(ID,t) + a(ID).\Delta t, \tag{25}$$

where:

$V(ID, t+\Delta t)$:	speed of vehicle ID in the next time step,
$V(ID, t)$:	speed of vehicle ID in the current time step,
Δt:	duration of time step,
$a(ID)$:	acceleration rate of vehicle ID.

If the current speed of vehicle ID is more than the maximum safe speed, it must decrease its speed, with respect to its deceleration rate. In this case, the speed of the vehicle at the next time step can be determined, using the following equation:

$$V(ID, t+\Delta t) = V(ID, t) - b(ID).\Delta t, \qquad (26)$$

where $b(ID)$ is the brake deceleration of vehicle ID.

DATA COLLECTION

Two sites, each one covering a basic segment of about 100 meters of the Tehran-Karaj freeway are used for videotaped observation. This freeway connects Tehran to a nearby city, Karaj. In selected segments of the freeway, there are four moving lanes in each direction and each lane is 3.65 meters wide. The camera was installed on a bridge over the freeway and a good view of the freeway was accessible. Simple plans of the two sites are shown in Figures 2 and 3.

The camera position on the bridge and other important distances are shown in Figures 2 and 3. The videotaped section is shown by a gray rectangle on the freeway.

The duration of observation was about 30 minutes at each site. The position, time and type of vehicle passing the segments during the observation, in every frame of the film, are detected and stored in a table, using an image processing system [16]. In the image processing program, the considered freeway has been divided into windows which are used for detecting the vehicles on the freeway. The size of these windows is determined such that they can provide a complete view of the considered freeway with enough resolution and accuracy. Windows are arranged in horizontal rows and each row contains 12 windows. Considering the non-linear projection of 3D images to 2D images, vertical distances between rows of windows are determined by a nonlinear equation [16]. Figure 4 shows the windows on a sample image of the Tehran-Karaj freeway.

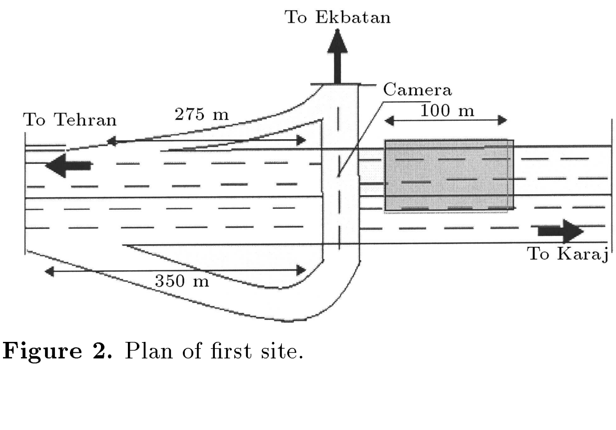

Figure 2. Plan of first site.

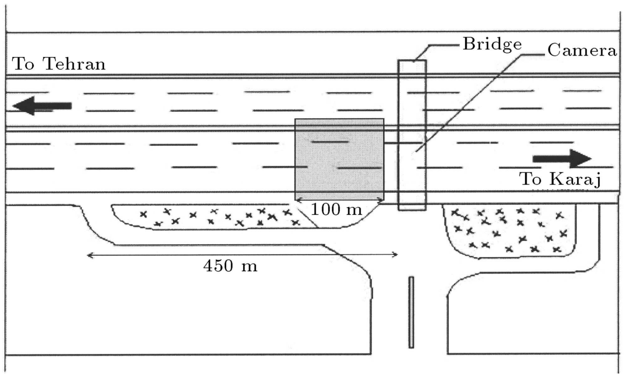

Figure 3. Plan of second site.

Windows positions are determined such that all distances between rows of windows show a specified distance on the freeway. So, each window can be matched to a specific space on the street. Each window is processed for detecting whether there is a vehicle on it or not. Vehicle positions in each frame are saved in a table in the form of (ID, X, Y, K, t) in which:

ID	identity number of the vehicle,
X, Y	longitudinal and lateral positions,
K	vehicle type according to size (truck or bus: $K=1$; private car: $K=0$)
t	time of observance.

Using the table of vehicle positions, the microscopic traffic characteristics of vehicles are determined. For example, the speed of a vehicle ID at time t could be calculated using the position and time of that vehicle by Equation 10. Macroscopic parameters like average speed are calculated by averaging the speed of vehicles.

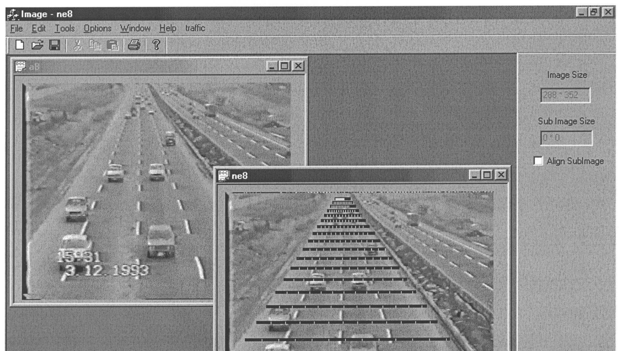

Figure 4. Windows on a sample image of Tehran-Karaj freeway.

The first site is observed from 15:30 to 16:00 and 1392 vehicles are detected. Average speed was about 83 km/hr and average density was about 32 veh/km. In the same way, the second site is observed from 15:30 to 16:00 and 1397 vehicles are detected. Average speed was about 82 km/hr and density was about 33 veh/km.

MODEL CALIBRATION

For calibrating the maximum safe speed; V_{MSS}, in Equation 3, the entire vehicle deceleration rate is set to 5 m/s^2, L is set to 5 and the reaction time is set to 2 seconds [20]. In this way, Equation 3 can be calibrated as below:

$$V_{\text{MSS}}\,(X,Y,ID,t) = -10 + \sqrt{50 + 10d + V_i^2}. \qquad (27)$$

Using the above equation, the maximum safe speed of cells can be calculated according to its distance from the nearest front vehicle, d, and the speed of that vehicle, V_i.

For calibrating the presence probability of vehicle i in longitudinal position $X, a_x(i,t)$ is calculated for different vehicles, i, at different times t. In this way, a sample of a_x is prepared, the frequency of different values of a_x in the sample is calculated and the following problem is solved to estimate the distribution function:

$$\min \sum (\text{frequency}(a_x) - \text{Normal}_x(a_x;\mu_{ax},\sigma_{ax}))^2$$

$$\text{for all observed} \quad a_x, \qquad (28)$$

where Frequency (a_x) is the frequency of observing longitudinal acceleration a_x in the sample data and $\text{Normal}_x(a_x;\mu_{ax},\sigma_{ax})$ is the value of a_x in a normal distribution with average μ_{ax} and standard deviation σ_{ax}.

It is shown that frequency of a_x can be best fitted as a normal distribution with average $\mu_{ax} = 0$ and standard deviation $\sigma_{ax} = 82$, with R^2=0.92. Considering l_v=5, Equation 12 is calibrated as the following equation:

$$P(i,X,t+\Delta t) =$$

$$\int_{X-2.5}^{X+2.5} \text{Normal}_x(x;x_i(t)+V_x(i,t).\Delta t, 41.\Delta t^2).dx. \qquad (29)$$

In the same way, for calibrating the presence probability of vehicle i in lateral position $Y, a_y(i,t)$ is calculated for different vehicles, i, at different times, t. In this way, a sample of a_y is chosen and the frequency of different values of a_y in the sample is calculated. It is shown that the frequency of a_y can be fitted as a normal distribution, with average $\mu_{ay} = 0$ and standard deviation $\sigma_{ay} = 3$, with $R^2 = 0.93$. Considering $W_v = 2$, Equation 17 is calibrated as Equation 30:

$$P(i,Y,t+\Delta t) =$$

$$\int_{Y-1}^{Y+1} \text{Normal}_y(y;y_i(t)+V_y(i,t).\Delta t, 1.5.\Delta t^2).dy. \qquad (30)$$

For calibrating the utility function in Equation 21, the frequency percentage of vehicles that choose front cell, Rank $(X+1,Y,ID,t+\Delta t)$, left cell, Rank $(X+1,Y-1,ID,t+\Delta t)$ and right cell, Rank $(X+1,Y+1,ID,t+\Delta t)$, in the next time step, $t+\Delta t$, is calculated, using the sample data as 0.97, 0.017 and 0.013, respectively. Considering the rank of the feasible cells to be equal to their frequency percentage, Equations 22, 23 and 24 are calibrated and the coefficients of Equation 21, α, β and ρ, are determined. Using a linear regression method, it is shown that Equation 21 can be calibrated in the following equation:

$$\ln(U(X,Y,ID,t)) = 16.4.\ln(V_{\text{MSS}}(X,Y,ID,t)) -$$

$$4.5.\ln(\text{Risk}\,(X,Y,ID,t)) + D\ln(7.89^*10^{13}). \qquad (31)$$

where all of the variables are defined in Equation 21 and $\text{R}^2 = 0.94$ for Equation 31.

Although acceleration and deceleration rates vary for different kinds of vehicles; vehicle acceleration is set to 1.2 m/s^2 and deceleration is set to 5 m/s^2 [20] in Equations 25 and 26.

PREPARED MICROSCOPIC TRAFFIC SIMULATION SOFTWARE

A time-based simulation software is prepared for simulating the movement behavior of vehicles, using the proposed model [16].

In the prepared simulation software, the freeway is divided into some cells, each cell almost equal to the length of a private car, with 5 meters length and 2 meters width. In each time step, each vehicle occupies a cell and the movement of vehicles is described by their movements in the cells. Vehicles are created in the time and location that they have been seen at for the first time in real data. Only the first position of a vehicle in the simulation program is the same as in real data and the positions in the next time steps are calculated using the proposed model.

Figure 5 shows a moment of the simulation software as it is executing the simulation. In Figure 5, the existence of vehicles in cells is specified by a change in cell color and the ID number of vehicles is shown on the cells.

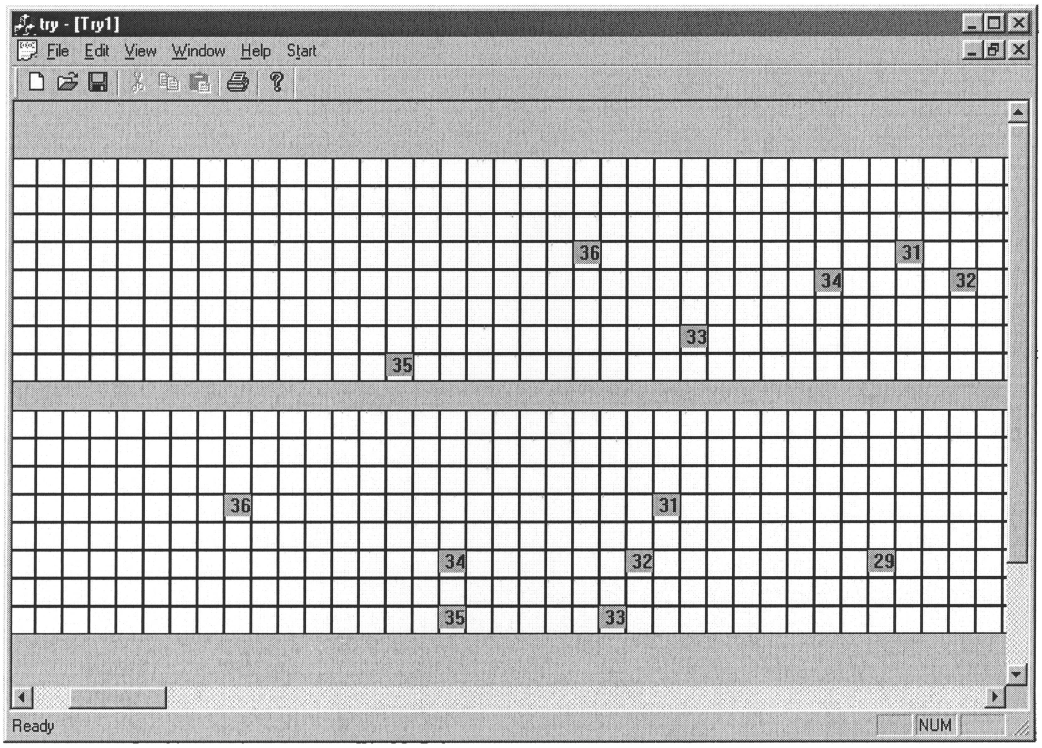

Figure 5. Simulation and real position of vehicles.

In the upper part of the figure, vehicles on the freeway are simulated and, in the lower part, vehicles on the freeway are positioned as they have been observed in the real world, using the collected data. In this way, the real world and the simulation model can be compared and verified visually.

AN EXAMPLE OF DRIVER BEHAVIOR

In Figure 6, an example of moving vehicles in a segment of freeway is shown, where vehicles are moving from left to right. The considered vehicle has been shown in the black cell by the number 1; other vehicles are shown in gray cells. Each cell is a 5 by 2 meter rectangle on the freeway surface. The considered section is a 4-lane freeway, each lane about 4 meters wide. In this way, the width of the freeway is separated into 8 cells.

In this example, the speed limit is 120 km/hr. The speed of vehicles 3 and 4 is 108 km/hr and the speed of vehicles 1 and 2 is 72 km/hr. Assuming the time step to be 0.2 seconds, vehicles 3 and 4 will move about 6 meters and vehicles 1 and 2 will move about 4 meters in the next time step. The cells of the freeway are 5 meters long, so in the simulation, all of the vehicles move one cell forward; but 1 meter will be added to the distance traveled by vehicles 3 and 4 and 1 meter will be differentiated from the distance traveled by vehicles 1 and 2 at the next time step. The maximum safe speed of three feasible cells of the considered vehicle are calculated using Equation 27 that is shown in Figure 6 in meters per second.

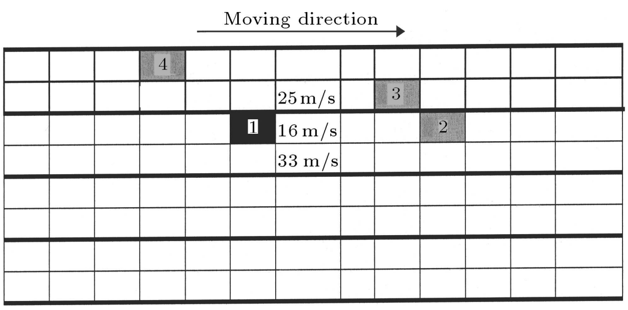

Figure 6. Maximum safe speed of feasible cells.

The presence probabilities of vehicles to be in the left, straight and right cells in front of the considered vehicle are calculated using Equation 7 and shown in Table 1. For the cells which are influenced by more than one vehicle, the maximum presence probability is chosen as the presence probability. Multiplying the impact and presence probability, the risk value of feasible cells has been calculated and shown in Figure 7. Vehicle 1 considers the rank of its feasible cells and

Table 1. Presence probability of vehicles in the feasible cells.

ID	2	3	4	Max
Left	1.96E-25	8.35E-17	1.77E-06	1.77E-06
Straight	3.72E-24	4.41E-18	1.06E-11	1.06E-11
Right	1.96E-25	2.65E-23	0.00E+00	2.65E-23

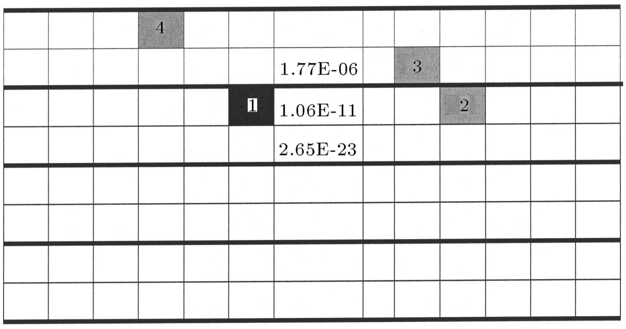

Figure 7. Risk values created for the considered vehicle.

chooses a cell with respect to its rank. The utility of each cell is calculated using Equation 31 as follows:

$$\ln(U_1) = 16.4.\ln(25) - 4.5.\ln(1.77^*10^{-6}) = 112, \quad (32)$$

$$\ln(U_2) = 16.4.\ln(16) - 4.5.\ln(1.06^*10^{-11}) + \ln(7.89^*10^{13}) = 191, \quad (33)$$

$$\ln(U_3) = 16.4.\ln(33) - 4.5.\ln(2.6^*10^{-23}) = 291. \quad (34)$$

The rank values of feasible cells are calculated, using Equation 18 and shown in Figure 8. As can be seen, the cell at the right of the vehicle has the higher rank and considering the low value of other cells, it will be chosen in the next time step.

The maximum safe speed of the chosen cell is 33 m/s or 120 km/hr. By considering that the speed of vehicle 1 is 72 km/hr or 20 m/s, vehicle 1 will increase its speed. Assuming the acceleration rate to be 1.2 m/s^2 and the time step to be 0.2 second, vehicle 1 will increase its speed about (0.2*1.2=) 0.24 m/s in the next time step. In this way, vehicle 1 will go to the right cell and will increase its speed to 20.24 m/s or 72.864 km/hr at the next time step.

VALIDATION OF THE PROPOSED MODEL

As mentioned in the data collection section, the duration of observation was about half an hour at each of the two sites. The collected data of the first site is used for calibrating the model and the second site is used for validating the model. Model validation is done both at the microscopic and macroscopic levels. At the macroscopic level, the average speed and number of lane changes in the simulation and in the real world are compared and, at the microscopic level, the longitudinal and lateral positions of vehicles are compared.

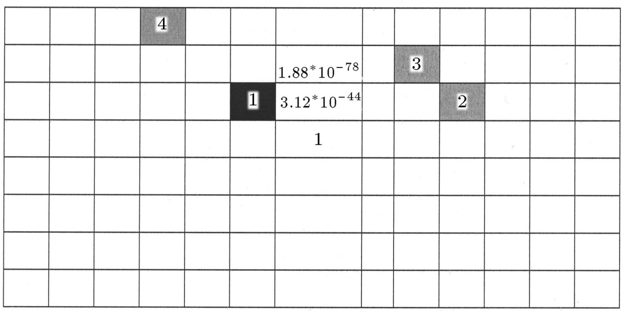

Figure 8. Rank values of feasible cells.

The duration of observation at the second site is divided into 6 intervals, each about 5 minutes. The simulation program is executed for each of the intervals individually. Vehicles are created in the simulation program at the time and position that they have been seen for the first time in the real world.

There are different methods for validating the microscopic models of traffic [21]. In this research, a paired observation test is used for validating the proposed model. Test statistics of the paired observation test are calculated by the following equation:

$$t_{po} = \frac{\overline{d}}{\frac{S_d}{\sqrt{n}}}, \quad (35)$$

where:

- t_{po}: test statistics,
- $\overline{d}$: average of difference between real and simulation pairs,
- S_d: standard deviation of difference between real and simulation pairs,
- n: number of pairs.

If $|t_{po}| > t_{\alpha/2,n-1}$, validation of the model is rejected at the α % level of confidence, otherwise there is no reason to consider the model invalid [22].

Macroscopic Validation

The average speed and number of lane changes in the real world and in the simulation program are calculated for each of the 6 intervals. Then, the difference of average speed and number of lane changes between the real world and the simulation is calculated.

The mean and standard deviation of difference of the average speed between the real world and the simulation is 0.3 and 5.04, consequently. In this way, t_{po} can be calculated as $t_{po} = \frac{0.3}{\frac{5.04}{\sqrt{6}}} = 0.36$. By considering $|t_{po}| = 0.36 < t_{\alpha/2,5} = 3.19$, there is no reason to consider the model invalid at the 95% level of confidence for simulating the average speed. In the same way, the average and standard deviation of the difference in the number of lane changes between the real world and the simulation is -1.5 and 5.3, consequently. In this way, t_{po} can be calculated as $t_{po} = \frac{-1.5}{\frac{5.3}{\sqrt{6}}} = -1.69$. By considering

$|t_{po}| = 1.69 < t_{\alpha/2,5} = 3.19$, there is no reason to consider the model invalid at the 95% level of confidence for simulating the number of lane changes of vehicles.

Microscopic Validation

The proposed model is verified visually for longitudinal positions. The real longitudinal position of some vehicles is compared with their longitudinal position in simulation. Figure 9 shows the longitudinal position of a vehicle in reality and in simulation. Although the graphs are not the same at every time step, they are accompanying each other.

Using the gathered data of the longitudinal position of vehicles in one of the 5-minute intervals and comparing them with the simulation data, t_{po} is calculated for the longitudinal position as $t_{po} = \frac{0.132}{\frac{0.486}{\sqrt{69}}} = -2.25$. Considering $|t_{po}| = 2.25 < t_{\alpha/2,n-1} = 2.7$, there is no reason to consider the model invalid at the 95% level of confidence for simulating the longitudinal position.

The proposed model is verified visually for lateral positions too. The real lateral positions of some vehicles are compared with their lateral positions in simulation. Figure 10 shows the lateral position of a vehicle in reality and simulation. Although the graphs are not the same at every time step, they are accompanying each other.

In the same way, using the gathered data of the lateral position of vehicles in one of the 5-minute intervals and comparing them with the simulation data, t_{po} is calculated for the lateral position as $t_{po} = \frac{-0.044}{\frac{0.61}{\sqrt{69}}} = -0.597$. Considering $|t_{po}| = 0.597 < t_{\alpha/2,n-1} = 2.7$, there is no reason to consider the model invalid at the 95% level of confidence for simulating the lateral position.

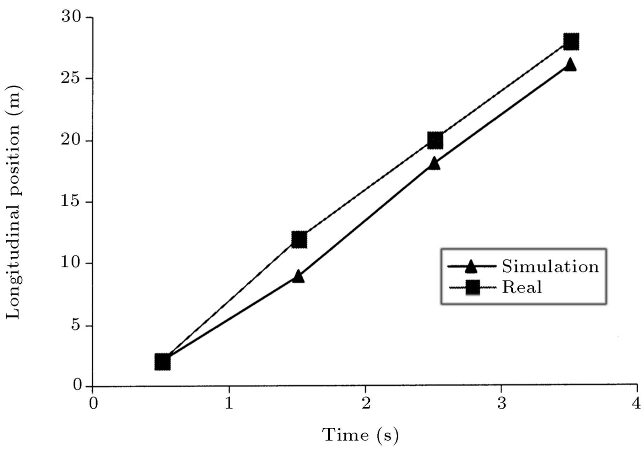

Figure 9. Real and simulation graphs of longitudinal position of a vehicle.

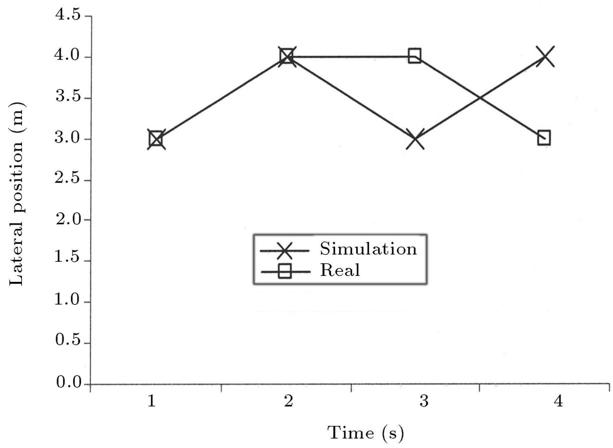

Figure 10. Real and simulation graphs of lateral position of a vehicle.

EFFECT OF VEHICLE CHARACTERISTICS ON CRASH RISK

A collision between vehicles is a rare event and in the collected sample data there is no collision to be observed. However, a driver's behavior is not based on the real collisions, but on the risk of collision which that driver considers. In this research, micro-simulation is used for studying the effect of differences between vehicle characteristics on the total risk of collision.

Total risk is the sum of the risks of the cells, which are occupied by the considered vehicles during the simulation and can be calculated using the following equation:

$$\text{Total Risk} = \sum_{ID=1}^{N} \sum_{t=1}^{T} \text{Risk}\,(x, y, ID, t), \tag{36}$$

where N is the number of considered vehicles and T is the number of time steps during the simulation.

In this research, total risk is used as a parameter for measuring crash risk in the considered segment of the freeway. In the previous sections, it is assumed that different vehicles have the same brake deceleration and the model has been calibrated based on this assumption. In this section, differences between vehicle characteristics and their effect on the total risk of collision are studied.

In Equation 27, it is assumed that $a_1 = a_2 = 5$ m/s^2. Here, brake deceleration is assumed to be different for different vehicles and distributed as a normal distribution function as shown in Equation 37:

$$b(ID) = \text{normal}\ (\mu_b, \sigma_b), \tag{37}$$

where:

$b(ID)$ brake deceleration of vehicle ID,
μ_b average of brake deceleration of vehicles,

σ_b standard deviation of brake deceleration of vehicles.

Considering $\mu_b = 5$ m/s^2, total risk is calculated in 10 runs of the simulation for different values of σ_b and total risk is calculated using the prepared simulation software. In Figure 11, the total risk average in 10 runs for different values of σ_b has been shown.

As can be seen in Figure 8, total risk average increases when the standard deviation of the brake deceleration increases. It means that, when the difference between vehicle brake deceleration increases, more crashes are expected.

CONCLUSION

A cellular model is proposed for simulating driving behavior on the freeways of Iran. This model, which is based on the utility maximization process of drivers for increasing speed and decreasing crash risk, is micro-simulated and it is shown that the proposed model is valid for describing driving behavior on basic segments of the freeway.

The maximum safe speed of vehicles is calculated using the laws of motion. For calibrating the maximum safe speed, reaction time and vehicle acceleration rate are selected from previous studies [20].

In this research, a new concept for estimating crash risk is introduced. Considering that the uncertainty of the position of vehicles is caused by their acceleration and deceleration at different times, a probability distribution function is calibrated for calculating the presence probability of vehicles on feasible cells of a freeway. By multiplying the presence probability and impact of a crash, the crash risk of cells is calculated.

Using the proposed model, total risk is introduced as a parameter for measuring the risk of collision and it has been shown that when differences between vehicle characteristics like brake deceleration increase, the crash risk also increases and vice versa. Therefore, it is advisable for a country to use vehicles with similar characteristics to decrease the rate of accidents.

This study deals with basic freeway segments where drivers' different destinations for exiting or entering the freeway or continuing straight on do not affect their movement behavior. For future studies, the destination factor could be added to the optimization process. In this way, the proposed model could be used, not only in the basic freeway segment, but also in other parts of the road network. In this research, driver behavior is modeled under not congested situations, where there are empty cells around the considered vehicle to be selected. For future studies, it is advisable to study traffic dynamics under congested situations.

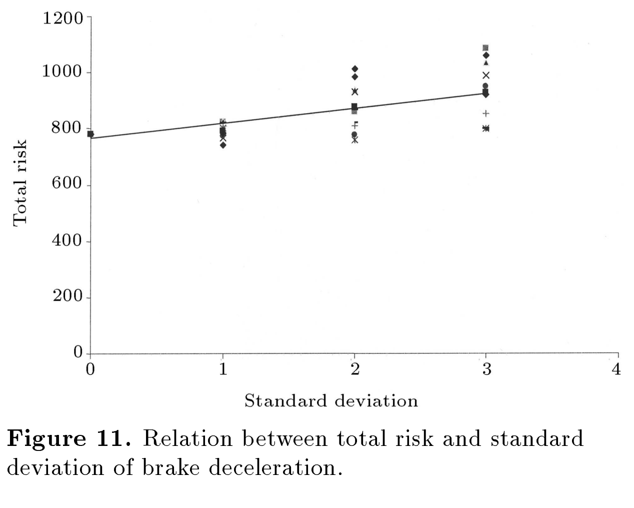

Figure 11. Relation between total risk and standard deviation of brake deceleration.

REFERENCES

1. Kachroo, P. and Ozbay, K. "Feedback control theory for dynamic traffic assignment", *Springer*, pp. 19-43 (1999).

2. Pursula, M. "Simulation of traffic systems", *Journal of Geographic Information and Decision Analysis*, **3**(1), pp. 1-8 (1999).

3. Kazi Iftekhar, A., *Modeling Drivers' Acceleration and Lane Changing Behavior*, Massachusetts Institute of Technology (1999).

4. Gipps, P.G. "A behavioral car-following model for computer simulation", *Transportation Research B*, **15**, pp. 105-111 (1981).

5. Hidas, P. "A car-following model for urban traffic simulation", *Traffic Engineering and Control*, **39**, pp. 300-305 (1998).

6. Kita, E., Tamaki, T. and Shimizu, H. "Cellular automata simulation of traffic flow through a road with a construction zone", *Transactions of the Wessex Institute*, http://library.witpress.com/pages/ PaperInfo.asp?PaperID =16769 (2008).

7. Denos, C.G., Herman, R. and Rothery, R. "Nonlinear follow-the-leader models of traffic flow", *Operations Research*, **9**(4), pp. 545-567 (1961).

8. Heng, W., Meyer, E., Lee, J. and Chuen, F. "Characterizing lane-choice behavior to build rules as part of lane-based traffic micro-simulation hierarchy", *Transportation Research Board (CD-ROM)*, 81st Annual Meeting (2002).

9. Sultan, B. "The lane changing process: Data analysis & modeling behavior", *Transportation Engineering and Control*, pp. 202-207 (2001).

10. Shahab Uddin, M. and Ardekani, S. "An observation study of lane changing on basic freeway segment", *Transportation Research Board (CD-ROM)*, 81st Annual Meeting (2002).

11. Gunay, b. "Modeling lane discipline on multilane uninterrupted traffic flow", *Traffic Engineering + Control*, pp. 440-447 (1999).

12. Nagel, K. and Schreckenberg, M. "A cellular automaton model for freeway traffic", *J. Physique*, **I**(2), p. 2221 (1992).

13. Esser, J., Neubert, L., Wahle, J. and Schreckenberg, M. "Microscopic online simulations of urban traffic using cellular automata", *Transportation and Traffic Theory*, pp. 535-554 (1999).

14. Barlovic, R., Esser, J. and Froese, K. "Online traffic simulation with cellular automata", http://www.traffic.uni-duisburg.de (2005).

15. Antonini, G., Bierlaire, M. and Weber, M. "Discrete choice models of pedestrian walking behavior", *Transportation Research Part B, Methodological*, **43**(1), pp 36-56 (2006).

16. Sadat Hoseini, S.M. and Vaziri, M. "Using image processing for microsimulation of freeways in developing countries", *Transportation Research Board*, TRB (CD-ROM) (2006).

17. Sadat Hoseini, S.M. and Vaziri, M. "Modeling drivers' behavior as an optimization process", *Transportation Research Board*, TRB (CD-ROM) (2007).

18. Hoogendoorn, S., Ossen, S. and Gorte, B. "Inter-driver differences in car following: A vehicle trajectory based study", *Transportation Research Board*, TRB (2006).

19. Ming-Chin Tsai and Chien-Chih Su "Scenario analysis of freight vehicle accident risks in Taiwan", *Accident Analysis & Prevention* (2004).

20. AASHTO *A Policy on Geometric Design of Highways and Streets*, p. 40 (1994).

21. Brockfeld, E., Kuhne, R. and Wagner, P. "Calibration and validation of microscopic models of traffic flow", *Transportation Research Board*, TRB (2005).

22. Chung, C., *Simulation Modeling Hand Book*, CRC Press, pp. 8.11-8.14 (2003).