Stochastic Modeling of Traffic Flow Breakdown Phenomenon: Application to Predicting Travel Time Reliability

Jing Dong and Hani S. Mahmassani

Abstract—This paper presents a modeling approach to generate random flow breakdowns on congested freeways and capture the subsequent wave propagation among heterogeneous drivers. The approach is intended for predicting travel time variability caused by such stochastic phenomena. It is assumed that breakdown may occur at different flow levels with some probability, and would sustain for a random duration. This is modeled at the microscopic level by considering speed changes that are initiated by a leading vehicle and propagated by the following vehicles with correlated-distributed behavioral parameters. Numerical results from a Monte Carlo simulation demonstrate that the proposed stochastic modeling approach produces realistic macroscopic traffic flow behavior and can be used to generate travel time distributions.

Keywords—travel time reliability, flow breakdown probability, duration model, heterogeneous drivers, car-following model, Monte Carlo simulation.

I. INTRODUCTION

Travel time reliability has been recognized as an important factor in traveler decisions [1][2], especially under congested traffic conditions where travel time is expected to be less predictable. Many factors account for the uncertainty associated with travel time, including accidents, work zones, adverse weather, special events, traffic control devices, fluctuations in demand, and inadequate base capacity [3]. Though significant progress has recently been made in understanding and measurement of travel time reliability [2][4]-[7], generating realistic reliability measures as output in traffic simulation models and planning tools remains an open question. To evaluate transportation network reliability performance measures and assess both short-term and long-run impacts of management interventions and policies aimed at improving travel time and service reliability, traffic models need to capture the causes of unreliability and the underlying physics of the associated processes and phenomena.

Among many sources that contribute to travel time variability, flow breakdown is one of the causes that reside primarily in traffic physics [8], and can be viewed as a collective phenomenon resulting from many individual decisions. Previous studies have shown that traffic breakdown may occur with some probability at various flow levels [9]-[11]. A systematic relation between the likelihood of breakdown and prevailing traffic conditions has been established, yielding the probability of flow breakdown as a function of the flow rate [10], [11]. In addition, the duration of the breakdown event can be viewed as a random variable and characterized by a hazard model [12]. Building on these research findings, this study proposes a methodology to produce random flow breakdowns, capture the subsequent wave propagation among heterogeneous drivers, and thus generate travel time distributions.

The rest of the paper is organized as follows. In the next section, the probabilistic characteristics associated with flow breakdown are studied based on field observations. After that, a stochastic modeling approach that produces flow breakdown endogenously in traffic simulation models is introduced in section III, followed by numerical results from a Monte Carlo simulation in section IV. Finally, conclusions and remarks on possible directions for future investigation are stated in section V.

II. PROBABILISTIC TRAFFIC FLOW BREAKDOWN

This section introduces the probabilistic flow breakdown phenomenon, based on the analysis of real world measurements.

To examine breakdown traffic characteristics, 5-minute averages of flow, speed and density data measured by the detector upstream from an on-ramp and averaged over all lanes on the I-405N Freeway (Irvine, California USA) at the Jeffrey section were collected (source: Caltrans Performance Measurement System, website http://pems.dot.ca.gov). Note that the choice of the aggregation time interval could affect the analysis results. In the literature, 5-minute [11]-[13] and 10-minute [2] aggregation intervals have been used in breakdown analysis of U.S., German and Dutch freeways.

Flow breakdown is detected when a substantial speed drop from the free mean speed occurs between two consecutive time intervals. The speed drop between two
consecutive time intervals (i.e. minimum speed difference), and the time duration over which the low speed is sustained (i.e. minimum breakdown duration), when used together, have been suggested as an appropriate set of criteria to detect breakdown occurrence [11]-[13]. The speed limit is 65 miles per hour (mph) for the I-405N Freeway. When the speed drops below 55 mph and low speed is sustained for at least three time intervals (i.e. 15 minutes), a traffic breakdown is identified. The minimum breakdown duration is used to eliminate “false alarms” due to flow fluctuation. The following quantities are defined in this context.

Pre-breakdown flow rate: The flow rate, expressed as a per-lane equivalent hourly rate, observed immediately before the onset of traffic breakdown.

Breakdown flow rate, density and speed: The average flow rate, density and speed observed after the onset of breakdown and before traffic recovery.

Breakdown duration: The time period between the onset of breakdown and the recovery of traffic flow.

By analyzing traffic data observed on work days over a period of seven months, we identified 227 occurrences of breakdown, most of which happened during the morning or evening rush hours. The occurrence of these flow breakdowns could be due to a local perturbation in traffic flow at the considered freeway section (i.e. spontaneous breakdown) or caused by an external disturbance associated with a queue spillback from a downstream location (i.e. induced breakdown) [2]. Fig. 1(a) plots the flow rate and density observed before and after breakdown. The scattering of observations indicates that flow breakdown occurs at various flow rates and results in a wide range of congested traffic states, associated with different speed, flow and density during breakdown. As the average breakdown speed and its duration largely determine the delay caused by a certain breakdown event, the distributions of these two variables are further examined and plotted in Fig. 1(b) and Fig. 1(c), respectively. The breakdown speeds are used to calibrate a left-truncated normal distribution function, from which a random breakdown speed can be drawn in the process of quantifying travel time reliability (see Equation (9)). The breakdown duration data, in conjunction with the breakdown speeds, are used to calibrate the hazard function (see Equation (2)) describing the probability that breakdown will end at a certain time, discussed in the next section.

This section presents a modeling approach to generate a random breakdown and its duration endogenously, capture
the wave propagation after the onset of breakdown, and consequently produce traffic flow patterns consistent with the observations described in the previous section. As illustrated in Fig. 2, the likelihood of breakdown is modeled as a function of the prevailing flow rate [11], the breakdown speed is drawn from a probability distribution (see Fig. 1(b)) and the breakdown duration is characterized by a hazard model [12]. At the start and the end of a breakdown event the leading vehicle changes the travel speed, which propagates among heterogeneous drivers according to a simplified car-following model with distributed behavioral parameters [14]. The sudden reduction in speed might be caused by different types of traffic flow behavior, including “bunching” of vehicles, intended interruptions and vehicle merging maneuvers [3], and has been observed by tracking individual vehicle trajectories [15]. Finally, the collective effects of individual drivers in terms of flow rates and densities are projected on the fundamental diagram.

A. Breakdown Probability and Duration

Higher flow rates are usually associated with greater likelihood of traffic breakdown. One way of implementing probabilistic breakdown in traffic models is to introduce random variables for breakdown occurrence (with flow-dependent probability) and breakdown duration. For each simulation run, namely one realization of the underlying stochastic process, a Monte-Carlo draw determines whether flow breakdown would occur at the prevailing flow level. If a breakdown occurs, a realization from a probabilistic duration model then determines how long the heavily-congested traffic state will sustain.

A systematic relation between the likelihood of breakdown and the pre-breakdown flow rate, namely the flow rate observed immediately before the onset of traffic breakdown, has been established [10][11]. The pre-breakdown flow distribution function expresses the probability that traffic breaks down in the next time interval (for a given time discretization). The analysis of data samples from freeway sections in California, USA [11] and Germany [13] indicated that the Weibull distribution provided a good description of the pre-breakdown flow rate.

\[ F(q_{bd}) = 1 - e^{-\frac{q_{bd}}{\sigma}} \]

where

\[ F() \text{ probability distribution function} \]
\[ q_{bd} \text{ pre-breakdown flow rate} \]
\[ \sigma \text{ scale parameter} \]
\[ s \text{ shape parameter} \]

Fig. 3 plots the cumulative distribution function (CDF) and probability density function (PDF) of a Weibull distribution calibrated using the data collected on I-405N Freeway at the Jeffrey section. The observed flow rate does not reach beyond 2300 vphp at this particular location.

![Weibull distribution](image)

In addition, breakdown duration is defined as the time period between the occurrence of breakdown and the recovery. The recovery of traffic flow could be a result of congestion elimination or owing to demand drop. In the latter case, the flow rate does not recover to the same level as the pre-breakdown flow rate. Thus, the recovery of the breakdown is identified when the speed is recovered to the free mean speed. Duration analysis is largely used to study the elapsed time until the occurrence of an event, or the duration of an episode. In particular, hazard-based models are applied to estimate the conditional probability that an event will occur in a time interval \( t \) given that the event has not occurred up to time \( t \). In the context of the flow breakdown problem, the hazard function represents the probability that breakdown will end at a duration \( t \) given that breakdown has continued up to a duration length \( t \). One of the significant factors that affect the duration of breakdown is the reduced speed during breakdown. Lower speed indicates heavier congestion, which usually requires a longer time to recover. A Cox proportional hazard function is used to represent the probability that breakdown will end at \( t \).

\[ h(t) = h_0(t) \cdot e^{\beta q_{bd}} \]

where

\[ h(t) \text{ hazard function at time } t \]
\( h_d(t) \) baseline hazard function
\( v_{bd} \) average speed during breakdown
\( \beta \) coefficient

B. Newell’s Car-Following Model with Heterogeneous Drivers

At the onset of a breakdown event, a sudden drop in speed is experienced. Similarly, when breakdown ends, traffic recovers to uncongested speed levels, typically the free mean speed. A stochastic version of Newell’s car-following model is adopted to characterize the wave propagation that results from this change in speed. The simple car-following model proposed by Newell [14] was shown to fit field data at a macroscopic scale [16] and relate to the Lighthill-Whitham-Richards (LWR) model [17]. It requires two parameters per vehicle, that is, a space displacement \( d_k \) and a time displacement \( t_k \). As shown in Fig. 4, the trajectory of vehicle \( k \) follows vehicle \((k-1)\) by a shift in the space-time domain.

\[
x_k(t + t_k) = x_{k-1}(t) - d_k
\]

where

- \( x_k(t) \): location of vehicle \( k \) at time \( t \)
- \( t_k \): time displacement
- \( d_k \): space displacement

The stochastic version of Newell’s car-following (NCF) model, namely NCF with distributed wave speed among vehicles, is adopted to account for heterogeneous driving behavior. The NCF parameters of a vehicle are constant in time. However, these parameters may vary from one vehicle to another [18]-[20]. Ahn et al. [16] verified, using data obtained by video-taping traffic, that the variation of drivers’ behavioral parameters \((t_k, d_k)\) were well described by a bivariate normal joint distribution. Different from Ahn et al.’s assumption, a left-truncated bivariate normal distribution is adopted in this paper, to assure that both \( t_k \) and \( d_k \) are positive.

\[
\varepsilon = \begin{pmatrix} \frac{\tau}{d} \\ -\frac{x}{d} \end{pmatrix} \sim N_L(b, \Omega), b = \begin{pmatrix} \bar{\tau} \\ \bar{x} \end{pmatrix}, \Omega = \begin{pmatrix} \sigma_x^2 & \sigma_{x\bar{\tau}} \\ \sigma_{\bar{x}\tau} & \sigma_{\bar{x}\bar{\tau}} \end{pmatrix}
\]

The Choleski decomposition approach and an accept-reject procedure [21] are used to draw car-following parameters of each driver from the truncated bivariate distribution. As a result, the wave speed, namely \( w_k = d_k / t_k \), varies from driver to driver. As shown in Fig. 5, the wave propagates at a microscopic scale like a random walk in the space and time dimensions.

C. Macroscopic Patterns

The macroscopic properties of the traffic stream can be derived as a collective effect of individual car-following behavior.

Under deterministic flow breakdown and homogenous driver assumptions, that is, breakdown occurs at the nominal maximal flow and propagates in the traffic at a fixed wave speed, the flow-density relation can be described by a triangular fundamental diagram in the following form.

\[
q = \begin{cases} k \cdot v_0, & 0 \leq k \leq k_{cr} \\ \frac{1}{\tau} - w \cdot k = w \cdot (k_{jam} - k), & k > k_{cr} \end{cases}
\]

where

- \( v_0 \): free mean speed
- \( k_{cr} \): critical density when the maximum flow rate is reached

Instead, the stochastic model assumes that (1) flow breakdown may occur at different flow levels with some probability, and (2) within a region in the space-time plane all vehicles are traveling at the same speed, but possibly with randomly distributed spacings and headways. The flow-density relation in the uncongested region, namely Equation (5), is modified so that the free mean speed is maintained until the pre-breakdown flow level is reached, which could be lower or higher than the traditionally accepted
(engineering) “capacity”.

\[ q = k \cdot v_b, 0 \leq q \leq q_{bd} \] (7)

The flow-density relation in the unstable congested region (i.e. during flow breakdown) cannot be represented as a deterministic function; rather, a Monte Carlo sampling method is used. For each replication, the flow rate before breakdown (i.e. \( q_{bd} \)) is determined by the pre-breakdown flow distribution function, defined in Equation (1); the average speed during breakdown (i.e. \( v_{bd} \)) is sampled from a left-truncated normal distribution; and the mean wave speed is calculated based on the behavioral parameters, namely \( (t_s, d_l) \), which are sampled from the bivariate distribution, defined in Equation (4).

\[
\bar{w} = \frac{\bar{d}}{\tau}, \text{ where } \bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_k \text{ and } \tau = \frac{1}{n} \sum_{i=1}^{n} t_k \] (8)

The breakdown flow rate and density can therefore be obtained by solving a system of linear equations, as follows.

\[
\begin{align*}
q &= k \cdot v_{bd} \\
q &= q_{bd} \cdot (1 + \frac{w}{v_0}) - \bar{w} \cdot k
\end{align*}
\] (9)

As illustrated in Fig. 6, in the two dimensional space of flow and density, the solution is the intercept of a line starting from the origin with a slope of \( v_{bd} \) (mathematically represented by the equation \( q = k \cdot v_{bd} \)) and a line starting from \( q_{bd} \) with a slope of \( -\bar{w} \), which corresponds to the second equation in the system of Equations (9).

IV. NUMERICAL EXPERIMENT

This section presents the numerical implementation of the modeling approach proposed in the previous section. A Monte Carlo simulation is performed to produce morning peak traffic conditions for the freeway section of I-405N over a one year period (i.e. 365 replications), where detector data were collected. Scatters of breakdown flow and density, as well as the resulting travel time distribution, can be generated based on the simulated samples.

Fig. 7 shows the simulated flow-density scatter plots. As pre-breakdown flow rate and density follow the deterministic relation defined in Equation (7), these simulated points fall on the same line with a slope equal to the free mean speed. The plotted points that correspond to 256 replications of breakdown are obtained by first sampling the pre-breakdown flow rate \( (q_{bd}) \), the breakdown speed \( (v_{bd}) \) and behavioral parameters \( (t_s, d_l) \), and then solving the system of Equations (9). Comparing the aggregated result with the real world observations shown in Fig. 1(a), we can see that the proposed approach can produce a realistic flow-density relation during breakdown. Note that the separation of the congested traffic into two clouds shown in the observed data cannot be replicated by the simulation, as a single congested regime is assumed in the proposed approach. This might be improved by adopting a multi-regime traffic flow model such as Kerner’s three-phase theory [26] in future study.

Fig. 7 plots the speed contour of an example breakdown event, which occurs at 7:18 AM and lasts 15 minutes. The shock wave propagates upstream along a 3-mile long freeway section at an average wave speed of 15 mph.
A novel feature of this approach is the combination of a stochastic macroscopic model of flow breakdown with a microscopic model of driver behavior. This reflects the recognition that certain phenomena are best modeled as probabilistic collective effects even in the context of microscopic simulation tools. Such integrated modeling approach also facilitates the use of data collected from multiple sources, at individual and aggregated levels, for the design, calibration and evaluation of ITS applications [27].

Future work includes further calibration and validation of the proposed methodology based on real-world measurements in a variety of network contexts. Other methods for modeling the stochasticity of flow breakdown, in lieu of the probability distribution function of pre-breakdown flow rate, including non-parametric methods, could be implemented and compared [8][28]. Alternative car-following models, especially ones that explicitly produce flow breakdown could also be considered [29][30]. In addition, lane-changing and over-taking behavior might also contribute to flow breakdown and wave propagation, and thus could also be incorporated in estimating travel time distributions. Indication as to which model or modeling approach might perform better under certain circumstances could be obtained based on validation against field observations, other driving behavioral models and through the application to an actual network.

ACKNOWLEDGMENT

This paper is based in part on work supported under National Science Foundation Civil Infrastructure Systems Grant #0928577 to the second author. The authors remain solely responsible for the content of this paper.

REFERENCES

Emerging Technologies versus Dynamic Control,
A. H. Ghods
Predictive Strategies: A lower order following theory,
K. E. Train, Transportation Research Board


Jing Dong received the B.S. degree in Automation and the M.S. degree in Systems Engineering from Tsinghua University in 2001 and 2003, respectively, and the Ph.D. degree in Civil and Environmental Engineering from Northwestern University in 2008. She works as a research staff at the Oak Ridge National Laboratory. Her research interests include network modeling, transportation energy analysis, and intelligent transportation systems. Dr. Dong is a member of the Transportation Research Board (TRB) Committee on Traffic Flow Theory and Characteristics and the TRB Committee on User Information Systems.

Hani S. Mahmassani holds the William A. Patterson Distinguished Chair in Transportation at Northwestern University, and serves as Director of the Transportation Center. Professor Mahmassani specializes in multimodal transportation systems analysis, planning and operations, dynamic network modeling and optimization, dynamics of user behavior, and real-time operation of logistics and distribution systems. He received the 2010 IEEE Outstanding Intelligent Transportation Systems Application Award, recognizing pioneering contributions to development of dynamic network modeling software (DYNASMART).