FORCED MERGING IN TRAFFIC

by

W. S. Jewell

OPERATIONS RESEARCH CENTER

INSTITUTE OF ENGINEERING RESEARCH

UNIVERSITY OF CALIFORNIA - BERKELEY
FORCED MERGING IN TRAFFIC

by

W. S. Jewell
Operations Research Center
University of California, Berkeley

15 January 1964

This research has been partially supported by the Office of Naval Research under Contract Nonr-222(83) with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.
ABSTRACT

A vehicle waiting at an intersection of a major road forces an entry into the main-stream traffic by requiring the oncoming traffic to slow down. Assuming that the main-stream traffic can be described as a renewal process, this paper examines the resulting disturbance which the forced entry creates in the main stream. After showing that it is formally equivalent to a busy period problem, explicit results are obtained in the case of Poisson traffic. It is shown that there is a minimal main-stream headway which should be forced in order to maximize the rate of entry into the major road by many waiting vehicles. Finally, two measures of accident potential are discussed.
FORCED MERGING IN TRAFFIC

The situation in which a vehicle on a secondary road at an uncontrolled intersection must wait for a large-enough gap in the major road traffic stream before entering has been extensively analyzed in the literature (See, for example, [6], [7], [11]). The purpose of this paper is to examine the effects of a forced merge or entry into the main stream. Attention will be focused on the resulting "compression" of the main stream as the entry disturbance propagates, rather than on the transient mechanism of the merge. A rule for deciding how small a headway should be forced is given, based on maximizing the efflux rate from the side road. Finally, some implications about necessary driver behavior will be given, and two simple measures of accident potential are discussed.

1. The Model

Consider a single lane road with vehicles traveling at a constant velocity, such that their successive headways (time spacings) are the intervals of a renewal process; i.e., the headway between the \( i^{th} \) and \((i + 1)^{st} \) vehicle, \( \tau_i \), is an independent sample from the d. f. \( A(t) \ (t \geq 0) \ (i = \ldots, -2, -1, 0, \, 1, 2, \ldots) \).

Suppose that at time zero, the \( 0^{th} \) vehicle passes a secondary road where there is a waiting vehicle. Just after passage, the secondary vehicle immediately begins to force a merger into the main stream, accelerating until he has reached the common velocity, and
is following the 0th vehicle at headway $\sigma_0$. This will of course force the 1st vehicle to slow down, and after some transient period, we assume that it will again be traveling at the common velocity, choosing to follow the merged vehicle at some headway $\sigma_1$. Clearly this effect may propagate upstream for many vehicles, as the 2nd, 3rd, ... vehicles are forced to slow down, choosing to follow the 1st, 2nd, ... vehicles at some minimal spacing $\sigma_2, \sigma_3, ...$

The assumptions of the model are shown in Figure 1, where the merged vehicle (dashed line) forces a "compression" of the first four vehicles. The trajectory of the 5th vehicle is unchanged, although its headway following the 4th vehicle has diminished, since it is still larger than some minimal spacing, $\sigma_5$, at which it would choose to follow. We shall not attempt to model the actual forcing mechanism, nor the transient period during which each of the drivers slows his vehicle and then readjusts his velocity and headway; some preliminary results on the first problem have been obtained by Bisbee and Conan[2]. Instead, we shall concentrate on the nature of the interaction between the arriving vehicles and those which have slowed down, and examine the behavior of this interaction as a function of the $\sigma_i$ ($i = 0, 1, ...$).

The $\sigma_i$ ($i = 1, 2, ...$) may be thought of as "jam" headways, or minimal time spacings which the drivers would choose in such a maneuver. We shall make the assumption that these compressed headways are independent samples from the same d. f., $B(t)$ ($t \geq 0$). The spacing generated by the merged vehicle, $\sigma_0$, could possibly be obtained from the geometry of the intersection, and the acceleration characteristics of the vehicle; we shall assume that it is a random variable.
with d.f. C(t) \((t \geq 0)\).

From the assumptions, vehicle 1 is delayed if \(\tau_1 \leq \sigma_0 + \sigma_1\), and vehicle \(n (n = 2, 3, \ldots)\) is delayed if the 1\(^{\text{st}}\) through the \((n - 1)\)^{\text{st}} vehicles are delayed, and \(\Sigma_{i=1}^{i=n} \tau_i \leq \Sigma_{i=0}^{i=n} \sigma_i\). If a total of \(N_D\) vehicles are delayed by the merging disturbance, then the \((N_d + 1)\)^{\text{st}} vehicle must have \(\Sigma_{i=1}^{i=N_d+1} \tau_i > \Sigma_{i=0}^{i=N_d+1} \sigma_i\).

In the analysis to follow, we shall be interested in the number of vehicles delayed by the forced merge, \(N_d\), and the duration of the merging disturbance, \(T_D\), which we define as \(T_D = \Sigma_{i=0}^{i=N_d+1} \sigma_i\). The reason for the latter definition will become apparent in the next section.

2. Busy Period Analogy

Upon examination, the problem just posed can also be thought of as a queueing problem. The headways, \(\tau_i\), are just the interarrival spacings of customers approaching a service facility: \(s_0 = \sigma_0 + \sigma_1\) is the service time of the 0\(^{\text{th}}\) (or the merged) arrival, and \(s_1 = \sigma_2\), \(s_2 = \sigma_3\), \ldots, \(s_n = \sigma_{n+1}\), \ldots are the service times of the 1\(^{\text{st}}\), 2\(^{\text{nd}}\), \ldots, \(n\)^{\text{th}} \ldots customers. The first customer must wait in queue if \(\tau_1 \leq s_0\), and \(n\)\(^{\text{th}}\) customer \((n = 2, 3, \ldots)\) must wait in queue if the 1\(^{\text{st}}\) through the \((n - 1)\)^{\text{st}} customer waited in queue, and \(\Sigma_{i=1}^{i=n} \tau_i \leq \Sigma_{i=0}^{i=n-1} \sigma_i\).

We see that the duration of the merging disturbance, \(T_D\), as defined, is identical with the length of a busy period generated in the queueing model; \(N_d\) is one less than the total number of customers served in a busy period. Thus, the problem reduces to the analysis of the busy period of a queue with interarrival d.f. \(A(t)\); a special service-time d.f. for the customer who arrives when the service facility
is empty, \( D(t) = C(t)B(t) \); and a regular service-time d.f. for the other customers (if any) in a busy period, \( B(t) \).

The analysis of the busy period when \( D(t) = B(t) \) has been carried out by many authors, including Borel, Kendall, Takács, Beneš, and Pollaczek (For discussion, see for example Cox [3]). The most general case of arbitrary \( A(t) \) and \( B(t) \) was theoretically solved by Politiczek [8], but the contour integration formulae he gives are extremely difficult to compute; the simplest formulae seem to result when either \( A(t) \) or \( B(t) \) is the negative exponential d.f. (See Takács [10]).

Accordingly, we shall examine only the case where \( A(t) \) is a negative exponential (Poisson mainstream traffic) in order not to obscure the main presentation. In this case, the analysis of the busy period has been made when the initial service of the busy period is from a different d.f. \( D(t) \) by Finch [4] and Yeo [12], using a method of Takács [9]. Because these papers are not easily accessible, we shall sketch in their results, as well as developing some additional formulae needed when selecting a headway to be forced.

In the case where \( B(t) \) is a constant (Poisson traffic), the merging problem is also analogous to a problem of "overflows" at a signalized intersection [5]. Formulae for this case were first developed by Borel [1].

3. Poisson Traffic

The assumption of Poisson mainstream traffic, \( (A(t) = 1 - \exp(-\lambda t), t \geq 0) \), allows us to treat the input in any interval of time as a homogeneous process.
First, assume that all of the customers have the same service-time d.f. \( B(t) \), and define \( G(t) = \Pr \{ T_D \leq t \mid D(t) = B(t) \} \). Suppose exactly \( j \) additional customers arrive during the \( 0^{th} \) service interval; i.e., \( \sum_{i=1}^{j} \tau_i \leq s_0 \), and \( \sum_{i=1}^{j+1} \tau_i > s_0 \). If the queue discipline is rearranged to be LIFO, instead of FIFO, the last of the new arrivals will generate his own "descendants" during his service time, who must be served before the other "first generation" arrivals; this will alter the individual waiting times, but can not affect the distribution of the total additional busy period, which must be the \( j \)-fold convolution of \( G(t) \), denoted by \( G^j(t) \).

But, if the \( 0^{th} \) service time were of length \( y \), then the probability of \( j \) additional first generation arrivals would be the Poisson probability, \((\lambda y)^j \exp(-\lambda y)/j!\). Since the total busy period is the sum of \( y \) and the total additional period described above, we must have

\[
G(t) = \sum_{j=0}^{\infty} \int_0^t e^{-\lambda y} \frac{(\lambda y)^j}{j!} G^j(t-y) \, dB(y) \quad (t \geq 0)
\]

A similar argument can then be made for the case where the \( 0^{th} \) service-time d.f. is \( D(t) \), instead of \( B(t) \). Letting \( H(t) = \Pr \{ T_D \leq t \} \) in this case, we obtain:

\[
H(t) = \sum_{j=0}^{\infty} \int_0^t e^{-\lambda y} \frac{(\lambda y)^j}{j!} G^j(t-y) \, dD(y) \quad (t \geq 0)
\]

The above formulae can be put into simpler form if we use (LaPlace-Stieltjes) transforms with the notation:
\[ \tilde{g}(s) = \int_{0}^{\infty} e^{-st} dG(t) \]

and similarly for the other distribution functions. From (1) and (2) we get the implicit relations

\begin{equation}
(3) \quad \tilde{g}(s) = \tilde{b}(s + \lambda - \lambda \tilde{g}(s))
\end{equation}

and

\begin{equation}
(4) \quad \tilde{\mu}(s) = \tilde{d}(s + \lambda - \lambda \tilde{g}(s))
\end{equation}

which are mostly useful for obtaining moments, although they can be inverted in special cases. In the forced merge example, of course, we will set \( \tilde{d}(s) = \tilde{b}(s) \tilde{c}(s) \).

Denote the first moment of a d.f. \( B(t) \) by \( \nu_B \) and its variance by \( \sigma_B^2 \), and similarly for the other distributions \( \nu_A = \lambda^{-1} \), \( \sigma_A^2 = \lambda^{-2} \). Then by differentiating (3) and (4), we find after some algebra:

\begin{equation}
(5) \quad \nu_G = \frac{\nu_B}{1 - \lambda \nu_B} ; \quad \sigma_G^2 = \frac{\sigma_B^2 + (\lambda \nu_B) \nu_B^2}{(1 - \lambda \nu_B)^3}
\end{equation}

\begin{equation}
(6) \quad \nu_H = \frac{\nu_D}{1 - \lambda \nu_B} ; \quad \sigma_H^2 = \frac{(1 - \lambda \nu_B) \sigma_D^2 + (\lambda \nu_B) \sigma_B^2 + \nu_B^2}{(1 - \lambda \nu_B)^3}
\end{equation}

Of course, in the traffic example:

\begin{equation}
(7) \quad \nu_D = \nu_B + \nu_C ; \quad \sigma_D^2 = \sigma_C^2 + \sigma_B^2
\end{equation}

Thus, the average duration of the disturbance period depends only
on $\lambda$, $\nu_B$, and $\nu_C$.

Similar arguments can be used to find the distribution of the additional number of vehicles delayed, $N_D$. Let $G_n = \Pr(N_D = n \mid D(t) = B(t))$, and $H_n = \Pr(N_D = n)$ in general. Then:

\[ G_n = \sum_{j=0}^{n} \int_0^\infty \frac{e^{-\lambda y} (\lambda y)^j}{j!} G_j^{\ast n-j} \mathrm{d}B(y) \quad (n = 0, 1, \ldots) \]

where $G_j^{\ast}$ is the $j$-fold convolution of $G_k$. Also:

\[ H_n = \sum_{j=0}^{n} \int_0^\infty \frac{e^{-\lambda y} (\lambda y)^j}{j!} G_j^{\ast n-j} \mathrm{d}D(y) \quad (n = 0, 1, \ldots) \]

By the use of generating functions, defined as:

\[ \hat{G}(z) = \sum_{n=0}^{\infty} z^n G_n \]

we find the implicit relations

\begin{align*}
(10) \quad & \hat{G}(z) = b(\lambda - \lambda z\hat{G}(z)) \\
(11) \quad & \hat{H}(z) = d(\lambda - \lambda z\hat{G}(z))
\end{align*}

Denote the mean and variance of $G_n$ by $m_G$ and $\nu_G^2$, respectively. Then:

\[ m_G = \frac{\lambda \nu_B}{1 - \lambda \nu_B}; \quad \nu_G^2 = \frac{\lambda^2 \nu_B^2 + \lambda \nu_B}{(1 - \lambda \nu_B)^3} \]
and of course (7) holds in the traffic problem.

It is important to note that in this model \( T \) may be large enough so that no mainstream vehicles are delayed.

4. A Condition for Stability

It is a well known result that for the solution of (3) to give an honest distribution for \( G(t) \), that as \( s \) approaches zero, the smallest root of \( x = \beta (\lambda - \lambda x) \) must be unity; one can easily show that this means that \( \lambda v_B \leq 1 \). This is not surprising, since this is just the utilization ratio of importance in queueing theory. Thus:

1. If \( \lambda v_B > 1 \), with probability \( 1 - x > 0 \), the merging disturbance period will never terminate.

2. If \( \lambda v_B = 1 \), the disturbance period will terminate with probability one, but from (12) and (13), it will have infinite mean length.

3. If \( \lambda v_B < 1 \), the disturbance period has finite mean length.

More simply stated, our model of driver behavior requires that, when a forced entry is made, the delayed cars must "compress," on the average, in order for the disturbance to eventually die out.

5. Selecting a Minimal Headway to Force

Suppose there are many vehicles on the secondary road. If the first driver forces a very small headway, this may hinder the subsequent merging of the next vehicle in line (assuming he cannot force
his way out during the disturbance interval). On the other hand, waiting until a large headway comes along will also delay cars behind him. In this section, we shall examine the question as to what choice of $T$, the minimal size headway to be forced, will maximize the rate at which merges are made from the secondary road. Successive drivers are supposed to have the same initial service-time d.f. $D(t)$.

We require that: (1) $\sigma_0 + \sigma_1 > T$, so that all merges will be forced, and (2) no entries are made during the disturbance interval, either because $T > \sigma_i (i = 1, 2, ...)$ or because of driver concern for accidents.

Let $F(t) = \Pr (T_D \leq t \mid T_1 > T)$ and note that the previous argument still applies, except there are new arrivals only during the interval $(T, y)$. Remembering assumption (1) above, it is not difficult to show that:

(14) \[ \hat{f}(s) = \hat{h}(s) \exp (\lambda T - \lambda T \hat{g}(s)) \]

and

(15) \[ \nu_F = \frac{\nu_D - \lambda \nu_B T}{1 - \lambda \nu_B} ; \quad \sigma_F^2 = \frac{(1 - \lambda \nu_B)^2 \sigma_D^2 + \lambda (\nu_D - T)[\sigma_B^2 + \nu_B^2]}{(1 - \lambda \nu_B)^3} \]

Defining $F_n = \Pr (N_D = n \mid T_1 > T)$, we obtain

(16) \[ \hat{F}(z) = \hat{H}(z) \exp (\lambda T - \lambda T \hat{g}(z)) \]

with:

(17) \[ m_F = \frac{\lambda \nu_D - \lambda T}{1 - \lambda \nu_B} ; \quad \nu_F^2 = \nu_H^2 - \lambda T \left[ \frac{\lambda^2 \sigma_B^2}{(1 - \lambda \nu_B)^3} \right] \]
Equation (15) gives the mean length of the disturbance interval when a secondary vehicle forces some headway > T. However, the next vehicle in line must wait an additional time past the end of this interval until a headway > T appears (he may wait zero time if the (N_D + 1)st main-stream vehicle arrives at an instant > T_D + T). This additional wait is just the problem of "waiting for a gap" which has been previously analyzed in great detail [6], [7], [11].

The mean wait in Poisson traffic for a gap greater than T is:

\[ \nu_w(T) = \frac{1}{\lambda} \left[ e^{\lambda T} - 1 - \lambda T \right] \]

Thus the total mean spacing L(T) between successive forced merges is:

\[ L(T) = \nu_F(T) + \nu_w(T). \]

Since the instants of merging constitute an imbedded renewal process, the mean rate of merging, \( \phi(T) \), is just \( L(T)^{-1} \). Figure 2 shows \( \lambda L(T) \) versus \( \lambda T \) for \( \lambda v_B = 0.8 \), and \( \lambda v_D = 4.0 \) (\( m_H = 20 \); \( v_H^2 = 500 \)). For small T the length of the disturbance interval keeps the merge rate low, and for large T the wait for a gap dominates.

An optimal choice of T can be found by calculus to be:

\[ \lambda T^* = - \ln(1 - \lambda v_B) \]

which gives the unique maximum \( \phi(T^*) \), provided that the assumption of \( T^* < \sigma_0 + \sigma_1 \) (and certainly < \( v_B + v_C \)) is satisfied. Note that the optimal \( T^* \) does not depend upon \( v_C \), so that the choice of minimal headways to force is independent of the acceleration characteristics.
of the merging vehicle. For the example of Figure 2, \( \lambda T^* = 1.61 \), indicating that headways at least 61% greater than the average headway should be forced.

Figure 3 shows \( T^*/v_B \) versus \( \lambda v_B \). For sparse traffic, \( T^* \) is very close to \( v_B \); as the main-stream traffic increases, the optimal minimal headway to force also increases, limited only by assumption (1) above. If \( \sigma_0 \) and the \( \sigma_i(i = 1, 2, \ldots) \) are fixed numbers, this limit is just \( T^* = v_B + v_C \), the point at which a secondary vehicle would choose to wait for a gap without attempting to force a merge.

6. Measures of Accident Potential

One of the reasons for not allowing a forced merge is the possibility of accidents caused by the "chain reaction" of vehicles which must deaccelerate suddenly in the main-stream. The actual causes of such accidents are very difficult to model until more is known about driver reactions under sudden stress. However, as a rough measure of accident potential with a forced merge, we shall consider two simple ideas which probably bound the possible damage. Assume that the probability that any successive pair of cars (including the merged vehicle) has an accident during a merging disturbance is a known constant, \( p \), a function of \( \lambda \), the velocity of the main-stream, the visibility, etc.

First let us assume that accidents are independent of one another, or, roughly speaking, that each following vehicle has an equal chance of avoiding a collision. The mean number of vehicles in a collision, \( M_C \), is twice the mean number of pairs colliding in
a disturbance period, or:

\[(21) \quad M_c = 2 m_p p\]

Another assumption might be that once the \(j^{th}\) and \((j+1)^{st}\) \((j = 0, 1, 2, \ldots)\) vehicles have an accident, then all of the cars \(j + 2, j + 3, \ldots, N_D\) will also be involved; this is the familiar "chain reaction" in poor visibility. In this case:

\[(22) \quad M_c = \sum_{j=0}^{\infty} F_j \sum_{k=2}^{j+1} k p (1 - p) \frac{j-k}{2} = \frac{1 - 2p}{p} \left\lfloor \frac{F(1 - p)}{1 - p} \right\rfloor + m_p\]

For very small \(p\),

\[(23) \quad M_c = \left[ \frac{3}{2} m_p + \frac{v_F^2 + m_F^2}{2} \right] p + O(p^2) \]

Under either assumption, the probability of at least one collision is:

\[(24) \quad P \geq 1 = 1 - \hat{F}(1 - p) = m_p p + O(p^2)\]

The reader may easily modify the distribution if it is known that a headway of exactly \(T\) units was forced (instead of only knowing it was \(> T\)).

\section{Extensions}

The formulas developed for optimal choice of a minimal headway do not, of course, take the delays in the main stream into account. This delay is just the usual waiting time in the system (queue + service) of the queueing model; by finding the average wait of those who wait (except the initial customer), one can then weight the total main-stream
delay in any desired combination with (19). This analysis has been carried out by the author, and will be reported in a subsequent paper. One new feature of interest is that it may be worthwhile to force a merge for several secondary vehicles.

Although the analysis has been carried through for Poisson traffic, it can also be done for other specific cases of interest, by simple recursive computations on the delay distributions of the $j^{th}$ vehicle. In particular, the case of shifted-exponential headway distributions, and the case of deterministic $\sigma_0$ and $\sigma_i$ ($i = 1, 2, \ldots$) recommend themselves as subjects for further study.
REFERENCES


Figure 1. Trajectories of vehicles during a forced merge.
Figure 2. Total mean spacing between forced merges as a function of the minimal size headway which is forced.
Figure 3. Optimal minimal size headway to be forced as a function of a function of main-stream flow rate.
BASIC DISTRIBUTION LIST
FOR UNCLASSIFIED TECHNICAL REPORTS

Head, Logistics and Mathematical Statistics Branch
Office of Naval Research
Washington 25, D. C.

C. O., ONR Office
Navy No. 100, Box 39, F. P. O.
New York, New York

ASTIA Document Service Center
Building No. 5
Cameron Station
Alexandra, Virginia

Institute for Defense Analyses
Communications Research Division
von Neumann Hall
Princeton, New Jersey

Technical Information Officer
Naval Research Laboratory
Washington 25, D. C.

C. O., ONR Branch Office
1030 East Green Street
Pasadena 1, California
ATTN: Dr. A. R. Laufer

Bureau of Supplies and Accounts
Code OW, Department of the Navy
Washington 25, D. C.

Professor L. W. Cohen
Mathematics Department
University of Maryland
College Park, Maryland

Professor Donald Bockman
Director, Systems Research Center
Case Institute of Technology
Cleveland, Ohio

Professor Lawrence E. Fouraker
Graduate School of Business
Harvard University
Cambridge, Massachusetts

Professor David Gale
Department of Mathematics
Brown University
Providence 12, Rhode Island

Dr. Murray Geisler
The RAND Corporation
1700 Main Street
Santa Monica, California

Professor L. Hurwicz
School of Business Administration
University of Minnesota
Minneapolis 14, Minnesota

Dr. James R. Jackson, Director
Western Management Sciences Institute
University of California
Los Angeles 24, California

Professor Samuel Karlin
Mathematics Department
Stanford University
Stanford, California

Professor C. E. Lemke
Department of Mathematics
Rensselaer Polytechnic Institute
Troy, New York

Professor W. H. Marlow
Logistics Research Project
The George Washington University
707 - 22nd Street, N. W.
Washington 7, C. D.

Professor L. Hurwicz
School of Business Administration
University of Minnesota
Minneapolis 14, Minnesota

Professor Samuel Karlin
Mathematics Department
Stanford University
Stanford, California

Professor C. E. Lemke
Department of Mathematics
Rensselaer Polytechnic Institute
Troy, New York

Professor W. H. Marlow
Logistics Research Project
The George Washington University
707 - 22nd Street, N. W.
Washington 7, C. D.
BASIC DISTRIBUTION LIST
FOR UNCLASSIFIED TECHNICAL REPORTS

Professor A. Charnes
The Technological Institute
Northwestern University
Evanston, Illinois

Professor R. Radner
Department of Economics
University of California
Berkeley 4, California

Professor Murray Rosenblatt
Department of Mathematics
Brown University
Providence 12, Rhode Island

Mr. J. R. Simpson
Bureau of Supplies and Accounts
Navy Department (Code W31)
Washington 25, D. C.

Professor A. W. Tucker
Department of Mathematics
Princeton University
Princeton, New Jersey

Professor J. Wolfowitz
Department of Mathematics
Lincoln Hall, Cornell University
Ithaca 1, New York

C. O., CNR Branch Office
346 Broadway
New York 13, New York
ATTN: J. Laderman

Professor Oskar Morgenstern
Economics Research Project
Princeton University
92 A Nassau Street
Princeton, New Jersey