The University of Michigan

IVHS

Intelligent Vehicle-Highway Systems

Nonlinear Disutility of Travel Delays: Implications in Diversion Benefits and Strategies

Kan Chen
Department of Electrical Engineering
and Computer Science

August 1990

IVHS Technical Report-90-3

UNIVERSITY OF MICHIGAN
TRANSPORTATION RESEARCH INSTITUTE • (313) 936-1066
2901 Baxter Road, Ann Arbor, MI 48109-2150, and

COLLEGE OF ENGINEERING • (313) 764-4333
4112 EECS, Ann Arbor, MI 48109-2122

Nonlinear Disutility of Travel Delays: Implications in Diversion Benefits and Strategies

Kan Chen
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan 48109-2122

August, 1990

Abstract

Most drivers tend to be oblivious about minor traffic delays but are disproportionately concerned about excessively long delays caused by traffic congestion or incidents. If this nonlinear disutility of traffic delays as subjectively assessed by the typical driver is taken into account, benefits of traffic diversion as could be implemented through Intelligent Vehicle-Highway Systems (IVHS) may be substantially above the traditional benefit assessment based on linear disutility or constant time value calculations. In addition, graphical analyses taking into account such nonlinear disutility can provide new insights that would suggest new strategies for traffic managers to divert traffic from incidents, as will be illustrated in this paper.

1 INTRODUCTION

Total traffic delays caused by highway congestion or incidents are traditionally measured in terms of total vehicle-hours. Social benefits obtained by applying advanced technology and/or advanced traffic management technique are usually assessed by multiplying the reduction of vehicle-hours of traffic delays resulting from these approaches by an average per-unit time value, such as the average wage rate in society. Although this time value in more sophisticated benefit analyses may vary according to the purpose of the vehicle trip -- i.e., commuting, delivery, shopping, leisure, etc. -- the time value even in such analyses is assumed to be fixed for any given trip purpose. In other words, the per-unit time value has been assumed to be independent of the duration of traffic delays.

Similarly, in traffic management, a common criterion for optimum traffic diversion is to minimize total vehicle-hours of traffic delays. Even though the distinction is well recognized between user optimization (in which the individual vehicle time delay is minimized) and system optimization (in which the total vehicle hours are minimized), the per unit time value is implicitly assumed to be constant in system optimization [Wardrop, 1952].

The recent literature has begun to pay attention to the fact that typical vehicle drivers (and passengers) are increasingly (i.e., disproportionately) concerned when the traffic delays experienced by them become excessive. This refers to the phenomenon that people are usually tolerant and even oblivious of minor traffic delays. However, when such delays become major (and especially if such delays are excessive or unexpected), people get very annoyed and experience tangible losses which cannot be properly represented by the constant per-unit time value. For example, there are a couple of recent papers that consider the impact of nonlinear disutility of travel delays on trip departure time [Hendrickson and Kocur, 1981; Hendrickson and Plank, 1984]. In these papers, the penalty for negative delays, which would result in unexpectedly early arrival in the destination, is also considered in optimizing the departure time from the trip origin. Another paper has considered route choice behavior from the perspective of multiattribute utility theorem [Keeney and Raiffa, 1976], especially as applied to the tripmaker's choice of bus routes [Tzeng *et al.*, 1989]. Interestingly, the nonlinear disutility function assumed in this example is concave rather than convex; that is, the per-unit time value of penalty for longer travel decreasing with travel time, which the authors believe may be typical in the case of bus riders. In any event, there has been nothing found in the literature which has analyzed the impact of nonlinear disutility of travel delays in the context of Intelligent Vehicle-Highway Systems (IVHS).

The purpose of this paper is to examine the impact of nonlinear disutility of travel delays on route guidance, both from the standpoint of assessing the social benefits of successful route guidance to avoid traffic delays, and from the standpoint of traffic management strategies for traffic diversion from confirmed incidents.

2 BASIC CHARACTERISTICS

All drivers dislike traffic delays, due to both recurrent traffic congestion and nonrecurrent incidents, which is one of the undesirable phenomena of automotive travel. Such delays are so common, especially in urban road travel, that most drivers (and passengers) factor minor traffic delays into their travel plans. Thus, for the relatively short and/or normally expected traffic delays, the typical driver would not mind very much. In fact, most of them would even be oblivious of such delays if they are reasonably short. However, when the delays become excessively or unexpectedly long, the driver would become at least annoyed, and often experience serious inconvenience (such as missing a flight), or take tangible economic losses (as in being late to work or late in delivery) or suffer from even more dire consequences (as in the case of a patient arriving too late to receive life-saving medical care).

The abhorrence of excessive traffic delays as discussed above suggests a convex shape of disutility of travel delay as shown in Figure 1. The curve indicates that, as far as the driver's subjective assessment is concerned, there is practically no penalty for short traffic delays. However, when the delay becomes longer than some threshold, the penalty increases with each successive unit of delay. If this penalty is translated into monetary terms, the per-unit value of time implied by the convex curve would not remain constant but would increase with traffic delays. Another way of saying this is that the loss function of traffic delays is nonlinear, and that the marginal loss increases with delays.

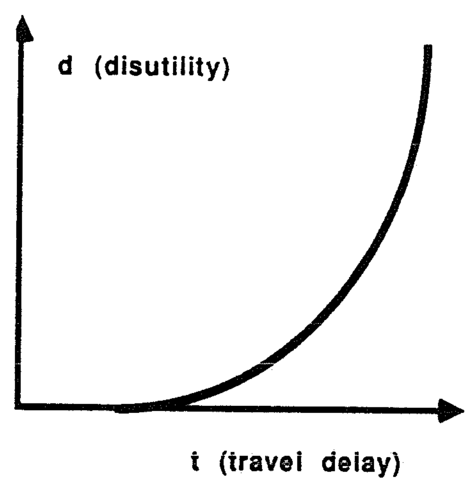

Fig. 1 A Typical Nonlinear Disutility of Travel Delays

Disutility instead of monetary penalty or loss function is used here in the general case for two reasons. Firstly, the subjectively felt penalty by drivers experiencing major traffic delays varies from one driver to another and from one situation to another. Thus, the penalty cannot be easily translated into monetary terms although it may be inferred from the tripmaker's choice of transportation modes and/or willingness to pay, as will be discussed later. Secondly, the concept of utility theory (disutility being negative utility) can be used to capture the driver's risk taking attitude. The convex shape of disutility as shown in Figure 1 corresponds to a concave utility function, and represents risk aversion of a typical driver over major traffic delays. Thus, Figure 1 suggests that the typical driver would opt for a route with one unit of delay with certainty rather than another route which has a 50-50 chance of incurring either two units of delay or no delay.

3 BASIC ANALYTICAL APPROACH

Many approaches have developed over the years for traffic analysis [May, 1990]. Some take the macroscopic look at the traffic in aggregate and others take the microscopic look at individual vehicles. Due to the nonlinearity of disutility of traffic delays, the subject matter in this paper, we must not lose sight in our analysis the traffic delay incurred by each individual vehicle. At the same time, for the purposes of assessing social benefits and developing traffic management strategies, we must also be able to consider all the vehicles impacted by an incident.

One traffic analytical approach which does satisfy the above criteria is the graphical portrayal of cumulative number (n) of vehicles passing a particular point on the highway versus time (t) as shown in Figure 2. For the purpose of analyzing traffic diversion around a nonrecurrent incident, the point on the highway should be the location (x) which is just upstream from the incident. If we consider the three-dimensional space of (n,x,t), the graph in Figure 2 is obtained by taking a cut parallel to the n-t coordinate plane at x=location just upstream from the incident [Makigami *et al.*, 1971].

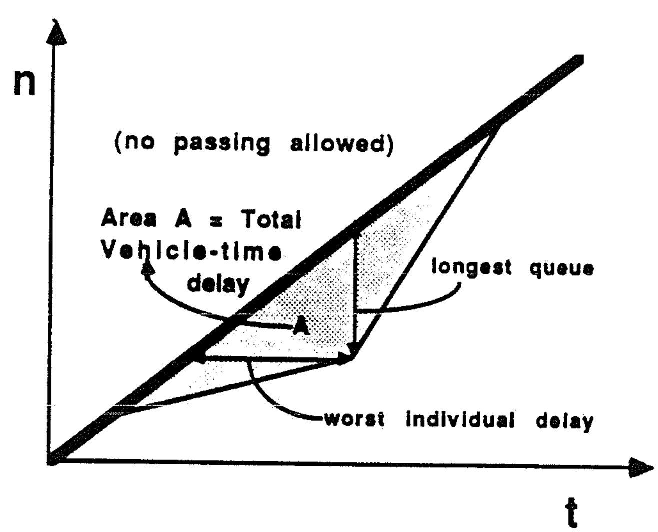

Fig. 2 Basic Analytical Graphics

The heavy slanted line in Figure 2 represents the traffic flow since the slope of that line is the number of vehicles passing point x per unit time t. The slope of the heavy line is medium as the actual flow is limited by traffic demand which is below the road capacity. Starting at the time of the incident, some of the lanes are blocked; the flow is constricted and is represented by the lower thinner line which has a minimum slope. A queue of vehicles builds up during this period. After the incident is cleared, the actual traffic flow, being fed by the queued-up traffic, goes up to the maximum level equal to the road capacity as represented by the upper thinner line which has a maximum slope. This continues until the queue is cleared and the traffic returns to the normal condition where the actual traffic is limited again by demand. An implied assumption in this model, which seems quite reasonable, is that vehicles are not allowed to pass each other and thus would remain in the same sequence throughout the incident.

One of the advantages of the graph in Figure 2 is that it provides both a microscopic view of each individual vehicle and a macroscopic view of the total traffic. For example, microscopically each horizontal line in the graph represents a particular individual passing point x. Thus, the horizontal time segment in the shaded triangle for a particular n is the delay incurred by that nth vehicle which gets stuck at x during that time segment. Macroscopically, the shaded area A gives the total vehicle-hours of traffic delays caused by the incident. Each vertical segment in the shaded triangle over time t represents the length of the queue of vehicles at time t. Thus, the worst individual vehicle delay and the longest queue can also be determined graphically as shown in figure 2.

The graphical analysis represented by Figure 2 has been used in traffic management handbooks [FHWA, 1983]. It is already well known to many traffic managers and it is fortunate that this approach, as pointed out previously, satisfies the criteria for traffic analysis which includes the consideration of nonlinear disutility of traffic delays. This basic analytical approach will be used for the remainder of this paper.

4 A SPECIAL CASE: PIECEWISE LINEAR DISUTILITY

For the sake of simplicity, let us consider the special case in which the nonlinear disutility of traffic delays may be represented by two linear segments as shown in Figure 3 below.

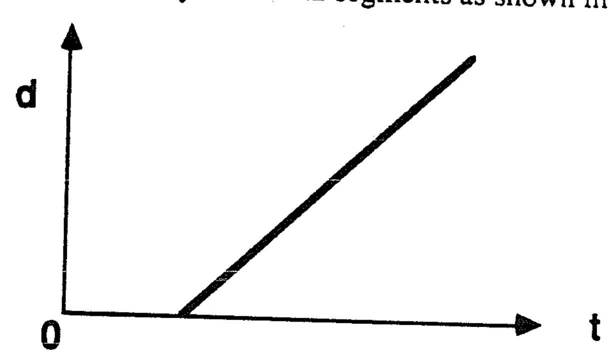

Fig. 3 A Simple Piecewise Linear Disutility

This simple hypothetical case implies that drivers would completely neglect the minor traffic delays and would attach a relatively high but constant per-unit penalty to all traffic delays above a given threshold. Furthermore, we assume that every driver's disutility function is identical and that the total disutility resulting from traffic delays due to an incident is the summation of all the individual drivers' disutilities. This latter assumption is equivalent to the common implicit assumption that every driver has the same time value (usually the average wage rate) in the traditional assessment of the monetary cost of traffic delays.

In this case, since the minor delays below the threshold incur no penalty, the total disutility resulting from the incident is represented by the shaded triangle D in Figure 4. Note that this triangle D is simply the original triangle A in Figure 2 less a strip with a constant width corresponding to the threshold of traffic delays shown in Figure 3. Graphically this is quite an easy solution. Two caveats are in order here. First, the accumulation of individual disutilities assumes the interpersonal comparison of utility (ICU) which is at least problematical [Luce and Raiffa, 1957]. However, as will discussed later, the total disutility may be construed as the aggregated value of time based on the concept of collective willingness to pay. Secondly, the fact that the triangle D is smaller than triangle A does not mean that the penalty is smaller in this evaluation. On the contrary, since the implied time value in the steep portion of the piecewise linear disutility is supposed to be higher than the average time value traditionally used in evaluating the dollar value corresponding to triangle A, the monetary penalty represented by D may be substantially higher than that represented by A.

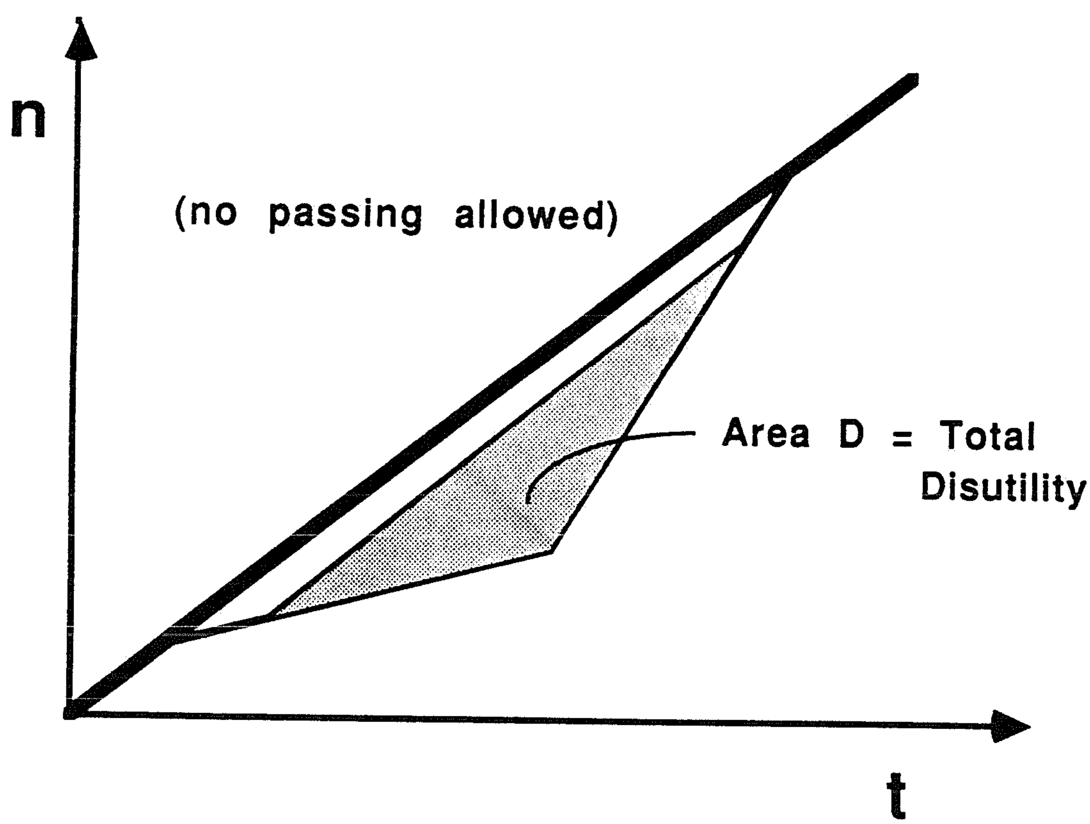

Fig. 4 Total Disutility for a Simple Piecewise Linear Case

5 THE GENERAL CASE OF NONLINEAR DISUTILITY

The graphical assessment of total disutility in the general case is more cumbersome than that in the special case discussed in the last section but is still rather straightforward. Since each horizontal segment in the shaded triangle in Figure 2 represents the traffic delay incurred by an individual vehicle, that segment may be used to compute the disutility for the driver in that vehicle. If the disutilities of all the drivers are thus portrayed graphically, the total disutility would simply be the area under that graph. This procedure is illustrated in Figure 5 for the case in which the disutility of each vehicle delay is assumed to be a quadratic function:

$$d = c\,t^2$$

where d = disutility of a driver

c = a constant

t = traffic delay incurred by an individual vehicle

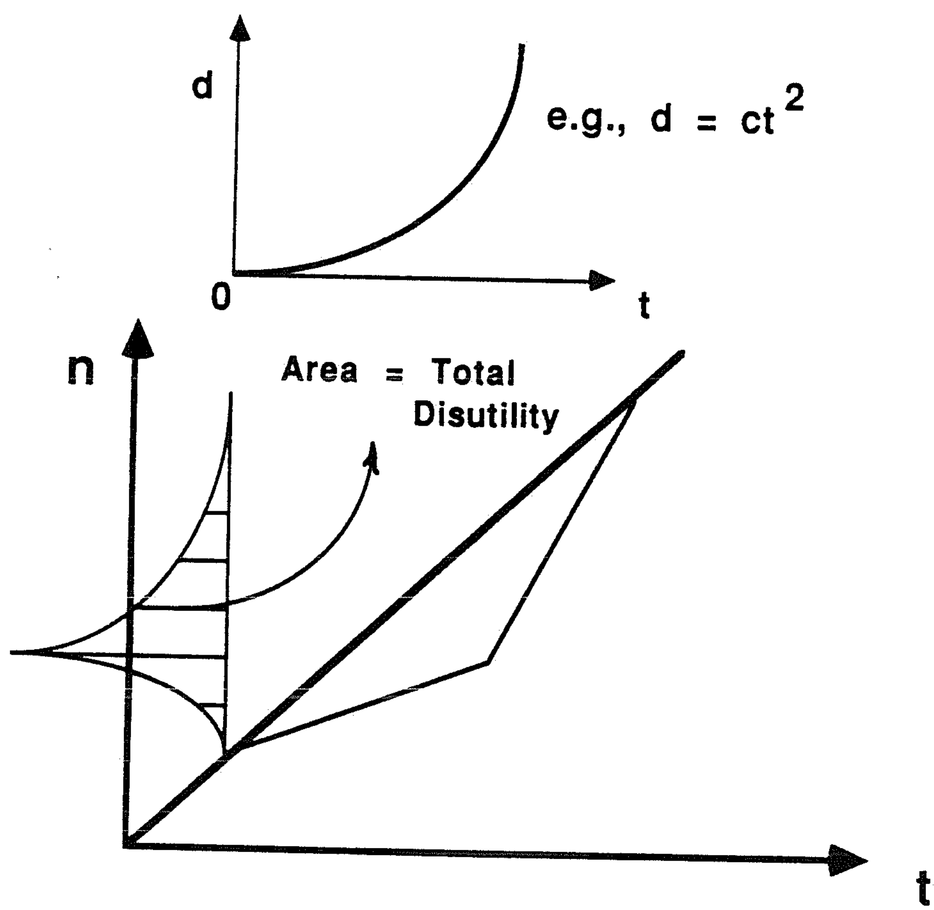

Fig. 5 Total Disutility for a Quadratic Case

Note that, in Figure 5, disutility begins to accumulate as soon as the incident occurs because there is no threshold in the quadratic case as in the piecewise linear case. The shape of the shaded area representing the total disutility has revealing values as it calls the attention of the traffic manager to the widest elements in that shaded area, as these elements are the worst contributors to total disutility. As will be illustrated later, this graphical insight could help traffic managers develop new time-varying strategies for diverting traffic from an incident.

6 BENEFIT ASSESSMENT

Social benefits of advanced technology such as IVHS are not merely aggregation of individual benefits resulting from the direct usage of the technology, but should include the external effects (both positive and negative benefits) which represent the impact of the users on nonusers. For example, the social benefit of route guidance through IVHS includes not only the time savings of all the IVHS users who have found faster routes to their destinations but also the time savings of the nonusers who travel faster on the original routes which get less congested as the users take alternate routes. On the debit side is the increase of travel time of the drivers who have been traveling on the alternate routes which now get more congested.

The net reduction of total travel time or travel delays have been traditionally converted to monetary social benefits by multiplying the total time reduction by the average national wage rate, in the order of $8 per hour. It has been estimated that the potential savings from IVHS on a national scale by the year 2010 may amount to $41 billion per year [Mobility 2000, 1990]. Even this apparently enormous figure may have been underestimated if the nonlinear disutility of travel delays is taken into account. One national study has shown a large ratio (1.27) of standard deviation to mean commuting time in the United States [Stafford and Duncan, 1985]. This large variance may be due to many unexpected traffic delays, as a high person-specific variance for commuters would contribute to a high overall commuting time variance from a sample of commuting trips. If so, the average wage rate should be at least shaded upward to include overtime work premium for the benefit assessment since long delays would create growing disruptions of commuters' work schedule equivalent to their having to work overtime [Stafford, 1990]. If this were done, the $41 billion estimated social benefit could easily get boosted to, say, $60 billion per year by the year 2010.

The use of wage rate, with or without the overtime work premium, for calculating benefits of reduction in traffic delays is more justifiable in the case of commuting to work than in other types of traffic. It has been suggested that the driver's willingness to pay for reduction of both the mean and the variance of his travel time may be a more rational approach. Unfortunately, no studies of this nature have been done. However, it may be speculated that, if this were done, the benefit figure may get even higher as previous surveys have shown that people generally tend to overestimate time spent in virtually all activities and that people are often willing to pay dearly to get out of the kind of dire consequences (missing flights, being late in critical situations, etc.) as discussed at the beginning of this paper.

7 TRAFFIC DIVERSION STRATEGIES

It was pointed out in Section 5 that the widest part of the triangular area D in Figure 5 is the part that contributes the most to total disutility of traffic delays. Thus, it behooves the traffic manager, who wants to reduce total disutility most effectively, to consider new ways of "cutting into" this widest part of area D. Of course, as discussed in Section 6, the traffic manager has to be wary also of the disutility or negative benefit incurred by the vehicles in the alternate route.

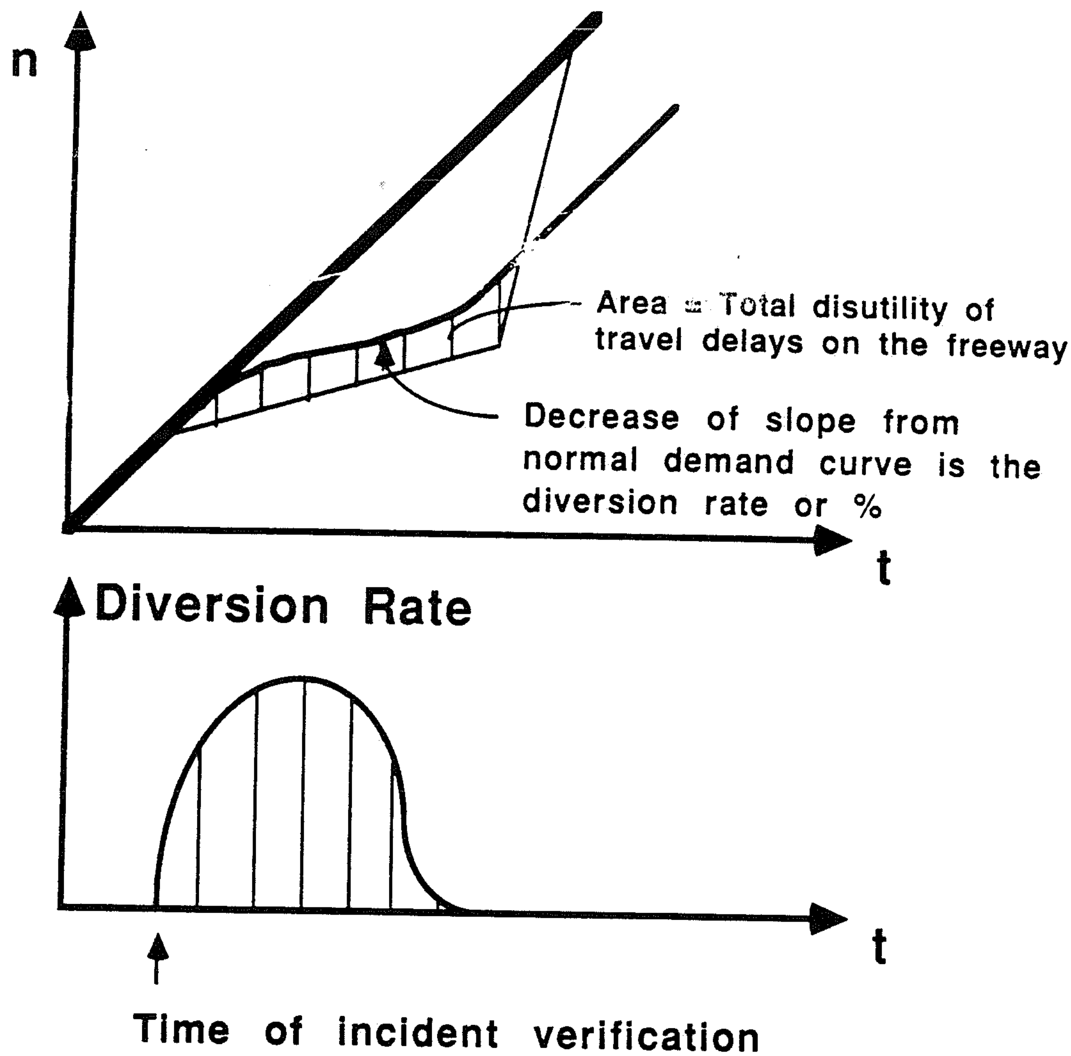

Fig. 6 A Time-Varying Diversion Strategy

For the moment, let us ignore the possible increase of disutility on the alternate route. An intuitive way of looking at the time-varying diversion rate is to think about how to limit or control maximum individual vehicle delay (the horizontal segments of the shaded area in Figure 6) through traffic diversion. A new strategy worth considering for traffic management is to not divert much traffic at the beginning of the incident since the small delays incurred by the first few vehicles in the short queue are relatively minor-- and thus with low disutility. Diversion should gradually increase toward a maximum as we try to "cut" deeply into the widest portion of the "triangle D." This graphical insight suggests that the optimum diversion rate may vary over time as shown in the lower portion of Figure 6. The upper portion shows the "cut", and the difference of the slope of the heavy line (normal traffic demand) and the slope of the smooth curve (actual flow) represents the flow rate of the diverted traffic. Also, as the time of incident clearance draws near, the diversion rate should be reduced gradually since a reduction of individual vehicle delays may be anticipated as the incident becomes clear.

One way to appreciate the effectiveness of the above-described diversion strategy is to consider the following example in which, with certain reasonable assumptions, the total disutility of traffic delays may be held to zero by using the strategy as shown in Figure 7. The two assumptions in this example are:

(1) The disutility of traffic delays for all drivers is piecewise linear as shown in Figure 3, and

(2) Increase of travel time of diverted vehicles on the alternate route is just below the threshold.

In this case, upon the detection and verification of an incident, the traffic manager should not divert any traffic until the queue builds up to the point where the rising individual vehicle delays reach the threshold in the piecewise linear disutility function. Then, sufficient traffic is diverted so that the traffic remaining in the original route is equal to the reduced capacity of the partially blocked route. Beyond this time, the vehicles staying in the original route will all incur minor delays just below the threshold, thus still having zero disutility according to the first assumption. The diverted vehicles will also experience zero disutility, according to the second assumption. The traffic manager does not need to wait until the incident clears before he stops the traffic diversion. As shown in Figure 7, he may bring diversion down to zero at a time equal to the predicted incident clearing time less the threshold (labeled as "scheduled normal delay" in the figure). This last step is taken so that more vehicles will reach their destinations sooner by keeping them in the original route than if they are diverted to the alternate route.

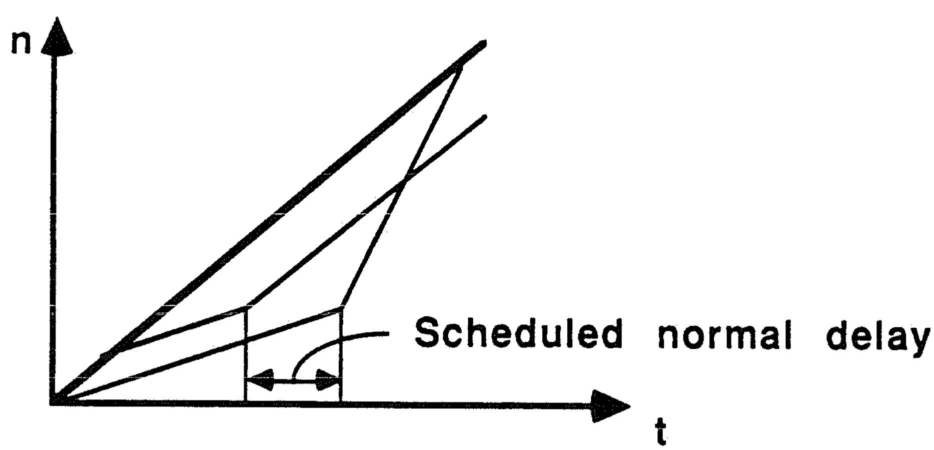

FIG. 7 A Case of Diversion with No Disutility

Although the above example is not very realistic, it should dramatically illustrate the potential advantage of new diversion strategies. To be realistic, we must of course consider the generally nonzero disutility of diverted vehicles that may take much longer travel time on the alternate route. This is particularly important when the diversion traffic is substantial, causing the alternate route to become excessively congested. In that case, optimum traffic control must be exercised on both the original route and the alternate route in a coordinated manner, which is not a trivial task.

8 CONCLUSIONS

This paper has discussed the concept and implications of nonlinear disutility of travel delays. When this common phenomenon is taken into account, the social benefits of route guidance resulting from the application of IVHS technology are believed to be substantially higher than what has been assessed on the basis of average wage rate. Also, with nonlinear disutility, the optimum strategy for traffic diversion from an incident is likely to be time-varying rather than some fixed volume or fixed ratio of vehicles as has been usually assumed.

What this paper has concentrated on is the traffic manager's goal of system optimization. If the individual driver's goal of user optimization is considered, then anticipatory route guidance, based on dynamic programming, would be the appropriate basic analytical approach [Kaufman *et al.*, 1990]. In that case, it appears that anticipatory route guidance can still be used to minimize driver's disutility instead of travel time, as long as the disutility function is convex as shown in Figure 1. If the convexity of disutility is not assured, then there may be problems of applying the principle of optimality in dynamic programming. This general statement may be applied to traffic simulation models to the extent that such models may include user-optimized route determination in their traffic simulation [Van Aerde and Voss, 1988].

9 REFERENCES

Federal Highway Administration (1983). *A Freeway Management Handbook* 2(23-31).

Hendrickson, C. and Kocur, G. (1981). "Schedule Delay and Departure Time Decisions in a Deterministic Model." *Transportation Science* **15** 1(62-77).

Hendrickson, C. and Plank, E. (1984). "The Flexibility of Departure Times for Work Trips." *Transportation Research* **18A** 1(25-36).

Kaufman, D.E., Lee, J., and Smith, R.L. (1990). "Anticipatory Traffic Modeling and Route Guidance in Intelligent Vehicle-Highway Systems." Intelligent Vehicle-Highway Systems Program, The University of Michigan.

Keeney, R.L. and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley & Sons, New York.

Luce, R.D. and Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*, John Wiley and Sons, New York.

Makigami, Y., Newell, G.F., and Rothery, R. (1971). "Three-Dimensional Representation of Traffic Flow." *Transportation Science* **5** 3(302-313).

May, A.D. (1990). *Traffic Flow Fundamentals*, Prentice-Hall, Englewood Cliffs, New Jersey.

Mobility 2000 (1990). "Executive Summary: Mobility 2000 Workshop on Intelligent Vehicles and Highway Systems." published by the Texas A&M University System.

Stafford, F.P. (1990). "Social Benefits of IVHS Systems," in Automated Highway/Intelligent Vehicle Systems: Technology and Socioeconomic Aspects, (SAE Publication SP-833).

Stafford, F.P. and Duncan, G. (1985). "The Use of Time and Technology by Households in the United States," in *Time, Goods and Well-Being*, (Juster, F.T. and Stafford, F.P., editors), Institute for Social Research, The University of Michigan.

Tzeng, G.H., Shieh, H.M., and Shiau, T.A. (1989). "Route Choice Behavior in Transportation--An Application of the Multiattribute Utility Theorem," *Transportation Planning and Technology* **13** 289-301.

Van Aerde, M. and Voss, J. (1988). *Integration 1 Users Manual*. Department of Civil Engineering, Queens University for the Ministry of Transportation, Ontario, Canada.

Wardrop, J. (1952). "Some Theoretical Aspects of Road Traffic Research," *Proceedings of the Institute of Civil Engineers*, Part II, **1** 352-378.