

Balancing Efficiency and Equity of Ramp Meters

Lei Zhang¹ and David Levinson²

Abstract: An efficient ramp control strategy may not be politically attractive due to its negative impacts on equity. This paper addresses this issue and presents a new objective for ramp metering—minimizing weighted travel time—which is able to balance efficiency and equity. A quasioptimization/simulation approach is used to achieve the new objective. BEEEX, a new family of control strategies with various degrees of equity consideration, are developed and used to demonstrate this approach.

DOI: 10.1061/(ASCE)0733-947X(2005)131:6(477)

CE Database subject headings: Ramps; Travel time; Traffic management.

Introduction

Previous evaluation studies (e.g., Cambridge Systematics 2001) show ramp meters can lower travel time, increase throughput, reduce accidents, and reduce fuel consumption and emissions. However, meters also have some negative effects. Sometimes, long queues created at entrance ramps frustrate drivers and the diverted traffic may place nearby intersections above capacity. It has also been a concern that ramp meters could exacerbate urban sprawl by encouraging longer trips. A spatial equity measure that summarizes the distribution of delays among drivers originating from different on ramps can capture those negative impacts of ramp meters. Kotsialos and Papageorgiou (2001) examine the spatial distribution of ramp delays with different on-ramp storage restrictions. Levinson and Zhang (2004) apply a Gini coefficient to measure distributional properties of trip delays. By one definition, a freeway system is perfectly equitable if all drivers experience the same delay per unit distance of travel. Most existing ramp control strategies aim to minimize total delay. This efficiency-oriented objective, however, leads to a control logic that is the least equitable, because the fewest on ramps are controlled to provide uncongested traffic conditions on the freeway mainline for all drivers (Zhang and Levinson 2004). The tradeoff between efficiency and equity has also been confirmed in the Twin Cities (Minneapolis, Minn.) ramp meter shut down experiment (Levinson and Zhang 2004). The most efficient ramp control strategy, therefore, is not politically attractive due to its negative

impacts on equity. The public also has substantial doubts on the overall system effectiveness after experiencing long delays at on ramps. In the case of the Twin Cities, dissatisfaction with long ramp delays (at times greater than 20 min) forced engineers to revise the algorithm to incorporate equity. All this suggests that, though improving efficiency is the original and still the most important goal of ramp metering, equity should be considered in more proactive ways to make the metering system viable.

Some elements in existing ramp control strategies, such as the constraint of maximum ramp queue length and on-ramp coordination, improve equity by equalizing delays to some extent. However, those elements were developed originally for other reasons (to prevent spillover effects and to prevent mainline congestion, respectively) and often operate through overriding mechanisms. Their effectiveness in balancing efficiency and equity is unclear.

This research redefines user benefits for freeway operations with drivers' perception in mind, and explores how the traditional efficiency and the emerging equity goals can be systematically balanced in light of this new definition. A practical simulation approach is recommended, instead of strict optimization for reasons explained below.

New Objective for Ramp Metering

In order to apply standard optimization methods to balance efficiency and equity, decision makers must either specify the relative importance of the two objectives, or determine their satisficing levels. However, in the case of ramp metering, the relationships between design variables (metering rates) and the objectives are not explicit, which makes it difficult to directly weight efficiency and equity. Furthermore, even if a biobjective formulation is available, solving for global optimality will be very difficult because the optimal control problem with the efficiency objective alone is already hard and requires heuristics (Lovell and Daganzo 2000) or search for local optimum (Kotsialos et al. 2002). Therefore, a more practical approach is adopted herein which starts with an inquiry into the reasons why the traditional metering objective fails to address equity.

Weighted Travel Time

Minimizing total travel time does not consider equity because all travel times (free flow, congested, stop and go) are counted

¹PhD Candidate, Dept. of Civil Engineering, Univ. of Minnesota, 500 Pillsbury Dr. SE, Minneapolis, MN 55455. E-mail: zhan0294@tc.umn.edu

²Assistant Professor, Dept. of Civil Engineering, Univ. of Minnesota, 500 Pillsbury Dr. SE, Minneapolis, MN 55455. E-mail: levin031@tc.umn.edu

Note. Discussion open until November 1, 2005. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this technical note was submitted for review and possible publication on August 21, 2003; approved on May 10, 2004. This technical note is part of the *Journal of Transportation Engineering*, Vol. 131, No. 6, June 1, 2005. ©ASCE, ISSN 0733-947X/2005/6-477-481/\$25.00.

equally. Under this objective, there is no particular incentive to reduce waiting times in queues that are usually perceived as more onerous by drivers. In order to improve the equity of ramp metering, drivers who are forced to wait in queue by the algorithm for overall system efficiency in a control interval should be considered more important in subsequent intervals so that extremely long delays can be eliminated. A straightforward way to incorporate this idea into traffic control is to weight travel times based on drivers' perception of their desirableness reflected in a weighting function (for instance, the longer the delay, the larger the weight). The weights are similar, but not necessarily identical, to values of travel time. Then the objective function becomes minimizing weighted travel time.

The objective of many ramp control strategies is to preclude freeway mainline queues by restricting mainline flows just below capacity. In this case, the new objective can be simplified to minimizing weighted ramp delay because: (1) the freeway mainline is uncongested and the total mainline travel time is constant (assuming fixed demand) and (2) uncongested travel times should have the same weights. It can be further shown by contradiction that for the new metering objective to be able to balance efficiency and equity, the weighting function must not be constant (i.e., longer delays have heavier weights). If all weights from ramp delays are the same, total weighted delay is minimized when total absolute delay is minimized because the weight is simply a constant and can be cancelled out from the objective function.

Weighting Function

The task of developing a new metering objective thus comes down to the identification of a nonconstant weighting function for ramp delays. One would naturally relate this weighting function to the value of travel time. Previous research focuses on how drivers perceive travel times classified by some rather coarse categories. For instance, Hensher (2000) shows that stop-and-go travel is 8–17 times more onerous than free-flow travel using logit models and data from stated preference surveys. However, such a value-of-time function is still constant in a sense that they would not reveal how much drivers perceive the first minute of delay differently from the second minute and so on. In order to obtain a truly nonconstant subjective value of time function for freeway travel and ramp delay, the authors have been working on data collected from stated preference surveys and driving simulation experiments designed explicitly for this purpose (Levinson et al. 2004). The most recent results which can be applied in operational ramp control strategies are summarized by Zhang et al. (private communication).

However, the difficulty of getting the value-of-time function does not completely hinder applications of the new control objec-

tive, because weights can also be appropriately specified by engineers to address practical problems. For instance, if long ramp delays are perceived as unbearable and it is important to ensure no one waits for more than 4 min. to gain public acceptance, the weighting function could be specified such that the weights for all delays shorter than 4 min are unity and for delays longer than 4 min infinity. This is the strategy adopted in the Twin Cities after the shut down experiment.

Implementation

Implementing the objective of minimizing weighted travel time requires the ability to track individual ramp delays. As mentioned above, a strict optimization is not pursued for its overwhelming complexity and uncertain benefits. An alternative procedure for implementation is to rely on traffic simulation, which consists of three steps: (1) various equity considerations are added to the standard efficiency-maximizing metering logic, and each equity consideration would then correspond to a candidate ramp control strategy; (2) all candidate control strategies are implemented in a traffic simulator and weighted travel times measured using simulation outputs (i.e., the task of mapping equity considerations to the objective is achieved via simulation); and (3) the best available control strategy yielding the lowest weighted travel time is determined. The following sections demonstrate this procedure.

Family of Ramp Control Strategies

There are two types of on-ramp coordination that should be distinguished. Coordination in many existing algorithms aims to prevent queue formation on the freeway mainline so as to minimize absolute travel time. We refer to it as efficiency coordination. In contrast, equity coordination is dedicated to equalizing delay. Equity coordination tends to meter more on ramps than "necessary." For instance, if at least two on ramps must be metered to relieve a downstream bottleneck, efficiency coordination would only coordinate these two ramps while equity coordination would restrict more than two on ramps to distribute delays more evenly among more drivers. As more ramps are metered, an improvement in system equity at the price of efficiency should be observed. The scope of equity coordination can be easily adjusted in response to traffic growth and drivers' perception.

The highest efficiency of the ramp metering system is achieved when only the on ramp(s) closest to every active freeway bottleneck are metered when freeway internal queues are not allowed in the objective (Zhang and Levinson 2004). A new family of ramp control strategies, balanced efficiency and equity

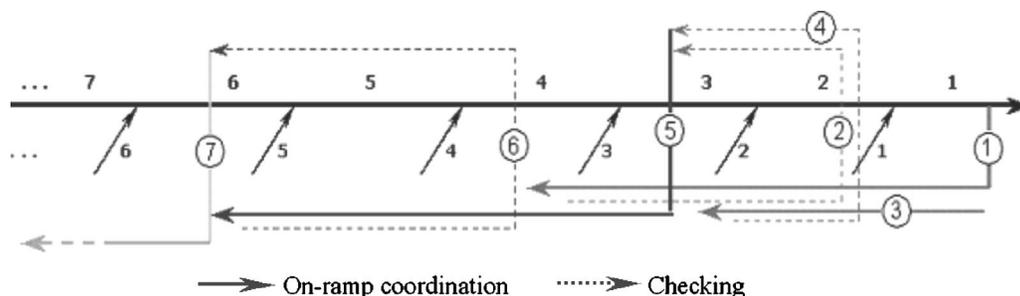


Fig. 1. On-ramp coordination and checking in BEE3

strategies with equity coordination factor X (BEEs) is developed in this study which implements this most-efficient ramp metering logic with various degrees of equity coordination. Active bottlenecks are identified dynamically in real time in BEEs. In BEEs, X on ramps will be metered upstream of an active bottleneck such that mainline queues are avoided (or minimized when queues inevitably form because of temporary demand surge) and delays are evenly distributed whenever feasible. However, coordination groups are not predetermined and adjusted in response to real-time traffic conditions. The mechanism of BEEs is illustrated in detail using the example of BEE3 in which the on-ramp coordination factor $X=3$ (see Fig. 1).

The control process starts from the furthest downstream section. In Step 1, Section 1 automatically becomes a critical section. Three on-ramps upstream of Section 1 (Ramps 1, 2, and 3) are coordinated to ensure the flow in Section 1 will not exceed its effective capacity. The coordination is executed in a way such that individual ramp delays at all on ramps in the same coordination group are equal for vehicles arriving at these on ramps at the same time. This is achieved by equalizing the ratio of the metering rates to ramp demands. Meter rates at these on ramps are determined tentatively according to the above criteria. For instance, if the extra capacity at Section 1 in a control interval (the difference between the capacity of Section 1 and the synchronized upstream mainline flow of the coordination group) is 12 vehicles/control interval, and the current queue lengths (demand) at On Ramps 1, 2, and 3 are 10, 10, and 20, respectively. The extra capacity should be allocated to the on ramps in proportion to the demand. Therefore, the tentative metering rates for the three ramps would be 3, 3, and 6, respectively, keeping a constant ratio of metering rates to demands (0.3) across all on ramps in the current coordination group.

In Step 2, a checking process assesses whether any mainline section (1, 2, and 3) in the current coordination group would become a new critical section under the tentative metering rates determined in the preceding coordination step. In the example, flow on Section 3 (the sum of upstream mainline flow and proposed On-Ramp 3 flow) will exceed its capacity if the tentative metering rates are implemented.

Therefore, in Step 3, the original critical section needs to be relieved by a new coordination group including only On Ramps 1 and 2. Metering rates for those two ramps are recalculated, which

is similar to Step 1. Following the numerical example in Step 1, the new tentative metering rates for On Ramps 1 and 2 should be 4 and 8, respectively.

Step 4 checks mainline sections (1 and 2) in the new coordination group, and no new critical section is identified. Therefore, the metering rates for On Ramps 1 (four vehicles) and 2 (eight vehicles) are finalized.

The next critical section is the next upstream mainline section (Section 3). Because $X=3$, three upstream on-ramps (3, 4, and 5) are coordinated in Step 5. Tentative metering rates are determined according to the equity coordination process previously described in Step 1.

Step 6 is a checking step similar to Steps 2 and 4, and concludes no section in the current coordination group (mainline Sections 3, 4, and 5) will be critical under the tentative rates (i.e., ramp flow determined in the previous coordination step would not cause over capacity flow), which are then finalized.

These alternating coordination and checking steps in BEE3 are executed until the metering rates for all on ramps are determined. The input data required by the BEEs strategies include mainline capacity and real-time traffic counts, both of which are widely available. These strategies do not require origin–destination information but assume that the probability a vehicle exits via a specific off ramp is independent of its origin.

Simulation and Results

The BEEs is a family of control strategies with different degrees of equity consideration. They can be evaluated and the strategy with the least weighted travel time identified via simulation. *Aimsun2* is selected to demonstrate this concept. A 20 km two-lane section of TH169 northbound from I-494 to I-94 in the Twin Cities is selected as the test site. There are ten weaving sections, 24 on ramps, 25 off ramps, four high occupancy vehicle bypasses, and two freeway-to-freeway ramps. The temporal (1400–1930 hrs) and spatial boundaries of the simulation experiment are free of congestion and traffic demand data were collected on March 21, 2000. This freeway segment was calibrated for *Aimsun2* by Hourdakakis and Michalopoulos (2002). The BEEs are coded as C++ extensions that read detector data from and feed metering

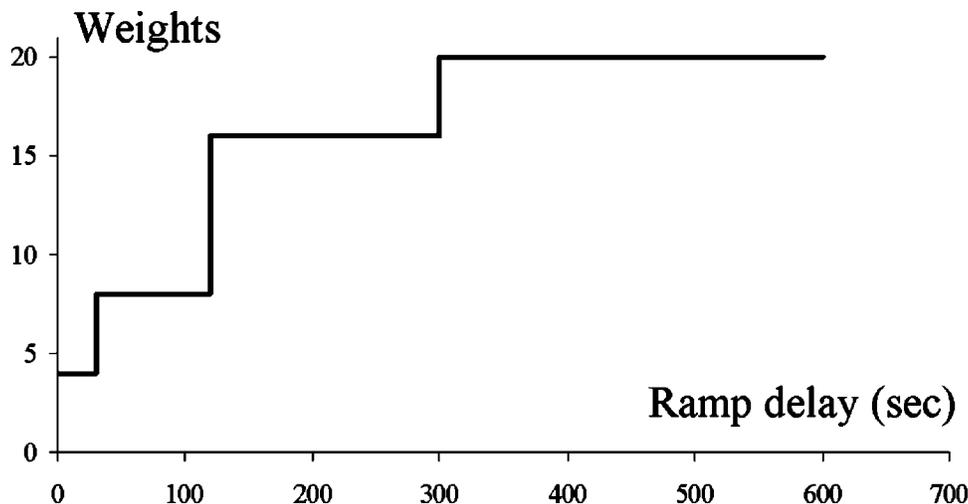


Fig. 2. Nonlinear weighting function

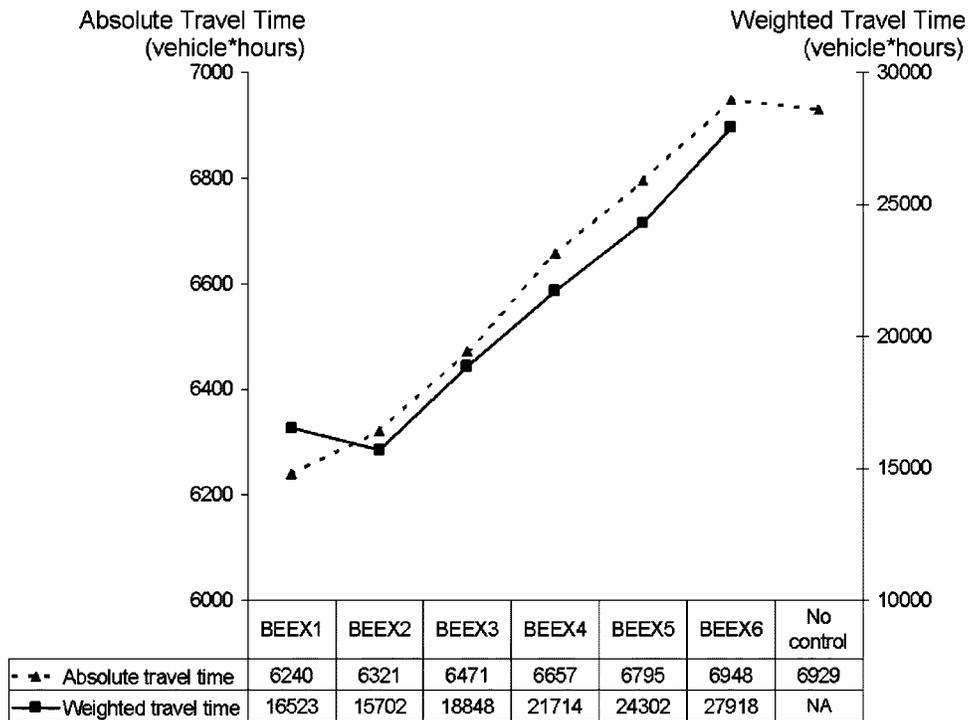


Fig. 3. Simulation results: absolute and weighted total travel time

rates back to the simulator every 30 s. After five runs for each BEEEX strategy and the base case (no metering), weighted travel times are calculated based on a plausible nonconstant weighting function specified in Fig. 2. In this function, uncongested travel times are weighted by unity. It is assumed that as delay increases, the weight also increases at a diminishing rate. The average weight for delays is between 8 and 17 based on the findings in

Hensher (2000). Queuing analysis is used to track individual ramp delays. Freeway mainline travel times are computed directly from *Aimsun2* statistical and detector outputs.

Results are summarized in Fig. 3. The BEE1, with no equity consideration, is the most efficient strategy with the minimum absolute travel time which is 10% lower than no-control. It is clear that equity is improved at the expense of efficiency which

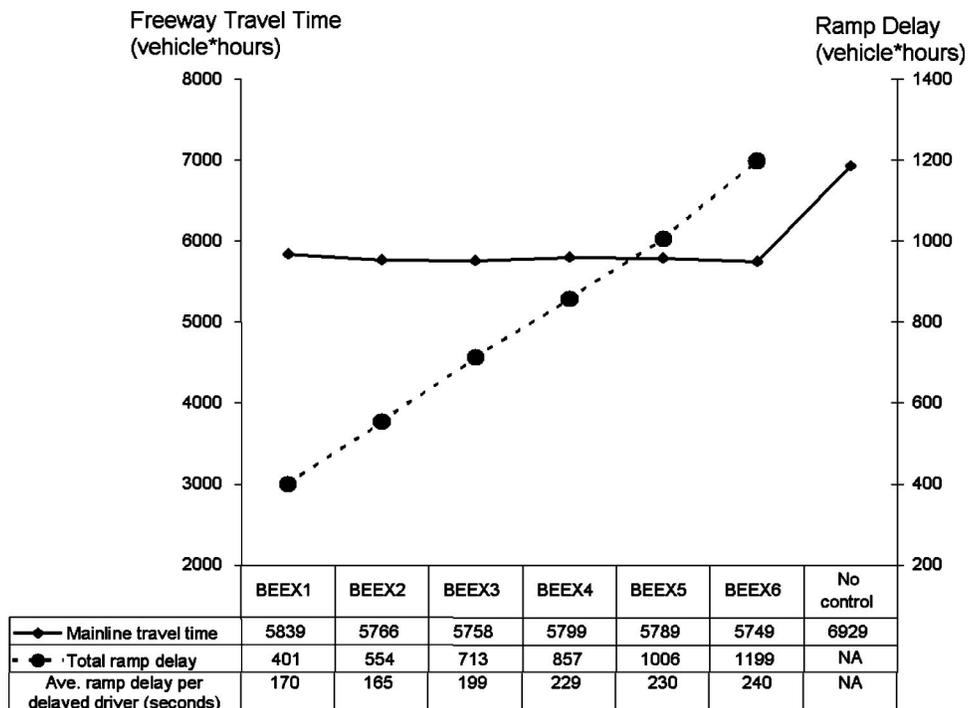


Fig. 4. Simulation results: freeway mainline travel time and ramp delay

drops consistently as X increases. Even coordinating one more ramp for equity consideration would cause a significant increase in total travel time (the difference between BEE1 and BEE2 is statistically significant at level 0.05). Up to a critical point ($X=5$), the remaining efficiency is even worse than the no-control scenario. If at this critical point the metering system is still considered as unacceptably inequitable, the best choice is not to deploy the metering system. However, when minimizing weighted travel time is the objective, BEE2 becomes the best strategy. All BEEEX strategies can be viewed as Pareto optimal solutions because no strategy is more efficient and more equitable than another. Weighted travel time naturally balances the two objectives and identifies the best-available strategy.

Since there is no congestion on the freeway mainline under BEEEXs, total absolute mainline travel time is about constant (see Fig. 4). Compared to the no-control scenario, BEEEXs on average save 1,146 vehicle h on the freeway mainline. As X increases more drivers are delayed at on ramps and total delay increases, but delays are more evenly distributed among drivers. For instance, BEE2, compared to BEE1, causes more total but shorter average delay per delayed driver. This is why BEE2 outperforms BEE1 when both are evaluated by weighted travel time.

Conclusions

An alternative objective for ramp metering, minimizing weighted travel time, is proposed to address the emerging need of balancing efficiency and equity in freeway operations. Drivers on metered freeways will become more and more aware of the ramp control inequity as traffic continues growing and delays at some on ramps increase. The new objective can be viewed as a compromise between engineers (efficiency) and politicians (equity and public acceptance). But more importantly, it comes from a better understanding of utility because the weights could represent drivers' real perception of different types of travel experience.

Ramp control strategies in the BEEEX family identify bottlenecks in real time and only require traffic data that are widely available. Decision makers can tune the balance between efficiency and equity to the desirable level by simply adjusting the value of one parameter in BEEEX—on-ramp coordination factor—without any equipment change.

Equity coordination improves system equity in a very system-

atic manner at an acceptable loss of efficiency. On the test site, BEE2–BEE5 all significantly reduce travel time. A constraint of maximum queue length or delay can also improve equity via overriding rules. But in that case, a sophisticated algorithm maximizing efficiency determines metering rates most of the time, and occasionally an extremely simple rule overrides the algorithm in the name of equity. Future research may compare the effectiveness of different methods.

Acknowledgments

This research was sponsored by the Minnesota Department of Transportation. The writers would like to thank the Center for Transportation Studies at the University of Minnesota for providing additional support. The opinions and errors remain those of the writers.

References

- Cambridge Systematics. (2001). *Mn/DOT Ramp Metering Study Final Rep.* (<http://www.dot.state.mn.us/rampmeterstudy>) (March 14, 2003).
- Hensher, D. (2000). "The valuation of travel time savings for urban car drivers: Evaluating alternative model specifications." *ITS-PWP-00-9*.
- Hourdakis, J., and Michalopoulos, P. (2002). Evaluation of ramp control effectiveness in two Twin Cities freeways. *J. Transp. Res. Board*, in press.
- Kotsialos, A., and Papageorgiou, M. (2001). Efficiency versus fairness in network-wide ramp metering, *2001 IEEE Intelligent Transportation Systems Conf. Proc.*, Oakland, Calif.
- Kotsialos, A., Papageorgiou, M., Mangeas, M., and Haj-Salem, H. (2002). "Coordinated and integrated control of motorway networks via non-linear optimal control." *Transp. Res., Part C: Emerg. Technol.*, 10C, 65–84.
- Levinson, D., Harder, K., Bloomfield, J., and Winiarczyk, K. (2004). "Weighting waiting: Evaluating the perception of in-vehicle travel time under moving and stopped conditions," in press.
- Levinson, D., and Zhang, L. (2004). "Evaluating the effectiveness of ramp meters." *Assessing the benefits and costs of ITS*, D. Gillen and D. Levinson, eds., Kluwer Academic, Dordrecht, The Netherlands.
- Lovell, D., and Daganzo, C. F. (2000). "Access control on networks with unique origin-destination paths." *Transp. Res., Part B: Methodol.*, 34B, 185–202.
- Zhang, L., and Levinson, D. (2004). Optimal freeway ramp control without origin-destination information. *Transp. Res., Part B: Methodol.*, 38B(10), 869–887.