Paper Entitled

"Traffic Estimation Based on Point Process Models"

From the Proceedings of

Engineering Foundation Conference on
Research Directions in Computer Control
of Urban Traffic Systems

pp. 266-285

Asilomar
February 1979

TRAFFIC ESTIMATION BASED ON POINT PROCESS MODELS

By J.S. Baras*, W.S. Levine*, A.J. Dorsey* and T.L. Lin*

Abstract

Two families of models for the relation between detector data and the flow of traffic on urban streets are developed. The first describes the relation between queues at a traffic signal and data from nearby detectors. The second models headway statistics for urban traffic thereby relating detector data to the passage of platoons of vehicles. Recent results in the theory of point processes then give the nonlinear minimum error variance filters/predictors corresponding to these models. It is then shown that these optimal estimators are computationally feasible in a microprocessor. The algorithms also perform very well when tested using data derived from the UTCS-1 traffic simulation.

1. Introduction

There is considerable current interest in the development of computer-based systems for the control of urban traffic. Over 25 such systems have been installed within the United States and approximately 125 others are in various stages of implementation [1]. Such systems have the potential to reduce traffic delay, fuel consumption, air pollution and accidents. It has been estimated that improved traffic signal systems could save 800 million gallons of fuel annually in the United States [2].

Generally, these computer-based systems consist of a collection of conventional looking traffic lights, a collection of vehicle detectors (usually inductive loops) and a computer that adjusts the traffic lights based to some extent on the signals from the detectors. However, the systems that are in use today are, from the control engineer's viewpoint, rather crude. The Federal Highway Administration recently built a system (the Urban Traffic Control System or UTCS) in Washington, D.C. to serve as a test of more advanced control procedures [3]. The UTCS was built in three versions. The first generation was a conventional system in which the computer was used to detect traffic patterns based on detector data. The computer then chose one

of 6 previously determined timing patterns and set the traffic lights accordingly. The timing pattern in use could be changed, at most, every 15 minutes. The first generation system produced only slight improvement over a completely open-loop system in which signal timing patterns were chosen according to time of day. The second and third generations of UTCS were designed to be successively more traffic responsive and to rely more heavily on the detector data and on-line signal optimization. The second and third generation systems performed worse than the first generation [4], the costs of data transmission escalated and UTCS was shut down in the fall of 1976. Tarnoff [4] argues that this degradation in performance was due to errors in surveillance and prediction of traffic.

As control engineers we could find several other reasons for the poor performance of the second and third generation systems but we also felt that the filtering and prediction could be substantially improved. Furthermore, since the models developed to solve the filtering and prediction problem could then be used in the control problem we felt that filtering and prediction was a good starting point.

We have since developed two families of filters/predictors. The first family consists of estimators, of increasing complexity and sophistication, of the queue at a traffic signal. The second family consists of estimators of the time at which a platoon of vehicles passes a detector and of the number of vehicles in the platoon. Since platoons of vehicles, in an urban network, are largely formed by traffic signals (one can view a platoon as a rolling queue or, conversely, a queue as a stopped platoon) the two families are closely interrelated. In actual operation, the queue estimator would initialize the platoon estimator. The platoon estimator would be used to correct errors in the queue estimate and to adjust parameters of the queue estimator (make it "adaptive").

In the following section we describe the queue estimators in detail. The third section describes the platoon estimators. The fourth and final section gives a brief description of our current research into control of urban traffic.

2. Queue Estimators

2.1 Models of Queue Formation and Dispersion: The modeling, and with it the filtering and prediction, of signals that are indirectly observed via a point process has progressed considerably in the past few years thanks to the efforts of many researchers (see [5] or [6] for references). In applications, modeling is the fundamental problem since some models lead directly to finite dimensional realizations of the optimal filter/predictor while other models do not. We present here a brief summary of results on the modeling of discrete time point processes following Segall [7].

Consider a sequence of observations $\{n(t)\}_{t=1}^{\infty}$ with $n(t)=1$ or $n(t)=1$

COMPUTER CONTROL OF TRAFFIC

being the only possibilities for each t. Suppose the probability that $n(t)=1$ is influenced by previous occurrences as well as by some other related sequence $\{\underline{x}(t)\}_{t=1}^{\infty}$ ($\underline{x}(t)$ may be vector-valued). The factors that may affect the occurrence probability at time t are the past observations denoted

$$\underline{n}^{t-1}=\{n(1), n(2), \ldots, n(t-1)\} \tag{2.1}$$

and the past and present of the related sequence

$$\underline{x}^t=\{\underline{x}(1), \underline{x}(2), \ldots, \underline{x}(t)\}. \tag{2.2}$$

It can then be shown that

$$\underline{x}(t+1)=\underline{f}(t, \underline{n}^{t-1}, \underline{x}^t)+\underline{u}(t) \quad \text{and} \quad n(t)=a(t, \underline{n}^{t-1}, \underline{x}^t)+w(t) \tag{2.3}$$

where $\underline{u}(t)$ and $w(t)$ are "unpredictable" given the information available at time $(t-1)$. For a precise mathematical statement of the meaning of "unpredictable" in this context see [5], [6] or [7].

We emphasize that this is *not* a "signal plus noise" model and that this model applies to *any* discrete time point process that is related to another time varying quantity. The equations simply reflect the fact that *any* observation sequence can be divided into the sum of a *predictable* part and an *unpredictable* part. Thus, the modeling problem reduces to finding the functional form of $a(t, \underline{n}^{t-1}, \underline{x}^t)$ and $\underline{f}(t, \underline{n}^{t-1}, \underline{x}^t)$.

We now apply the above modeling procedure to produce models for traffic queues.

The simplest practical traffic flow estimation problem occurs in the case of the single, isolated, intersection of two one-way, single lane streets. In order to adjust the traffic light to, in some sense, optimize (or even improve) the flow of traffic it is necessary to obtain fairly good estimates of the traffic queues upstream from the intersection. In practical systems, the estimate needs to be based on a minimal amount of historical data and on the signals from one, or more, detectors positioned as shown in Fig. 1. Assume, for simplicity, that the light operates on a simple, known, red-green cycle (no amber), that there is only one detector and that the detector is located N car lengths from the stop line.

The observed signal from the detector will be denoted by $n^a(t)$,

$$n^a(t) = \begin{cases} 0 & \text{if no vehicle is over the detector} \\ 1 & \text{if a vehicle is over the detector.} \end{cases} \tag{2.4}$$

In practice, time is discretized with a small enough discretization interval (1/32 second in UTCS) for each vehicle to be over the detector for several samples. For simplicity, it is assumed here that the data are sampled so that each vehicle produces exactly one pulse (one 1).

There are many factors which affect the rate process associated with $n^a(t)$. We believe that two of the most important are the upstream traffic signal and the number of vehicles in the queue. Thus, in this simplest model we let

$\lambda(k,t)$ = volume (in vehicles/second) at the detector given k vehicles are in the queue

$z(t)$ = number of vehicles in the queue at time t

Equivalently,

$$\lambda(k,t) = \Pr[n(t)=1|z(t)=k, t]$$
$$\lambda(N,t) = \Pr[n(t)=1|z(t)=N, t]=0. \tag{2.5}$$

Also assume

$$\lambda(k,t) = \begin{cases} \lambda_{kr} & \text{when upstream traffic light is red} \\ \lambda_{kg} & \text{when upstream traffic light is green.} \end{cases} \tag{2.6}$$

Similarly, let

$\mu(k,t)$ = rate at which vehicles depart from the queue given that the queue length is k.

So,

$$\mu(k,t) = \Pr[1 \text{ departure} | z(t)=k, t]$$
$$\mu(0,t) = 0 \text{ for all } t \tag{2.7}$$

and

$$\Pr[\text{more than 1 departure}] = 0$$

$$\mu(k,t) = \begin{cases} \mu_{kr} & \text{when downstream traffic light is red} \\ \mu_{kg} & \text{when downstream traffic light is green} \end{cases} \tag{2.8}$$

Furthermore, assume that arrivals and departures conditioned on knowledge of the queue, are independent.

This actually completes the construction of a model for the point process $n^a(t)$. To see this, and to put the model in the form of Eqs. (2.3), define

$$Q_{ij}^T(t)=\Pr\left[\begin{array}{l} \text{queue at time } t+1 \text{ contains } i \text{ vehicles}| \\ \text{queue at time } t \text{ contains } j \text{ vehicles} \end{array} \right] \tag{2.9}$$

At this point we make the approximation that a vehicle joins the queue the instant that it crosses the detector. Thus, strictly speaking, our "queue" is the number of vehicles between the detector and the stopline. With this approximation, it is elementary that

$$Q_{i,i}^T(t)=(1-\lambda(i))(1-\mu(i))(1-\mu(i-1))+\lambda(i)\mu(i)$$
$$Q_{i,i-1}^T(t)=\lambda(i-1)(1-\mu(i-1)),$$
$$Q_{i,i+1}^T(t)=\mu(i+1)(1-\lambda(i+1)), \qquad i=0,1,\ldots,N \tag{2.10}$$
$$Q_{i,j}^T = 0 \text{ elsewhere}$$

where the argument t has been suppressed in both λ and μ. Introduce the row vector

$$\underline{\lambda}^T(t)=[\lambda(0,t), \lambda(1,t), \ldots, \lambda(N-1,t), 0] \tag{2.11}$$

the model, one has to specify the function

$$\underline{\lambda}^t(\underline{x}(t), t, \underline{\sigma}) = \Pr[\text{one arrival at the tail of the queue at time } t \mid \underline{x}(t), \underline{\sigma}] \tag{2.16}$$

Once this is done, we write

$$\underline{x}(t+1) = \underline{Q}^t(t, \underline{\sigma})\underline{x}(t) + \underline{u}(t) \tag{2.17}$$

where $\underline{u}(t)$ is "unpredictable" and $\underline{Q}^t(t, \underline{\sigma})$ is the same as in Eq. (2.10) except that $\underline{\lambda}^t$, which depends on $\underline{\sigma}$, replaces $\underline{\lambda}$. Then Eqs. (2.15) and (2.17) comprise model B.

The only problem involved in augmenting Model B to utilize the information available from the additional detector is to characterize the point process at the new detector. We assume the new detector provides only occurrence times (no velocity data), since there is relatively little information in the velocity at the stop line. Proceeding, let

$$n^d(t) = \begin{cases} 1 & \text{if a vehicle crosses the stop line detector at time } t \\ 0 & \text{otherwise.} \end{cases}$$

We assume the associated rate

$$\lambda^h(\underline{x}(t), t)$$

This is obviously an approximation to the reality but we believe it is an adequate approximation. The model (Model C) now becomes

$$\underline{x}(t+1) = \underline{Q}^T(t, \underline{\sigma})\underline{x}(t) + \underline{u}(t) \tag{2.18}$$

$$\underline{n}^a(t) = \underline{V}\underline{x}(t)\underline{x}^T(t)\underline{\lambda}^a(t) + \underline{w}(t)$$
$$n^d(t) = \underline{x}(t)^T\underline{\lambda}^h(t) + w_{J+1}(t) \tag{2.19}$$

where $\underline{Q}(t, \underline{\sigma})$ is as in (2.17) (the corresponding equation in Model B) and $\underline{w}(t)$ and $w_{J+1}(t)$ are "unpredictable".

In summary, it should be clear that many other, similar, queuing models could be constructed. In fact, it is hoped that the ones constructed here demonstrate the technique so that the reader can, if he wishes, construct a model of his own.

2.2 Estimators of Queue Size: The detailed calculations are rather long and can be found in our report [5] so they are omitted. However, it can be shown that the minimum error variance estimate of $\underline{x}(t)$ in model A given data up to time t (n^t) is given by

and following Segall [7], define

$$x_k(t) = \begin{cases} 1 & \text{if there are } k \text{ vehicles in the "queue" } k=0,1,\ldots,N \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

$$\underline{x}^T(t) = [x_0(t), x_1(t), \ldots, x_N(t)].$$

It is now straightforward to establish that (Model A)

$$\begin{aligned} \underline{x}(t+1) &= \underline{Q}^T(t)\underline{x}(t) + \underline{u}(t) \\ n(t) &= \underline{\lambda}^T(t)\underline{x}(t) + \underline{w}(t) \end{aligned} \tag{2.13}$$

where $\underline{u}(t)$ and $\underline{w}(t)$ are "unpredictable" processes.

There are two changes that, on theoretical grounds, ought to improve the above model of traffic queueing at a signal. The first arises because many detectors give velocity (or a signal related to velocity) in addition to occurrence time for each vehicle that crosses the detector. The second involves defining the queue more accurately as the number of vehicles that are actually stopped at the traffic signal. The model described here incorporates both of these improvements.

First, we include the velocity information in the detector signal by defining the vector observation process $n_j^a(t)$ by

$$n_j^a(t) = \begin{cases} 0 & \text{if no vehicle crosses the detector at time } t \\ 1 & \text{if a vehicle crosses the detector at time } t \\ & \text{with velocity } v_j, j=1,2,\ldots,J. \end{cases} \tag{2.14}$$

Note that:
(a) we discretize velocity with a fairly coarse discretization so that $j=1,2,\ldots,J$ (J small, around 5);
(b) this model applies to the detector model incorporated in the UTCS-1 simulation. It needs to be modified slightly for a real detector.

With these definitions, we show in [5] and [6] that

$$\underline{n}^a(t) = \underline{V}\underline{x}(t)\underline{x}^T(t)\underline{\lambda}^a(t) + \underline{w}(t) \tag{2.15}$$

where
$\underline{\lambda}^a(t)$ is identical to $\underline{\lambda}(t)$ in model A (See Eq. 2.11)
$v_{ji}=\Pr[\text{vehicle crosses detector with velocity } v_j \text{ given a vehicle crosses the detector and } x_i(t)=1]$

$\underline{w}(t)$ is "unpredictable"
$\underline{x}(t)$ is as before.

Next we have to model the arrival process at the tail of the queue. It is obvious that there is a delay between the appearance of a vehicle with velocity v_j at the detector and the time that vehicle joins the queue. This delay clearly depends on the number of vehicles in the queue ($\underline{x}(t)$) and the velocity with which the vehicle crossed the detector. In the report [5], we give a detailed derivation for this delay based on reasonable assumptions about the way vehicles decelerate to join a queue. As a practical matter, all the information not contained in $\underline{x}(t)$ is stored in a three component vector denoted by $\underline{\sigma}$. Then, to complete

$$\hat{x}_i(t|t) = \begin{cases} \dfrac{(1-\lambda(i,t))\hat{x}_i(t|t-1)}{\displaystyle\sum_{i=0}^{N}(1-\lambda(i,t))\hat{x}_i(t|t-1)} & \text{if } n(t)=0 \\[3ex] \dfrac{\lambda(i,t)\hat{x}_i(t|t-1)}{\displaystyle\sum_{i=0}^{N}\lambda(i,t)\hat{x}_i(t|t-1)} & \text{if } n(t)=1 \end{cases} \tag{2.20}$$

where

$\hat{x}_i(t|t) = Pr[x_i(t)=1|n,t] = $ "filtered" output

$\hat{x}_i(t+1|t)=Pr[x_i(t+1)=1|n,t]=$ "one step predicted" output

The filter is initialized by

$\hat{\underline{x}}(1|0)=\underline{\pi}(0)=$ a priori probability distribution of queue length at $t=0$ (2.21)

Eqs. (2.20) and (2.21) are useless without the equation for updating the predictor. That is Eq. (2.22) below

$$\hat{\underline{x}}(t+1|t) = \begin{cases} \underline{M}^T(t)\begin{bmatrix}\hat{\underline{x}}(t|t)\\0\end{bmatrix} & \text{if } n(t)=1 \\[3ex] \underline{M}^T(t)\begin{bmatrix}0\\\hat{\underline{x}}(t|t)\end{bmatrix} & \text{if } n(t)=0 \end{cases} \tag{2.22}$$

where $\underline{M}^T(t)$ depends only on $\mu(t)$ and can be found in [5] or [6].

It should be apparent that Eqs. (2.20)-(2.22) represent an algorithm for filtering/predicting that can be easily realized in a microprocessor or special purpose hardware, see [5]or[6] This is especially so since $\underline{\lambda}(t)$, $\underline{Q}^T(t)$ and $\underline{M}^T(t)$ are all piecewise constant and periodic.

It is shown in [5] and [6] that the single detector filter/predictor using velocity and occurrence time can be reduced to:

$$\hat{\underline{x}}(t+1|t)=\underline{Q}^T(t,\underline{\sigma})\hat{\underline{x}}(t|t) \tag{2.23}$$

and

$$\hat{x}_i(t|t)=\begin{cases}\dfrac{(1-\lambda^a(i,t))\hat{x}_i(t|t-1)}{\displaystyle\sum_{j=0}^{N}(1-\lambda^a(j,t))\hat{x}_j(t|t-1)}, & \text{if } n_\ell^a(t)=0,\ \ell=1,\dots,5 \\[3ex] \dfrac{v_\ell\lambda_i^a(i,t)\hat{x}_i(t|t-1)}{\displaystyle\sum_{j=0}^{N}v_\ell\lambda_j^a(j,t)\hat{x}_j(t|t-1)} & \text{if } n_\ell^a(t)=0,\ \ell=1,\dots,5 \\ & \text{some } \ell\epsilon[1,2,\dots,5]\end{cases} \tag{2.24}$$

By comparing these equations to Eqns. (2.20) and (2.22) which

describe the simpler filter/predictor it is seen that the numerical complexity of the second filter/predictor is similar to that of the first filter predictor. Thus, a micro processor realization is feasible again.

The two detector filter/predictor is based on Model C. Since Model C is so similar to Model B, it is quite straightforward to derive the new filter/predictor.

The only complication is that it is possible to have $n^d(t)=1$ and one of the components of $n^a(t)$ also equal to one. However, it is reasonable to assume $n^d(t)$ and $n^a(t)$ are uncorrelated whenever the queue is not empty. When the queue is empty there is some relation between the two measurements, which is very difficult to describe and model. Thus, we make the simplifying assumption that $n^a(t)$ and $n^d(t)$ are uncorrelated.

Once this assumption is made, the derivation of the new filter/predictor goes through easily [5]. The most convenient way to express the result is in terms of a correction to Eqns. (2.23) and (2.24). See [5] or[6] for the expressions.

2.3 Test and Evaluation: The ultimate test of an algorithm for estimating traffic flow parameters is to include it in the software for an operating computer controlled traffic network. If, under those circumstances, the filter/predictor algorithm performs well then it is a good algorithm regardless of its performance on any other tests. Unfortunately, we do not have an operational computer controlled traffic network for use as a test. However, the Urban Traffic Control System Number One (UTCS-1) simulation model provides a reasonable and comparatively inexpensive means to test our filter/predictors.

The UTCS-1 simulation model is a very detailed simulation of urban traffic, developed under the auspices of the Federal Highway Administration. It is believed to be a fairly accurate simulation of urban traffic [11]. Furthermore, it is based on a model of traffic flow that is very different from any of our models [12]. For our tests, we simulated two simple urban networks, of which only the simplest is included in this paper. This network is shown in Fig. 2, where all streets are one way, single lane streets on which traffic flows in the arrow direction. The rectangles represent detectors which give occurrence time and correct velocity for each vehicle crossing the detector.

There are a number of detailed assumptions regarding street lengths, signal timing and detector locations incorporated in the simulation. These can be found in [5] or [6].

Four tests were run using the above network. Test 1 corresponds to "moderate" traffic flow in which the traffic signals dominate the traffic. Test 2 is again moderate traffic but is more "random" than Test 1. Test 3 corresponds to moderate to heavy traffic while Test 4 corresponds to heavy traffic.

The first queue estimation scheme that we tested was based on Model A. This filter/predictor, hereinafter denoted F/P I, depends only on two functions, $\lambda(i,t)$ and $\mu(i,t)$, where i expresses the dependence on queue length and t the dependence on time. For simplicity, we eliminated the dependence on i in the actual implementation of F/P I. Thus, F/P I depends only on three parameters

$$\lambda(i,t) = \begin{cases} \lambda_g & \text{upstream traffic signal green} \quad i=0,1,\ldots,N-1 \\ \lambda_r & \text{upstream traffic signal red} \quad i=0,1,\ldots,N-1 \end{cases}$$

$$\mu(i,t) = \begin{cases} \mu & \text{downstream traffic single green for 5 sec. or more} \\ & \quad i=1,2,\ldots,N \\ 0 & \text{otherwise} \quad i=0,1,2,\ldots,N \end{cases}$$

In all of the tests, the detector was 210 ft. from the stopline so N=10 was the value used. It will be seen that, once or twice, there were actually 11 vehicles in the segment between detector and stopline. The value for λ_g was estimated by averaging the number of vehicles crossing the defector during the upstream green over the upstream green time. λ_r was estimated similarly. The results are fairly insensitive to the values for λ_g and λ_r. On the other hand, μ is very important. Thus, several different values of μ were tried for each simulation. We expect that, because of this sensitivity to μ, it will be possible to use an adaptive procedure to compute it in an actual implementation of F/P I. The performance evaluation needed for adaptivity would be based on data from downstream detectors and the correlation between queues and downstream platoons discussed at length below. We have not had sufficient time to do this as yet.

We compared the performance of F/P I with ASCOT [9], [13], one of the "queue" estimators currently in use. ASCOT is regarded as one of the best single detector queue estimators [10]. However, it uses the velocity data from the detector, it gives only a single number as its estimate and this estimate is given only at the instant of red to green transition of the traffic signal. Thus, F/P I can give much more information, based on less data but more computation than ASCOT. The comparisons are summarized in tables 1 and 2.

We did a great many additional tests of all three filter/predictors using the basic simulation described earlier. Some of these are reported in [6] and most in [5]. Unfortunately there is not enough room here to describe the results.

3. Platoon Estimators

3.1 Models for Headway Statistics: It has been recognized that one of the most important components in the description of traffic flow is the distribution of headways. Although several definitions of headway exist, we will always mean the time difference between the passage of the leading edges of successive vehicles. The statistical distribution of headways has been studied extensively since the early

days of traffic control. It is natural for our work for two reasons:
(a) it is relatively easy to collect headway data from the existing detectors.
(b) the statistical description of headways (interarrival times in the point process jargon) is the essential part in modelling the underlying point process and is the point of departure for the modern theory of estimation for point processes (see [5] or [8] for references).

For a complete description of the traffic process we need to include the speed measurements provided by the detectors [5]. This is the mark (in point process jargon) of the point process that characterizes traffic detector output. Such measurements and a model similar to ours have been effectively utilized in [15] to describe freeway traffic. It is worth emphasizing that [15] provides a substantial validation of the aforementioned model. In this paper, due to simplifying considerations and space limitation, we consider only headway statistics. Speed statistics and a more complete development based on a mixed headway-speed model will be given elsewhere.

Most of the prior work on headway statistics was concerned only with the probability density for headways. Our report [5] and [14] contain a detailed survey.

It has been known for some time that, because of the different statistical behavior of short and long headways, so-called composite density models give better fit to data than single density models, [15]. These models assumed a structure of traffic consisting of two subpopulations: one corresponding to following traffic (i.e., traffic grouped in platoons) and one corresponding to nonfollowing traffic (i.e., freely flowing vehicles or leaders of platoons). The headway probability density assumes then the form

$$p(h) = \psi p_f(h) + (1 - \psi)p_{nf}(h) \tag{3.1}$$

where

p_f = following headway probability density function (short headways),

p_{nf} = nonfollowing headway probability density function (longer headways). Usually a displaced negative exponential density.

ψ = degree of interaction.

Since headway is dependent on traffic flow, the degree of interaction incorporates this dependency. For light traffic for example, ψ equals zero yielding a composite density that is a displaced negative exponential. There are several interpretations one can give to ψ and we shall return to this point later. It has been found, [15], that p_f does not depend on the position of the vehicle within the platoon and on the size of the platoon.

People have tried several different densities to model following or platooned vehicles. From these, the lognormal density

$$p_f(h) = \begin{cases} \dfrac{1}{\sigma h \sqrt{2\pi}} \exp\left(-\dfrac{(\ln h - \mu)^2}{2\sigma^2}\right) &, h \geq 0 \\ 0 &, h < 0 \end{cases} \quad (3.2)$$

(where μ, σ^2 are the mean and variance of $\ln h$), or shifted lognormal density gave the best results in fitting observed data from platooning vehicles. There are various justifications for these findings about the lognormal density. The primary reason is the multiplicative, independent, identically distributed errors by various drivers attempting to follow each other combine to give a lognormal density.

It is also well known that a shifted exponential distribution gives a good fit to non-following headway data. As a result of these considerations the model adopted for the first order headway density is given by (3.1) where p_f is as in (3.2) and p_{nf} has the form

$$p_{nf}(h) = \begin{cases} \lambda e^{-(h-\tau)\lambda} &, h \geq \tau \\ 0 &, h < \tau. \end{cases} \quad (3.3)$$

The model requires five parameters for the headway density, $\psi, \lambda, \tau_\mu, \mu$ and σ. To completely specify the model for a particular link or section of a link in a traffic network, it is important to understand the variation of these parameters with respect to traffic flow and speed. Others (see [5] or [8] for references) report that μ, σ are fairly insensitive to traffic flow level while varying from lane to lane and different links. The parameters ψ and λ depend on traffic flow and are rather easily estimated if one utilizes velocity (speed) statistics as well. Finally τ varies between .25 to 4.00 sec and can be easily estimated [15].

The probability density given does not provide in general a complete description of the headway stochastic process at a particular point in a traffic network. Higher order probability density functions are also needed because there may exist correlation between successive headways. On the other hand, we know from point process theory that interarrival time statistics completely characterize the process and, in particular, can be used to determine the "rate" of the process. This rate plays a central role in estimation. To simplify computations and based on evidence provided in [15] we analyze for the balance of the paper a model which employs uncorrelated following headways.

We developed two interpretations for the mixed headway model. The first model is intended for use in estimating gross traffic patterns for the slow updating of traffic flow parameters (both in urban and in particular freeway traffic). In such a case ψ, which should be interpreted as the probability that a particular headway is a following headway, should be constant for long time intervals. The second model is intended for use in urban nets with small average link lengths for use in urban nets with small average link lengths and traffic signals. In such cases it is crucial to model the periodic formulation and

---

propagation of platoons or queues as modulated by traffic lights. Then ψ is modelled as a time function with values 1, corresponding to passage of a platoon or a queue discharge and 0 corresponding to nonfollowing freely flowing traffic.

We call the first model average mixed headway model. The point process it characterizes has rate

$$\lambda_t(N_t = n, T_n) = p(t-T_n) \frac{1}{1 - \int_0^{t-T_n} p(x)dx} \quad (3.4)$$

where p is given by (3.1). The function

$$h(t) = \frac{p(t)}{1 - \int_0^t p(x)dx} \quad (3.5)$$

is sometimes referred to as the hazard function in birth or renewal process jargon. Our results indicate that filters/predictors really behave well if the hazard function is chosen appropriately. This suggests the alternative: derive filter/predictors by appropriate choice of the hazard function and make them adaptive by tuning the hazard function to the traffic flow pattern.

We call the second model switching rate mixed headway model. This model is based on the switching of ψ between 0 and 1. As a result the point process will have two rates. The following headway rate is

$$\lambda_f(N_t = n, T_n) = \frac{g(t-T_n)}{1 - erf\left[\frac{\ln(t-T_n) - \mu}{\sigma}\right]} \quad (3.6)$$

where g is the lognormal density (3.2). For the nonfollowing headway process the rate is given by (using 3.3)

$$\lambda_{nf}(N_t = n, T_n) = \begin{cases} \lambda & \text{if } t-T_n \geq \tau \\ 0 & \text{if } t-T_n < \tau \end{cases} \quad (3.7)$$

Some of these computations are used later in the disorder problem for point processes. These computations complete the description of the headway process model.

A model can now be developed for urban traffic flows based on the headway model adopted. Each link is divided in sections in accordance with the detectorization of the link. For each section of the link the input and output traffic flows will have headway distributions as described above. Notice that the headway distribution model can vary (and it should) from lane to lane. The required parameters of the model will be estimated at appropriate intervals from actual data, or from historical data as required. The effect of the link will be to alter the parameter value as traffic moves down stream.

3.2 Platoon Structure Estimation: Although several filtering/prediction problems of relevance to urban traffic control problems can be formulated, we concentrate on the estimation of traffic patterns (i.e., passage time of platoon or queue). From section 3.1 the point process observed by a traffic detector is a mixture of two point processes each (i.e., in platoons or queues) (3.6) and a different one associated with nonfollowing vehicles (3.7). The rate of the overall process switches between these two rates (switching rate mixed headway model). Estimates of the switching times can be very useful for the following reasons: a) they determine the traffic flow pattern and if transmitted to downstream detectors and traffic light controllers will lead to improvement in filtering/prediction and control of subsequent links; b) a common problem with queue estimators is the errors from traffic cycle to traffic cycle due to vehicles trapped by the red light or vehicles passing during the amber to red transition. By effectively estimating from the first downstream detector (i.e., the second from the left in figure 2) the time when the last queueing vehicle has passed that detector a reinitialization of the upstream queue estimator can be implemented to correct cycle by cycle propagation of cumulative errors.

In a different, traffic oriented problem, we are often interested in estimating or detecting the times when large changes in the rate process occur. This is often related to an incident in a freeway (or urban traffic link). This is the incident detection problem and will be treated elsewhere.

All the above problems can be formulated in the context of the so called point process "disorder" problem. Namely, we observe a point process N_t which is governed by a rate process λ_t^0 until some random time T (called the "disorder" time), and by a different rate λ_t^1 after this time. The problem is then to estimate the switching time T from the observations of N_t only. This problem has been studied by Shiryaev, Galchuk and Rozowsky, Davis and in complete generality by Wan and Davis, see [5], [8], or [14] for references. We follow Davis in the development presented here.

We first need to establish the structure of the problem. Define

$$x_t = \begin{cases} 0 & t<T, \text{ i.e., if the rate process is } \lambda_t^0 \text{ ("disorder" has not occurred)} \\ 1 & t\geq T, \text{ i.e., if the rate process is } \lambda_t^1 \text{ ("disorder" has occurred)} \end{cases}$$

It is slightly more convenient and precise to denote this by

$$x_t = I_{\{t>T\}} \tag{3.8}$$

where $I_{\{t\geq T\}}$ is the characteristic function of the set $\{t\geq T\}$. So x_t indicates by switching from 0 to 1 the "disorder" time. Of the several cases considered in the literature, the appropriate one for the traffic problems discussed earlier is the following: the switching time T

coincides with one of the detector activation times T_i (occurrence times). In general, and in particular for traffic problems, the events $\{T=T_i\}$ may not be independent from the underlying point process N_t. Let

$$p_i = Pr\{T=T_i\}, \quad q_i = p_i/\sum_{k\geq i} p_k \tag{3.9}$$

$$q_t = \sum_i q_i I_i \{T_{i-1} \leq t < T_i\} \tag{3.10}$$

By some calculations which can be found in [5] or [8], one can then show that

$$dx_t = (1-x_t)q_t\hat{\lambda}_t^0 + dv_t \tag{3.11}$$

and

$$dN_t = ((1-x_t)\lambda_t^0 + x_t\lambda_t^1)dt + dw_t \tag{3.12}$$

where dv_t and dw_t are unpredictable increments and dx_t, dN_t denote increments in x_t and N_t. We will see momentarily that, practically, we can eliminate the increments by writing explicit solutions for the stochastic differential equations. The mathematical details can be found in [5]. The filter then becomes:

$$d\hat{x}_t = -(\lambda_t^1-\lambda_t^0)\hat{x}_t(1-\hat{x}_t)dt + \frac{(\lambda_t^1-\lambda_t^0)\hat{x}_t(1-\hat{x}_t)+q_t\lambda_t^0(1-\hat{x}_t)}{(\lambda_t^1-\lambda_t^0)\hat{x}_t+\lambda_t^0} dN_t \tag{3.13}$$

$$\hat{x}_0 = E\{x_0\}=P_0.$$

where $\hat{x}_t = E(x_t$ given all the data up to time t)
$= Pr[T\leq t|$ all the data up to time t]

Thus (3.13) computes the probability that the switch has occurred prior to time t given the detector data up to time t. When there is dependence between the events $\{T=T_i\}$ and $\{N_t\}$ some simple arguments [5] lead to Eq. (3.13) with the exception that

$$q_t = \sum_i q_i\{t, T_1, T_2, \ldots, T_{i-1}\} I\{T_{i-1} \leq t < T_i\}. \tag{3.14}$$

Thus, the only change needed to accomodate dependence between $\{N_t\}$ and $\{T=T_i\}$ is to let q_i be a function of t and the prior T_i.

Given explicit expressions for the two rates λ_t^0, λ_t^1 then (3.13) is an implementable nonlinear filter. Using then expressions (3.6), (3.7) we proceed to derive explicit equations for the filter. Between detector activations $dN_t=0$)

$$\frac{d\hat{x}_t}{dt} = -(\lambda_{nf}-\lambda_f)\hat{x}_t(1-\hat{x}_t), \quad T_{i-1}<t<T_i. \tag{3.15}$$

This equation can be solved explicitly [5] [14] to give

with a uniform probability for any number of vehicles in the platoon up to twenty. The uniform distribution was chosen because it provides essentially no a priori information. Thus, the performance of PE in these tests depends only on the data from the detector and is not biased by either accurate or erroneous foreknowledge. In a real application the performance would almost certainly be better.

A detailed description of the simulation and of the tests can be found in our report [5] and in [14]. For our purpose here, it is sufficient to note that the detector is 290 feet downstream from the traffic light that causes the platoon to form.

In order to evaluate PE conveniently, the conditional distribution is reduced to a scalar estimate in Table 3. Two estimates are obtained:

a) the estimate is the number of vehicles that have passed the detector at the first instant that the estimated probability that the platoon has passed the detector exceeds 0.7. This is called the threshold estimate.

b) The estimate is the number of vehicles that have passed the detector at the time of the largest increase in the estimated probability that the platoon has passed the detector. This is called the maximum jump estimate.

The errors in Table 3 are due to a platoon from upstream joining the end of the platoon formed by the traffic signal and then the combined platoon crossing the detector (the actual error is only one vehicle).

Table 4 summarizes the results of a much more favorable traffic situation. The upstream traffic signal is 800 ft. away so that there is a relatively large gap between successive platoons. Furthermore, the detector is located near the downstream stop line so the flow over the detector is in clearly defined platoons.

These results indicate that PE is a fairly accurate estimator of the number of vehicles in a platoon. Furthermore, it accurately determines whether or not the queue emptied on a cycle by cycle basis.

3.4 Parameter Sensitivity and Estimation: To complete the analysis of the estimator presented we need to determine methods which compute adaptively the filter parameters and we need to know the response sensitivities to these parameters. These are rather hard analytical problems and some partial results have been obtained in [14], where we refer for further details. A more complete analysis will appear in [8]. The parameters λ, τ are determined by fitting the tail of the observed headway density. Such methods were successful in [15]. Once an appropriate separation of short headways is available μ, σ can be easily estimated since the natural logarithm of the following headway is gaussian. We tried several techniques which automatically tried to separate the data (employing outlier tests). The results were

$$x_t = \left[1 + \frac{1 - \hat{x}_{T_{i-1}}}{\hat{x}_{T_{i-1}}} \exp\lambda(t-T_{i-1}-\tau)u(t-T_{i-1}-\tau)\right](1 - erf[\frac{\ell n(t-T_{i-1})-\mu}{\sigma}])\Big]^{-1}$$ (3.16)

$$\text{for } T_{i-1} < t < T_i.$$

where u is the unit step function. On the other hand when $t=T_i$ (i.e., at detector activation times) the estimate has a jump discontinuity with size equal to the coefficients of dN_t in (3.13).

$$\hat{x}_t - \hat{x}_{t-} = (1-\hat{x}_{t-})\frac{(\lambda_{nf}(t-) - \lambda_f(t-))\hat{x}_{t-} + q_{t-}\lambda_f(t-)}{(\lambda_{nf}(t-)-\lambda_f(t-))\hat{x}_{t-} + \lambda_f(t-)}$$ (3.17)

Thus, the filter is actually implemented as follows: a) between detector activation times (3.16) is used, b) at detector activation times the jump discontinuity is computed from (3.17), c) the error function appearing in (3.16) is computed by a five term series expansion.

Finally, for the implementation of the filter we need to determine the the deterministic function q_t, which in our case is given by (3.11) and therefore we need to specify the p_i's in (3.10). It is clear from the definition of the p_i's that the information carried by them is identical to the probability density for queue length. That is the output of queue estimators that computes the probability density for queue length can be used to compute the values for p_i. For simplicity and to obtain a "worst case" type evaluation of the filter performance we used a uniform probability density over the maximum possible queue length. That is

$$p_i = 1/N, \quad i=1,\ldots,N$$ (3.18)

where N was the maximum queue length allowed (i.e., the distance in car lengths of the upstream detector from the traffic light). Further details can be found in our report [5] and [14].

3.3 Platoon Estimator Evaluation: The estimator depends on four parameters. The first parameter is λ, the mean arrival rate for free-flowing traffic. The second parameter is τ the displacement of the negative exponential density. The third and fourth parameters μ, σ define the lognormal distribution associated with following headways.

In all of the tests, the parameters were held at $\lambda=.10$, $\tau=0.5$ secs. $\mu=1.0$ and $\sigma=.1681$. The value of σ was chosen to match Branston's value which was obtained for freeway traffic. He showed that σ did not vary very much over different traffic flow levels. The value of μ was chosen so that the mean headway between successive vehicles in a platoon, as given by the lognormal distribution, would be 2.9 secs. The value of λ was chosen so that the mean headway for nonfollowing vehicles would be 10 secs. The estimator, which we will denote by PE, must be given an initial estimate of the probability of each feasible number of vehicles in the platoon. In all of our tests PE was initialized

not very satisfactory. Convergence was not a serious problem however. The development of completely adaptive parameter estimation techniques remains an open question. However, computation of the filter output (i.e., the conditional distribution of the switching time) showed negligible variation with large variations in the filter parameters $\lambda,\mu,$ σ (we tried variations as large as 50%!). Furthermore, since the quality of the estimator is judged by the conditional error variance

$$V_t = \hat{x}_t (1 - \hat{x}_t) \qquad (3.19)$$

we studied variations in V_t under similar variations in the parameters. Again the observed variations in V_t were minute. In particular the result of the maximum jump estimator was almost unaffected. Several bounds and analytical expressions of the sensitivity of V_t with respect to λ,μ,σ can be found in [14]. The filter appears to be very robust, although we have not as yet obtained a complete mathematical proof.

4.   Application to Control

The results summarized here suggest that our queue estimators are probably an improvement over existing ones. Similarly, although we are not aware of previous attempts to estimate platoon size or passage time, our platoon estimator appears to be effective at doing this. One's first thought is to incorporate these estimators in a modified version of UTCS 2nd and 3rd generations. We are currently doing this. Our basic approach is to use our queue estimators to simply replace the UTCS filtering and prediction algorithms. The platoon estimator is being used to try to improve a green wave scheme.

It is more interesting to try to use the models developed for estimation to derive "optimal" control laws. There are two immediate lessons to be learned from this. First, the data from the detectors should be used as soon as possible. It is a general rule in control systems that delay in a feedback loop is bad. We are attempting to quantify this statement using our models but there is no doubt that it is true. Second, a related point is that there is very little advantage, from the control viewpoint, to complete centralization. We are trying to determine quantitative measures of the value of signal and detector data versus distances from a controller.

We are also currently trying to compute bounds on the possible performance of a controller at a single intersection. We have shown that $\hat{x}(t|t)$ is a sufficient statistic for the optimal control. However, this still leaves a very difficult to implement optimal control. Thus, we are more interested in sub-optimal controls that have nearly optimal performance and are easy to implement.

Finally, we are trying to make our estimators and control laws "adaptive" in the following precise sense. Our models depend, for each detector, on a small number of parameters. These parameters change slowly with time as the urban area and its traffic patterns changes. These parameters determine the parameters of our estimators and

controllers. We are trying to devise algorithms which will readjust these parameters on-line ("self-tune"). We believe this is crucial for the successful and economic operation of traffic control systems.

References

1.   Stockfisch, C.R., "Selecting Digital Computer Systems", Report of the U.S. Dept. of Trans., Federal Highway Administration, Offices of Research and Development, December 1972.

2.   Torrevo, E.A., "Unjamming Traffic Congestion", IEEE Spectrum, Nov. 1977, pp. 77-80.

3.   "Urban Traffic Control and Bus Priority System, Vol. 1, Design and Installation", Prepared by Sperry Systems Management Div., Sperry Rand Corp. for Federal Highway Administration, Office of Research and Development under Contract FH-11-7605, 1972.

4.   Tarnoff, P.J., "The Results of FHWA Urban Traffic Control Research: An Interim Report", Traffic Engrg., Vol. 45, April 1975, pp. 27-35.

5.   Baras, J.S., W.S. Levine, A.J. Dorsey, T.L. Lin, "Advanced Filtering and Prediction Software for Urban Traffic Control Systems", Final Report, D.O.T., contract No. DOT-OS-60134, Dec. 1977.

6.   Baras, J.S., W.S. Levine and T.L. Lin, "Discrete Time Point Processes in Urban Traffic Queue Estimation" (to appear) IEEE Trans. on Automatic Control, Vol. AC-24, No. 1, Feb. 1979.

7.   Segall, A., "Recursive Estimation of Discrete Time Point Processes", IEEE Trans. on Info. Theory, Vol. IT-22, No. 4, July 1976.

8.   Baras, J.S., Levine, W.S., and Dorsey, A.J., "Estimation of Traffic Platoon Structure from Headway Statistics", to appear IEEE Trans. on Autom. Control.

9.   Ross, D.W., et al., "Improved Control Logic for Use with Computer Controlled Traffic - Interim Report for NCHRP Project 3-18(1)", Stanford Res. Inst., Stanford, Calif., prepared for the Highway Res. Board, Washington, D.C., July 1972.

10.   MacGowan, C.J., "A Single Detector Queue Determination Algorithm", M.S. Thesis, Dept. of Civil Engrg., Univ. of Md., College Park, Md., 1975.

11.   "Network Flow Simulation for Urban Traffic Control System - Phase II Vol. 5 Applications Manual for UTCS-1 Network Simulation Model", prepared by Peat, Marwick, Mitchell & Co. and KLD Assoc., Inc. for Fed. Hwy. Admin., Office of Res. & Dev. under Contract DOT-FH-11-7885 (1974).

12.   "Network Flow Simulation for Urban Traffic Control System - Phase II Vol. 4 User's Manual for UTCS-1 Network Simulation Model" by E. Lieberman, D. Wicks and J. Woo for Fed. Hwy. Admin., Office of Res. and Dev., under Contract DOT-FH-11-8502 (1977).

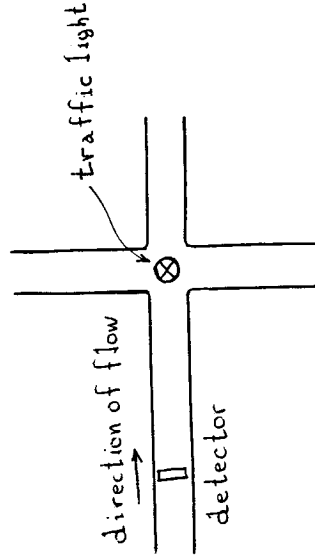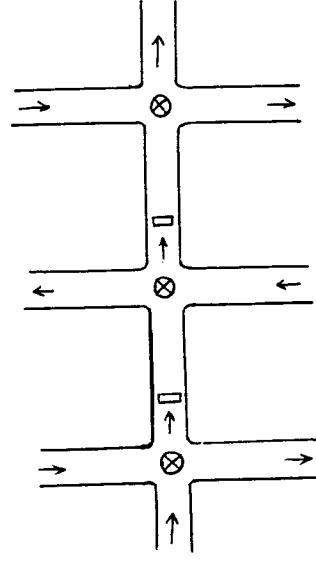| Cycle Number	Threshold Estimate	Maximum Jump Estimate	Actual Queue
1	6	7	6
2	6	8	8
3	7	7	7
4	7	7	7
5	9	9	9
6	1	6	7

Table 4   Performance of Platoon Estimator

Figure 1.   Detector location

Figure 3.   Street Configuration Test Network #I.

---

13. Ross, D.W., "Single Detector Queue Measurement for Computerized Traffic Systems" DARO Associates, Inc., Sunnyvale, Calif., (unpublished) 1975.

14. Dorsey, A., "Point Process Estimation Derived from Statistical Description of Vehicle Headways", M.S. Thesis, Dept. of Electrical Engrg., Univ. of Md., College Park, Md., May 1978.

15. Brieman, L., Lawrence, R., Goodwin, D., and Bailey, B., "The Statistical Properties of Freeway Traffic", Transp. Res., Vol. 11, pp. 221-228, 1977.

Time	Actual Queue	ASCOT Estimate	F/P I Estimate
80	2	3	2.7
160	2	2	2
240	5	5	5
320	5	7	4.5
400	5	7	3.8

Table 1   Comparison of ASCOT with F/P I Data is Test 3, link 5-6

Time	Actual Queue	ASCOT Estimate	F/P I Estimate
80	8	3	5.8
160	9	7	9
240	10	7	9.5
320	10	12	9.5
400	7	7	9.5

Table 2   Comparison of ASCOT with F/P I Data is Test 4, link 5-6

Cycle Number	Threshold Estimate	Maximum Jump Estimate	Actual Queue
1	11	11	2
2	3	2	2
3	3	3	3
4	10	10	1
5	0	11	1
6	3	3	2

Table 3   Performance of Platoon Estimator