JB 76-03

Filtering and Prediction Algorithms for Urban Traffic Control Systems

Proceedings:

4th Intersociety Conference on Transportation

July 18-24, 1976
Los Angeles, CA.

Filtering and Prediction Algorithms for Urban Traffic Control Systems

by

John S. Baras* and William S. Levine*

Abstract

There is evidence that the algorithms for estimating traffic flows from sensor data need to be improved before computer controlled traffic responsive urban traffic control systems can reach their full potential effectiveness. A large part of the problem appears to be that the data from traffic sensors is, in the statistical jargon, a marked point process. It is only very recently that the theoretical techniques for estimation based on point process data have reached the sophistication needed for traffic problems. Thus, this paper describes and illustrates the way in which these techniques are being applied to the design of algorithms for filtering and prediction of urban traffic flows,

*Department of Electrical Engineering, University of Maryland, College Park, Maryland 20742.

Introduction:

There is some evidence that a major difficulty in implementing traffic responsive, computer controlled, urban traffic control systems is that the existing algorithms for estimating and predicting traffic parameters are not accurate enough. For example, recent attempts to implement one such system (the UTCS in Washington, D.C.) resulted in, at best, only marginal improvement over the previously implemented system based on time of day and historical data [1]. This paper describes some techniques and traffic models that are being utilized in the development of improved filtering and prediction algorithms for use in urban traffic control.

The basic problem is to take the signals from traffic sensors (normally loop detectors) scattered throughout the network and to process this data to obtain good estimates of traffic volume, occupancy, queue length, stops, delay, average speed and travel time. One difficulty is that the data are either:

1) a sequence of times $t_1, t_2, \ldots (t_i < t_{i+1})$ representing the activation times of the detector or,

2) the data in (1) together with some auxiliary observations, such as the characteristics of each pulse (e.g. duration).

Case (1) is a random point process; case (2) is a marked point process. In the urban traffic situation, neither of these processes is a Poisson process.

Theoretically, an optimal estimator for traffic can be developed via a purely Bayesian approach. To do this one first determines (experimentally or theoretically) a statistical model for traffic flow. Specifically, it is necessary to obtain the conditional probability of the sensor data given the state of traffic in the network (the apriori probabilities). Straightforward application of Bayes' Rule then allows the calculation of the conditional probability of the traffic state given the sensor data (the a posteriori probability). This approach is illustrated in Section Three of this paper and leads to a set of equations that are difficult, if not impossible, to solve in a practical traffic situation. Thus, the major practical problem is to determine means to either reduce the complexity of these equations, or approximate them in a useful way, or calculate the estimates without having to first calculate the entire posterior probabilities.

Section Four introduces some recent results from the theory of point processes which reduce the complexity of the equations for the a posteriori probabilities. These results are then applied, primarily for illustrative purposes, to a very simple and highly idealized traffic problem. The basic idea utilized is to represent the detector signals as "generated" by certain other stochastic processes intrinsically associated with them (effectively, to model the probabilities in a more convenient way). Once this representation has been found it is fairly straightforward to produce a dynamical system, driven by the sensor signal, that produces the time trajectory of the estimators.

2. Traffic Models

It should be apparent from the discussion above that our approach to the design of filtering and prediction algorithms requires the preliminary development of a model for the statistics and, even more important, the dynamics of traffic flow. The model described below is believed to be adaptable to a wide variety of traffic situations, convenient to parameterize, reasonably accurate and analytically tractable.

Consider a link of the traffic network (assumed to be one-way without loss of generality), say link i between intersections k, ℓ. We divide each link into a maximum of three sections, with the division points (some or all) being detector locations.

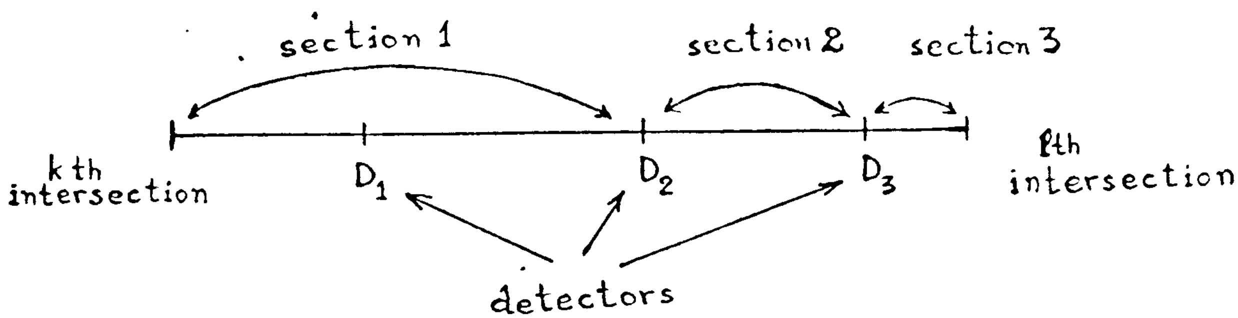

Figure 1. Illustrating the sections of a link of the network.

For $1 \leq j \leq 3$ we let d_{ij} be the length of the j^{th} section of link i, $x_{ij}(t)$ be the number of vehicles in section j of link i at time t (i.e., aggregate density), $u_{ij}(t)$ be the mean speed of vehicles in section j of link i at time t and $\phi_{ik}(t)$ the number of vehicles entering link i at the upstream intersection k during the time interval $[t, t+1]$. The determination of the time unit will be based on the state of the traffic flow (light, moderate, heavy) and on the effect it will have on the filter/predictor performance. Finally let $\alpha_{ij}(t)$ denote the percentage of cars in section j of link i that stay in that section during one unit of time. Then the equations for link i are:

$$
\left.\begin{aligned}
x_{i1}(t+1) &= \alpha_{i1}(t)\, x_{i1}(t) + \phi_i(t) \\
x_{i2}(t+1) &= [1 - \alpha_{i1}(t)]\, x_{i1}(t) + \alpha_{i2}(t)\, x_{i2}(t) \\
x_{i3}(t+1) &= [1 - \alpha_{i2}(t)]\, x_{i2}(t) + \alpha_{i3}(t)\, x_{i3}(t)
\end{aligned}\right\} \qquad (1)
$$

Let $q_i^j(t)$ be the output of the j^{th} detector on the i^{th} link. As was previously indicated $q_i^j(t)$ is a marked point process [2]. Let $N_i^j(t)$ be the associated counting process per unit time. That is, $N_i^j(t)$ denotes the number of vehicles passing detector j on link i during the time interval $[t, t+1]$. Then if detector D_{j+1} is located at the downstream end of section j of link i we have that

$$
(1 - \alpha_{ij}(t))\, x_{ij}(t) = N_i^{j+1}(t) \; ; \; j = 1, 2 \qquad (2)
$$

We let $\theta_i(t)$ denote the flow entering the network at the input intersection and proceeding in link i during $[t, t+1]$, and $\gamma_m(t)$ denote the flow that was formerly in branch m (which is entering node k) and leaves the network at the k^{th} intersection. Then

$$
\phi_i(t) = \sum_{m=1}^{k_i} [1 - \alpha_{m3}(t))\, x_{m3}(t) - \gamma_m(t)] + \theta_i(t) \qquad (3)
$$

Equations (1), (2), (3) describe the model for link i. The summation in (3) is overall links entering node k and such that traffic can proceed to link i. Normally $k_i \leq 3$. In general the α_{ij} will depend randomly on the traffic conditions in the section and the downstream section and on traffic signals. This can be expressed as an initial functional relationship

$$
\alpha_{ij}(t) = f_{ij}(x_{ij}(t),\ x_{i(j+1)}(t),\ u_{ij}(t),\ u_{i(j+1)}(t),\ \nu_{ij}(t)) \qquad (4)
$$

(where j+1 should be interpreted as "next"). Simplicity of the model requires simplicity in the f_{ij}. Typically $f_{i1}=f_{i2}=f_{i3}=f$ in terms of their functional form.

It is generally possible, by means of one of the various traffic flow theories [3, 4] to eliminate the dependence on velocity and reduce Eq. (4) to:

$$\left.\begin{aligned} \alpha_{ij}(t) &= f(x_{ij}, x_{i(j+1)}, \nu_{ij}(t)); \; j = 1, 2 \\ \alpha_{i3}(t) &= 1 - u_{\ell}(t) [1 - f(x_{i3}(t), x_{m1}(t), \nu_{ij}(t))] \end{aligned}\right\} \quad (5)$$

where $u_{\ell}(t)$ denotes the time variation of the traffic signal at intersection ℓ and is a time function taking values 1 (corresponding to green) or 0 (corresponding to red). The subindex m in (5) refers either to the next link of the same roadway downstream or to the link of a crossing roadway.

Equations (1)-(3) and (5) can then be combined to give a complete model of an urban traffic network. This model is used in the next two sections to derive procedures for estimating traffic parameters.

3. Bayesian Traffic Estimation

In this section a straightforward Bayesian approach to the estimation of traffic parameters is illustrated. It is essential to first interpret the traffic flow model of the previous section in a slightly different way.

For simplicity, assume we are dealing with a single street segment (drop the subscript i in Eqs. (1)-(5)). In addition assume that:

A1) Δt, the time interval between t and t+1 is so small that $N^j(t)$ and $\phi(t)$ are 0-1 processes, e.g. $N^j(t) = 0$ or $N^j(t) = 1$

A2) $\phi(t)$ has a known deterministic rate. That is, $\phi(t)$ is a sequence of Bernoulli trials

$$Pr[\phi(t) = 1] = p_\phi = \lambda_\phi \text{ ; } Pr[\phi(t) = 0] = 1 - p_\phi$$

p_ϕ assumed constant for simplicity.

A3) The observations from the detectors are:

D_1 - ignored

D_2 - $z_1(t) = N^2(t) + \xi_1(t) = (1 - \alpha_1(t))x_1(t) + \xi_1(t)$

D_3 - $z_2(t) = N^3(t) + \xi_2(t) = (1 - \alpha_2(t))x_2(t) + \xi_2(t)$

where ξ_1 and ξ_2 are noise terms. Again, we are ignoring velocity data to simplify the presentation. Detector 1 is primarily for velocity estimates and so is ignored.

A4) $Pr[x_j(0) = m] = \pi_{mj}(0)$ is known $j = 1, 2, 3$; $m = 0, 1, 2 \ldots$

A5) $Pr[\alpha_j(t)x_j(t) = m \mid x_j(t) = m] = P_{mj}$ is known ; $m = 0, 1, 2 \ldots$

then $Pr[\alpha_j(t)x_j(t) = m-1 \mid x_j(t) = m] = 1 - P_{mj} = \lambda_{mj}$ by assumption (1)

A6)
$$\left.\begin{aligned}
&Pr[z_j(t) = 0 \mid N^{j+1}(t) = 0] = P_{j0}\\
&Pr[z_j(t) = 1 \mid N^{j+1}(t) = 0] = (1 - P_{j0})\\
&Pr[z_j(t) = 1 \mid N^{j+1}(t) = 1] = P_{j1}\\
&Pr[z_j(t) = 0 \mid N^{j+1}(t) = 1] = (1 - P_{j1})
\end{aligned}\right\} ;\ j = 1, 2$$

The above assumptions provide a complete statistical description of the traffic model and, in a moment, will be used to derive an estimator. However, several remarks about the model are in order first.

It is clear that several of the assumptions simplify the model beyond the point where it would be useable in a real traffic environment. In particular, $\phi(t)$ would really have a rate function $\lambda(t)$ that is time-varying, dependent on traffic elsewhere in the network and unknown. Similarly, (A5) makes $\alpha_j(t)$ dependent only on $x_j(t)$ and ignores the dependence on $x_{j+1}(t)$ (the downstream traffic).

It is not difficult to incorporate improvements to the statistical model that would eliminate these objections without changing the essential features of the model. Furthermore, the theoretical traffic estimator developed below can be developed for the more complex and more realistic traffic model. We leave the assumptions as they are in the interest of simplifying and clarifying the exposition of the techniques.

It should be clear that the traffic flow model, Eqs. (1)-(3) and (5) and assumptions (A1)-(A6) result in a traffic model that is a discrete time and discrete space Markov Process. Similar models have proven useful before in both theoretical and practical traffic problems [5, 6] . Given the model, one poses the following problem.

Given a sequence of measurements $z_j(t)$; $j = 1, 2$; $t = 0, 1, \ldots, T$

Estimate:

a) $x_j(T)$ = number of vehicles in segment j at time T

b) the rate at which vehicles leave segment j in interval T to T + 1

The minimum variance estimates can be computed as follows [7, 8] .

Let

$$\hat{\pi}_{\underline{n}}(t) = \Pr[\,\underline{x}(t) = \underline{n} \mid \underline{\varepsilon}(t)] = \Pr[\,x_1(t) = n_1, x_2(t) = n_2, x_3(t) = n_3 \mid \underline{\varepsilon}(t)] \qquad (6)$$

where $\underline{\varepsilon}(t)$ = total information available to the estimator at time t.

Then, since the only information acquired at time t is $\underline{z}(t)$

$$\underline{\varepsilon}(t) = [\, \underline{z}(t), \underline{\varepsilon}(t-1) \,] \tag{7}$$

We next perform a sequence of straightforward calculations of conditional probabilities.

$$\hat{\pi}_{\underline{n}}(t) = \Pr[\underline{x}(t) = \underline{n}, \underline{z}(t) = \underline{\ell} \,|\, \underline{\varepsilon}(t-1)] \,/\, \Pr[\underline{z}(t) = \underline{\ell} \,|\, \underline{\varepsilon}(t-1)] \tag{8}$$

where $\underline{z}(t) = \underline{\ell}$ means $z_1(t) = \ell_1$ and $z_2(t) = \ell_2$ for known ℓ_1 and ℓ_2 ($\ell_i = 0$ or 1).

But,

$$\Pr[\underline{x}(t) = \underline{n}, \underline{z}(t) = \underline{\ell} \,|\, \underline{\varepsilon}(t-1)] = \sum_{\underline{m}} \{ \Pr[\underline{z}(t) = \underline{\ell} \,|\, \underline{x}(t) = \underline{n}, \underline{x}(t-1) = \underline{m}, \underline{\varepsilon}(t-1)] \times$$

$$\Pr[\underline{x}(t) = \underline{n} \,|\, \underline{x}(t-1) = \underline{m}, \underline{\varepsilon}(t-1)] \times \Pr[\underline{x}(t-1) = \underline{m} \,|\, \underline{\varepsilon}(t-1)] \} \tag{9}$$

where $\sum_{\underline{m}}$ equals the sum over all possible triples m_1, m_2, m_3. However, under the assumpitons made about the model

$$\Pr[\underline{x}(t) = \underline{n} \,|\, \underline{x}(t-1) = \underline{m}, \underline{\varepsilon}(t-1)] = \Pr[\underline{x}(t) = \underline{n} \,|\, \underline{x}(t-1) = \underline{m}] \tag{10}$$

$$\Pr[\underline{z}(t) = \underline{\ell} \,|\, \underline{x}(t) = \underline{n}, \underline{x}(t-1) = \underline{m}, \underline{\varepsilon}(t-1)] = \Pr[\underline{z}(t) = \underline{\ell} \,|\, \underline{x}(t) = \underline{n}] \tag{11}$$

$$\Pr[\underline{x}(t-1) = \underline{m} \,|\, \underline{\varepsilon}(t-1)] = \hat{\pi}_{\underline{m}}(t-1) \tag{12}$$

Combining Eqs. (8)-(12) gives

$$\hat{\pi}_{\underline{n}}(t) = \frac{\sum_{\underline{m}} \Pr[\underline{x}(t) = \underline{n} \,|\, \underline{x}(t-1) = \underline{m}] \cdot \Pr[\underline{z}(t) = \underline{\ell} \,|\, \underline{x}(t) = \underline{n}]\, \hat{\pi}_{\underline{m}}(t-1)}{\sum_{\underline{n}} \sum_{\underline{m}} \Pr[\underline{x}(t) = \underline{n} \,|\, \underline{x}(t-1) = \underline{m}] \cdot \Pr[\underline{z}(t) = \underline{\ell} \,|\, \underline{x}(t) = \underline{n}]\, \hat{\pi}_{\underline{m}}(t-1)} \tag{13}$$

All of the quantities in Eq. (13) can be calculated from the initial assumptions, before any data is acquired, with the exception of $\hat{\pi}_{\underline{m}}(t-1)$. The $\hat{\pi}_{\underline{m}}(t)$ are calculated recursively, as data arrives, from Eq. (13) and the known value for $\hat{\pi}_{\underline{m}}(0)$ (obtained from A4). Once $\hat{\pi}_{\underline{n}}(t)$ is known, the minimum variance estimates of traffic parameters are calculated as follows:

$$\hat{x}_1(t) = \sum_{n_1} n_1 \left(\sum_{n_2} \sum_{n_3} \hat{\pi}_{\underline{n}}(t) \right) \tag{14}$$

$$\hat{\lambda}_1(t) = \text{rate estimate for section } 1 = \sum_{n_1} (1 - P_{n_1}) \left(\sum_{n_2} \sum_{n_3} \hat{\pi}_{\underline{n}}(t) \right) \tag{15}$$

The other estimates are computed analogously.

These results can be summarized as follows:

1) The estimates $\underline{\hat{x}}(t)$ and $\underline{\hat{\lambda}}(t)$ are minimum variance estimates. That is, no filter or estimator can be built, using the same data, that gives a lower error variance provided traffic actually matches the model.

2) Most of the defects in the model can be corrected. That is, at the expense of some additional complication, the model can be made into a very good mathematical model of real traffic.

3) It is necessary to posses a good deal of information about traffic flow on the network to build the estimator. However, the probabilities that are needed are probabilities of vehicle flow and are thus relatively easy to measure.

4) This is not a practical estimator as it stands. The reason is almost entirely that the computational burden is too large. One must calculate $\hat{\pi}_{\underline{n}}(t)$ for all possible values of $\underline{n}$ at each instant of time. This is an enormous number for a realistic network.

5) Thus, one approach to the development of practical traffic estimators is to try to find good approximations to the above calculations. An alternate approach is described in the next section.

4. Techniques Based on Point Processes Theory

From the previous section, it is apparent that efficient optimal filtering or prediction algorithms should either produce a solution to the conditional density equation with simpler calculations, or directly compute conditional expectations in a more efficient manner. Since data from traffic detectors (see Introduction) are point processes, recent advances in point process theory [9, 10, 11, 12] hold promise for simple recursive solutions to various estimation problems of interest to traffic engineers. To illustrate these techniques, we give a somewhat oversimplified application to traffic below.

Suppose we have a perfect detector located at the upstream end of an urban street (by a perfect detector we mean that the detector accurately counts every vehicle passing it). The output of the detector is then denoted by $N(t)$ and is a counting observation process. In fact $N(t) = \phi(t)$ from the previous sections. However, we now make the more realistic assumption that the rate $\lambda_\phi(t)$ is, in fact, time varying. To keep the problem as simple as possible we assume that

$$\lambda_\phi(t) = \lambda(t) = \begin{cases} \lambda_0 & 0 < t \le T \quad ; \; T \text{ random} \\ \lambda_1 & T < t \end{cases} \tag{14}$$

The problem is to estimate the time, T, at which the rate changed, based on the signal from the detector.

This model, while somewhat arbitrary, is reasonable in situations where it is important to estimate the time when the rate of passing cars switches between markedly different values. This is certainly important for detectors located at entrance points of the network, since the switching in the rate of arrivals of cars, determines a change in the user demand for the network. On the other hand, most control algorithms are based on fixed rates for periods of time, and change as the rates change. Thus determination of switching times for the rates of passing cars, influences the switching times for various control policies. By comparing data from two detectors the motion of large platoons of cars can be followed and traffic lights adjusted accordingly. Finally the detection of switches in the rates of passing cars clearly has applications in the incident detection problem.

As was the case in the previous section, it would be possible to solve a more realistic traffic estimation problem. For example, point process theory produces a recursive solution to the problem of optimally estimating the time when the distributions of a general counting observations process change. The problem has been solved in a general situation, allowing arbitrary apriori distributions for the switching time T and dependence of T on the underlying general counting process in [13]. The latter case is of obvious importance in closed loop traffic control systems. It is shown in [13] that the solution of the above filtering problem can be easily computed on a computer. We present here, for simplicity, the solution when the counting process N(t) is a Poisson counting process whose rate changes from λ_0 to λ_1 (positive constants) at a certain time T. T is a random variable that is zero with probability π; and given that $T>0$ is exponentially distributed with parameter λ. We follow the analysis as presented in [10]. For the solution of this filtering problem one wants to calculate the time evolution of the a posteriori probability

$$\hat{\pi}(t) = \Pr(T \leq t \,|\, \underline{\varepsilon}(t)) \tag{15}$$

where $\underline{\varepsilon}(t)$, is as before, the information available at time t (i.e. N(s), $s \leq t$).

In the filtering theory for point process [9, 10] observations one usually needs a dynamical model for a "signal" process and for an "observation" process. In the following paragraph the present filtering problem is brought into an appropriate form for the application of the general filtering equation [10]. This involves the introduction of some artificial random variables (for details and further explanations we refer to [10]). So, let α be a 0, 1 random variable with probabilities π and $1-\pi$ and $p(t)$, $p^{(0)}(t)$, $p^{(1)}(t)$ be three Poisson processes with constant rates λ, λ_0, λ_1 such that α, p, $p^{(0)}$, $p^{(1)}$ are mutually independent. Then if T_1 is the first jump of $p(t)$, one defines

$$\left.\begin{aligned} T &= \alpha T_1 \\ f(t) &= \alpha \lambda I_{\{t<T_1\}} \\ y_t(t) &= (1-\alpha) + \alpha p(t \wedge T_1) \end{aligned}\right\} \tag{16}$$

where $I_{\{t < T_1\}}$ is the indicator function for the interval $(0, T_1)$, $t \wedge T_1 = \min(t, T_1)$. The random process $y(t)$ plays the role of the "signal process" and we want a dynamical model for this process. Observe that $\{t<T\} = \{y(t) = 0\}$, $\{T \le t\} = \{y(t) = 1\}$ and therefore $y(t) = I_{\{y(t)=1\}} = I_{\{T \le t\}}$. Therefore $f(t) = \lambda(1 - y(t))$ and the signal equation becomes

$$\left.\begin{aligned} dy(t) &= \lambda(1 - y(t))\, dt + dv(t) \\ y(0) &= 1-\alpha \end{aligned}\right\} \tag{17}$$

where

$$v(t) \triangleq y(t) - \int_0^t f(s)ds = (1-\alpha) + \alpha[p(t \wedge T_1) - \lambda(t \wedge T_1)].$$

Then the counting observation process is

$$N(t) = \int_0^t (1-y(s))dp^{(0)}(s) + \int_0^t y(s)dp^{(1)}(s) \tag{18}$$

Therefore the dynamical model for the observation process is

$$dN(t) = [(1-y(t))\lambda_0 + y(t)\lambda_1]dt + dw(t) \tag{19}$$

where

$$w(t) \triangleq N(t) - \int_0^t [(1-y(s))\lambda_0 + y(s)\lambda_1]ds .$$

Since $\hat{y}(t) = E\{y(t)|\underline{\varepsilon}(t)\} = E\{I_{\{y(t)=1\}}|\underline{\varepsilon}(t)\} = Pr(y(t)=1|\underline{\varepsilon}(t)) = Pr(t \geq T|\underline{\varepsilon}(t)) = \hat{\pi}(t)$ (20)

one applies to the "signal" and observation model (17), (19) the general filtering equation to obtain [10]:

$$\left.\begin{aligned} d\hat{\pi}(t) &= (\lambda - (\lambda_1 - \lambda_0))\hat{\pi}(t)(1-\hat{\pi}(t))dt + \\ &+ \frac{(\lambda_1 - \lambda_0)\hat{\pi}(t) - (1-\hat{\pi}(t-))}{\lambda_0(1-\hat{\pi}(t-)) + \lambda_1\hat{\pi}(t-)} \quad dN(t) \\ \pi(0) &= \pi \end{aligned}\right\} \tag{21}$$

This is a simple equation, driven by the counting process N(t), whose solution completely specifies the posterior distribution of T. Here $\hat{\pi}(t-)$ denotes the left hand limit of $\hat{\pi}(t)$ at time t. Observe that when no car is passing the second term vanishes, leaving a simple differential equation to solve, while the second discontinuous term enters into the computation only when counts occur. It is easy to see that given a record of the counting process one can solve the above equation. Trajectories of $\hat{\pi}(t)$ are therefore easily computed. For further results on trajectories for other cases (i.e. different counting processes and (or) different apriori distributions for T) we refer to [13].

In more complicated situations in traffic, we can have more than one possibility for the rate process. Point process theory can handle this case too. Furthermore in several instances in traffic control problems the "signal" process y(t) will generally be a finite state Markov process satisfying

$$dy(t) = f(t)dt + du(t), \; y(0) = 0$$

with discrete state space $\{0, 1, \ldots N\}$. Then point process filtering theory allows for the derivation of the time evolution of the N-vector process $\underline{\hat{\pi}}(t)$ where

$$\hat{\pi}_i(t) = Pr(y(t) = i \mid \varepsilon(t))$$

Further applications of point process methods to traffic control problems will be reported elsewhere.

5. Conclusions

The basic idea in all of the preceeding discussion is that the effective estimation of traffic parameters is contingent upon the effective mathematical modeling of the underlying stochastic processes that give rise to the signals from the traffic sensors on the streets. In fact, the same detector statistics can be mathematically described in many ways, each of which gives rise to a different estimator. The basic thrust of this research is to find that representation that gives rise to the "best" estimator in the practical sense.

6. Acknowledgement

The computer time for this project was supported in full through the facilities of the Computer Science Center of the University of Maryland.

References

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

[2] D. L. Snyder, Random Point Processes, John Wiley and Sons, New York, 1975.

[3] D. C. Gazis, Traffic Science, John Wiley and Sons, New York, 1974.

[4] G. F. Newell, "Mathematical Models of Freely Flowing Highway Traffic", Operations Res. 3, pp. 176-186.

[5] E. C. Carter and S. P. Palaniswamy, "Study of Traffic Flow on a Restricted Facility", Interim Report: Phase I, Dept. of Civil Engineering, University of Maryland, College Park, Md., June 1973.

[6] M. C. Dunne, R. W. Rothery and R. B. Potts, "A Discrete Markov Model of Vehicular Traffic", Transport. Sci. 2, No. 3, p. 167-183, 1968.

[7] R. D. Smallwood and E. J. Sondik, "The Optimal Control of Partially Observable Markov Processes Over a Finite Horizon", Oper. Res., Vol. 21, No. 5, pp. 1071-1088, 1973.

[8] H. J. Kushner, Introduction to Stochastic Control, Holt, Rhinehart and Winston Inc., 1971.

[9] R. Boel, P. Varaiya and E. Wong, "Martingales on Jump Processes, Part I: Representation Results; Part II: Applications", SIAM J. Control, Vol. 13, pp. 999-1061, 1975.

[10] A. Segall, M. H. Davis and T. Kailath, "Nonlinear Filtering with Counting Observations", IEEE Trans. on Information Theory, IT-21, pp. 143-149, 1975.

[11] A. Segall and T. Kailath, "The Modeling of Randomly Modulated Jump Processes", IEEE Trans. on Information Theory, IT-21, pp. 135-143, 1975.

[12] M.H. Davis, "The Representation of Martingales of Jump Processes", SIAM J. Control and Optimization, Vol. 14, No. 4, pp. 623-638, July 1976.

[13] C. B. Wan and M. H. A. Davis, "The General Point Process Disorder Problem", preprint.