TIME-DELAY ESTIMATION FOR COMPOUND POINT-PROCESSES USING HIDDEN MARKOV MODELS

Jens Wohlers, Alfred Mertins, and Norbert Fliege

Hamburg University of Technology, Telecommunications Dept., Eissendorfer Str. 40, D-21071 Hamburg, Germany, e-mail: J.Wohlers@tu-harburg.d400.de

ABSTRACT

In this paper a new time-delay estimation algorithm for compound point-processes is presented. Compound point-processes, a generalization of temporal point-processes, describe processes with discrete events, where each occurrence time is associated with certain features. It is shown that, although the events are not observable, the time delays from events at one location to the same events at a second location can be estimated using hidden Markov models based on the associated features. We demonstrate the performance of this time-delay estimation algorithm with an application to the estimation of section-related traffic data in road traffic monitoring and control systems.

1. INTRODUCTION

While the time-delay estimation for continuous processes and the estimation of time-of-arrival for determined events or signals have been studied intensively in the past, solutions for the time-delay estimation for compound point-processes give only estimations for the average time-delay of point clusters between two locations using correlation methods [1, 2].

We propose a new solution to the estimation problem which is based on the re-identification of single events at the output location with the aid of the feature vectors from the input location. In general, both, input and output feature vectors, are disturbed by noise.

As we will show, a hidden Markov model [3]–[5] can be used for the re-identification. In the model, the feature vectors of the events are combined to a parametric random process similar to speech recognition problems [6], and the statistical properties of the mixture of the unobservable events between input and output location are taken into account. The efficient determination of the most likely sequence of events is based on the Viterbi algorithm.

The algorithm is implemented to estimate section-related traffic data in traffic monitoring and control systems on freeways. Section-related traffic data including the travel time of vehicles in a road section improves the estimation of traffic states and the incident detection in traffic control systems. The vehicles on the road can be regarded as points of a point process and the traffic flow process can be modeled as a compound point-process. The estimation is based on vehicle signals measured with inductive loops. It will be shown via simulations based on measured traffic data that the new approach gives a good estimation of individual and average travel times of vehicles.

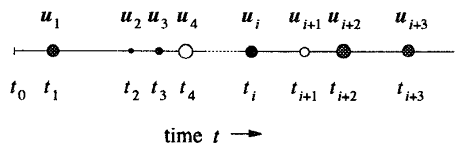

Figure 1: Compound point-process in time

Compound point-processes and the estimation problem are described in Section 2. In Section 3, hidden Markov models are introduced. A hidden Markov model for the re-identification of events is presented in Section 4. The choice of the model parameters and the derivation of the model probability distributions will be discussed. The application of the algorithm is explained in Section 5. Finally, simulation results will be shown.

2. PROBLEM FORMULATION

A random point process is a mathematical model for a physical phenomenon characterized by highly localized events distributed randomly in a continuum. Each event is represented in the model by an idealized point to be conceived of as identifying the position of the event. The space of the process is usually a semi-infinite real line representing time, a subset of Euclidean space representing a spatial region, or a combination of these.

Compound point-processes are obtained from temporal point processes by associating an auxiliary random variable, called a mark, with each point occurrence. Each occurrence time t_i is associated with a mark u_i having values in a specified space $\mathcal{U}$. A marked point-process is called a compound Poisson-process, if the marks are independent from the occurrence-time sequence, and if the occurrence-time sequence is an inhomogeneous Poisson-process [7, 8]. Figure 1 shows a marked point-process and its notation.

We consider a point process with discrete events in time and space as shown in Figure 2. The points are in motion with different velocities under the restriction $x_2(t_2) > x_1(t_1)$ with $t_2 > t_1$. Each point only occurs once at the locations x_1 and x_2.

The point processes at the locations x_1 and x_2 can be regarded as temporal Poisson-processes. Each occurrence time of an event $t_{i,1}$ and $t_{i,2}$ at x_1 and x_2 is associated with

0-7803-3192-3/96 $5.00©1996 IEEE

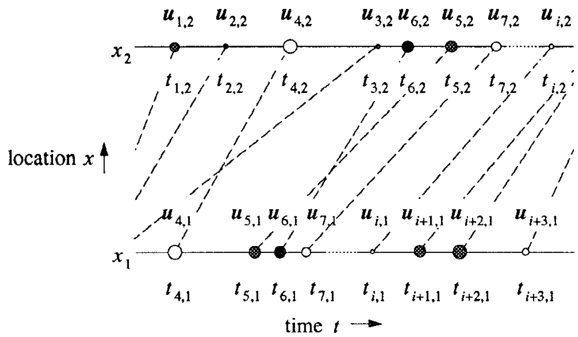

Figure 2: Compound point-process in time and space

a mark $u_{i,1}$ and $u_{i,2}$, respectively. In our approach the marks are assumed to be feature vectors of the events detected at location x_1 or x_2 in a region of a finite-dimensional Euclidean space. The resulting processes are compound Poisson-processes.

The problem is to estimate the time delay of one point i with occurrence time $t_{i,1}$ at location x_1 and $t_{i,2}$ at location x_2:

$$\tau_i = t_{i,2} - t_{i,1}. \tag{1}$$

Since the events are not observable, the feature vectors at the input and the output location have to be combined to a parametric random process to assign the feature vector of an event at x_2 to a feature vector of the same event at x_1. The assignment can be realized using hidden Markov models. After the assignment, the individual time delay of the event can be determined.

3. HIDDEN MARKOV MODELS

A hidden Markov model is a doubly stochastic process with an underlying stochastic process that is not directly observable, but can be observed through another set of stochastic processes that produce the sequence of observed symbols. The symbols can be countably or continuously distributed, they can be scalars or vectors.

The hidden Markov model is characterized by the number N of states in the model. The individual states are labeled as $\{1, 2, \ldots, N\}$ and the state at time t is labeled as q_t. Based upon the state-transition probability distribution $\mathbf{A} = \{a_{ij}\}$, the new state q_{t+1} at $t+1$ is entered, where

$$a_{ij} = P(q_{t+1} = j | q_t = i), \quad 1 \le i, j \le N. \tag{2}$$

With the initial state distribution $\boldsymbol{\pi} = \{\pi_i\}$, $\pi_i = P(q_1 = i)$, $1 \le i \le N$, the probability of a state sequence $\mathbf{q} = (q_1\, q_2\, \ldots\, q_N)$ can be written as

$$P(\mathbf{q}) = \pi_{q_1} \prod_{t=2}^{T} a_{q_t - 1 a_{q_t}}, \tag{3}$$

where T is the length of the observation sequence $\mathbf{O} = (\mathbf{o}_1\, \mathbf{o}_2\, \ldots\, \mathbf{o}_T)$.

After each transition, an observation output symbol is produced according to a probability distribution which depends on the current state. The observation symbol probability distribution $\mathbf{B} = \{b_j(k)\}$, in which

$$b_j(k) = P(\mathbf{o}_t = \mathbf{v}_k | q_t = j), \quad 1 \le k \le M, \tag{4}$$

defines the symbol distribution in state j, $j = 1, 2, \ldots, N$. $\mathbf{v} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M\}$ is the set of M possible discrete symbol observations. For observations of continuous symbols or continuous vectors, the probability density $b_j(k)$ is replaced by the continuous density, $b_j(\mathbf{o})$, $1 \le j \le N$. A special form for $b_j(\mathbf{o})$ is the Gaussian M-component mixture density

$$b_j(\mathbf{o}) = \sum_{k=1}^{M} c_{jk} \mathcal{N}[\mathbf{o}, \boldsymbol{\mu}_{jk}, \mathbf{R}_{jk}], \tag{5}$$

where c_{jk} is the mixture weight, $\mathcal{N}$ is the normal density and $\boldsymbol{\mu}_{jk}$ and $\mathbf{R}_{jk}$ are the mean vector and covariance matrix associated with state j and mixture k.

It can be seen that a complete specification of an HMM requires the specification of two model parameters, N and M, the specification of observation symbols $\mathbf{v}$, and the specification of the three sets of probability densities $\mathbf{A}$, $\mathbf{B}$ and $\boldsymbol{\pi}$. To indicate the complete parameter set of the model, the compact notation

$$\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}) \tag{6}$$

is used [3].

The probability of the observation sequence $\mathbf{O}$ given the state sequence $\mathbf{q}$ is

$$P(\mathbf{O}|\mathbf{q}) = \prod_{t=1}^{T} P(\mathbf{o}_t | q_t) = b_{q_1}(\mathbf{o}_1) \cdot b_{q_2}(\mathbf{o}_2) \ldots b_{q_T}(\mathbf{o}_T). \tag{7}$$

The joint probability of $\mathbf{O}$ and $\mathbf{q}$ is

$$P(\mathbf{O}, \mathbf{q}) = P(\mathbf{O}|\mathbf{q}) \cdot P(\mathbf{q}). \tag{8}$$

Given the observation sequence $\mathbf{O}$ and the model λ, the probability of the observation sequence $\mathbf{O}$ is obtained by summing the joint probability over all possible state sequences $\mathbf{q}$:

$$\begin{aligned} P(\mathbf{O}|\lambda) &= \sum_{\text{all } \mathbf{q}} P(\mathbf{O}|\mathbf{q}, \lambda) P(\mathbf{q}|\lambda) \\ &= \sum_{q_1, \ldots, q_T} \pi_{q_1} b_{q_1}(\mathbf{o}_1) a_{q_1 q_2} b_{q_2}(\mathbf{o}_2) \ldots a_{q_{T-1} q_T} b_{q_T}(\mathbf{o}_T). \end{aligned} \tag{9}$$

The model parameters $\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$ that maximize the probability $P(\mathbf{O}|\lambda)$ can be adjusted with a set of training samples, e.g. by an iterative procedure called the Baum-Welch method [9].

Given an observation sequence $\mathbf{O} = (\mathbf{o}_1 \ldots \mathbf{o}_T)$, one has to find the optimal state sequence $\mathbf{q}^* = (q_1^* \ldots q_T^*)$, that is, to maximize the probability $P(\mathbf{q}|\mathbf{O})$. A formal technique for finding this single best state sequence is the Viterbi algorithm [10].

4. PROBLEM SOLUTION

The estimation of the time delay τ_i of an event i with occurrence times $t_{i,k}$ at locations x_k, $k = 1, 2$, is based on the associated feature vectors $\boldsymbol{u}_{i,k}$. The feature vectors of a point group of size T can be interpreted as the observation vectors $\mathbf{O} = (\mathbf{o}_1\, \mathbf{o}_2 \ldots \mathbf{o}_T)$ of a hidden Markov model. The hidden state sequence $\mathbf{q} = (q_1\, q_2 \ldots q_T)$ of the model corresponds to the unknown positions of the points at x_1. The position at location x_1 of an event occurring at time $t_{i,2}$ at x_2 represents the actual state q_t of the model.

Given the positions at x_1, each event can be classified and the time delay of the event can be determined. The actual number of points between locations x_1 and x_2 gives the number N of states of the hidden Markov model.

The state-transition probability distribution for the transition from state i at time t to state j at time $t+1$, $a_{ij} = P(q_{t+1} = j | q_t = i)$, $1 \leq i, j \leq N$, depends on the mix-up of the points between x_1 and x_2. This leads to a non-ergodic model with forbidden state transitions. Since multiple assigns of one event are forbidden, the transition from state i to state $j = i$ is impossible, i.e. $a_{ij} = 0$ for $i = j$.

Since the statistical dependence of the point positions have to be independent from the detected point at time t, the state-transition matrix $\mathbf{A} = \{a_{ij}\}$ results in a band matrix with zeros on the main diagonal.

The continuous observation probability density function in state j, $b_j(\mathbf{o}) = p(\mathbf{o}|q_t = j)$, is based on the comparison of the feature vectors $\boldsymbol{u}_{i,2}$, $1 \leq i \leq T$, with the feature vectors $\boldsymbol{u}_{j,1}$, $1 \leq j \leq N$.

Given the hidden Markov model for the re-identification and the observation sequence of the feature vectors $\boldsymbol{u}_{1,2} \ldots \boldsymbol{u}_{T,2}$ from location x_2, one has to find the most likely point group of the N points associated with the feature vectors $\boldsymbol{u}_{1,1} \ldots \boldsymbol{u}_{N,1}$ from location x_1, that is, to find the optimal sequence $\mathbf{q}^* = (q_1^* \ldots q_T^*)$ for a given observation sequence $\mathbf{O} = (\mathbf{o}_1 \ldots \mathbf{o}_T)$.

After re-identification, the individual time delays of the events can be obtained.

5. APPLICATION

Due to optimal use of existing road networks an optimization of traffic flow with traffic monitoring and control systems is essential. The desired information for these systems can be extracted from model based traffic state estimation procedures. Present traffic data acquisition systems are usually based on local traffic data like the speed or flow rate of passing vehicles. An improvement of the traffic state estimation on freeways can be obtained with section-related traffic data including the travel time of vehicles in a road section.

The vehicles on a freeway can be represented as points and, in case of free traffic flow, the motion of the vehicles on the road can be modeled as a Poisson process [8, 11].

A block-diagram of the section-related traffic data estimation is shown in Figure 3. The estimation is based on vehicle signals measured with inductive loops at the entry and the exit of a freeway section, so the resulting process is a compound Poisson-process.

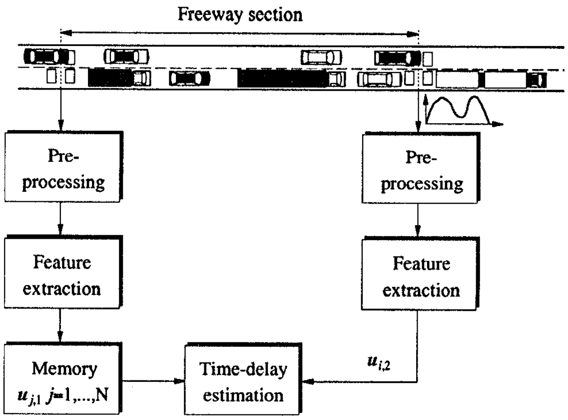

Figure 3: Section-related traffic data estimation

A hidden Markov model for the section-related traffic data estimation is characterized by the following elements:

The position of the vehicles entering the road section is described by the state sequence $Q = \{q_1, q_2, \ldots, q_T\}$.

The position of the vehicle leaving the road section is denoted as the discrete observation time t.

The unknown position of the vehicle t at the entry of the road section is denoted as q_t.

The actual number of vehicles in the road section is the number of states N in the model. The number of states depends on the observation time t.

The random process of the feature vectors of the vehicles from the second measurement point can be interpreted as an observation sequence $\mathbf{O} = \{\mathbf{o}_1, \mathbf{o}_2, \ldots, \mathbf{o}_T\}$.

The size of the vehicle group leaving the section corresponds to the number of objects T of the hidden Markov model.

The state-transition probability distribution for the transition from state i at time t to state j at time $t+1$, a_{ij} depends on the mixture of the vehicles in the road section. Figure 4 shows the state-transition probability distribution a_{ij} estimated with a sample of 8000 vehicles in a 2.5 km road section.

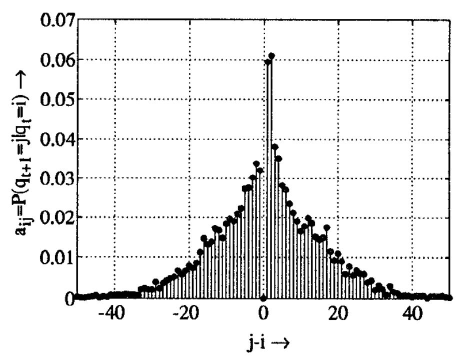

Figure 4: Estimated state-transition probability distribution a_{ij} for the transitions from state i at time t to state j at time $t+1$

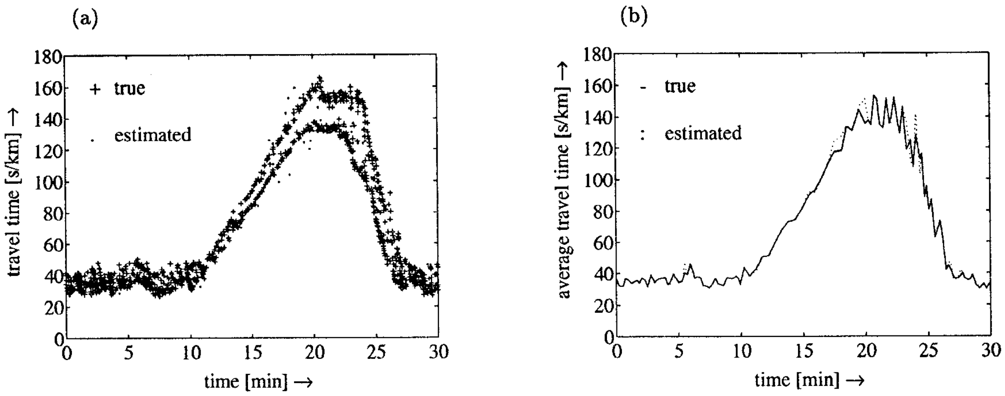

Figure 5: True and estimated individual travel time (a) and average travel time (b) in a 30 minute simulation of a 2.5 km road section with included incident

The continuous observation probability density function in state j, $b_j(\mathbf{o}) = p(\mathbf{o}|q_t = j)$, is based on the comparison of the feature vectors $\boldsymbol{u}_{i,2}$ representing the vehicles from the exit of the road section with the feature vectors $\boldsymbol{u}_{j,1}$ representing the vehicles from the entry of the road section.

The most likely vehicle group of the N vehicles is estimated with the Viterbi algorithm.

6. SIMULATION RESULTS

The simulation results have been obtained using measured vehicle signals from the German freeway A7 Hamburg-Flensburg. A microscopic computer simulation model for freeway corridors was used to test the algorithm for different kinds of road sections and traffic states. Especially the estimation of section-related data in critical traffic states for incident detection has been investigated.[1] Figure 5 shows both the true and the estimated individual and average travel time of vehicles in a 2.5 km two-lane freeway section for a 30 minute simulation. After 10 minutes one lane of the freeway section was closed for the time interval of 10 minutes. Because of a high traffic flow rate the travel time increases spontaneously. After clearing the incident the travel time decreases.

7. CONCLUSION

We have proposed a new approach to the time-delay estimation for compound point-processes. Based on a hidden Markov model and the Viterbi algorithm a good estimation of the time delay is possible. The performance of the algorithm was demonstrated by applying it to the estimation of vehicle travel times in road traffic monitoring and control systems.

[1] The traffic flow process can only be regarded as Poisson process for low flow conditions if the movements of different vehicles are independent. However, the estimation of section-related traffic data for higher flow conditions is possible.

8. REFERENCES

[1] E. Pfannerstill: "Measuring section-related traffic data by correlation methods: A new approach to traffic monitoring and control," in *Proc. 2nd International Conference on Applications of Advanced Technologies in Transportation Engineering*, Minneapolis, MN, Aug. 1991.

[2] R. Kühne and S. Immes: "Freeway control systems using section-related traffic variable detection," in *Proc. Pacific Rim TransTech Conference*, Seattle, WA, pp. 56–62, Jul. 1993.

[3] L. R. Rabiner and B. H. Juang: "An introduction to hidden Markov models," *IEEE ASSP Magazine*, vol. 3, pp. 4–16, Jan. 1986.

[4] A. B. Poritz: "Hidden Markov models: A guided tour," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, New York, pp. 7–12, 1988.

[5] R. J. Elliott, L. Aggoun, and J. B. Moore: *Hidden Markov Models.* New York: Springer, 1995.

[6] L. R. Rabiner and B.-H. Juang: *Fundamentals of Speech Recognition.* Englewood Cliffs, NJ: Prentice Hall, 1993.

[7] D. L. Snyder and M. I. Miller: *Random Point Processes in Time and Space.* New York: Springer, 1991.

[8] J. F. Kingman: *Poisson Processes.* Oxford: Clarendon Press, 1993.

[9] L. E. Baum: "An inequality and associated maximization technique in statistical estimation for probabilistic functions of a Markov process," *Inequalities*, vol. 3, no. 1, pp. 1–8, 1972.

[10] G. D. Forney: "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.

[11] A. D. May: *Traffic Flow Fundamentals.* Englewood Cliffs, NJ: Prentice Hall, 1990.