European Journal of Operational Research 198 (2009) 275–286

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Computational Intelligence and Information Management

Mathematical justification of a heuristic for statistical correlation of real-life time series

Evgeny Agafonov [a], Andrzej Bargiela [b,*], Edmund Burke [b], Evtim Peytchev [c]

[a] INFOHUB Ltd., 278-290 Huntingdon Street, Nottingham NG1 3LY, UK
[b] University of Nottingham, School of Computer Science and IT, Jubilee Campus, Nottingham NG8 4BB, UK
[c] Nottingham Trent University, School of Computing and Informatics, Clifton Campus, Nottingham NG11 8NS, UK

ARTICLE INFO

Article history:
Received 20 November 2007
Accepted 30 June 2008
Available online 18 July 2008

Keywords:
Programming: algorithms, heuristic
Statistics: correlation, time series
Transportation: models, traffic

ABSTRACT

Many of the analyses of time series that arise in real-life situations require the adoption of various simplifying assumptions so as to cope with the complexity of the phenomena under consideration. Whilst accepting that these simplifications lead to heuristics providing less accurate processing of information compared to the solution of analytical equations, the intelligent choice of the simplifications coupled with the empirical verification of the resulting heuristic has proven itself to be a powerful systems modelling paradigm. In this study, we look at the theoretical underpinning of a successful heuristic for estimation of urban travel times from lane occupancy measurements. We show that by interpreting time series as statistical processes with a known distribution it is possible to estimate travel time as a limit value of an appropriately defined statistical process. The proof of the theorem asserting the above, supports the conclusion that it is possible to design a heuristic that eliminates the adverse effect of spurious readings without loosing temporal resolution of data (as implied by the standard method of data averaging). The original contribution of the paper concerning the link between the analytical modelling and the design of heuristics is general and relevant to a broad spectrum of applications.

1. Introduction

The analysis of complex systems poses two interrelated challenges: the challenge of coping with the complexity of data and the challenge of coping with the complexity of the data processing algorithms. Although this general insight has motivated the development of scientific disciplines for centuries, the formalisation of it, in the context of information theory, is much more recent. In his pioneering contribution, Zadeh introduced the fuzzy set representation of information and the concept of information granularity (Zadeh, 1979). Subsequently, he postulated the development of a computational framework for processing information granules and has called this framework Granular Computing (Zadeh, 1997, 1999). The vigorous research activity that ensued (Bargiela and Pedrycz, 2003; Bargiela and Pedrycz, 2005; Bargiela et al., 2006; Pedrycz and Bargiela, 2002; Reformat et al., 2004; Yao, 2001), addressed both the abstraction (granulation) of data and the development of methods for processing information granules. Insights gained from this research point to the convergence of methodological approaches between our development of granular algorithms and our development of heuristics, as methods designed to solve complex problems in a pragmatic way (Burke et al., 2003, 2004, 2006a,b,c; Agafonov et al., 2006). To highlight this convergence we investigate, in this study, the theoretical basis of our recent successful heuristic for the estimation of urban travel times.

The problem of estimating urban travel times from the lane occupancy readings has been investigated by various researchers and has been found to be representative of a significant class of real-life time series analysis problems. Taking our recently published algorithm (Agafonov et al., 2006) as a point of departure, we discuss here a mathematical justification of the proposed heuristic. However, to put this discussion in the right context we recall briefly the early attempts to estimate travel times. These were based on general notions of traffic flow and lane occupancy to estimate speed and consequently, travel time along a link. The relationship used was

$$\text{speed} = \frac{\text{flow}}{\text{occupancy}} * \text{effective\_car\_length}. \quad (1)$$

* Corresponding author. Tel.: +44 115 8467279.
E-mail address: abb@cs.nott.ac.uk (A. Bargiela).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2008.06.040

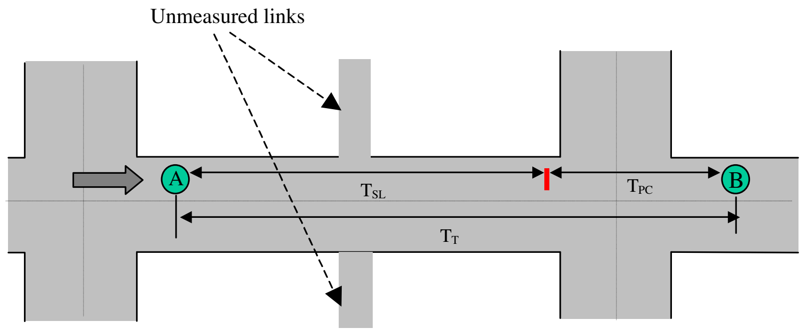

Fig. 1. Urban link model with upstream measurement (A), downstream measurement (B) and two unmeasured links.

In Hall and Persaud (1989), Pushkar et al. (1994), the authors investigated the relationship of various factors in Eq. (1) and showed that the accuracy of this equation depends on such things like physical location and weather conditions. They also suggested that the relationship is prone to a systematic bias with respect to occupancy. Attempts to quantify such errors have been made by various researchers (Hall and Persaud, 1989; Pursula, 1995) and have prompted alternative approaches in which stochastic nature of the measurements was taken explicitly into account (Dailey, 1997, 1999; Pursula, 1998; Bargiela et al., 2006).

The approaches based on Kalman filtering (Dailey, 1997, 1999), implied the need for the identification of several parameters which, in themselves, were not trivial to calculate. The problems associated with this approach seem to have been rooted in the implicit rather than explicit identification of the sources of errors in the fundamental traffic flow relationship (1). Refinements proposed by Coifman (2001) have produced some improvement but did not resolve the root problem.

In a more recent work, Petty et al. (1998) has proposed a more general model of traffic flow. This model has some similarity to the Dailey's cross-correlation approach described in (Dailey, 1993). The model is using stochastic variables and suggests practical procedures that can potentially lead to accurate travel time estimates. The underpinning assumption in this model is that travel times can be considered to

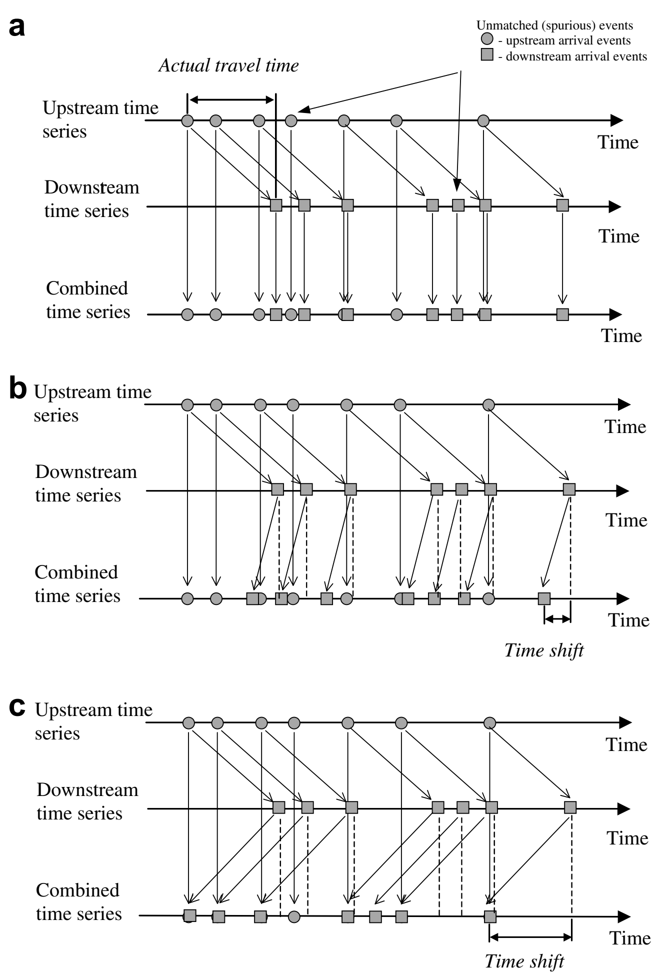

Fig. 2. Heuristic for estimating travel time by varying "Time shift" and evaluating the resulting matching of the two time series.

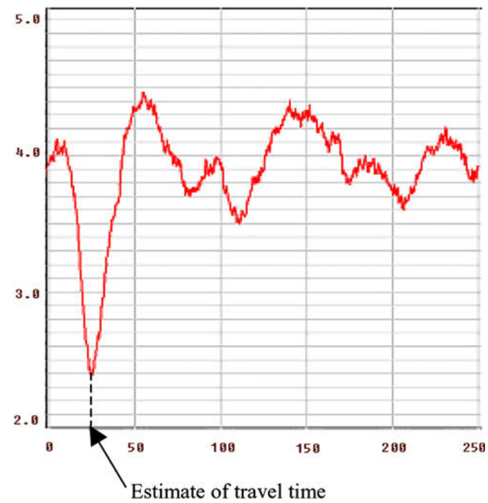

Fig. 3. Indicator of the quality of matching evaluated for a typical of real-life data set.

be drawn from the same probability distribution. This enables an estimation of probability distribution from the cumulative upstream and downstream arrival processes. However, there are several difficulties associated with the above methodology and these are widely recognised by transportation researchers. Firstly, the choice of the time windows, in which the upstream and downstream processes are considered, is not trivial. The results of empirical study (Agafonov, 2003), intended to inform such a choice, indicate that adaptive choice of parameters (depending on the traffic conditions) is necessary. Secondly, it is difficult to find "stationary" periods in traffic behaviour from which to draw measurement data. It is clear that the larger, such a period is the better, is the precision of the proposed model. However, there does not seem to be an easy way of identifying such a period other than monitoring transitions in traffic regime through some measurements of traffic density. Thirdly, the choice of the appropriate level of data aggregation represents a significant challenge. Although it is possible to assess the effects of different levels of aggregation retrospectively, what is needed is the *a-priori* information about the right aggregation level.

2. The heuristic

In (Agafonov et al., 2006), we proposed a heuristic designed to overcome the shortcomings of previous travel time estimation techniques. The heuristic is based on the intuitive notion of maximising the overlap between the upstream and downstream arrival processes by iterative refinement of the time shift of the downstream arrival process while discarding individual items in the respective time series that are deemed to represent traffic originating from- or destined to- the unmeasured links (as indicated in Fig. 1).

The iterative refinement of time shift is represented graphically in three stages in Fig. 2. Fig. 2a represents the initial state with no time shift applied to the downstream arrival process. Fig. 2b represents an intermediate state where the time shift is smaller than the actual travel time between the two measurement points and Fig. 2c represents the time shift which is equal to the travel time.

While the match of the upstream and downstream time series can never be perfect, due to the variation of *actual travel times* of individual vehicles, the unmatched (spurious) events corresponding to traffic entering/leaving unmeasured links are clearly discarded resulting in a well-defined minimum of the performance index. A typical result is illustrated in Fig. 3.

3. Mathematical justification of the heuristic

In order to explore analytical underpinning of the travel time estimation heuristic outlined in Section 2, we adopt the representation of the upstream and downstream time series as homogeneous Poisson processes. Since these two processes represent the combined traffic between the measurement points and the traffic entering/leaving through unmeasured links respectively, we can take advantage of the property of homogeneous Poisson processes implying that all three stochastic processes have the same distribution (Feller, 1968). The combination of the three processes into upstream and downstream measurements is illustrated in Fig. 4.

The observable processes $N(t)$ and $M(t)$ can be expressed as follows:

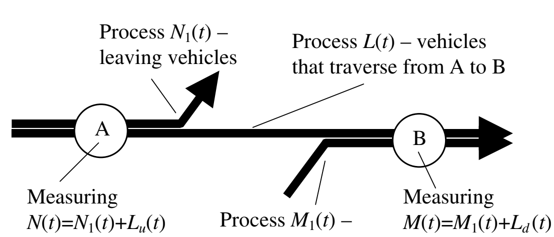

Fig. 4. Structure of the considered traffic processes.

$$N(t) = N_1(t) + L_u(t),$$
$$M(t) = M_1(t) + L_d(t). \tag{2}$$

Since $N_1(t), M_1(t), L_u(t)$ and $L_d(t)$ are Poisson processes with parameters λ_1, μ_1, α and α respectively, the processes $N(t)$ and $M(t)$ are also Poisson processes with parameters $\lambda = \lambda_1 + \alpha$ and $\mu = \mu_1 + \alpha$ respectively. The arrival times of processes $L_u(t)$ and $L_d(t)$ are subject to the conditions

$$\forall a_i \in L_u(t) \quad \exists b_j \in L_d(t): \quad b_j = a_i + T \quad \text{and also}$$
$$\forall b_k \in L_d(t) \quad \exists a_l \in L_u(t): \quad b_k = a_l + T, \tag{3}$$

where T represents the estimate of travel time and a_i, a_l are arrival times of events belonging to $L_u(t)$ and b_j, b_k are arrival times of events belonging to $L_d(t)$. It then follows that process $L_d(t)$ is equal to $L_u(t)$ shifted in time by a constant T, or $L_d(t) = L_u(t - T)$. With the above notation, the algorithmic procedure for estimating travel time T can be expressed as follows:

Algorithm 1

1.1 Combine the upstream and downstream processes as

$$R(t, \tau) = N(t) + M(t - \tau). \tag{4}$$

This is subject to the condition that if the arrival time of an upstream event equals to the arrival time of a downstream event, then the upstream arrival time is given a preference and is placed in front of the downstream arrival time. $N(t)$ and $M(t - \tau)$ are Poisson processes but the compound process $R(t)$ is obviously not a Poisson process since the processes $N(t)$ and $M(t)$ are not independent. We denote the arrival times of the compound process $R(t)$ by $r_i(\tau)$. In order to distinguish arrival times $r_i(\tau)$ by their original process ($N(t)$ or $M(t)$), let us introduce the indicator function $I(r_i(\tau))$ as

$$I(r_i(\tau)) = \begin{cases} 0 & \text{if } r_i \text{ originates from process } N(t), \\ 1 & \text{if } r_i \text{ originates from process } M(t). \end{cases} \tag{5}$$

1.2 Extract interarrival times $r_{i+1} - r_i$ according to the following procedure

Step	Action	Comment
1	$k = 1$; $i = 1$	Initialisation
2	While $I(r_i(\tau)) = 1$; $i = i + 1$	Looking for an event from the upstream process
3	if $\mathrm{I}(r_{i+1}(\tau)) = 0$; $i = i + 1$; go to 2	Skip this event if next is not from downstream process
4	$\Delta_k(\tau) = r_{i+1}(\tau) - r_i(\tau)$; $k = k + 1$; $i = i + 1$; go to 2	Save the found interarrival time at k-s position

The above procedure is infinite but, in practice, it can stop whenever either processes $N(t)$ or $M(t)$ are exhausted, or the necessary precision is reached.

1.3 Calculate the limit of the average value of interarrival times calculated in 1.2.

To formalise and justify the step 1.3 of the algorithm we express it as the following theorem.

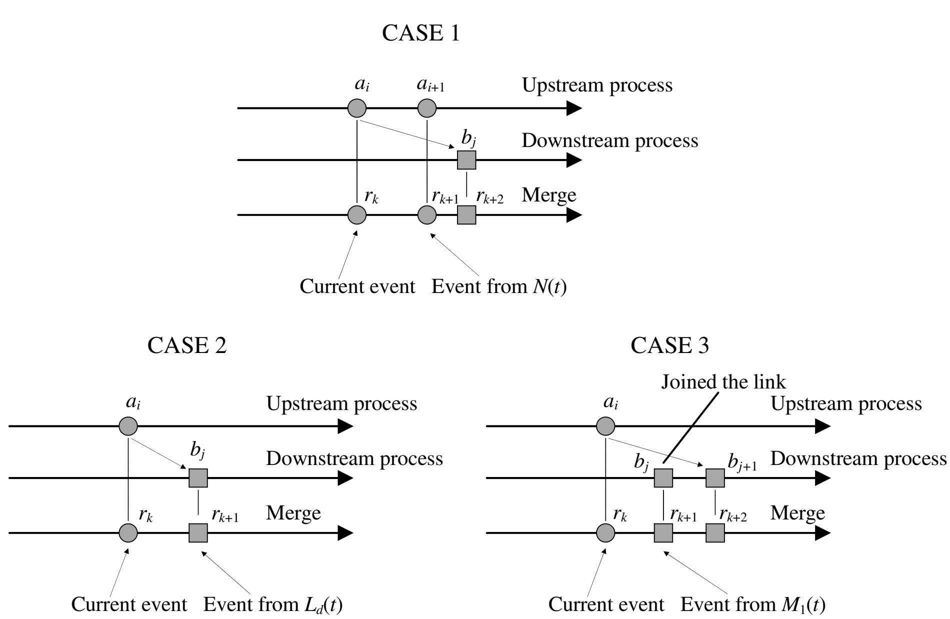

Fig. 5. Illustration of possible outcomes for the next event after currently chosen one.

Theorem 1. *Consider processes N, N_1, L_u, M, M_1, L_d defined in (2). Assuming that $\lambda > 0$ and $\mu > 0$ we consider the average of N values $\Delta_k(\tau)$ (a sample of interarrival times of process $R(t,\tau)$) evaluated as*

$$F_N(\tau) = \frac{1}{N}\sum_{k=1}^{N}\Delta_k(\tau). \tag{6}$$

The travel time estimate T can be found as

$$T = \arg\min F^*(\tau) \tag{7}$$

with $F^*(\tau)$ representing a finite limit of $F_N(\tau)$ for N increasing to infinity, i.e.

$$F^*(\tau) = \lim_{N\to\infty} F_N(\tau) \tag{8}$$

Proof. Assume that $T - \tau \geqslant 0$. In order to find the analytical form of function $F^*(\tau)$, we need to investigate the properties of interarrival times extracted by Algorithm 1. We observe that the shifting of the downstream process by a constant time τ in step 1.1 does not change the statistical properties of the downstream process. The extraction of interarrival times in step 1.2 implies identifying a corresponding upstream event. Since processes $L_u(t)$ and $N_1(t)$, forming the observed upstream process $N(t)$, are independent it follows that the probability of picking an event corresponding to part $L_u(t)$ of $N(t)$ is $\frac{\alpha}{\alpha+\lambda_1} = \frac{\alpha}{\lambda}$ and the probability of picking an event corresponding to part $N_1(t)$ is $\frac{\lambda_1}{\alpha+\lambda_1} = \frac{\lambda_1}{\lambda}$. Once the algorithm has chosen the next event from the upstream process, it checks if the following event of the merged process $R(t)$ corresponds to the downstream process. There are three possibilities for the next event of process $R(t)$:

1. Next event originates from upstream.
2. Next event originates from downstream from its $L_d(t)$ part.
3. Next event originates from downstream from its $M_1(t)$ part.

Fig. 5 illustrates the above outcomes. Let us introduce an extended indicator function in order to distinguish the above combinations.

$$I(r_i(\tau)) = \begin{cases} 0 & \text{if } r_i \text{ originates from process } N_1(t), \\ 1 & \text{if } r_i \text{ originates from process } L_u(t), \\ 2 & \text{if } r_i \text{ originates from process } M_1(t), \\ 3 & \text{if } r_i \text{ originates from process } L_d(t). \end{cases} \tag{9}$$

Thus, the algorithm basically deals with 6 types of pairs of events. Let us represent the type of a pair $(r_i(\tau), r_{i+1}(\tau))$ of consequent events of process R(t) by the pair of their indicator functions $(I(r_i(\tau)), I(r_{i+1}(\tau)))$. Then the types of pairs the algorithm deals with are given in Table 1.

The purpose of differentiating the types of the pairs is that the corresponding time differences form random variables of different types and in order to derive the properties of the sum of them, they need to be investigated separately. Pairs of type 3 and type 6 are of no interest and are reported here only for completeness.

Let us denote by Q_1 the choice of an event from $N_1(t)$ by and by Q_2 the choice of an event from $L_u(t)$. It has already been noted that

$$P\{Q_1\} = \frac{\lambda_1}{\lambda} \quad \text{and} \quad P\{Q_2\} = \frac{\alpha}{\lambda}. \tag{10}$$

Assuming that event Q_1 has taken place we can evaluate probabilities for the algorithm to encounter pairs of types 1, 2 and 3. Fig. 6 helps to visualise the evaluation. The circles on the time axis represent events of the merged process and the numbers within the circles show the

Table 1
Types of pairs classified by the values of indicator functions

Type number	Pair of indicator functions	Comment
1	(0,2)	First event is from $N_1(t)$ and second event is from $M_1(t)$
2	(0,3)	First event is from $N_1(t)$ and second event is from $L_d(t)$
3	(0,0), (0,1)	First event is from $N_1(t)$ and second event is from $N(t)$
4	(1,2)	First event is from $L_u(t)$ and second event is from $M_1(t)$
5	(1,3)	First event is from $L_u(t)$ and second event is from $L_d(t)$
6	(1,0), (1,1)	First event is from $L_u(t)$ and second event is from $N(t)$

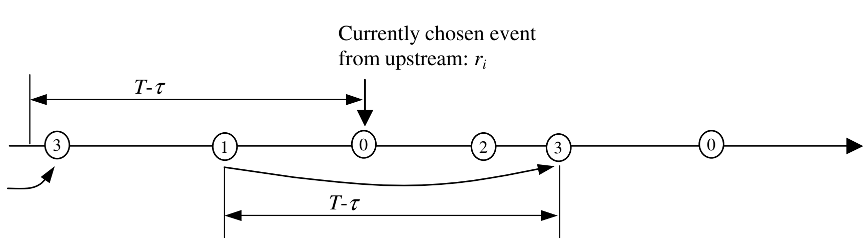

Fig. 6. Illustration of the resulting process and procedure of choosing the pair of events.

value of indicator function (the source process) of the event. The arrows emphasise the dependency between co-related events from $L_u(t)$ and $L_d(t)$. The distance between corresponding events from $L_u(t)$ and $L_d(t)$ is $T-\tau$ due to the shift τ introduced by the algorithm.

With the above assumption, the probability that the event following the currently chosen one will be from $M_1(t)$ (indicator function equals to 2) can be calculated as

$$P\{I(r_{i+1})=2|Q_1\}=\int_0^{T-\tau}\mu_1 e^{-\mu_1 t}\int_t^{+\infty}(\alpha+\alpha+\lambda_1)e^{-(\alpha+\alpha+\lambda_1)x}\,dx\,dt++e^{-\alpha(T-\tau)}\int_{T-\tau}^{+\infty}\mu_1 e^{-\mu_1 t}\int_t^{+\infty}(\alpha+\lambda_1)e^{-(\alpha+\lambda)x}dxdt. \tag{11}$$

The first component of the above sum represents the fact that on the interval $[0,T\text{-}\tau]$ four independent exponentially distributed random variables are active. The process $L_d(t)$ has effect at the distance from r_i that does not exceed $T-\tau$, since its generating event from $L_u(t)$ has to be before r_i and not further from r_i than $T-\tau$. The probability of the event that an event originated from $L_u(t)$ lies within the interval $[r_i-(T-\tau),r_i]$ equals $1-e^{-\alpha(T-\tau)}$ (following from properties of exponentially distributed random variable) and thus the second component of the sum (11) represent the case when this does not happen. If $L_u(t)$ does not lie within the interval $[r_i-(T-\tau),r_i]$, there is no process $L_d(t)$ acting and the second component of the sum represents this fact by omitting one alpha-parameter (that corresponds to process $L_d(t)$). Bearing in mind that $\lambda=\alpha+\lambda_1$ and $\mu=\alpha+\mu_1$, Eq. (11) can be evaluated as

$$\begin{aligned}P\{I(r_{i+1})=3|Q_1\}&=\int_0^{T-\tau}\mu_1 e^{-(\mu+\lambda)t}dt+e^{-\alpha(T-\tau)}\int_{T-\tau}^{+\infty}\mu_1 e^{-(\lambda+\mu_1)t}dt=\frac{\mu_1}{\lambda+\mu}(1-e^{-(\lambda+\mu)(T-\tau)})+e^{-\alpha(T-\tau)}\frac{\mu_1}{\lambda+\mu_1}e^{-(\lambda+\mu_1)(T-\tau)}\\&=\frac{\mu_1}{\lambda+\mu}\left(1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}\right).\end{aligned} \tag{12}$$

Using the same arguments, the following probability for pairs of type 3 can be calculated

$$\begin{aligned}P\{I(r_{i+1})=0\quad\text{or}\quad I(r_{i+1})=1|Q_1\}&=\int_0^{T-\tau}\lambda e^{-\lambda t}\int_t^{+\infty}(\alpha+\mu_1)e^{-(\alpha+\mu_1)x}\,dx\,dt+e^{-\alpha(T-\tau)}\int_{T-\tau}^{+\infty}\lambda e^{-\lambda t}\int_t^{+\infty}\mu_1 e^{-\mu_1 x}\,dx\,dt\\&=\frac{\lambda}{\lambda+\mu}\left(1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}\right).\end{aligned} \tag{13}$$

Table 2
Probabilities of occurrence of different types of pairs

Type number	Pair of indicators $(I(r_i),I(r_{i+1}))$	Full probability
1	(0,2)	$\frac{\lambda_1}{\lambda}\frac{\mu_1}{\lambda+\mu}\left(1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}\right)$
2	(0,3)	$\frac{\lambda_1}{\lambda}\frac{\alpha}{\lambda+\mu}(1-e^{-(\lambda+\mu)(T-\tau)})$
3	(0,0), (0,1)	$\frac{\lambda_1}{\lambda}\frac{\lambda}{\lambda+\mu}\left(1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}\right)$
4	(1,2)	$\frac{\alpha}{\lambda}\frac{\mu_1}{\lambda+\mu}(1-e^{-(\lambda+\mu)(T-\tau)})$
5	(1,3)	Both: $\frac{\alpha}{\lambda}\left(\frac{\alpha}{\lambda+\mu}+\frac{\lambda+\mu_1}{\lambda+\mu}e^{-(\lambda+\mu)(T-\tau)}\right)$
5a		Independent: $\frac{\alpha}{\lambda}\frac{\alpha}{\lambda+\mu}(1-e^{-(\lambda+\mu)(T-\tau)})$
5b		Dependent: $\frac{\alpha}{\lambda}e^{-(\lambda+\mu)(T-\tau)}$
6	(1,0), (1,1)	$\frac{\alpha}{\lambda}\frac{\lambda}{\lambda+\mu}(1-e^{-(\lambda+\mu_1)(T-\tau)})$

The probability of an event of choosing a pair of type 2 is different in that the effect of process $L_d(t)$ lies in the interval $[r_i, r_i + (T - \tau)]$ only, due to the previously discussed reasons. So,

$$P\{I(r_{i+1}) = 3|Q_1\} = \int_0^{T-\tau} \alpha e^{-\alpha t} \int_t^{+\infty} (\alpha + \lambda_1 + \mu_1) e^{-(\alpha+\lambda_1+\mu_1)x}\, dx\, dt = \frac{\alpha}{\lambda + \mu}\left(1 - e^{-(\lambda+\mu)(T-\tau)}\right). \tag{14}$$

Eqs. (12)–(14) form a partition of the space that consist of outcomes of the experiment of trying the type of the next event that follows the chosen r_i. It can be easily checked that

$$P\{I(r_{i+1}) = 2|Q_1\} + P\{I(r_{i+1}) = 3|Q_1\} + P\{I(r_{i+1}) = 0 \quad \text{or} \quad I(r_{i+1}) = 1|Q_1\} = 1$$

as required.

Now, assume that Q_2 has taken place while the algorithm was choosing the upstream event. Then together with the following event r_{i+1} the pair can be of types 4, 5 or 6 as given in the table. Let us calculate the probabilities of the above types.

Consider the outcome that leads to a pair of type 4. Since the current upstream even r_i originates from process $L_u(t)$, it is known that within the interval of length $T - \tau$ there is a related event originated from $L_d(t)$. Therefore r_{i+1} are of the $M_1(t)$ process if and only if they happen before the expected event from $L_d(t)$. It is also known that on the interval $[r_i, r_i + (T - \tau)]$ events from $L_u(t)$, $N_1(t)$ and uncorrelated $L_d(t)$ can take place. Then the probability that the outcome is of type 4 will be

$$P\{I(r_{i+1}) = 2|Q_2\} = \int_0^{T-\tau} \mu_1 e^{-\mu_1 t} \int_t^{+\infty} (\alpha + \lambda_1 + \alpha) e^{-(\alpha+\lambda_1+\alpha)x}\, dx\, dt = \int_0^{T-\tau} \mu_1 e^{-(\mu_1+\alpha+\lambda_1+\alpha)t}\, dt = \frac{\mu_1}{\lambda + \mu}\left(1 - e^{-(\lambda+\mu)(T-\tau)}\right). \tag{15}$$

Now consider the outcome that leads to a pair of type 5. In this case there are two possible outcomes. First one is when the events that form a pair are not related. In this case the difference is a random variable. Its distribution will be studied later. The second outcome is when the events that form the pair are related. Then their difference equates to $T - \tau$ is a constant provided that τ is fixed. Let us study the probabilities of the above outcomes.

The events r_i and r_{i+1} will be independent if $r_{i+1} - r_i < T - \tau$. Therefore, the probability of such outcome, denoted by P_1, is

$$P_1\{I(r_{i+1}) = 3|Q_2\} = \int_0^{T-\tau} \alpha e^{-\alpha t} \int_t^{+\infty} (\alpha + \lambda_1 + \mu_1) e^{-(\alpha+\lambda_1+\mu_1)x}\, dx\, dt = \frac{\alpha}{\lambda + \mu}\left(1 - e^{-(\lambda+\mu_1)(T-\tau)}\right). \tag{16}$$

The probability for r_i and r_{i+1} that correspond to independent events to be $r_{i+1} - r_i = T - \tau$ is zero and thus can be neglected. Therefore, if no events of processes $L_u(t)$, $M_1(t)$, $N_1(t)$, $L_d(t)$ have happened in the interval $[r_i, r_i + T - \tau)$ then r_i and r_{i+1} are dependent and the probability for this event is

$$P_2\{I(r_{i+1}) = 3|Q_2\} = e^{-(\lambda+\mu)(T-\tau)}. \tag{17}$$

Together P_1 and P_2 from (16) and (17) give the total probability for the pair to be of the type 5

$$P\{I(r_{i+1}) = 3|Q_2\} = P_1\{I(r_{i+1}) = 3|Q_2\} + P_2\{I(r_{i+1}) = 3|Q_2\} = \frac{\alpha}{\lambda + \mu}\left(1 - e^{-(\lambda+\mu_1)(T-\tau)}\right) + e^{-(\lambda+\mu)(T-\tau)} = \frac{\alpha}{\lambda + \mu} + \frac{\lambda + \mu_1}{\lambda + \mu} e^{-(\lambda+\mu)(T-\tau)}. \tag{18}$$

The last possible outcome for the pair is to be of type 6. In this case the events with arrival times r_i and r_{i+1} are both produced by the upstream and thus the algorithm will skip over them. The probability for the pair to be of type 6 equals the probability of an exponentially distributed random variable with parameter λ (upstream) to be greater than another independent exponentially distributed random variable with parameter μ (downstream). Therefore

$$P\{I(r_{i+1}) = 0 \quad \text{or} \quad I(r_{i+1}) = 1|Q_2\} = \int_0^{T-\tau} \lambda e^{-\lambda t} \int_t^{+\infty} \mu e^{-\mu x}\, dx\, dt = \frac{\lambda}{\lambda + \mu}\left(1 - e^{-(\lambda+\mu_1)(T-\tau)}\right). \tag{19}$$

All the above probabilities have been calculated conditional on the occurrence of Q_1 and Q_2. Then the probabilities for the algorithm to encounter a certain type of a pair in a step of its cycle is given as full probabilities

$$P\{Q_1, A\} = P\{Q_1\}P\{A|Q_1\} \quad \text{and} \quad P\{Q_2, A\} = P\{Q_2\}P\{A|Q_2\},$$

where A is one of the following outcomes $\{I(r_{i+1}) = 0 \text{ or } I(r_{i+1}) = 1, I(r_{i+1}) = 2, I(r_{i+1}) = 3\}$. This is summarised in Table 2. Probabilities $P\{Q_1\}$ and $P\{Q_2\}$ have been given in (10).

From the description of the algorithm it follows that it extracts only pairs of types 1, 2, 4 and 5 and sends the corresponding difference $r_{i+1} - r_i$ to the resulting series. Thus it is required to calculate the probability with which a certain type of differences goes into the resulting series. Let us denote by A_k, an event that the current type of a pair encountered by the algorithm is k. Then the difference of the pair will go to the resulting series if and only if one of A_1, A_2, A_4 or A_5 has taken place. Then among outcomes A_1, A_2, A_4 or A_5 the probability for the difference of type k to go into the resulting series is given by the conditional probability

$$P_k = P\left\{A_k \Big| \bigcup_{i\in\{1,2,4,5a,5b\}} A_i\right\} = \frac{P\{A_k\}}{P\left\{\bigcup_{i\in\{1,2,4,5a,5b\}} A_i\right\}}, \quad k = 1, 2, 4, 5a, 5b. \tag{20}$$

Since the A_k's are disjoint, (20) can be rewritten as

$$P_k = \frac{P\{A_k\}}{\sum_{i\in\{1,2,4,5a,5b\}} P\{A_i\}}, \quad k = 1, 2, 4, 5a, 5b. \tag{21}$$

It can be easily calculated that

$$\sum_{i\in\{1,2,4,5a,5b\}} P\{A_i\} = \frac{\mu}{\lambda+\mu}\left(1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}\right). \tag{22}$$

Finally, the calculated probabilities P_k are given as

$$P_1 = \frac{\lambda_1\mu_1}{\lambda\mu}; \quad P_2 = \frac{\alpha\lambda_1}{\lambda\mu}\frac{1-e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}}; \quad P_4 = \frac{\alpha\mu_1}{\lambda\mu}\frac{1-e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}};$$
$$P_{5^a} = \frac{\alpha^2}{\lambda\mu}\frac{1-e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}}; \quad P_{5^b} = \frac{\alpha(\lambda+\mu)}{\lambda\mu}\frac{e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}}.$$

Now consider Eq. (6) representing the average of the elements of the resulting series. We note that the algorithm gives four types of differences as elements of the series. Let us denote these differences by Δ^k where k is its type (as in Table 2). The algorithm does not recognise dependent and independent differences of type 5 but they will be distinguished here and thereafter as Δ^{5a} for independent and Δ^{5b} for dependent differences. Thus, the sum (6) can be expanded onto a four sub-sums that contain elements only of certain type

$$F_N(\tau) = \frac{1}{N}\sum_{k=1}^{N}\Delta_k(\tau) = \frac{1}{N}\left(\sum_{k\in\{1,2,4,5a,5b\}}\sum_{i\in\Omega_k}^{|\Omega_k|}\Delta_i^k(\tau)\right), \tag{23}$$

where Ω_k is a set of indices that correspond to differences Δ of type k, $k \in \{1,2,4,5a,5b\}$ upper index of the Δ-s shows the type of this difference and $|\Omega_k| = N_k$ is the cardinality of Ω_k. Further

$$F_N(\tau) = \frac{1}{N}\left(\sum_{k\in\{1,2,4,5a,5b\}}\left(\frac{N_k}{N_k}\sum_{i\in\Omega_k}^{N_k}\Delta_i^k(\tau)\right)\right) = \sum_{k\in\{1,2,4,5a,5b\}}\left(\frac{N_k}{N}\left[\frac{1}{N_k}\sum_{i\in\Omega_k}^{N_k}\Delta_i^k(\tau)\right]\right). \tag{24}$$

As $N\to\infty$ according to the strong law of large numbers the terms $\frac{1}{N_k}\sum_{i\in\Omega_k}^{Nk}\Delta_i^k(\tau)$ converges with probability 1 to the expectation $E[\Delta^k(\tau)]$ of the random variable Δ^k, and according to the weak law of large numbers (Bernoulli's theorem), frequencies $\frac{N_k}{N}$ converge (in probability) to the probabilities $\Pr_k$ of a certain type of differences to get into the resulting series (Feller, 1968). It is well-known that if sequence X_N converges to X with probability 1 and Y_M converges to Y in probability, then X_NY_M converges to XY in probability. Therefore $F_N(\tau)$ converges to $F^*(\tau)$ in probability and

$$F^*(\tau) = \lim_{N\to\infty}F_N(\tau) = \sum_{k\in\{1,2,4,5a,5b\}}\Pr_k E\left[\Delta_i^k(\tau)\right]. \tag{25}$$

In order to calculate the expectations of the above random variables, their properties need to be investigated.

The first type of the differences, Δ^1, is produced by r_i and r_{i+1} such that $I(r_i) = 0$ and $I(r_{i+1}) = 2$, that is r_i and r_{i+1} correspond to the events that belong to processes $N_1(t)$ and $M_1(t)$ accordingly. These two processes are independent and also independent of $L_u(t)$ and $L_d(t)$. Let us denote the random variable whose realizations are Δ^1 by ξ_1 and calculate its probability distribution function. Assume that r_i is fixed and $I(r_i) = 0$ and consider a random variable χ that represents the time from r_i to the occurrence of an event from either of the processes, that is $\chi = r_{i+1} - r_i$. Then ξ_1 can be considered as being conditional on χ such that it is produced by a pair of type 1 (or $I(r_{i+1}) = 2$ subject to $I(r_i) = 0$). This fact will be denoted as $\chi\in(0,2)$. Formally, it can be written as follows:

$$F_{\xi_1}(t) = \Pr\{\chi < t|\chi\in(0,2)\} = \frac{\Pr\{\chi < t\cap\chi\in(0,2)\}}{\Pr\{\chi\in(0,2)\}} \tag{26}$$

The numerator in the above formula represents the probability that from the given time r_i first arrival event will correspond to the process $M_1(t)$ and the distance from r_i to that event is less than t. Taking into account that process $L_d(t)$ is active only on the interval $[r_i, r_i+T-\tau]$, and does not realise on that interval with probability $e^{-\alpha(T-\tau)}$, that probability can be calculated as follows (similar arguments to those used in derivation of Eq. (12) can be applied)

$$\begin{aligned}
\Pr\{\chi < t\cap\chi\in(0,2)\} &= \int_0^t \mu_1 e^{-\mu_1 x}\int_x^{+\infty}(\lambda_1+\alpha+\alpha)e^{-(\lambda_1+\alpha+\alpha)y}\,dy\,dx \\
&\quad\times\frac{\mu_1}{\lambda+\mu}(1-e^{-(\lambda+\mu)t}) \quad t\leqslant T-\tau \\
\Pr\{\chi < t\cap\chi\in(0,2)\} & \\
&= \frac{\mu_1}{\lambda+\mu}(1-e^{-(\lambda+\mu)(T-\tau)}) + e^{-\alpha(T-\tau)}\int_{T-\tau}^t \mu_1 e^{-\mu_1 x}\int_x^{+\infty}(\lambda_1+\alpha)e^{-(\lambda_1+\alpha)y}\,dy\,dx \\
&= \frac{\mu_1}{\lambda+\mu}(1-e^{-(\lambda+\mu)(T-\tau)}) + e^{-\alpha(T-\tau)}\frac{\mu_1}{\lambda+\mu_1}(e^{-(\lambda+\mu_1)(T-\tau)} - e^{-(\lambda+\mu_1)t}) \quad t > T-\tau
\end{aligned} \tag{27}$$

It can be easily seen that the denominator in the Eq. (26) is the same as (12), that is (Q_1 represents an event that $I(r_i) = 0$)

$$\Pr\{\chi\in(0,2)\} = \Pr\{I(r_{i+1}) = 2|Q_1\} = \frac{\mu_1}{\lambda+\mu}\left(1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}\right). \tag{28}$$

Finally, substitution of (27) and (28) in (26) yields the following probability distribution function for random variable ξ_1

$$F_{\xi_1}(t) = \begin{cases} \frac{1-e^{-(\lambda+\mu)t}}{1+\frac{\alpha}{\lambda+\mu}e^{-(\lambda+\mu)(T-\tau)}}, & t\leqslant T-\tau, \\ \frac{1-e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu}e^{-(\lambda+\mu)(T-\tau)}} + e^{-\alpha(T-\tau)}\frac{(\lambda+\mu)(e^{-(\lambda+\mu_1)(T-\tau)} - e^{-(\lambda+\mu_1)t})}{(\lambda+\mu_1)\left(1+\frac{\alpha}{\lambda+\mu}e^{-(\lambda+\mu)(T-\tau)}\right)}, & t > T-\tau. \end{cases} \tag{29}$$

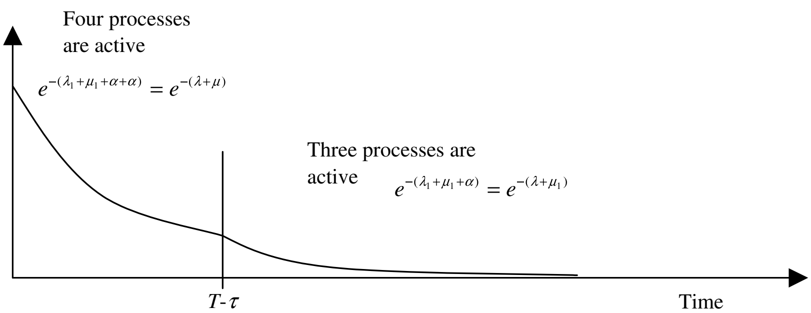

Fig. 7. Structure of probability density function of random variable ξ_1.

In order to obtain the expectation of ξ_1, its probability density function is calculated

$$f_{\xi_1}(t) = \begin{cases} \frac{\lambda+\mu}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}}e^{-(\lambda+\mu)t}, & t \in [0, T-\tau), \\ \frac{(\lambda+\mu)e^{-\alpha(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}}e^{-(\lambda+\mu_1)t}, & t \in [T-\tau, +\infty). \end{cases} \quad (30)$$

An illustration of the above probability density function is given in Fig. 7. It is worth noticing that the probability density function is continuous in the point of $t = T - \tau$.

Having calculated the probability density function of the random variable ξ_1 it is possible to calculate its expectation using the definition

$$E[\xi_1] = \int_{-\infty}^{+\infty} t f_{\xi 1}(t)\,dt \Rightarrow E[\xi_1] = \frac{1}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}}\int_0^{T-\tau}(\lambda+\mu)te^{-(\lambda+\mu)t}\,dt + \frac{\frac{\lambda+\mu}{\lambda+\mu_1}e^{-\alpha(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}}\int_{T-\tau}^{+\infty}(\lambda+\mu_1)te^{-(\lambda+\mu_1)t}\,dt \quad (31)$$

After integration by parts and some elementary calculation and simplification we obtain

$$E[\xi_1] = \frac{\frac{1}{\lambda+\mu} + \frac{\alpha}{\lambda+\mu_1}\left[T - \tau + \frac{1}{\lambda+\mu} + \frac{1}{\lambda+\mu_1}\right]e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}e^{-(\lambda+\mu)(T-\tau)}} \quad (32)$$

Now we consider differences Δ^2, Δ^4, Δ^{5a}. Despite the fact that these differences are produced by events that belong to different Poisson processes, they can be considered as realisations of the same random variable, denoted by ξ_2. This result is implied by the fact that events, whose arrival times produce the above differences belong to either of the processes $L_u(t)$ or $L_d(t)$ (see Fig. 6, events of types 1 and 3). For example, consider events that generate differences of the 2nd class. The indicator functions for such a pair of events are $I(r_i) = 0$ and $I(r_{i+1}) = 3$. Due to the relationship between the processes $L_u(t)$ and $L_d(t)$, it is known, that at the time $r_{i+1} - T + \tau$ there is an event from $L_u(t)$ that has taken place. Therefore, the event from $N_1(t)$ that happened at time r_i and produced the difference cannot be located into the past further then $r_{i+1} - T + \tau$. From this argument it follows that all differences of the second type do not exceed the value of $T - \tau$. Now consider time interval $[r_i; r_{i+1} - T + \tau]$. On this interval four independent exponentially distributed random variables are acting. Thus, random variable ξ_2 can be seen as being exponentially distributed but with the restriction that its values are not greater than $T - \tau$. Therefore, its probability density function is the head of the exponential density cut at the point $T - \tau$. Fig. 8 illustrates such a distribution.

It can also be shown that random variables that produce differences Δ^4 and Δ^{5a} have the same form of density function. Since ξ_2 is exponentially distributed but conditional on not being greater than $T - \tau$, it can be defined via a proper exponentially distributed random variable as having a conditional probability distribution function as follows:

$$F_{\xi_2}(t) = \Pr\{\psi < t | \psi \leqslant T - \tau\}, \quad (33)$$

where ψ is an exponentially distributed random variable with parameter $\lambda + \mu$. Then, following from the definition of conditional probability:

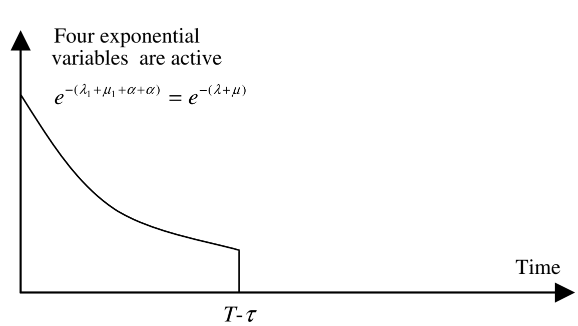

Fig. 8. Structure of probability density function of variables Δ^2, Δ^4 and Δ^{5a}.

$$\Pr\{\psi < t|\psi \leqslant T-\tau\} = \frac{\Pr\{\psi < \min(t, T-\tau)\}}{\Pr}\{\psi < T-\tau\} = \frac{1-e^{-(\lambda+\mu)t}}{1-e^{-(\lambda+\mu)(T-\tau)}}, \quad \forall t \in [0, T-\tau). \tag{34}$$

Finally, the random variable that corresponds to differences Δ^2, Δ^4, Δ^{5a} has the following probability density function:

$$f_{\xi_2}(t) = \begin{cases} \frac{\lambda+\mu}{1-e^{-(\lambda+\mu)(T-\tau)}} e^{-(\lambda+\mu)t}, & t \in [0, T-\tau), \\ 0, & t \in [T-\tau, +\infty) \end{cases} \tag{35}$$

and the expectation of ξ_2 is evaluated as

$$E[\xi_2] = \int_{-\infty}^{+\infty} t f_{\xi_2}(t) dt \Rightarrow E[\xi_2] = \frac{1}{1-e^{-(\lambda+\mu)(T-\tau)}} \int_0^{T-\tau} (\lambda+\mu) t e^{-(\lambda+\mu)t} dt = \frac{1}{1-e^{-(\lambda+\mu)(T-\tau)}} \left[-e^{-(\lambda+\mu)t} \left(t + \frac{1}{\lambda+\mu} \right) \right] \Bigg| \begin{matrix} T-\tau \\ 0 \end{matrix} E[\xi_2]$$
$$= \frac{\frac{1}{\lambda+\mu} - \left(T-\tau+\frac{1}{\lambda+\mu}\right) e^{-(\lambda+\mu)(T-\tau)}}{1-e^{-(\lambda+\mu)(T-\tau)}}. \tag{36}$$

The last type of the differences left to be considered is Δ^{5b}. The differences of this type correspond to the events of the processes $L_u(t)$ and $L_d(t)$ that are bound by the relationship given in (3). Therefore, the arrival time r_i that corresponds to an event of process $L_u(t)$ and arrival time r_{i+1} that corresponds to the related event of process $L_d(t)$ are related as follows: $r_{i+1} = r_i + T - \tau$.

Therefore, the differences Δ^{5b} are not random variables but a constant

$$\Delta^{5b} = r_{i+1} - r_i = T - \tau. \tag{37}$$

We have now expressed analytically all components for the Eq. (24). This allows for the calculation of the limit of the average $F(\tau)$, as given in (8).

$$F^*(\tau) = \sum_{k \in \{1,2,4,5a,5b\}} \Pr_k E[\Delta_i^k(\tau)] = E[\xi_1]P_1 + E[\xi_2] \sum_{k \in \{2,4,5a\}} P_k + (T-\tau)P_{5b} \tag{38}$$

The first component of the sum is

$$E[\xi_1]P_1 = \frac{\frac{1}{\lambda+\mu} + \frac{\alpha}{\lambda+\mu_1}\left[T-\tau+\frac{1}{\lambda+\mu}+\frac{1}{\lambda+\mu_1}\right] e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1} e^{-(\lambda+\mu)(T-\tau)}} \frac{\lambda_1\mu_1}{\lambda\mu}. \tag{39}$$

It can be shown, that

$$\sum_{k \in \{2,4,5a\}} P_k = \frac{\alpha(\lambda+\mu_1)}{\lambda\mu} \frac{1-e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1} e^{-(\lambda+\mu)(T-\tau)}}.$$

Therefore, the second component of the sum (38) is

$$E[\xi_2] \sum_{k \in \{2,4,5a\}} P_k = \frac{\frac{1}{\lambda+\mu} - \left(T-\tau+\frac{1}{\lambda+\mu}\right) e^{-(\lambda+\mu)(T-\tau)}}{1-e^{-(\lambda+\mu)(T-\tau)}} \frac{\alpha(\lambda+\mu_1)}{\lambda\mu} \frac{1-e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1} e^{-(\lambda+\mu)(T-\tau)}} = \frac{\alpha(\lambda+\mu_1)}{\lambda\mu} \frac{\frac{1}{\lambda+\mu} - \left(T-\tau+\frac{1}{\lambda+\mu}\right) e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1} e^{-(\lambda+\mu)(T-\tau)}} \tag{40}$$

and the last component of the sum (38) is simply

$$(T-\tau)P_{5b} = (T-\tau) \frac{\alpha(\lambda+\mu)}{\lambda\mu} \frac{e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1} e^{-(\lambda+\mu)(T-\tau)}}. \tag{41}$$

After substitution of Eqs. (39)–(41) into (38) and some simplifications $F^*(\tau)$ can be obtained as

$$F^*(\tau) = \frac{\frac{1}{\lambda+\mu} + \frac{\alpha}{\lambda\mu}\left[\frac{\lambda\mu}{\lambda+\mu_1}(T-\tau) + \frac{\lambda_1\mu_1}{\lambda+\mu_1}\left(\frac{1}{\lambda+\mu_1}+\frac{1}{\lambda+\mu}\right) - \frac{\lambda+\mu_1}{\lambda+\mu}\right] e^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1} e^{-(\lambda+\mu)(T-\tau)}}. \tag{42}$$

The above calculation have been conducted with the assumption that $T - \tau \geqslant 0$. Since the algorithm does not know T a priori, it can choose τ such that $T - \tau < 0$. Assume now that $T - \tau < 0$ and calculate $F^*(\tau)$. Since $T - \tau < 0$ the dependency picture on Fig. 9 will change as follows.

Comparing Figs. 6 and 9 it is clear that every event of $L_u(t)$ has a corresponding event $L_d(t)$ located away by time $T - \tau$ in the past. Since the algorithm looks ahead for a candidate for the pair, the type of the currently chosen event will not affect the algorithm's choice, that is

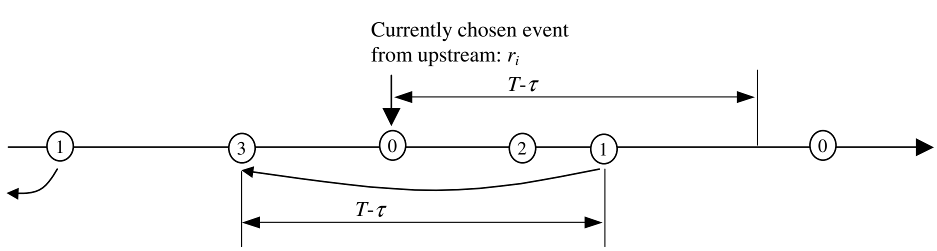

Fig. 9. The events of types 1 and 3 have swapped their roles.

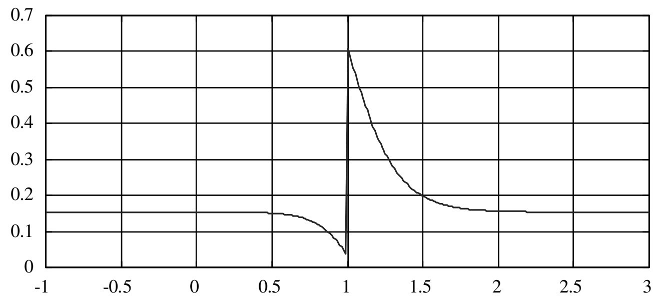

Fig. 10. A typical plot of function (44).

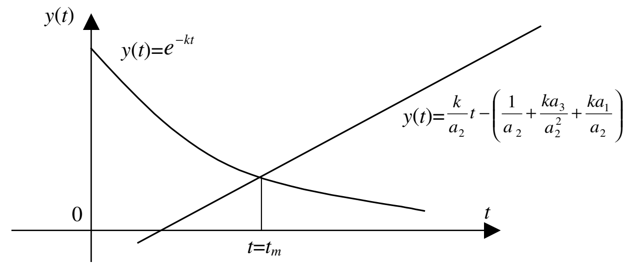

Fig. 11. A unique solution to Eq. (47).

the following events of $M(t)$ will be known to be independent of the currently chosen event of $N(t)$. Therefore, a separate consideration of $N_1(t)$ and $L_u(t)$, as was undertaken for the case of $T-\tau \geqslant 0$, is not required.

Taking into account the above argument, the pairs of events taken out by the algorithm will be identically distributed random variables with the probability density function shown in Fig. 7 and thus having the mean function given by Eq. (32) with $T-\tau$ replaced with $\tau - T$. Since the algorithm deals only with one type of random variables in the case of $T-\tau < 0$, the Eq. (8) takes the following form:

$$F^*(\tau) = \frac{\frac{1}{\lambda+\mu} + \frac{\alpha}{\lambda+\mu_1}\left[\tau - T + \frac{1}{\lambda+\mu} + \frac{1}{\lambda+\mu_1}\right]\mathrm{e}^{-(\lambda+\mu)(\tau-T)}}{1 + \frac{\alpha}{\lambda+\mu_1}\mathrm{e}^{-(\lambda+\mu)(\tau-T)}}. \tag{43}$$

Finally, (42) and (43) together give the full solution for the limit version of the cost function (6) as follows:

$$F^*(\tau) = \begin{cases} \dfrac{\frac{1}{\lambda+\mu}+\frac{\alpha}{\lambda\mu}\left[\frac{\lambda\mu}{\lambda+\mu_1}(T-\tau)+\frac{\lambda_1\mu_1}{\lambda+\mu_1}\left(\frac{1}{\lambda+\mu_1}+\frac{1}{\lambda+\mu}\right)-\frac{\lambda+\mu_1}{\lambda+\mu}\right]\mathrm{e}^{-(\lambda+\mu)(T-\tau)}}{1+\frac{\alpha}{\lambda+\mu_1}\mathrm{e}^{-(\lambda+\mu)(T-\tau)}}, & \tau \leqslant T, \\ \dfrac{\frac{1}{\lambda+\mu}+\frac{\alpha}{\lambda+\mu_1}\left[\tau-T+\frac{1}{\lambda+\mu}+\frac{1}{\lambda+\mu_1}\right]\mathrm{e}^{-(\lambda+\mu)(\tau-T)}}{1+\frac{\alpha}{\lambda+\mu_1}\mathrm{e}^{-(\lambda+\mu)(\tau-T)}}, & \tau > T. \end{cases} \tag{44}$$

A typical plot of the above function is given in the following Fig. 10.

It has therefore been shown that the limit (8) exists and is finite with probability 1 and its analytical form has been identified in the form of the Eq. (44).

In order to prove the statement of the theorem concerning the travel time estimate (7) it is required to show that (44) has a single global minimum at the point $\tau = T(T-\tau = 0)$. To do so we introduce the following parameters.

$$a_1 = \frac{1}{\lambda+\mu}, \quad a_2 = \frac{\alpha}{\lambda+\mu_1}, \quad a_3 = \frac{\alpha}{\lambda\mu}\left[\frac{\lambda+\mu_1}{\lambda+\mu} - \frac{\lambda_1\mu_1}{\lambda+\mu_1}\left(\frac{1}{\lambda+\mu_1}+\frac{1}{\lambda+\mu}\right)\right], \quad k = \lambda+\mu \quad \text{and} \quad t = T-\tau. \tag{45}$$

The function (44) can now be written as

$$f(t) = \frac{a_1 + (a_2 t - a_3)\mathrm{e}^{-kt}}{1 + a_2\mathrm{e}^{-kt}}, \tag{46}$$

where $a_1 > 0$, $a_2 > 0$, $a_3 > 0$, $k > 0$ and $t \geqslant 0$. The condition $f'(t) = 0$ leads to the following equation:

$$\mathrm{e}^{-kt} = \frac{k}{a_2}t - \left(\frac{1}{a_2} + \frac{ka_3}{a_2^2} + \frac{ka_1}{a_2}\right) \tag{47}$$

representing an intersection of a line and an exponential function as illustrated in Fig. 11.

It has therefore been shown that function (44) has a single extremum at $t = t_m$ on the interval $t \geqslant 0$. By checking the sign of the derivative $f'(t)$ we can see that the extremum is maximum. This completes the proof. □

4. Conclusions

Heuristics have proven themselves to be effective in solving complex real-life problems but they suffer from the credibility problem if they are designed on a trial-and-error basis. This paper presents a derivation of a heuristic dealing with processing of time series data as a simplification (abstraction) of the statistical correlation of random variables. While the heuristic discussed in this paper is derived in the context of an urban traffic analysis, the bridge between the mathematical analysis of the correlation of random variables and the pattern matching in experimental data is general and applicable to a broad spectrum of practical problems.

The paper supports also a more general conclusion about the nature of heuristics as abstractions of analytical algorithms in a similar way as information granules are considered abstractions of numerical data in the context of Granular Computing (Bargiela and Pedrycz, 2002). It is interesting to note that this algorithmic abstraction is prompted by the abstraction of the data (ignoring its statistical properties) so that the two abstraction paradigms can cross-validate each other.

On the application specific level, the adequateness of modelling the road traffic as Poisson processes has been established empirically by other researchers (Gerlough and Barnes, 1971; Cowan, 1975) and has served as a basis of our analytical model. In this paper we take full advantage of the property that the sum of two independent Poisson processes is also a Poisson process and derive on that basis an analytical formula for the estimation of travel times by the correlation of upstream and downstream traffic counts. The uniqueness of the travel time estimate calculated in this way has been asserted as a theorem and has been formally proven.

The simplifying assumption about constant travel time between upstream and downstream measurement points is well justified in the context of urban road networks (with short travel times between detectors) and it ensures that a convolution of gamma-distributed arrival times is also gamma-distributed. It is worth emphasising however that although the assumption makes the analytical evaluation of the correlation much easier it does not exclude the possibility of applying the same approach to a situation where the travel times are different for different categories of vehicles.

The complexity of the analytical evaluation of the travel time highlights the advantages of the heuristic outlined in Section 2. The less detailed view of the lane occupancy time series (abstracting from the actual probability distributions) leads to an algorithm that replaces the notion of statistical correlation with the notion of a simple match of upstream and downstream time series. The close correspondence of analytical and heuristic results confirms the value of the heuristic.

Acknowledgements

The authors gratefully acknowledge the provision of real-time telemetry data by the Nottingham Traffic Control Centre and the financial support for the project from the EPSRC.

References

Agafonov, E., 2003. Fuzzy and multi-resolution data processing for advanced traffic and travel information. PhD Dissertation, The Nottingham Trent University.
Agafonov, E., Bargielam A., Tanaka, M., Peytchev, E., 2006. Probabilistic travel time estimation. In: 38th International Symposium on Stochastic Systems Theory and its Applications, November 2006, Suwa, Japan.
Bargiela, A., Pedrycz, W., 2002. Granular Computing – An Introduction. Kluwer Academic Publishers.
Bargiela, A., Pedrycz, W., 2003. Recursive information granulation: Aggregation and interpretation issues. IEEE Transactions on Systems Man and Cybernetics SMC-B 33 (1), 96–112.
Bargiela, A., Pedrycz, W., 2005. Granular mappings. IEEE Transactions on Systems Man and Cybernetics SMC-A 35 (2), 288–301.
Bargiela, A., Kosonen, I., Pursula, M., Peytchev, E., 2006. Granular analysis of traffic data for turning movements estimation. International Journal of Enterprise Information System 2 (2), 13–27.
Burke, E.K., Kendall, G., Soubeiga, E., 2003. A tabu search hyper-heuristic for timetabling and rostering. Journal of Heuristics 9 (6), 451–470.
Burke, E.K., Kendall, G., Whitwell, G., 2004. A new placement heuristic for the orthogonal stock cutting problem. Operations Research 52 (4), 655–671.
Burke, E.K., Decausmaecker, P., Petrovic, S., Vanden Berghe, G., 2006a. Metaheuristics for handling time interval coverage constraints in nurse scheduling. Applied Artificial Intelligence 20 (9), 743–766.
Burke, E.K., Hellier, R., Kendall, G., Whitwell, G., 2006b. New bottom-left-fill heuristic algorithm for the 2D irregular packing problem. Operations Research 54 (3), 587–601.
Burke, E.K., Landa Silva, J.D., 2006c. The influence of the fitness evaluation method on the performance of multiobjective optimisers. European Journal of Operational Research 169 (3), 875–897.
Coifman, B., 2001. Improved velocity estimation using single loop detectors. Transportation Research: Part A 35 (10), 863–880.
Cowan, R.J., 1975. Useful headway models. Transportation Research 9 (6), 371–375.
Dailey, D.J., 1993. Travel time estimation using cross-correlation techniques. Transportation Research, Part B 27B (2), 97–107.
Dailey, D.J., 1997. Travel time estimates using a series of single loop volume and occupancy measurements. In: Transportation Research Board 76th Annual Meeting, Washington.
Dailey, D.J., 1999. A statistical algorithm for estimating speed from single loop volume and occupancy measurements. Transportation Research B 33, 313–322.
Feller, W., 1968. An Introduction to Probability Theory and its Applications. John Wiley & Sons.
Gerlough, D.L., Barnes, F.C., 1971. The Poisson and other probability distribution in highway traffic. In: Poisson and Other distributions in Traffic. The Eno Foundation for Transportation, Saugatuck, Conn.
Hall, F.L., Persaud, B.N., 1989. Evaluation of speed estimates made with single-detector data from freeways traffic management systems. Transportation Research Record (1232), 9–16.
Pedrycz, W., Bargiela, A., 2002. Granular clustering: A granular signature of data. IEEE Transactions on Systems Man and Cybernetics 32 (2), 212–224.
Petty, K., Bickel, P., Jiang, J., Ostland, M., Rice, J., Ritov, Y., Schoenberg, F., 1998. Accurate estimation of travel times from single-loop detectors. Transportation Research, Part B: Methodological 32 (1), 1–17.
Pursula, M. 1995. The approximation of percentage time delay with local measurements. Transportation Research Record 1484, Washington, DC, pp. 58–65.
Pursula, M., 1998. Two dimensional estimation of speed flow relationships with backpropagation neural networks. In: Himanen, V., Nijkamp, P., Reggiani, A., Raitio, J. (Eds.), Neural Networks in Transport Applications. Ashgate, Aldershot, England, pp. 263–294.
Pushkar, A., Hall, F.L., Acha-Daza, J.A., 1994. Estimation of speeds from single-loop freeway travel times in real time from flow measurements. Journal of Transportation Engineering 122, 185–191.
Reformat, M., Pedrycz, W., Pizzi, N., 2004. Building a software experience factory using granular-based models. Fuzzy Sets and Systems 145, 111–139.
Yao, Y.Y., 2001. Information granulation and rough set approximation. International Journal of Intelligent Systems 16 (1), 87–104.
Zadeh, L., 1979. Fuzzy sets and information granularity. In: Gupta, M., Ragade, R., Yager, R. (Eds.), Advances in Fuzzy Set Theory and Applications. North-Holland Publishing Co., Amsterdam, pp. 3–18.
Zadeh, L.A., 1997. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets and Systems 90, 111–127.
Zadeh, L., 1999. From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions. IEEE Transactions on Circuits and Systems 45, 105–119.