Queues with server vacations in urban traffic control

Maria de Lurdes Simões[1], Paula Milheiro Oliveira[1], and Américo Pires da Costa[2]

[1] Faculdade de Engenharia da Universidade do Porto
DEC – CEC
Rua Dr. Roberto Frias, s/n P-4200-465 PORTO, Portugal
(e-mail: maires@fe.up.pt, poliv@fe.up.pt) Partially supported by PRODEP III-5.3

[2] Faculdade de Engenharia da Universidade do Porto
DEC – CITTA
Rua Dr. Roberto Frias, s/n P-4200-465 PORTO, Portugal
(e-mail: aapc@fe.up.pt)

Abstract. A queuing system resulting from a semaphorized intersection regulated by semi-actuated control in a network urban traffic is considered. Modelization of the queue length and of the delay of vehicles is crucial in the study of the performance of intersections equipped with traffic signals. In these systems, the server (green signal) is desactivated (red signal) during a random period of time. Due to this particularity, models for classic queues such as $M/M/1$, $M/G/1$ and $G/M/1$ are not appropriate. In the urban traffic literature, the frequent desactivation of the server as well as the variation of the service period are not well formulated. In the present work a $M/G/1$ queue where the server occasionally takes vacations and the service discipline is a non-gated time-limited policy is analyzed. The present analysis follows [Leung and Eisenberg, 1991] who consider an application of these models in telecommunications. Their implementation, given its complexity, is made possible by using Laguerre functions when looking for an approximate solution of the differential equations involved. One concludes that the mean delays of vehicles given by this model are slightly smaller than those obtained by simulation procedures, but they are able to give us a good approximation for larger flows, which is of interest for traffic engineers, since, in that case, the approximations one can find in the traffic literature are known not to be adequate.
Keywords: Queues, Server vacations, Traffic models.

1 Introduction

Waiting systems that admit interruptions of service often appear when the server uses idle periods of time of one queue or one task to serve clients in another queue or to perform another task. What matters is that, for these idle periods, the server is not available nor operational for new arrivals to the system (see *e.g.* [Doshi, 1986] for an interesting briefing on the subject). Among other applications these waiting systems appear in the literature as models for computer networks and telecommunications, production and quality control.

Models with interruptions of the server have been analyzed for different waiting systems, as the M/GI/1 or the GI/GI/1 queues with a single server, no restrictions existing on the arrivals process or the service time distribution, as long as the stationarity is maintained. In what regards the pause of the server, the model may fall into different classes, depending on the situations that trigger the pause (or vacation) and on the service policy, when the server returns from a pause and is available for service again.

In the context of urban traffic, modeling the queue length and the waiting time (delay) of vehicles is fundamental if one wants to study the performance of semaphorized intersections. Here we are concerned with semi-actuated intersections, which means that there are a main street and a secondary street and a sensor is placed in the secondary street, enabling the activation of the green signal and thus of the vehicles in this street to go through the intersection. The main difficulties involved in the analysis by means of the queuing theory come from the need of a good characterization of the circulating vehicles and drivers and from the fact that the desactivation of the server for random periods of time (red signal) has to be incorporated in the behavior of the queue. Due to this, essentially, the $M/M/1$, $M/G/1$ and $G/M/1$ models do not satisfactorily fit the waiting phenomena in these kinds of traffic intersections.

A detailed study of semaphorized intersections with a fixed period of green signal, which is not the case of semi-actuated signals, can be found in [Webster, 1958] where a formula of the delay of traffic which is much used in the traffic engineering practice is given. The traffic flow that reaches the intersection is assumed to follow a Poisson distribution and several parameters of the model are reduced to mean values which are obtained from the results of the $M/D/1$ and $M/D^X/1$ queues. Nevertheless, with such models, the regular but random desactivation of the served can not be well described. Indeed, as the signal alternates between red and green, modeling a semaphorized intersection is a problem lying in the class of queuing systems with server vacations [Doshi, 1986], with the particularity that the server remains inactive for random time durations. [Heidemann, 1994] proposes an analytic model that includes server vacations, starting from the assumption that the arrival process is Poissonian, that the intersection has a fixed cycle regulation, that the interval between departure of vehicles is constant and the traffic capacity is one way only. With these restrictions the probability generating functions for the measures of performance queue length and delay of a vehicle can be derived from the associated Markov chains. More recently [Alfa and Neuts, 1995] suggested the use of discrete time Markov arrival processes to describe the nature of platoons in the traffic flow.

In the present work a M/G/1 model for which the server occasionally takes a vacation and the service policy is non-gated time-limited is analyzed.

The term time-limited refers to the fact that the server is available to the queue for a maximum time duration at each visit (constant T_m). The term non-gated refers to the fact that clients that arrive while the server is active are candidates for service during this visit of the server in as much as the maximum service time T_m is not achieved. Clients are served in a FIFO regime and the server starts a vacation as soon as all clients in the queue are served or T_m expires, whatever occurs first. If the queue is empty when the server returns from a vacation it immediately starts a new vacation.

Our goal is to explore the theory of queues with server vacations, particularly the work by [Leung and Eisenberg, 1991], to find an approximate expression for the mean delay of a vehicle in the context of semi-actuated traffic using the comparison with the results obtained by numerical simulation of an intersection in [Simões *et al.*, 2002] to judge on the appropriateness of the proposed method.

2 An equation for the amount of work

For the class of models introduced above, the probability density function (pdf) of the amount of work at an arbitrary instant during a vacation period of the server is obtained by solving a functional equation that characterises the amount of work at the exact time the server starts a service period. Solving this equation, due to its complexity, is done by means of a numerical technique analogous to the one of [Weeks, 1966], based on the numerical inversion of the Laplace Transforms (LT).

The complementary of the distribution function of the duration of a service period (time between the beginning of service and the instant the queue becomes empty, assuming that T_m is never achieved) is approximated by a sum of Laguerre functions. Using the relation between the amount of work at the beginning of a service period and the duration of the server busy interval, the functional equation in transformed into a set of linear equations, from which the solution corresponds to the coefficients of the Laguerre functions in the expansion just mentioned.

Thus the amount of work in the queue at an arbitrary instant can be obtained from the equation that runs the amount of work at the instants the service starts serving the clients. From the decomposition of the amount of work and the PASTA property [Wolff, 1982], the mean waiting time can be deduced.

Notation:

$\bar{x}$, $\bar{x}^2$, $X^*(\cdot)$: mean, second moment and LT of the service time;

$\bar{\nu}$, $\bar{\nu}^2$, $V^*(\cdot)$: mean, second moment and LT of the duration of the vacation;

$\bar{u}_p$, $f_p(\cdot)$, $U_p^*(\cdot)$: mean, pdf and LT of the amount of work at the beginning of a service period;

$P_0(t)$: probability of the queue being empty at time t.

The following assumptions are made:

i) Clients arrive according to a Poisson process with parameter λ and the service time follows a general distribution for which the first two moments are finite;

ii) The system has an infinite waiting room;

iii) The system is in equilibrium and $\rho(=\lambda\bar{x}) < \dfrac{T_m}{T_m+\bar{\nu}}$;

iv) The duration of a vacation (random variable) is independent from the amount of work at the beginning of a service period.

The main theoretical result that we need when dealing with queues with server vacations is the stochastic decomposition property [Boxma and Groenendijk, 1987]: if the queue is in equilibrium, the LT of the amount of work at the beginning of a service period may be written as the product of the LT of the amount of work at the end of a service period, $U^*(s,T_m)$, by the LT of the amount of work that arrives during a vacation, $U_v^*(s)$. Making use of this property the major difficulty in the analysis of models that have a limited service time lies in the characterization of the amount of work at an arbitrary instant during a vacation period. In order to overcome this difficulty performing the following steps is required:

i) Set up the functional equation that characterizes the amount of work at the beginning of a service period. The stochastic decomposition property states that

$$U_p^*(s) = U_v^*(s)\cdot U^*(s,T_m)\,. \tag{1}$$

On the other hand one has $U_v^*(s) = V^*(\lambda - \lambda X^*(s))$ and

$$U^*(s,T_m) = e^{\hat{s}T_m}\left\{U_p^*(s) - \hat{s}\int_{y=0}^{T_m} e^{-\hat{s}y}P_0(y)dy\right\}\,, \tag{2}$$

where $\hat{s} = s - \lambda + \lambda X^*(s)$.

ii) Equation (2) is solved numerically, given that $1 - P_0(t)$ can be approximated by a weighted sum of Laguerre functions:

$$P_0(t) \overset{\text{sim}}{=} 1 - \sum_{n=0}^{N} a_n e^{-\frac{t}{2T}} L_n\left(\frac{t}{T}\right)\,.$$

Thus

$$P_0^*(s) \overset{\text{sim}}{=} 1 - \sum_{n=0}^{N} a_n \frac{s\left(s-\frac{1}{2T}\right)^n}{\left(s+\frac{1}{2T}\right)^{n+1}}\,. \tag{3}$$

The LT $U_p^*(s)$ is also approximated by means of Laguerre functions:

$$U_p^*(s) \overset{\text{sim}}{=} 1 - \sum_{n=0}^{N} a_n \frac{\hat{s}\left(\hat{s}-\frac{1}{2T}\right)^n}{\left(\hat{s}+\frac{1}{2T}\right)^{n+1}}\,. \tag{4}$$

The approximations given in (3) and (4) are used in equation (2), from which, using (1), one gets:

$$U_p^*(s) \stackrel{\text{sim}}{=} U_v^*(s).e^{\hat{s}T_m}\left\{e^{-\hat{s}T_m} + U_p^*(s) - \sum_{n=0}^{N} a_n e^{-(\hat{s}+\frac{1}{2T})T_m} L_n\left(\frac{T_m}{T}\right)\right.$$
$$\left. - \int_{y=0}^{T_m} e^{-\hat{s}y} \sum_{n=0}^{N} a_n e^{-y/2T}\left[\frac{1}{2T}L_n\left(\frac{y}{T}\right) - L_n'\left(\frac{y}{T}\right)\right] dy\right\} \quad (5)$$

Notice that this equation is linear in the a_n for a given s with $Re(s) \geq 0$.

iii) The functional equation (5) is transformed into a linear system of equations, since, by taking $s = i\omega$ and using $N+1$ appropriated values for ω in equation (5), a set of $N+1$ linear equations is obtained (see [Weeks, 1966]). The coefficients a_n are known by solving this system.

iv) To end with, by using the decomposition of the amount of work and the PASTA[1] property [Wolff, 1982], the mean amount of work in the system as seen by a Poisson arrival is given by

$$\bar{u} = \frac{\lambda \bar{x^2}}{2(1-\rho)} + \sum_{n=0}^{N}(-1)^n(2T)(1-\rho)a_n - \rho\bar{\nu} + \rho\frac{\bar{\nu^2}}{2\bar{\nu}}\,. \qquad (6)$$

The mean waiting time of a client is obtained by applying Little's formula.

3 Application to the control of semi-actuated traffic

As mentioned in the introduction, traffic signals with semi-actuated regulation are frequently used in intersections which consist of a main street and a secondary street. The actuated phase serves the movement of vehicles in the secondary street. The control variables that lead the efficiency of a semi-actuated operation are the regulation plan of the semaphore and the placement of the sensor. The difficulty in applying the semi-actuated control is in the selection of an optimum combination of these operations. In the absence of a service call (non activation of the sensor) the green signal is always given to the non-actuated phase. As soon as the sensor is activated a change in the signals occurs. The time interval for this change to occur includes a yellow period followed by a period of "all red" (cleaning time). During the activation of the sensor the arrival of a vehicle in the actuated street extends the interval of green signal of this phase by an amount of time so that the minimum of green time is exceeded but not the maximum. It means that, in semi-actuated traffic, the green time is adapted to the demand, having a minimum and a maximum value. In this way, a larger number of vehicles is able to pass through the intersection per unit of time.

[1] *Poisson Arrival See Time Average*

In the present work the intersection illustrated in Fig. 1 is considered. The sensor is placed $5\,m$ away from the stopping line of the secondary street. The times given to the regulation of the two phases are shown in Table 1.

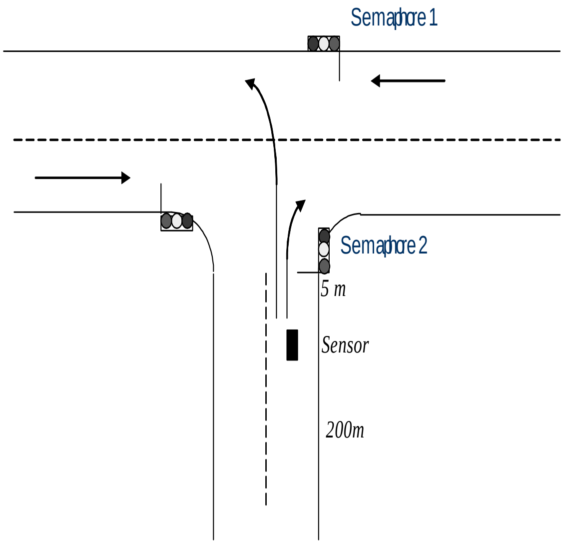

Fig. 1. Scheme of a semi-actuated intersection.

Table 1. Times given to the regulation of the two phases.

Time (sec.)	Semaphore 1	Semaphore 2
Green	20 to ∞	7 to 40
Yellow	3	3
Extension of green	–	4
All red	2	2

The degree of saturation, $x_{sat} = \rho \dfrac{T_m + \bar{\nu}}{T_m}$, represents the ratio between the mean number of vehicles that arrive during a cycle and the maximum number of vehicles that may pass through the intersection during that period of time. In the terminology of the queuing systems this parameter is known as the congestion index.

The mean waiting times estimated by the model presented in Section 2 (referred to as the analytical model) are shown in Fig. 2 as well as the average delays experienced by drivers according to the simulation (see [Simões *et al.*, 2002] for the detailed simulation study). A Dirac function with a mass at the point 27 and a Gaussian distribution with mean value equal to 2 and variance 0.04 are considered in the analytical model as the laws of the duration of a vacation (red period) and of the service time, respectively, since these are the best fit distributions in the case of semi-actuated urban traffic intersections.

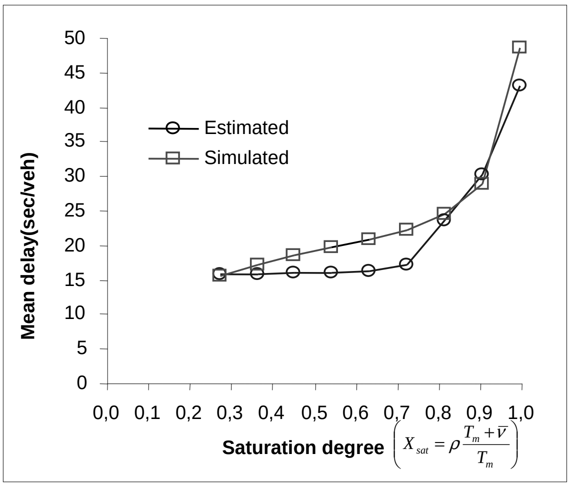

Fig. 2. Comparison between the mean delay estimated by the analytical model and the simulated mean delay ($\bar{\nu} = 27$ sec., $\bar{x} = 2$ sec. and $T_m = 43$ sec.).

The results suggest that, for approximately $x_{sat} > 0.7$, the analytical model gives good estimates of the mean delay of drivers. For $0.3 < x_{sat} < 0.7$, however, the estimates given by this model are smaller than those obtained by numerical simulation. This fact may be due to the diversity of reactions that is typical of drivers behavior and of interactions between vehicles but, most of all, the fact is that the duration of a vacation (red signal) is not really bounded, since it is extended until the activation of the sensor, which means it has no maximum value although it has a minimum.

It is important to remark that when dealing with the analytical model one should be aware of the importance of choosing adequate values for N and of the need of a high precision in the computations, as the numerical method explained here is very sensitive to precision errors. Difficulties in making these numerical procedures converging are also reported in the literature[Leung and Eisenberg, 1991] in the case of probability density functions with jumps or discontinuities (service times or durations of the vacations that are deterministic). In practice it is very much recommended to validate the outputs of the numerical procedures ensuring that the amplitudes of the a_n's are smaller than 10^{-8}.

4 Final comments

An analytical expression for the evaluation of an approximation of the mean delay of vehicles in semi-actuated traffic was found by applying systems of queues with server vacations theory, while previous expressions were known to be inappropriate for the semi-actuated case.

This procedure gives good approximations when the arrival flow is large, which was not possible with heuristic expressions commonly used in traffic engineering that had been developed for the fixed control case. The expressions that we give here provide realistic estimates of the mean delay particularly when the saturation index is below 70%, while for large traffic flows (congestion scenarios) the estimates they provide appear to be smaller than the real mean delays.

Having in mind improving the reliability of the results presented here and others that will be obtained in the future, the numerical properties of the relationship between N and T deserves a careful investigation, aiming to establish, for different distributions of the service durations, which values should be given to N and T in order to ensure good results when this method is applied.

References

[Alfa and Neuts, 1995]A.S. Alfa and M.F. Neuts. Modelling vehicular traffic using the discrete time Markovian Arrival Process. *Transportation Science*, 29(2), pages 109–117, 1995.

[Boxma and Groenendijk, 1987]O.J. Boxma and W.P. Groenendijk. Pseudo-conservation laws in cyclic-service systems. *J. Appl. Prob.*, 24, pages 949–964, 1987.

[Doshi, 1986]B.T. Doshi. Queueing systems with vacations - A survey. *Queueing Systems*, 1, pages 29–66, 1986.

[Heidemann, 1994]D. Heidemann. Queue length and delay distributions at traffic signals. *Transportation Research-B*, 28(5), pages 377–389, 1994.

[Leung and Eisenberg, 1991]K.K. Leung and M. Eisenberg. A single-server queue with vacations and non-gated time-limited service. *Performance Evaluation*, 12, pages 115–125, 1991.

[Simões *et al.*, 2002]M.L. Simões, P.M. Oliveira, and A.P. Costa. Análise probabilística do fluxo de tráfego num cruzamento semi-actuado. *Actas do IX Congresso Anual da SPE*, pages 111–124, 2002.

[Webster, 1958]F.V. Webster. *Traffic Signal Settings*. Road Research Laboratory, Road Research Laboratory 39, HMSO, London, 1958.

[Weeks, 1966]W.T. Weeks. Numerical inversion of Laplace transforms using Laguerre functions. *J. ACM*, 13, pages 419–426, 1966.

[Wolff, 1982]R.W. Wolff. Poisson arrivals see time averages. *Operations Research*, 30, pages 223–231, 1982.