

Continuous polling models and application to ferry assisted WLAN

Veeraruna Kavitha · Eitan Altman

© Springer Science+Business Media, LLC 2011

Abstract In this paper we focus on a class of polling systems encountered while modeling the ferry based wireless local area network (FWLAN). A moving ferry, while walking in a predetermined cyclic path, communicates with the static nodes (or users) of the network via a wireless link. The ferry is assumed to stop and communicate with a node that has a packet to send or to receive, when it is closest to that node. The location distribution of the node to which or from which a packet arrives is assumed to have a support of positive Lebesgue measure. These features imply that polling models with finite number of queues cannot be used to model the system. We study in this paper the continuous polling systems with service disciplines that model the use of the FWLAN (and that are more complex than the classical exhaustive or gated services). Our approach is based on discretization of the continuous polling model. We propose a special way of discretizing the continuous system such that: (1) the known Pseudo conservation laws can be applied to obtain the stationary expected workload of the discrete systems; (2) the limit, of these ‘discretized’ expected workloads, equals the stationary expected workload of the continuous system. Our results rely heavily on fixed point analysis of infinite dimensional operators.

1 Introduction

Polling systems are a class of queueing systems wherein a single server attends to a number of queues. Systems with a finite number of queues (called discrete polling systems) have

This work was sponsored by The Indo-French Center for the Promotion of Advanced Research under CEFIPRA Project# 4000-IT-1.

V. Kavitha (✉) · E. Altman
INRIA, Sophia Antipolis, France
e-mail: kavitha.veeraruna@gmail.com

E. Altman
e-mail: Eitan.Altman@sophia.inria.fr

V. Kavitha
LIA, University of Avignon, Avignon, France

been analyzed to a good extent and there is good understanding of their stability conditions and on the Pseudo conservation laws for a variety of discrete polling systems in Takagi (1986), Boxma et al. (1992), Boxma (1989), Boxma and Groenendijk (1987), Foss and Chernova (1996), Khamisy et al. (1992), Sidi et al. (1992), Georgiadis and Szpankowski (1992), Altman et al. (1992) etc.

Continuous polling systems are the ones in which the location of arrivals as well as of the service can occur anywhere in a continuum of locations. Continuous polling were first introduced by Fuhrmann and Copper (1985), further studied, explored by Coffman and Gilbert (1986, 1987) and Kroese and Schmidt (1992, 1993, 1994, 1996) in a series of works. Stability of such polling systems is discussed in Altman and Levy (1994), Leskela and Unger (2010), Robert (2010) and a review of such results are in Rojas-Nandayapa et al. (2011).

It is relatively more complicated to study continuous systems and these are usually analyzed under simplified conditions, which we refer as ‘symmetric conditions’: every arrival picks up an uniformly-distributed landing site on the circle, settles down there and awaits service while the server is moving at a constant speed and in a fixed direction around the circle. Further the service time requirements of the users do not depend upon the position of the user.

The work on snowplowing systems in Eliazar (2003, 2005) generalizes many of the above assumptions and studies a more general continuous system. For example, the incoming work-flow to the system is taken to be a general Levy random measure and the walking times are assumed to be random.

In previous literature so far, the continuous polling systems have been mostly analyzed with standard gated/exhaustive service. In such systems, the server, while moving in a cyclic path, attends the users waiting anywhere on its path the first time it encounters them. But there are other important service disciplines (e.g., globally gated, elevator service etc.) which are studied in a discrete polling system. In a globally gated service discipline (Boxma et al. 1992), the server attends the users the first time it encounters them after they are tagged and all the users are tagged whenever the ferry reaches a global point. Elevator service (Altman et al. 1992) is similar to globally gated service except that the server reverses its direction at the end of every cycle. There can be applications, wherein the users need to be supported via a mixture of the above service disciplines in a continuous polling system. For example, in a Ferry assisted Wireless LAN all uplink transmissions (data uploaded to BS via the Ferry from static sources) use an exhaustive discipline while all downlink transmissions (data downloaded to a sink from the Ferry which had earlier received the same data from BS) have to be modeled using a globally gated or an elevator regime.

The central idea of this paper is to use the known Pseudo conservation laws (known for a discrete system) for obtaining the stationary expected virtual workload of a continuous system (which supports mixture of services) via discretization. We discretize the continuous polling system in a special way such that the known Pseudo conservation laws can be applied for the resulting discrete systems and the limit, of these ‘discrete’ stationary workloads, equal that of the continuous system. We achieved this by utilizing the fixed point equations that can result because of stationarity. We basically utilize the periodic behavior in the system, identify certain quantities whose moments must be equal at the beginning and at the end of a cycle because of stationarity (stationarity implies palm stationarity) and hence obtain fixed point equations for the stationary moments. We identify those quantities, which can further be utilized for expressing the stationary expected workload in the system. We obtained the required limit majorly via continuous parameter (the number of discretization levels is the parameter here) dependence of the fixed points. Further, we obtain an expression for the expected stationary virtual workload abandoning the standard symmetric as-

sumptions. Discretization approach is used for mixed service systems, while, the stationary workload for Globally gated and elevator service disciplines is obtained directly.

In this paper, incoming jobs arrive to the system following a Poisson process. Our approach can be used for studying other types of continuous polling whenever one can solve the related discrete problem and use similar limiting arguments. In Kavitha (2011) we used the same approach to study a continuous system with rerouting.

Polling systems are used to obtain analysis in variety of applications (see for example a recent survey paper (Boona et al. 2011) and the references therein). The second part of the paper applies the Polling system results to study a Ferry assisted Wireless Local Area Network (FWLAN). Message Ferry is a mobile relay station that serves as “postman” to deliver (collect) messages to (from) the static or dynamic wireless nodes in a sparse network, where direct connectivity is not possible. Mobile BS have been proposed in the context of Mobile Ad Hoc Networks (Tariq et al. 2006), Vehicular Ad-Hoc Networks (Vanets) (Yousefi et al. 2007) and in Wireless (static) Sensor Networks (Shi and Hou 2008). In the UmassDiesel project, computers have been installed in 30 out of 40 buses and these then serve as Message Ferry to deliver messages to throw boxes.¹ In a related paper (Tariq et al. 2006), the message ferry serves some fixed finite number of nodes and the routes are designed based on the mobility models of those finite number of nodes. In contrast, in our work the ferry serves all the nodes that arrive anywhere in the given area as and when they arrive and the optimal routes are designed based on the arrival process, service requirements, network architecture, radio conditions, etc.

The ferry moves periodically in the cyclic route and either halts at predetermined finite number of stops or halts on its way only when it encounters a user with a request. *We refer the former as BUS mode while the later one as TAXI mode.* BUS mode has been studied using discrete polling systems in Kavitha and Altman (2009, 2010a, 2010b), while this paper focuses on TAXI mode, using the continuous polling results. In the TAXI mode, every point on the ferry path is a potential stop and each stop is assigned an area that contains all points closer to that stop than to other stops. Upon reaching a stop, the ferry collects (uplink) and dumps (downlink) data from or to the area assigned to it.

An important performance measure for FWLAN design is the expected waiting time, i.e., the average time an arrival has to wait before it’s service starts. The waiting time depends on: (1) the over-all service times (the time required to complete an uplink or downlink request using wireless medium); (2) the walking time, i.e., the time taken by the ferry to traverse the cyclic path once. The over-all service times reduce if one uses longer ferry routes as, then, the users can be served by the ferry standing relatively at a shorter distance. However the walking time increases with longer routes. The objective of this paper is to study these trade-offs using stationary expected workload performance. By this, we minimize the expected waiting times at all users in the Pareto sense.

One of the important contributions of this paper is to show how theory of polling systems can be mapped to obtain interesting performance measures for some interesting configurations of FWLAN. These performance measures are used for designing optimal ferry routes (in a given class of routes) or for designing optimal number of stops (for a given type of ferry route Kavitha and Altman 2009) and or for designing the location of base station(s). However one can think of many more applications of this mapping. For example, in Saad et al. (2009) the authors consider a wireless scenario with many ferries and user groups. Their aim was to group ferries and users into coalitions in a optimal way. Each coalition in their

¹Umassdieselnet. <http://prisms.cs.umass.edu/diesel/>.

paper was a simple example of ferry based wireless LAN. Our results can probably be used to form these optimal coalitions for more complicated architectures.

The continuous polling models are studied in Sect. 2 while ferry based wireless LAN examples are analyzed in Sects. 3 and 4. Some proofs are available in Appendices.

2 Continuous polling systems

In a continuous polling system, the server moves continuously and stops at point q only when it finds a user with a request. The server is moving continuously on a circle \mathcal{C} of radius γ with speed α . The arrival process is modeled by a Poisson process with intensity λ and every arrival is associated with two marks, the position $Q \in \mathcal{C}$ distributed as P_Q and the service times B . The service times in general can depend upon the position Q of the arrival. Let $b(q)$ $b^{(2)}(q)$ represent the conditional first two moments of the service time B conditioned on the event that the position of arrival is at point q . However the service times of different users are independent of each other.

The server either provides pure or mixed type of service to the users. In a pure *globally gated service* system, the server closes a fictitious gate each time it arrives at a fixed point in the circle (we refer it as 0) and tags all the users that arrived before the closure of the gate. The *time period between two consequent arrivals of the server at 0 is referred as cycle*. In every cycle, the server moves along the circular path and serves only the tagged users as and when it encounters one. Note here that some times an user may be served only the second time the server encounters it after it's arrival. On the contrary in a pure *gated/exhaustive service* polling system, the users are served more immediately after their birth: they are served the first time, the server encounters them. In a *discrete polling system*, there is a difference between *gated service* (the server closes a local gate at every queue and leaves the queue only after serving the users that appeared before the gate closure) and *exhaustive service* (the server also attends the users that arrived during the service of already existing users and it leaves the queue only when the queue is empty). But the two converge to the same service in a continuous polling system with P_Q continuous: at any queue (which is a point in the continuous case), the probability of a new user appearing at the same point is zero.

In a *mixed service* polling system, the server attends some of the users with globally gated service and the rest with exhaustive/gated service. Let p_{gg} , $p_g = 1 - p_{gg}$ respectively represent the probability that an arrival is a globally gated user and a gated user. Let $b_g(q)$, $b_g^{(2)}(q)$ represent the conditional first and second moments of the service time given that the arrival is at point q and further given that the arrival is a gated user. Let $b_{gg}(q)$, $b_{gg}^{(2)}(q)$ represent same quantities for a globally gated user. Let $\bar{b} := E[b(Q)]$, $\bar{b}_{gg} := E[b_{gg}(Q)]$ and $\bar{b}_g := E[b_g(Q)]$ (expectations are with respect to P_Q) represent the unconditional moments. Note that $\bar{b} = \bar{b}_g p_g + \bar{b}_{gg} p_{gg}$. Similarly define the second moments, $\bar{b}_g^{(2)} := E[b_g^{(2)}(Q)]$, $\bar{b}_{gg}^{(2)} := E[b_{gg}^{(2)}(Q)]$ and $\bar{b}^{(2)} := E[b^{(2)}(Q)]$.

Notations The variables like b , b_g , b_{gg} , f_Q , τ , v represent the functions on interval $[0, |C|]$ while terms like $b(q)$ or $\tau(q)$ represent the function values at a point $q \in [0, |C|]$. The bar of the same variable, like \bar{b}_g represent the average w.r.t. to the position distribution P_Q .

²Right now the position of arrival is distributed the same way for gated as well as globally gated users. But the results can easily be generalized to separate $P_{Q,g}$ and $P_{Q,gg}$.

i.e., $\bar{b}_g = E[b_g(Q)]$. Similarly variables like $\tau^{(2)}$, $\bar{\tau}^{(2)}$ etc. are functions on $[0, |\mathcal{C}|] \times [0, |\mathcal{C}|]$ while $\tau^{(2)}(q, q')$ represents the function value at (q, q') . Let $q \wedge |\mathcal{C}|$ represent the minimum value, $\min\{q, |\mathcal{C}|\}$. The expectation with respect to the stationary measure of the process under consideration is given by E while E^0 represents the expectation with respect to Palm stationary measure.

Virtual workload of a polling system is defined as the total workload corresponding to all the waiting users, i.e., the sum of the service times of all the waiting users. It is an important performance measure and will be used in later sections for analyzing interesting application in delay tolerant networks. Not much theory is available for calculating the expected virtual workload of polling systems with arrivals in a continuum. In this section we derive new (stationary expected workload) results related to continuous polling systems. Throughout we consider *stationary and ergodic systems*. An exhaustive/gated continuous polling system is analyzed in Kroese and Schmidt (1992) under symmetric conditions and we summarize the same in Sect. 2.1. The main contribution of this paper towards polling systems are available in Sects. 2.2, 2.3 and 2.4 in the form of stationary expected virtual workload derivations for polling system respectively with a globally gated service, elevator and mixed service. These results also cover the gated/exhaustive service system under asymmetric conditions.

2.1 Exhaustive/gated service

An exhaustive service continuous polling system under symmetric conditions is analyzed in Kroese and Schmidt (1992). They have many more results but we only present the ones relevant to this paper, i.e., the expression for expected virtual workload under stationarity. The users arrive uniformly on the circumference with $|\mathcal{C}| = 1$ as a Poisson process with intensity λ , i.e., P_Q is uniform. Further the service time moments $b_g(q)$, $b_g^{(2)}(q)$ are assumed to be same at all points q and are equal to \bar{b}_g and $\bar{b}_g^{(2)}$ respectively. Then the stationary expected number of customers on the circle is given by Kroese and Schmidt (1992, Theorem 5.1)

$$E[N]_g = \lambda \bar{b}_g + \frac{\lambda(\alpha^{-1} + \lambda \bar{b}_g^{(2)})}{2(1 - \lambda \bar{b}_g)}.$$

In this case they also considered the user under service. Excluding the user under service, i.e., the stationary expected number of waiting customers equal:

$$E[N]_g^w = \frac{\lambda(\alpha^{-1} + \lambda \bar{b}_g^{(2)})}{2(1 - \lambda \bar{b}_g)}. \quad (1)$$

From (1), by Wald's lemma the virtual workload, the workload due to the waiting customers, equals:

$$V_g^{sym} := \bar{b}_g E[N]_g^w = \frac{\lambda \bar{b}_g (\alpha^{-1} + \lambda \bar{b}_g^{(2)})}{2(1 - \lambda \bar{b}_g)}. \quad (2)$$

2.2 Globally gated service

In this subsection we derive the corresponding results for a globally gated service. Further we don't assume symmetric conditions and work under more general conditions as explained before. *We map the circumference to a line of length $|\mathcal{C}| := 2\pi\gamma$ with the starting point (the point at which the global gate is closed) mapped to 0.*

2.2.1 Computation of the moments of the Cycle time in stationary regime

Let C_n represent the time taken to complete the n th cycle. Let $\mathcal{N}(C)$ represent the number of Poisson arrivals in time C . With B_i^n representing the service time of the i th user arrived during the $n - 1$ th cycle, the adjacent cycles times can be related using

$$C_n = \sum_i^{\mathcal{N}(C_{n-1})} B_i^n + |C|\alpha^{-1}.$$

By first conditioning on $\mathcal{N}(C_{n-1})$ and then on C_{n-1} , it is easy to see that,

$$E \left[\sum_i^{\mathcal{N}(C_{n-1})} B_i^n \right] = \lambda E[C_{n-1}] \int_0^{|C|} b_{gg}(q) P_Q(dq).$$

Thus under (Palm) stationarity,³

$$c_* := E^0[C_n] = \frac{|C|\alpha^{-1}}{1 - \lambda \bar{b}_{gg}}. \quad (3)$$

From the above it is clear that the system is *stable* only if the factor $\rho := \lambda \bar{b}_{gg}$ is strictly less than 1. This is the usual condition for stability and throughout this paper we assume stability. In a similar way as,

$$\left(\sum_i^{\mathcal{N}(C_{n-1})} B_i^n \right)^2 = \sum_i^{\mathcal{N}(C_{n-1})} (B_i^n)^2 + \sum_{i \neq j} B_i^n B_j^n$$

and because $E[\mathcal{N}(C_n)^2 - \mathcal{N}(C_n)] = \lambda^2 E^0[C_n^2]$ (with $c_*^{(2)} := E^0[C_n^2]$),

$$c_*^{(2)} = \lambda c_* \bar{b}_{gg}^{(2)} + \lambda^2 c_*^{(2)} \bar{b}_{gg}^2 + |C|^2 \alpha^{-2} + 2|C|\alpha^{-1} \lambda c_* \bar{b}_{gg}.$$

Thus,

$$c_*^{(2)} = \frac{|C|c_*^{-1}}{1 - \lambda \bar{b}_{gg}} \frac{|C|\alpha^{-1} + \lambda \bar{b}_{gg}^{(2)} + |C|\alpha^{-1} \lambda \bar{b}_{gg}}{1 - \lambda^2 \bar{b}_{gg}^2}. \quad (4)$$

³ Let ϕ_n represent the time at which the server visits the global gate, 0, in the n th cycle. Then $C_n = \phi_n - \phi_{n-1}$. For any stationary point process, for example in our case $\{\phi_n\}$, there will be two associated probabilities: Stationary and Palm Stationary (Baccelli and Brémaud 1991). In general, $\{\phi_n\}$ are defined such that $\phi_0 \leq 0 < \phi_1$. Palm probabilities are the stationary probabilities obtained after conditioning on the event that $\{\phi_0 = 0\}$ (see Baccelli and Brémaud 1991). Throughout the paper, the expectation under Palm stationary measure is represented by E^0 and the corresponding moments are usually denoted with a * as sub-script. In Baccelli and Brémaud (1991), the stationary moment of the residual cycle C_1 (which we refer as C_R) as well as the past cycle C_0 (which we refer as C_P) is obtained in terms of Palm probabilities as

$$E[C_R] = E[C_P] = \frac{E^0[C_1^2]}{2E^0[C_1]}.$$

This result is also explicitly derived specifically for cycle times in a special polling system in Sect. 3.1 of Boxma et al. (2008). This result is used throughout.

2.2.2 Stationary expected virtual workload

An user arriving at q and at an arbitrary time epoch (which is modeled as occurred at time (0) has to wait, for his service to start, on average for a time equal to the sum of three components: (1) the average residual cycle time of the ongoing cycle C_R ; (2) the time taken by the server to reach the point q from 0, $q\alpha^{-1}$; (3) the average time taken by the server to attend the users belong to the segment $[0, q]$ that arrived in the entire ongoing cycle $C_R + C_P$ (C_P represents the past cycle time of the ongoing cycle), $\sum_i^{N(C_P+C_R)} B_i 1_{\{q_i \in [0, q]\}}$. Thus the stationary expected waiting time of a user given that the user arrives at point q equals (see footnote 3),

$$\begin{aligned}
 E[W](q) &= E[C_R] + q\alpha^{-1} + \lambda P_Q(Q \in [0, q])E[C_R + C_P]E[B|Q \in [0, q]] \\
 &= \frac{c_*^{(2)}}{2c_*} + q\alpha^{-1} + \lambda \frac{c_*^{(2)}}{c_*} \widehat{b}_{gg}(q) \\
 \text{where } \widehat{b}_{gg}(q) &:= \int_0^q b_{gg}(y) P_Q(dy).
 \end{aligned}$$

By Little’s law (Whitt 1991) the stationary expected number of waiting users that belong to infinitesimal segment $[q - dq, q + dq]$ equals,

$$E[N](dq) = \lambda E[W](q) P_Q(dq)$$

and thus by Wald’s Lemma, the stationary expected virtual workload (by virtue of independence) due to users that belong to infinitesimal segment $[q - dq, q + dq]$ equals

$$V_{gg}(dq) = b_{gg}(q) \lambda E[W](q) P_Q(dq).$$

Hence the stationary expected virtual workload in the entire system equals

$$\begin{aligned}
 V_{gg} &= \lambda \int_0^{|\mathcal{C}|} E[W](q) b_{gg}(q) P_Q(dq) \\
 &= \lambda \left(\frac{c_*^{(2)}}{c_*} \left(\frac{\bar{b}_{gg}}{2} + \lambda E[\widehat{b}_{gg}(Q) b_{gg}(Q)] \right) + \alpha^{-1} E[Q b_{gg}(Q)] \right).
 \end{aligned}$$

By interchanging the order of the two integrals in $E[\widehat{b}_{gg}(Q) b_{gg}(Q)]$,

$$\begin{aligned}
 \int_0^{|\mathcal{C}|} \int_0^{q'} b_{gg}(q) P_Q(dq) b_{gg}(q') P_Q(dq') &= \int_0^{|\mathcal{C}|} \int_q^{|\mathcal{C}|} b_{gg}(q') P_Q(dq') b_{gg}(q) P_Q(dq) \\
 &= \int_0^{|\mathcal{C}|} \left(\bar{b}_{gg} - \int_0^q b_{gg}(q') P_Q(dq') \right) b_{gg}(q) P_Q(dq) \\
 &= \bar{b}_{gg}^2 - \int_0^{|\mathcal{C}|} \int_0^q b_{gg}(q') P_Q(dq') b_{gg}(q) P_Q(dq)
 \end{aligned}$$

and so, $E[b_{gg}(Q) \widehat{b}_{gg}(Q)] = \bar{b}_{gg}^2 / 2$. From (3), (4) the virtual workload simplifies to:

$$V_{gg} = \lambda \left(\frac{|\mathcal{C}| \alpha^{-1} (1 + \lambda \bar{b}_{gg}) + \lambda \bar{b}_{gg}^{(2)} \bar{b}_{gg}}{1 - \lambda \bar{b}_{gg}} \frac{\bar{b}_{gg}}{2} + \alpha^{-1} E[Q b_{gg}(Q)] \right). \tag{5}$$

Symmetric Case: We specialize to the symmetric case where $b(q) \equiv b$, $b^{(2)}(q) \equiv b^{(2)}$ and $P_Q(dq)$ is uniform with $|C| = 1$. In this case $E[Q\bar{b}_{gg}(Q)] = b/2$ and

$$V_{gg}^{sym} = \lambda \bar{b}_{gg} \left(\frac{\lambda \bar{b}_{gg}^{(2)}}{2(1 - \lambda \bar{b}_{gg})} + \frac{\alpha^{-1}}{1 - \lambda \bar{b}_{gg}} \right).$$

2.3 Elevator service

This polling system is similar to that in Sect. 2.2 except that the server moves in the opposite direction every time it reaches the point at which the global gate closes. That is, the server moves in counter clockwise direction in alternate, say odd numbered cycles and in the reverse direction in the even numbered cycles. In Altman et al. (1992), authors describe the scheme for discrete polling system. The analysis of the continuous counterpart will be similar to that in Sect. 2.2 and we mention only the points of difference.

Stationary moments of the Cycle time: The adjacent cycles times can once again be related using

$$C_n = \sum_i^{N(C_{n-1})} B_i^n + |C|\alpha^{-1} \quad \text{and} \quad C_{n-1} = \sum_i^{N(C_{n-2})} B_i^{n-1} + |C|\alpha^{-1}.$$

Under stationarity C_n will be distributed same as C_{n-2} and C_{n-1} as C_{n-3} . Because of the IID nature of the services times $\{B_i^n\}$ across any cycle, for the above relations to hold C_n should be distributed same as C_{n-1} . Thus again under stationarity,

$$c_* = \frac{|C|\alpha^{-1}}{1 - \lambda \bar{b}_{gg}},$$

$$c_*^{(2)} = \frac{|C|c_1^{-1}}{1 - \lambda \bar{b}_{gg}} \frac{|C|\alpha^{-1} + \lambda \bar{b}_{gg}^{(2)} + |C|\alpha^{-1} \lambda \bar{b}_{gg}}{1 - \lambda^2 \bar{b}_{gg}^2}.$$

Expected waiting time for a random user: Here again the time the user arriving at q has to wait for his service will be equal to the sum of three components as in Sect. 2.2 but the server will be moving either in clockwise direction or in anticlockwise direction with half probability. Thus the expected waiting time at any point q equals:

$$E[W](q) = E[C_R] + \frac{1}{2} (q\alpha^{-1} + \lambda P_Q(Q \in [0, q]) E[C_R + C_P] E[B|Q \in [0, q]])$$

$$+ \frac{1}{2} ((|C| - q)\alpha^{-1} + \lambda P_Q(Q \in [q, |C|]) E[C_R + C_P] E[B|Q \in [q, |C|]])$$

$$= \frac{c_*^{(2)}}{2c_*} + \frac{|C|\alpha^{-1}}{2} + \lambda \frac{c_*^{(2)}}{2c_*} \bar{b}_{gg} = \frac{\lambda \bar{b}_{gg}^{(2)}}{2(1 - \lambda \bar{b}_{gg})} + |C|\alpha^{-1}$$

which is independent of the position q . As anticipated, the expected waiting times are independent of the position at which the user waits. The stationary expected workload can be derived as before and equals

$$V_{elevator} = \lambda \bar{b}_{gg} \left(\frac{\lambda \bar{b}_{gg}^{(2)}}{2(1 - \lambda \bar{b}_{gg})} + |C|\alpha^{-1} \right). \quad (6)$$

2.4 Mixed service

This is a system in which the server attends some of the users with globally gated service and the rest with exhaustive service. For users with globally gated service the server closes the gate when it touches upon 0 and serves in the current cycle only those which arrived before the gate closure as in Sect. 2.2. The rest of the users will be served, immediately, as in Sect. 2.1. We obtain the stationary expected virtual workload of the mixed system as a limit of the same for a ‘certain’ set of the discrete polling systems, under an extra assumption: the service times are bounded with probability one by a constant that is less than infinity.

Theorem 1 Assume $P_Q(dq) = f_Q(q)dq$, i.e., f_Q is the density of P_Q . Assume further that b_g, b_{gg} and f_Q are continuous. Also, there exists a $\hat{B} < \infty$ such that, the service times (irrespective of the position of the arrival) are bounded by \hat{B} with probability one, i.e., $P(B < \hat{B}) = 1$. Then there exists a threshold $\hat{\mathbf{B}}$ (discussed in the proof) and for all service times with upper bound $\hat{B} \leq \hat{\mathbf{B}}$, the expected stationary virtual workload for the mixed service continuous polling system, V_{mix} , is given by the limit of the corresponding one, V_{mix}^σ , for a “certain” discrete polling system with 2σ number of queues:⁴

$$\begin{aligned}
 V_{mix} = \lim_{\sigma \rightarrow \infty} V_{mix}^\sigma &= \lambda \bar{b} \frac{\lambda \bar{b}^{(2)}}{2(1 - \lambda \bar{b})} + p_{gg} \lambda \alpha^{-1} E[Qb_{gg}(Q)] + \frac{\lambda \bar{b} |\mathcal{C}| \alpha^{-1}}{2} \\
 &\quad + \frac{\lambda^2 |\mathcal{C}| \alpha^{-1}}{2(1 - \lambda \bar{b})} \left(\bar{b}^2 + (p_{gg} \bar{b}_{gg})^2 + p_{gg} p_g \bar{b}_{gg} \hat{\bar{b}}_g \right), \tag{7}
 \end{aligned}$$

where $|\mathcal{C}| = 2\pi \gamma, \hat{b}_g(q) := E[b_g(Q)1_{\{Q \leq q\}}]$ and $\hat{\bar{b}}_g := E[\hat{b}_g(Q)]$. □

Before we proceed with proving the above theorem we compare the formula (7) with the special cases considered so far. We note that, the general formula (7) matches with (2) of Sect. 2.1, when specialized to the symmetric conditions and to the pure gated service (i.e., with $p_g = 1 = 1 - p_{gg}$). It also matches with globally gated workload (5) derived in Sect. 2.2, when specialized to this pure service (i.e., $p_g = 0 = 1 - p_{gg}$).

The proof of Theorem 1 is obtained using the following 3 major steps:

- (1) *Discretization*: Continuous polling system is converted to an appropriate discrete polling system in Sect. 2.4.1, for which the Pseudo conservation laws and hence the expected virtual workload is known (see Boxma et al. 1992; Boxma 1989 etc.).

Let $\delta^\sigma(q)$ for every point q on \mathcal{C} represent the point, in the discrete system with σ discretization levels, standing at which the server attends the possible users of point q . Let δ^∞ represent the same for continuous system. Note that $\delta^\infty(q) = q$ for all q , i.e., that δ^∞ is the identity map.

- (2) *Fixed point equations*: We express the stationary moments of the time to reach $\delta^\sigma(q)$, for every q on the circumference \mathcal{C} , starting from 0 (global gate point) as a fixed point (in the space of left continuous and right limit functions) of an affine linear operator in Sect. 2.4.2. We obtain a common operator (\mathcal{F} defined in Sect. 2.4.2), which is further parametrized by a map, $\delta : [0, |\mathcal{C}|] \rightarrow [0, |\mathcal{C}|]$. The fixed point of the common operator

⁴The formula (7) was conjectured in Kavitha and Altman (2010a, 2010b) itself. However its proof is derived in this paper. Further, we note that there is a small error in the formula published in Kavitha and Altman (2010a, 2010b) (in the last term $\hat{\bar{b}}_g$ is wrongly misplaced with \bar{b}_g) and is corrected here.

at $\delta = \delta^\sigma$ gives the required stationary first moments for the discrete system while that at $\delta = \delta^\infty$ (the identity map) corresponds to the continuous system. We show the continuity of these fixed points with respect to δ via contraction mapping theorem and hence show that the stationary moments of the discrete system converge to that of the continuous system. We obtain the convergence of second moments using similar but more complicated logic.

- (3) *Alternate expression for Virtual Workload:* We express the expected virtual workload in terms of the stationary moments of the step (2). Note, that this common expression (given by (19) in Sect. 2.4.3) cannot be computed easily and is used only for the proof. Using this common expression, we show the convergence of the stationary expected virtual workload of the discrete system to that of the continuous system in Sect. 2.4.3.

2.4.1 Discretization

For each integer σ , consider a discrete polling system with 2σ queues, half of which experience globally gated service while the rest half experience gated service. The circumference $|\mathcal{C}|$ is divided to σ equal segments $\{I_i\}_{i=1}^\sigma$ with $I_1 = [0, \mathcal{C}/\sigma)$. All the users arriving in area I_i are treated as though arriving in queue numbered $2i - 1$ (for gated users) or $2i$ (for globally gated users). Note that the globally gated users are the tagged ones, i.e., the ones that arrived before the server reached 0 in the current cycle. For every i , the server stops upon reaching the starting point, $i^\sigma := (i - 1)|\mathcal{C}|/\sigma$, of I_i and serves the users of I_i (that arrived/tagged before the server reached the stop i^σ) before moving further. Hence, $\delta^\sigma(q)$, the point standing at which the server attends the users at q , equals

$$\delta^\sigma(q) = \begin{cases} \sum_{i=1}^\sigma i^\sigma 1_{\{q \in I_i\}} & \text{with } I_i := [i^\sigma, (i+1)^\sigma), i^\sigma := \frac{(i-1)|\mathcal{C}|}{\sigma} \quad \text{if } \sigma < \infty, \\ q & \text{if } \sigma = \infty. \end{cases} \quad (8)$$

Within a queue, the server attends the users in a special order which we call as *arrival position order*. In this special order, the users within a queue are served in the order of their distance from the stop i^σ of the server, i.e., *the user at minimum distance is served first*. So, the users are almost served in the same order as done in a continuous system. The main difference b/w the continuous system and the discretized system is that some of the users are postponed to the next cycle in the discretized system. This is mainly because of the combined effect of discretization and the gated service. But we will see that the effect of the users, whose service is postponed, reduces to zero as σ tends to ∞ .

Define the following terms related to each stop and or each queue with obvious meanings as below:

$$\begin{aligned} b_{2i-1} &:= E[B|Q \in I_i, \text{gated}] = \frac{E[b_g(Q)1_{\{Q \in I_i\}}]}{P_Q(I_i)}, & b_{2i-1}^{(2)} &:= \frac{E[b_g^{(2)}(Q)1_{\{Q \in I_i\}}]}{P_Q(I_i)}, \\ b_{2i} &:= \frac{E[b_{gg}(Q)1_{\{Q \in I_i\}}]}{P_Q(I_i)}, & b_{2i}^{(2)} &:= \frac{E[b_{gg}^{(2)}(Q)1_{\{Q \in I_i\}}]}{P_Q(I_i)}, \\ l_i &:= P_Q(I_i), & \lambda_{2i-1} &:= \lambda l_i p_g, & \lambda_{2i} &:= \lambda l_i p_{gg}, \\ \rho_i &:= \lambda_i b_i & \text{and } \rho &:= \sum_i^{2\sigma} \rho_i. \end{aligned}$$

In the above, $\{\lambda_{2i-1}\}_{i=1}^\sigma$, $\{b_{2i-1}\}$, $\{b_{2i-1}^{(2)}\}$ respectively represent the arrival rates, the first and the second moments of the service times at the gated queues while $\{\lambda_{2i}\}_{i=1}^\sigma$, $\{b_{2i}\}$, $\{b_{2i}^{(2)}\}$

represent the same for globally gated queues. Factor ρ represents the total work load in the system, note that ρ is same for all σ and equals $\lambda\bar{b}$. The discrete system is stable only when $\rho < 1$. The continuous system is assumed to be stable and hence the stationary moments of Theorem 2 (given in later subsections) exist and hence by the same theorem $\rho < 1$. Thus the discrete systems are stable for all σ .

- *Expected Virtual Workload for discretized polling system:*

This type of a mixed service stable discrete polling system is considered in Boxma et al. (1992), Boxma (1989), Boxma and Groenendijk (1987). By Pseudo Conservation Laws of Boxma and Groenendijk (1987), Boxma (1989) the expected stationary virtual workload of the σ -polling model with mixed (globally gated and gated) services is:

$$V_{mix}^\sigma = \rho \frac{\sum_{j=1}^\sigma \lambda_j b_j^{(2)}}{2(1-\rho)} + \frac{\rho s^{(2)}}{2s^{(1)}} + \frac{s^{(1)}}{2(1-\rho)} \left[\rho^2 + \sum_{i=1}^\sigma \rho_i^2 \right] + \frac{s^{(1)}}{1-\rho} \sum_{i \in gg} \rho_i \sum_{j \leq i} \rho_j + \sum_{i \in gg} \rho_i \sum_{j < i} s_j.$$

In the above, $s^{(1)}$, $s^{(2)}$ represent the first and the second moment of the total walking time, $\{s_i\}$ represent the first moments of the walking times between consecutive queues and the remaining terms have similar definitions as before. The results of Boxma et al. (1992), Boxma (1989), Boxma and Groenendijk (1987) are valid for any work conserving order at each queue and hence the results are also valid for our *arrival position order*.

In our case we have a 2σ stable polling system with even numbered queues experiencing the globally gated service and the odd numbered queues experiencing the gated service. Further we have fixed walking times between queues, that between the queues of the same stop is zero, $s_{2i-1} = 0$ and $s_{2i} = |C|\alpha^{-1}/\sigma$. Thus the expression for expected stationary virtual workload for our discrete polling system is:

$$V_{mix}^\sigma = \rho \frac{\sum_{j=1}^{2\sigma} \lambda_j b_j^{(2)}}{2(1-\rho)} + \frac{\rho|C|\alpha^{-1}}{2} + \frac{|C|\alpha^{-1}}{2(1-\rho)} \left[\rho^2 + \sum_{i=1}^{2\sigma} \rho_i^2 \right] + \frac{|C|\alpha^{-1}}{1-\rho} \sum_i \rho_{2i} \sum_{j \leq 2i} \rho_j + \sum_i \rho_{2i} \frac{(i-1)|C|\alpha^{-1}}{\sigma}. \tag{9}$$

We will prove that the limit of the ‘discrete’ expected stationary virtual workload, V_{mix}^σ , indeed equals that of the continuous system. This basically forms the proof of the Theorem 1 and this proof is given in the next two subsections. We conclude this subsection by computing the limit of (9) (proof in Appendix A):

Lemma 1 *The limit of V_{mix}^σ (9) equals V_{mix} given by (7) of Theorem 1.*

2.4.2 Fixed point equations

Let $T_n^\sigma(q)$ represent the time to reach the point, $\delta^\sigma(q)$ of (8), starting from the global point 0, during the n th cycle. Let $V_n^\sigma([a, c])$ represent the workload present at the beginning of the n th cycle in the segment $[a, c] \subset [0, |C|]$. Let $\mathcal{T}_g([a, c], T)$ (or $\mathcal{T}_{gg}([a, c], T)$ in case of globally gated system) represent the total workload of the gated (or globally gated) Poisson

arrivals that arrived in segment $[a, c] \subset [0, |C|]$, such that a user at point $q \in [a, c]$ can arrive during a period of time $T(q)$. They are related in a continuous polling system by (i.e., with $\sigma = \infty$):

$$T_n^\sigma(q) = q\alpha^{-1} + V_n^\sigma([0, q]) + \mathcal{T}_g([0, q], T_n^\sigma), \quad (10)$$

$$V_n^\sigma([a, c]) = \mathcal{T}_{gg}([a, c], T_{n-1}^\sigma(|C|)) + \mathcal{T}_g([a, c], T_{n-1}^\sigma(|C|) - T_{n-1}^\sigma). \quad (11)$$

In the above, the first term of (10) represents the time taken to walk the distance q . The second term, $V_n^\sigma([0, q])$, represents the time taken to complete service of the globally gated and gated users (placed geographically before q) that arrived before the server reached 0. The third term represents the time taken to complete service of the gated users (placed before q) that arrived during the current cycle, before the server reached their respective δ^σ points.

In a discrete polling system, $T_n^\sigma(q)$ is the time till the server attends all the users belonging to the line segments placed before the point q , $\{I_j\}_{j < q\sigma/|C|}$, and hence is same for all the points of the same line segment and is given by:

$$T_n^\sigma(q) = T_n^\sigma(\delta^\sigma(q)) = \delta^\sigma(q)\alpha^{-1} + V_n^\sigma([0, \delta^\sigma(q)]) + \mathcal{T}_g([0, \delta^\sigma(q)], T_n^\sigma). \quad (12)$$

Note that the same equation at $\delta = \delta^\infty$ matches with the continuous time equation (10). The equation (11) remains the same for the discrete system also.

First moments By the lemma below, we obtain integral representation of the first moments of the workloads like, $\mathcal{T}_g([0, q], T_n^\sigma)$ using which we obtain the integral presentation of the first moments of T_n^σ (proof in Appendix B).

Lemma 2 Let $T : [a, c] \mapsto \mathcal{R}^+$ be either monotone (increasing or decreasing) or constant nonnegative random function on interval $[a, c]$. Assume that for any $q \in [a, c]$, the service times of the new arrivals and the arrival process at around point q is independent of the system evolution before⁵ time $T(q)$. Let $\mathcal{T}([a, c], T)$ represent the total workload of the Poisson arrivals that arrived in segment $[a, c] \subset [0, |C|]$, such that a user at point $q \in [a, c]$ can arrive during a period of time $T(q)$. Then with $\tau(q) := E[T(q)]$ for any q ,

$$E[\mathcal{T}([a, c], T)] = \lambda \int_a^c b(q) f_Q(q) \tau(q) dq.$$

Let $\mathbf{N} := \{1, 2, \dots, \infty\}$. Let $\tau_n^\sigma(q) := E[T_n^\sigma(q)]$ represent the first moment of $T_n^\sigma(q)$ when the number of stops equal σ and let $\tau_n^\infty(q)$ represent the same for continuous system. Substituting (11) in (12) and then applying Lemma 2, we get for any $n \in \mathbf{N}$:

$$\begin{aligned} \tau_n^\sigma(q) &:= \delta(q)\alpha^{-1} + \int_0^{\delta(q)} v_n^\sigma(y) dy + \lambda p_g \int_0^{\delta(q)} b_g(y) f_Q(y) \tau_n^\sigma(y) dy \\ &\text{where } v_n^\sigma(q) := \lambda p_{gg} b_{gg}(q) \tau_{n-1}^\sigma(|C|) f_Q(q) \\ &\quad + \lambda p_g b_g(q) (\tau_{n-1}^\sigma(|C|) - \tau_{n-1}^\sigma(q)) f_Q(q). \end{aligned} \quad (13)$$

⁵Note that this assumption is satisfied by our polling systems (discrete as well as continuous ones) because the system evolution is effected by an arrival only after it is served and here $T_n^\sigma(q)$ (for which this lemma would be applied) represents the time instance at which it starts serving the users of the segment in which q resides.

Let \mathcal{D} represent the space of left continuous functions with right limits on $[0, |\mathcal{C}|]$ equipped with supremum norm:

$$\text{for any } \tau \in \mathcal{D}, \quad \|\tau\|_\infty := \sup_{q \in [0, |\mathcal{C}|]} |\tau(q)|.$$

Consider a parametrized function $\mathcal{F} : \mathcal{D} \times \mathcal{D} \mapsto \mathcal{D}$ defined point-wise as below:

$$\mathcal{F}(\tau; \delta)(q) := \delta(q)\alpha^{-1} + \lambda\tau(|\mathcal{C}|) \int_0^{\delta(q) \wedge |\mathcal{C}|} b(y)f_Q(y)dy \quad \text{for all } q \in \mathcal{C}. \quad (14)$$

Let $\tau_*^\sigma(q)$, for $\sigma \in \mathbb{N}$, represent the (Palm) stationary moment corresponding to $\tau_n^\sigma(q)$. By *stationarity*⁶ from (13) the stationary first moments, τ_*^σ , of the discrete system is a fixed point of the parametrized function \mathcal{F} , at δ^σ with $\sigma < \infty$, while that of the continuous system is fixed point of the same function at δ^∞ . Note here that $b(q) = p_g b_g(q) + p_{gg} b_{gg}(q)$.

We obtain the continuity of the stationary moments τ_*^σ , in uniform/sup norm, in the following (proof in Appendix B):

Theorem 2 *Assume b_g, b_{gg}, f_Q are continuous in q . For any $\delta \in \mathcal{D}$ the map \mathcal{F} has an unique fixed point, τ_*^δ , if and only if $\rho < 1$. Further these fixed points are continuous in supremum norm w.r.t. to parameter δ . From this continuity we get that, $\tau_*^\sigma = \tau_*^{\delta^\sigma}$, the stationary moments corresponding to discrete system with σ stops, converge to that of the continuous system as σ increases to infinity:*

$$\sup_{q \in \mathcal{C}} |\tau_*^\sigma(q) - \tau_*^\infty(q)| \rightarrow 0.$$

$$\text{Indeed: } \tau_*^\delta(|\mathcal{C}|) = \frac{\delta(|\mathcal{C}|)\alpha^{-1}}{1 - \widehat{\rho}(\delta(|\mathcal{C}|) \wedge |\mathcal{C}|)} \quad \text{and}$$

$$\tau_*^\delta(q) = \delta(q)\alpha^{-1} + \tau_*^\delta(|\mathcal{C}|)\widehat{\rho}(\delta(q) \wedge |\mathcal{C}|) \quad \text{for all } q$$

$$\text{where } \widehat{\rho}(q) := \lambda \int_0^q b(y)f_Q(y)dy.$$

Second moments Let $C_n^\sigma(q)$ represent the time taken between the n th and $n - 1$ th visit of the point q ($\delta^\sigma(q)$ in case of discrete systems), i.e., the cycle time w.r.t. to the point q . Note that

$$C_n^\sigma(q) = T_n^\sigma(q) + (T_{n-1}^\sigma(|\mathcal{C}|) - T_{n-1}^\sigma(q)). \quad (15)$$

These cycle times are the most important quantities required for further analysis and in this subsection we study its moments. We obtain the convergence of stationary second moments corresponding to $C_n^\sigma(q)$ as $\sigma \rightarrow \infty$. Let $c_*^{2\sigma}(q)$ represent the (Palm) stationary moments corresponding to $E[C_n^\sigma(q)^2]$. We obtain their convergence and the exact result is given in the following theorem (proof in Appendix B):

⁶In Palm stationary regime, τ_n^σ are same for all n , the common values are represented by τ_*^σ and hence we get a fixed point operator, \mathcal{F} , to represent (13).

Theorem 3 Assume b_g, b_{gg}, f_Q are continuous. Then there exists a threshold $\hat{\mathbf{B}}$ on the workload threshold \hat{B} , i.e., for service time requirements satisfying $B \leq \hat{\mathbf{B}}$ with probability one, as $\sigma \rightarrow \infty$ the stationary second moments converge uniformly:

$$\sup_{q \in [0, |C|]} |c_*^{(2\sigma)}(q) - c_*^{(2\infty)}(q)| \rightarrow 0.$$

2.4.3 Proof of Theorem 1

By Theorems 2 and 3, the first stationary moments of T_n^σ and the second order stationary moments of C_n^σ , of the discretized polling system, converge towards the corresponding ones of the continuous polling system. The convergence is in supremum norm, i.e., uniformly of the positions q on the circle. We obtain the expression for the expected virtual workload using these moments for both the discrete and continuous polling systems and then using the convergence results of the Theorems 2 and 3 we complete the proof of Theorem 1.

The first two (Palm) stationary moments $E^0[C_n^\sigma(q)]$ and $E^0(C_n^\sigma(q))^2$ of C_n^σ defined in (15) are:

$$\begin{aligned} c_*^\sigma(q) &= E^0[C_n^\sigma(q)] = \tau_*^\sigma(|C|), \\ c_*^{(2\sigma)}(q) &= E^0(C_n^\sigma(q))^2. \end{aligned}$$

Convergence of $\tau_*^\sigma(|C|)$ is obtained in Theorem 2 while that of $c_*^{(2\sigma)}$ is obtained in Theorem 3. Thus the stationary first moment of the residual of the cycle $C_n^\sigma(q)$ as seen by a random user in terms of the two Palm moments is given by (see footnote 3),

$$E[C_R^\sigma(q)] = \frac{E^0(C_n^\sigma(q))^2}{2(E^0[C_n^\sigma(q)])^2} = \frac{c_*^{(2\sigma)}(q)}{2(\tau_*^\sigma(|C|))^2}. \quad (16)$$

Workload due to Globally Gated users A randomly arriving globally gated user, arriving at point q , will have to wait on average for a sum of the following three:

- (1) the residual of the current cycle (w.r.t. to the globally gated point 0), $E[C_R^\sigma(|C|)$;
- (2) $\tau_*^\sigma(q)$, average time taken by the server to travel from 0 to $\delta^\sigma(q)$;
- (3) in a discretized system the user has to wait an extra amount of time to complete service of:⁷ (a) the other globally gated users in $[\delta^\sigma(q), q]$ (that would be served before him while the server is standing at $\delta^\sigma(q)$) and (b) the service of all the gated users belonging to the same segment $[\delta^\sigma(q), \delta^\sigma(q) + |C|/\sigma]$, which by Lemma 2 equals:

$$\begin{aligned} & E \left[\mathcal{T}_g \left(\left[\delta^\sigma(q), \delta^\sigma(q) + \frac{|C|}{\sigma} \right], C_n^\sigma \right) \right] + E \left[\mathcal{T}_{gg} \left([\delta^\sigma(q), q], T_{n-1}^\sigma(|C|) \right) \right] \\ &= \tau_*^\sigma(|C|) \left(\rho_g \left(\delta^\sigma(q), \delta^\sigma(q) + \frac{|C|}{\sigma} \right) + \rho_{gg}(\delta^\sigma(q), q) \right) \\ & \text{with } \rho_m(a, c) := \lambda p_m \int_a^c b_m(y) f_Q(y) dy \text{ for } m = g \text{ or } gg. \end{aligned}$$

⁷See arrival position order discussed in Sect. 2.4.1 and further note that the gated users are served before the globally gated ones.

Thus the expected waiting time of such a user for any $\sigma \in \mathbf{N}$ is

$$E[W_{gg}^\sigma](q) = E[C_R^\sigma(|\mathcal{C}|)] + \tau_*^\sigma(q) + 1_{\{\sigma < \infty\}} \tau_*^\sigma(|\mathcal{C}|) \left(\rho_g \left(\delta^\sigma(q), \delta^\sigma(q) + \frac{|\mathcal{C}|}{\sigma} \right) + \rho_{gg}(\delta^\sigma(q), q) \right). \tag{17}$$

Hence, by Little’s law (Whitt 1991) the density of the stationary number of waiting globally gated users is $\lambda p_{gg} f_Q(q) E[W_{gg}^\sigma](q)$ and by Wald’s lemma the density of the expected stationary workload corresponding to globally gated users is

$$\lambda p_{gg} f_Q(q) E[W_{gg}^\sigma](q) b_{gg}(q).$$

Workload due to Gated users A randomly arriving gated user, arriving at q , has to wait on average for:

- (1) residual of his own cycle, $E[C_R^\sigma(q)]$;
- (2) in case of the discretized system till the users placed geographically above him are served which equals (by calculating as done for a globally gated user) $\tau_*^\sigma(|\mathcal{C}|) \times \rho_g(\delta^\sigma(q), q)$.

Therefore the expected waiting time of a gated user is:

$$E[W_g^\sigma](q) = E[C_R^\sigma(q)] + 1_{\{\sigma < \infty\}} \tau_*^\sigma(|\mathcal{C}|) \rho_g(\delta^\sigma(q), q). \tag{18}$$

Thus by Little’s law (Whitt 1991) and Wald’s lemma the expected stationary workload due to gated users is:

$$\int_0^{|\mathcal{C}|} \lambda p_g f_Q(q) E[W_g^\sigma](q) b_g(q) dq.$$

Total expected stationary virtual workload Therefore, the total expected stationary virtual workload is given by:

$$V^\sigma = \lambda \int_0^{|\mathcal{C}|} \left((\tau_*^\sigma(q) + E[C_R^\sigma(|\mathcal{C}|)]) b_{gg}(q) p_{gg} + E[C_R^\sigma(q)] b_g(q) p_g \right) f_Q(q) dq + 1_{\{\sigma < \infty\}} \lambda \tau_*^\sigma(|\mathcal{C}|) \int_0^{|\mathcal{C}|} v_q f_Q(q) dq$$

with $v_q := \rho_g(\delta^\sigma(q), q) p_g b_g(q) + \left(\rho_g \left(\delta^\sigma(q), \delta^\sigma(q) + \frac{|\mathcal{C}|}{\sigma} \right) + \rho_{gg}(\delta^\sigma(q), q) \right) \times p_{gg} b_{gg}(q).$ (19)

In the above V^∞ represents the expected virtual workload of the random user for the continuous system while V^σ with $\sigma < \infty$ represents the same for a discrete system. By the Theorems 2 and 3 the terms inside the first integral of (19) (see also (16), (17) and (18)), defining the expected virtual workload, converge towards that of the continuous polling system point-wise and uniformly in q . Terms inside the second integral, v_q , converge to zero point-wise and uniformly in q as $\sigma \rightarrow \infty$:

$$\sup_{q \in \mathcal{C}} v_q \leq \lambda \left(p_g^2 \|b_g\|_\infty^2 + p_{gg} p_g \|b_g\|_\infty \|b_{gg}\|_\infty + p_{gg}^2 \|b_{gg}\|_\infty^2 \right) \|f_Q\|_\infty \frac{|\mathcal{C}|}{\sigma}.$$

Thus by bounded convergence theorem, the expected stationary virtual workload of the discretized polling system converges towards that of the continuous polling system as $\sigma \rightarrow \infty$.

Thus the limit of the discrete stationary expected virtual workload, V_{mix}^σ of (9), given by (7) represents the expected stationary virtual workload corresponding to the continuous system. This completes the proof of Theorem 1.

In the coming sections, we analyze an FWLAN using the results of this section.

3 Ferry based wireless local area network (FWLAN)

Static users are scattered in a geographical area Δ . The network is sparse, there is no direct global connectivity and a message ferry facilitates the basic communication. FWLAN can operate in various architectures (in Kavitha and Altman 2009 various architectures are discussed) and we consider here a basic architecture, referred as Sensor Area Network (SAN) in Kavitha and Altman (2009). This configuration⁸ supports only data transfer between the users of the area and the outer world via a static gateway, base station (BS). The data transfer from the users of the area to the outer world is called uplink, while that from the outside world to the users of Δ is called downlink. Every cycle begins at the BS where: (1) it dumps to BS the uplink data, collected from the users of LAN in the previous cycle; (2) collects the downlink data from BS, to transfer in the ongoing cycle to the destined users (when it reaches a point on its path near the intended user).

The ferry “serves” the users: by “serves” a message we mean that the ferry transmits it if the connection is downlink (i.e. the message is destined to a user), or receives it if it is an uplink message.

Ferry's route The ferry moves in a closed cyclic path \mathcal{C} repeatedly with constant speed α . It stops as and when it encounters an user with a request. Every time it stops and restarts it spends an extra time t_a for acceleration. The shuttle continues with its journey after serving the encountered user. Each point in the cyclic path is a potential stop and every stop q in the path is assigned with a set of points $I(q) \subset \Delta$. These sets are assigned based on nearest point basis.

Arrival process Uplink/Downlink traffic arrives according to an independent marked point processes $\{\mathcal{T}_n, M_n\}$, where \mathcal{T}_n is the arrival time of the n th point and $M_n = [X_n, \eta_n]$ are the corresponding i.i.d. marks:

- \mathcal{T}_n is a Poisson point process with parameter λ ,
- X_n is the location of an arrival. It's distribution is given by P_X ,
- η_n is the size of the message. Its distribution can depend upon X the position of arrival and also on whether it is a uplink or a downlink arrival. It is finite with probability one and its first and second moments (conditioned on the arrival being at point x and further for uplink and downlink arrivals) are given by $\eta_b^u(x)$, $\eta_b^d(x)$, $\eta_b^{(2u)}(x)$, $\eta_b^{(2d)}(x)$.

⁸In a more complicated architecture (referred as Autonomous Network, AUN, in Kavitha and Altman (2009)), wherein the ferry also supports the local communication, i.e., the data transfers between users of Δ , is analyzed using continuous polling systems with rerouting in our paper (Kavitha 2011).

These arrivals⁹ can be uplink, downlink with probabilities $p_d, p_u = 1 - p_d$ respectively. That is, an uplink request arrives at rate $\lambda^u := p_u\lambda$ while the downlink at rate $\lambda^d := p_d\lambda$.

We use the following notations. The points in the two dimensional area Δ are represented by x (if it is a sample point) or X (if it is a random position). The points on the cyclic path \mathcal{C} are represented by q or Q . We shall use the superscript u or d to denote uplink or downlink.

Radio channel and service time The Ferry uses a wireless link to serve the users. It can receive/transmit the messages from/to the users at a distance of d from it at a rate $\kappa(d)$ for some decreasing function $\kappa(\cdot)$. Thus the total time required for transmitting a message of size η , when the user is located at $x \in \Delta$ and is associated with $q(x) \in \mathcal{C}$ is equal to its size divided by the service rate:

$$B(x) = \frac{\eta}{\kappa(\|q(x) - x\|)}.$$

The point $x \in \Delta$ is mostly associated with the closest point on the circular path \mathcal{C} , i.e., $q(x) := \arg \min_{q \in \mathcal{C}} \|q - x\|$, with $\|\cdot\|$ representing the Euclidean distance.

The ferry standing at $q \in \mathcal{C}$ has to serve all the users arriving in $I(q)$. Thus the first two moments of the overall time required for uplink service at a point $q \in \mathcal{C}$, $b_u(q)$ and $b_u^{(2)}(q)$ are given by (for all $q \in \mathcal{C}$):

$$\begin{aligned} b_u(q) &:= E [B(X)|X \in I(q), \text{uplink}], \\ b_u^{(2)}(q) &:= E [B(X)^2|X \in I(q), \text{uplink}]. \end{aligned}$$

The corresponding for downlink, $b_d(q)$ and $b_d^{(2)}(q)$, can be defined in a similar way.

Pareto optimality All the users are served as the ferry moves on \mathcal{C} . Any arrival at a point $x \in I(q) \subset \Delta$ has to wait for time $W(q)$ and the stationary expected values of these waiting times $E[W](q)$ in general can depend upon the point $q \in \mathcal{C}$ standing at which the ferry serves them. We are interested in designing an optimal route (among a given class of trajectories) which minimizes integral of the weighted expected waiting times at all the points in the cyclic path \mathcal{C} , as this would give a Pareto optimal solution (see definition of Pareto optimality and other details in Appendix C). Towards this we consider the problem of minimizing

$$\int_0^{|\mathcal{C}|} E[W](q)\zeta(q)dq$$

where $\zeta(q)$ is a positive weight.

We need to choose appropriate weights $\{\zeta(q)\}$. If $\zeta(q)$ is chosen to be the arrival density to the points that are handled by q , i.e. $\lambda P_{I(q)}$, then the integral corresponds to the total expected number of waiting users (by Little’s Theorem). If we further multiply the integrand at each point q by the expected service time at that point then we obtain the total expected virtual workload in the system. In this paper *the objective function to be minimized will always be the expected virtual workload (which is obtained in Sect. 2 as Theorem 1) and a ferry path which minimized the expected virtual workload is a Pareto optimal path for the multi-objective problem where the expected waiting times at different locations are to be minimized.*

⁹The analysis of Sect. 2.4 requires that at any time there can be at maximum one user waiting at a point. The examples considered in this section satisfy this assumption because of two reasons: (i) P_X is continuous, (ii) The paths considered are sufficiently smooth and hence (iii) we will have (will be seen in the coming paragraphs) a P_Q that would be continuous.

Optimizers We discuss design of optimal ferry paths \mathcal{C} and optimal partitioning of the area into line segments $\{I_q; q \in \mathcal{C}\}$. The aim of this subsection is to obtain these objects in an optimal way that minimizes the virtual workload. Solving this problem in complete generality will be a very difficult task. Hence we instead obtain optimal ferry path among a special class of ferry paths.

Prior to discussing the optimality issues one first needs to map the FWLAN to a continuous polling system, so that Theorem 1 can be used. This task is taken up immediately.

3.1 Mapping to a Mixed service Continuous polling system

We begin with identifying the components of the continuous polling system.

Server and path of the polling system: The ferry represents the server of the polling system. The ferry stops at a point q in its path only when there is a user with a (downlink/uplink) request anywhere on the segment $I(q)$. Thus the entire segment $I(q)$ is modeled as a point on the server's path, in an equivalent continuous polling system.

Service times: The time t_a for acceleration is required only when the ferry stops and hence is added to the equivalent service time. An arrival (η, X) is associated with the point $q(X) := q$ of the ferry route if X lies in the corresponding segment, i.e., if $X \in I(q)$. Thus service time of an arrival (η, X) is given by:

$$B(X, \eta) = \frac{\eta}{\kappa(|X - q(X)|)} + t_a.$$

The expected service times (and the corresponding second moment), for example, for uplink arrivals corresponding to a point $q \in \mathcal{C}$ is

$$\begin{aligned} b_u(q) &= E[B(X, \eta) | q(X) = q] = E_X \left[\frac{\eta}{\kappa(|q - X|)} \middle| X \in I(q), \text{uplink} \right] + t_a, \\ b_u^{(2)}(q) &= E[B(X, \eta)^2 | q(X) = q] \\ &= E_X \left[\frac{\eta^2}{\kappa(|q - X|)^2} \middle| X \in I(q), \text{uplink} \right] - t_a^2 + 2t_a b_u(q). \end{aligned}$$

Service types: Ferry collects all the downlink messages from the BS and distributes the same in the current cycle. Thus the downlink arrivals experience globally gated service, with BS being the global gate point, 0. On the other hand, the ferry collects uplink data from the users as and when it encounters one on its path. Thus the uplink arrivals experience exhaustive/gated service and therefore $p_{gg} = p_d, p_g = p_u$.

Position of arrival in the 'polling system': The position of arrival in FWLAN is given by P_X which is a distribution over Δ . Every arrival in the segment $I(q)$ marks the arrival at point q of \mathcal{C} in the equivalent polling system. The probability distribution,

$$P_Q(A) := P_X \left(\bigcup_{q \in A} I(q) \right) \quad \text{for any Borel set } A \subset \mathcal{C},$$

represents this arrival distribution. For example, if $\Delta := \{x \in \mathcal{R}^2 : h_1^2 \leq |x| \leq h_2^2\}$, $\mathcal{C} = \{q : \|q\| = l\}$ for some $h_1 \leq l \leq h_2$ as in Fig. 1, $P_X \sim \mathcal{U}(\Delta)$ (Uniform distribution) and $I(q) = \{x : \angle x = \angle q\}$ (with $\angle x$ representing the angle made by the line joining 0, x with the x axis), then P_Q will be uniform distribution on \mathcal{C} .

Fig. 1 One Ferry in an annular ring

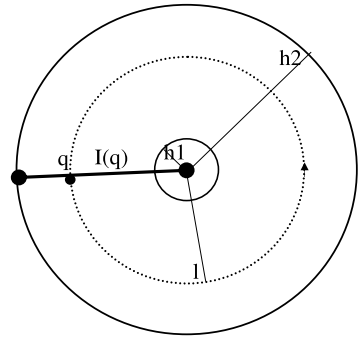
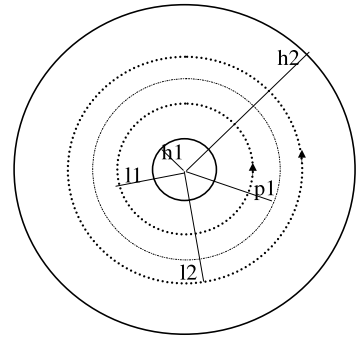


Fig. 2 Two Ferries in an annular ring



Thus the FWLAN can be modeled by a mixed service continuous polling system, Theorem 1 can be applied and the stationary expected virtual workload of the ferry with combined uplink and downlink arrivals can be calculated using (7) for any given cyclic path \mathcal{C} and the corresponding line segments $\{I_q\}_{q \in \mathcal{C}}$. The theorem can be applied only for those cases which satisfy the hypothesis of the theorem, like for example the service times should be bounded with probability one by $\hat{\mathbf{B}}$ of Theorem 1. The service times will be bounded as required if: (1) the function κ is continuous in d (which usually is the case) and hence its image of a bounded set will be bounded; (2) we will need $P(\eta < \hat{\eta}) = 1$ with an $\hat{\eta}$ that satisfies, $\hat{\eta} / \min \kappa(d) + t_a \leq \hat{\mathbf{B}}$ of Theorem 1.

In the following subsections, we design optimal FWLAN for a simple example. The ferry moves inside an annular ring $\Delta := \{x \in \mathcal{R}^2 : h_1^2 \leq |x| \leq h_2^2\}$ of Figs. 1 and 2. Ferry moves in a predefined closed cyclic path \mathcal{C} and each point q in the cyclic path is assigned with a line segment $I(q) \subset \Delta$, which is either obtained using the nearest point basis (in Sect. 3.2) or is computed in some optimal sense (in Sect. 3.3). In Sect. 3.3, we consider N circular paths with N ferries each moving in its own circular path. All the paths are optimized jointly so as to further obtain the segments $\{I(q)\}$ optimally.

3.2 SAN-TAXI example 1: Ferry moving in one circular path

The ferry moves on one concentric circle of radius l in the annular ring (Fig. 1), i.e., $\mathcal{C}_l = \{q \in \mathcal{R}^2 : \|q\| = l\}$. In this case, the best possible thing is to associate every arrival (η, X) with the nearest point of the ferry route, i.e., with the point $Q(X) := \arg \min_{q \in \mathcal{C}_l} |q - X|$ of ferry route. Hence the line segments will be angular segments, $I(q) = \{x \in \Delta : \angle(x) =$

$\angle(q)$ for all $q \in C_l$, (see Fig. 1). We would now like to find the optimal radius of this circle. The variable to be optimized, the radius, takes any value in $[h_1, h_2]$ and we are interested in

$$l^* = \arg \min_{l \in [h_1, h_2]} V_{SAN-TAXI}^1(l; h_1, h_2)$$

where $V_{SAN-TAXI}^1(l; h_1, h_2)$ represents the stationary expected workload in FWLAN when ferry moves in C_l , which can be computed using (7). To simplify the computations, in this section, we restrict ourselves to symmetric conditions, i.e.,

- uniform arrivals in annular area, $P_X \sim \mathcal{U}(\Delta)$,
- the rate function resulting from the losses in wireless medium considers only the direct path for attenuation (assuming a height difference of 1 unit between the transmitting and the receiving antennae and with path loss factor β):

$$\kappa(d) = (1 + d^2)^{-\beta/2},$$

- the distribution of η is independent of the location of arrival and the type of arrival (uplink or downlink) with $E[\eta|X] \equiv \eta_b$ and $E[\eta^2|X] \equiv \eta_b^{(2)}$ and such that $P(\eta < \hat{\eta}) = 1$ with an $\hat{\eta}/\kappa(d_{max}) + t_a < \hat{B}$, where $d_{max} := (h_2 - h_1)/2$,
- The uplink and downlink arrivals occur with equal probabilities $p_{ggs} = p_{gs} = 1/2$.

We note that the hypothesis of the Theorem 1 are satisfied under the symmetric conditions. Since the line segments $\{I(q)\}$ in this case are angular segments, for calculating the moments of the equivalent expected service times of a polling system $\{b(q; l)\}_{q \in C_l}$ one will require the radius $R = \|X\|$ of the arrival. Under the assumptions of this section, radius R is distributed as $R \sim 2rdr/(h_2^2 - h_1^2)$ and further it is easy to see that the moments of the overall service time will be independent of the position $q \in C_l$ but depend upon the ferry path radius l . The service moments are:

$$b_l = b(q; l) = \eta_b \int_{h_1}^{h_2} (1 + (r - l)^2)^{\beta/2} \frac{2rdr}{h_2^2 - h_1^2} + t_a,$$

$$b_l^{(2)} = b^{(2)}(q; l) = \eta_b^{(2)} \int_{h_1}^{h_2} (1 + (r - l)^2)^\beta \frac{2rdr}{h_2^2 - h_1^2} - t_a^2 + 2t_a b_l, \quad \text{for all } q \in C_l.$$

For example for $\beta = 2$ the above can be computed as,

$$b_l = \eta_b \left(\frac{1}{2}(h_1^2 + h_2^2 + 2) + l^2 - \frac{4l(h_1^2 + h_1 h_2 + h_2^2)}{3(h_1 + h_2)} \right) + t_a,$$

$$b_l^{(2)} = \frac{\eta_b^{(2)}}{15(h_2^2 - h_1^2)} \left(-5h_1^6 + 24h_1^5 l - 15h_1^4(3l^2 + 1) + 40h_1^3(l^3 + l) - 15h_1^2(l^2 + 1)^2 \right. \\ \left. + h_2^2(5h_2^4 - 24h_2^3 l + 15h_2^2(3l^2 + 1) - 40h_2(l^3 + l) + 15(l^2 + 1)^2) \right) - t_a^2 + 2t_a b_l.$$

Under the simplified assumptions of this section the stationary expected workload for this example can be calculated using (7):¹⁰

$$V_{SAN-TAXI}^1(l; h_1, h_2) = \frac{\lambda b_l \lambda b_l^{(2)}}{2(1 - \lambda b_l)} + \frac{\lambda b_l |C_l| \alpha^{-1} (12 - \lambda b_l)}{16(1 - \lambda b_l)}. \quad (20)$$

¹⁰The equation below is different from the one in Kavitha and Altman (2010a), because of the correction in (7).

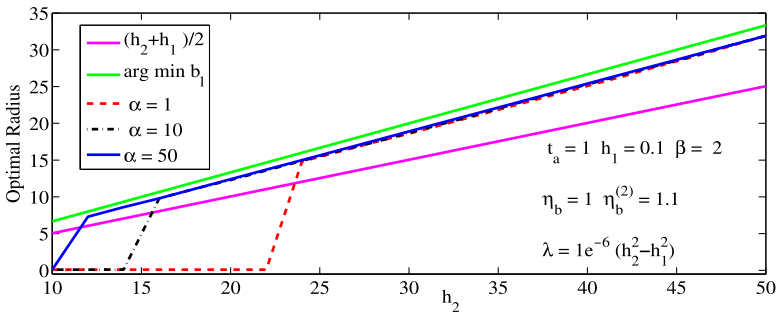


Fig. 3 Optimal orbit radius for a ferry moving continuously in an annular ring (Uplink and Downlink)

One has to perform numerical computations using the above formula to obtain the optimal radius and this is done below. However one can get the following asymptotic characteristics of l^* from the formula itself: (1) as the propagation coefficient β tends to zero, both the service moments b_l and $b_l^{(2)}$ become independent of l , stationary expected workload essentially depends upon l only via $|C_l|$ and hence l^* tends to h_1 , i.e., the optimal path for the ferry is the inner circle; (2) as the speed of the ferry, α , increases to infinity, the second term in the formula becomes negligible and hence optimal radius will be determined only by the service time moments and so the optimal radius will be above the middle of the ring, i.e., larger than $(h_1 + h_2)/2$.

Mixed (Uplink + Downlink) service example: In Fig. 3 we consider a FWLAN which supports both uplink and downlink services. The parameters of the FWLAN are mentioned in the figure itself and we plot the optimal radius as a function of the outer radius h_2 of the annular ring. We notice again that the optimal radius tends towards the inner circle either as the speed of the ferry decreases or as h_2 reduces. We also plot the optimizer of the expected service time,

$$l_b^* := \arg \min_{l \in [h_1, h_2]} b_l,$$

in the same figure and we note that the optimal radius is close to this optimizer whenever either the annular ring is large or whenever the speed of the ferry is large. This can also be understood by studying the expression for virtual workload (20). Hence in these scenarios one can obtain the optimal radius as (for $\beta = 2$)

$$l^* \approx \frac{2(h_1^2 + h_2^2 + h_1 h_2)}{3(h_1 + h_2)}.$$

3.3 SAN-TAXI example 2: N servers moving simultaneously in the annular area

When the annular ring spans over a large area, a ferry moving in one circular path may not be optimal. Hence we would like to consider a ferry moving in N circular paths before completing the area. To make things simpler we assume that there will be N ferries moving simultaneously (but independent of each other) with the same speed α respectively on circular paths of radii l_1, l_2, \dots, l_N (see Fig. 2). Without loss of generality we also assume that the BS is placed on the entire segment $\{x : \angle x = 0\}$, i.e., any ferry communicates with the BS instantaneously the moment it touches the angle 0. Throughout this paper we are neglecting the time taken for communication (e.g., the communication happens at high speed using a wired link) between the BS and the ferry(ies). Let $C_n := \{q : \|q\| = l_n\}$ and let

Table 1 Example: 2 Ferries moving in annular area with $\beta = 2$, $h_1 = 5$, $t_a = 10$, $\eta_b = 2$, $\eta_b^{(2)} = 4.4$, $\lambda = 1e^{-2}$

h_2	$\alpha = 100$ (l_1^*, p^*, l_2^*)	$\alpha = 10$ (l_1^*, p^*, l_2^*)	$\alpha = 1$ (l_1^*, p^*, l_2^*)
10	6.3, 7.7, 8.8	6.1, 7.7, 8.6	6, 7.8, 8.6
15	7.9, 10.4, 12.7	7.8, 10.6, 12.7	7.8, 10.8, 12.8
20	9.5, 13.1, 16.7	9.6, 13.5, 16.8	9.6, 13.8, 16.9
25	11.1, 15.8, 20.6	11.3, 16.3, 20.8	11.4, 16.7, 20.9
30	12.8, 18.5, 24.6	13, 19, 24.7	13.2, 19.5, 24.9
35	14.4, 21.1, 28.5	14.6, 21.5, 28.6	14.9, 22.1, 28.8

$\mathcal{B}(x, p) := \{y : \|y - x\| < p\}$ represent the open ball, while $\bar{\mathcal{B}}(x, p)$ represent its closure. The first ferry caters to the users in the annular ring $\bar{\mathcal{B}}(0, p_1) - \mathcal{B}(0, h_1)$ while the last one covers the annular ring $\bar{\mathcal{B}}(0, h_2) - \mathcal{B}(0, p_{N-1})$. Any intermediate n th ferry covers the area $\bar{\mathcal{B}}(0, p_n) - \mathcal{B}(0, p_{n-1})$. Let $p_0 := h_1$ and $p_N := h_2$. In this subsection we choose optimally $\mathbf{p} := [p_1, p_2, \dots, p_{N-1}]^T$ and $\mathbf{l} := [l_1, l_2, \dots, l_N]^T$, which in turn optimally defines the line segments $\{I_n(q); q \in \mathcal{C}_n, \text{ for all } n\}$ as

$$I_n(q) = \{x \in \bar{\mathcal{B}}(0, p_n) - \mathcal{B}(0, p_{n-1}) : \angle q = \angle x\}.$$

We map each ferry by a separate continuous polling system. With these, the moments of the overall service times for the n th ferry is

$$b_{n,\mathbf{p},\mathbf{l}} = \eta_b \int_{p_{n-1}}^{p_n} (1 + (r-l)^2)^{\beta/2} \frac{2rdr}{p_n^2 - p_{n-1}^2} + t_a \quad \text{and}$$

$$b_{n,\mathbf{p},\mathbf{l}}^{(2)} = \eta_b^{(2)} \int_{p_{n-1}}^{p_n} (1 + (r-l)^2)^\beta \frac{2rdr}{p_n^2 - p_{n-1}^2} - t_a^2 + 2t_a b_{n,\mathbf{p},\mathbf{l}}.$$

Using these moments one can estimate $\{V_{SAN-TAXI}^1(l_n; p_{n-1}, p_n)\}$, the N virtual workloads as a function of the vectors \mathbf{p}, \mathbf{l} . We now propose to choose the optimal ferry paths (and line segments $\{I(q)\}$) by:

$$(\mathbf{p}^*, \mathbf{l}^*) = \arg \min_{\mathbf{p}, \mathbf{l}} \sum_{n=1}^N \frac{p_n^2 - p_{n-1}^2}{h_2^2 - h_1^2} V_{SAN-TAXI}^1(l_n; p_{n-1}, p_n).$$

In the following we consider an example with 2 ferries moving in the annular area and supporting mixed service. The results are in Table 1. We obtain the optimal \mathbf{l}^*, p^* for different values of the outer radius h_2 and the common speed of the ferries α while the rest of the parameters are kept constant. We see again that the optimal radii of the ferry routes move towards the inner circle as the speed of the vehicle reduces and this effect is pronounced as the radius of the outer circle, h_2 , decreases. This effect is also seen in the case of single server and can be partially explained using the expression $V_{SAN-TAXI}$ (20) for the virtual workload:

- $V_{SAN-TAXI}$ depends upon the radius l via the moments of the effective service times $b^{(1)}$, $b^{(2)}$ and the circumference $|C|$. Virtual workload $V_{SAN-TAXI}$ (20) is directly proportional to all three factors.
- The influence of l on the two moments depends upon the path loss factor β . If β is close to zero the moments are almost independent of l while the optimal l (optimal for the service

moments) will be above the center of the annular ring for practical path loss factors. For practical values of path loss factors, the two moments reduce as l increases from the inner radius.

- However the circumference $|C|$ has a contrast effect, it increases linearly with l .
- For large values of speed α , the influence of $|C|$ on $V_{SAN-TAXI}$ reduces and hence the optimal radii will be above the center of the annular ring.
- For smaller values of speed α , we see the reverse effect.
- The area of the annular ring also influences, with larger areas the moments (especially the second) will be large and hence the influence of the moments on $V_{SAN-TAXI}$ will be larger than that of the circumference $|C|$. And so, we see from the bottom rows of the table that, the optimal radii are away from the inner circle even for lower speeds.

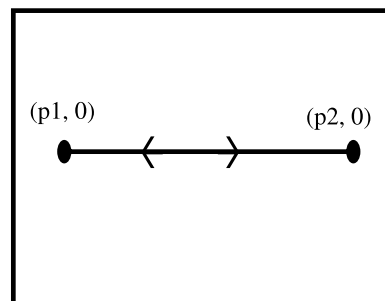
4 FWLAN with two base stations

We now consider a LAN equipped with two base stations and a moving ferry. These base stations have also to be placed in optimal locations along with the design of optimal ferry routes. In this section, we consider downlink service alone. Before we proceed with FWLAN examples with two base stations, we note here that many of the two base station examples can result in a special case of service called elevator polling (see Altman et al. 1992 which describes the scheme for discrete polling system). This polling system will have ‘fair’ behavior: the expected waiting times at all the points in the server route will be equal irrespective of whether the arrivals are symmetric or non symmetric. This ‘fair’ behavior is a welcome feature and it might be advantageous to consider the paths and the BS locations that can result in this behavior. In the Sect. 2.3 we derived the expression for the stationary expected waiting time in a continuous polling system with elevator gating. These results are applied to some interesting FWLAN examples with two base stations in the following.

4.1 Ferry moving in a straight line in Rectangular area

We consider a rectangular area $\Delta = [-l, l] \times [-d, d]$. A ferry moves in the horizontal straight line as shown in Fig. 4. There are two base stations, each one of them is located at the end of the ferry path. The ferry, say, starts its cycle by collecting all the downlink data from the BS on the left, tags the users for whom the downlink data has to be distributed, moves towards right with speed α , distributes the data to all the tagged users. Once reaching the BS on the right it collects all the data from the BS and then moves left. This procedure repeats. We analyze this FWLAN using the elevator polling system of Sect. 2.3.

Fig. 4 One Ferry in Rectangular area with two base stations



Let the base stations be placed respectively at $(p_1, 0)$, $(p_2, 0)$. Let $\mathbf{p} := (p_1, p_2)$. In this case the ferry path is given by $C_{\mathbf{p}} = \{(q, 0) : p_1 \leq q \leq p_2\}$. The partitions are calculated based on the nearest distance criterion.

$$I(q) = I(q, 0) = \begin{cases} \{q\} \times [-d, d] & p_1 < q < p_2, \\ [-l, p_1] \times [-d, d] & q = p_1, \\ [p_2, l] \times [-d, d] & q = p_2. \end{cases}$$

Note in this case that, both the ends of the ferry path cover not just a line segment but a rectangular segment. We make simplifying assumptions as in Sect. 3. Under these assumptions

$$b(q; \mathbf{p}) = \begin{cases} \eta_b \int_{-d}^d (1 + y^2)^{\beta/2} \frac{dy}{2d} + t_a & p_1 < q < p_2, \\ \eta_b \int_{-l}^{p_1} \int_{-d}^d (1 + y^2 + (p_1 - x)^2)^{\beta/2} \frac{dx dy}{4ld} + t_a & q = p_1, \\ \eta_b \int_{p_2}^l \int_{-d}^d (1 + y^2 + (p_2 - x)^2)^{\beta/2} \frac{dx dy}{4ld} + t_a & q = p_2, \end{cases}$$

$$b^{(2)}(q; \mathbf{p}) - 2b(q; \mathbf{p})t_a + t_a^2 = \begin{cases} \eta_b^{(2)} \int_{-d}^d (1 + y^2)^{\beta} \frac{dy}{2d} & p_1 < q < p_2, \\ \eta_b^{(2)} \int_{-l}^{p_1} \int_{-d}^d (1 + y^2 + (p_1 - x)^2)^{\beta} \frac{dx dy}{4ld} & q = p_1, \\ \eta_b^{(2)} \int_{p_2}^l \int_{-d}^d (1 + y^2 + (p_2 - x)^2)^{\beta} \frac{dx dy}{4ld} & q = p_2. \end{cases}$$

Note here that even under symmetric assumptions the overall service moments are not equal at all points, thus resulting in an asymmetrical polling system. With the above the overall service moments are given by,

$$b_{\mathbf{p}} = 2d(l - p_2)b(p_1; \mathbf{p}) + 2d(p_1 + l)b(p_2; \mathbf{p}) + 2d(p_2 - p_1)b(p_1+; \mathbf{p}),$$

$$b_{\mathbf{p}}^{(2)} = 2d(l - p_2)b^{(2)}(p_1) + 2d(p_1 + l)b^{(2)}(p_2) + 2d(p_2 - p_1)b^{(2)}(p_1+; \mathbf{p}).$$

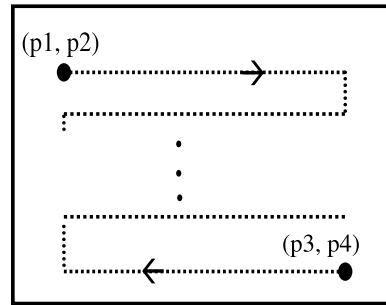
In the above by $b(p_1+; \mathbf{p})$ we meant the expected value at any point with its x component greater than p_1 . As the expected weighting times in this case are equal at all the points, it will be Pareto optimal to optimize the expected weighting times themselves. Further in this symmetric case it is easy to see that $p_1^* = -p_2^*$. Thus,

$$p_2^* = \arg \min_{0 < p_2 \leq d} \frac{\lambda b_{-p_2, p_2}^{(2)}}{2(1 - \lambda b_{-p_2, p_2})} + 2p_2 \alpha^{-1}.$$

Note here that optimal choice of p_2^* optimally places the two base stations as well as optimally designs the ferry path. In this example, it is easy to see that: (1) As the speed of the ferry α increases, the p_2^* moves towards l , i.e., if the ferry can move at large speeds, it is optimal to place the base stations at both the edges of the area and allow the ferry to traverse the entire width of Δ . This conclusion is also true whenever the path loss factor β is large. (2) If the ferry moves at moderate speeds and or the path loss factor is small p_2^* reduces. Thus for small path loss factors, the ferry can move in shorter line segments and cover the entire area optimally.

We simplified the problem substantially by considering that the ferry moves in a horizontal line placed at the center of the area. This simplification is good as long as the arrivals are uniform and the file sizes are distributed the same way at all points. For asymmetrical arrivals this choice of ferry route may not be good. In this case it might be better to consider any general straight line passing through points (p_1, p_2) and (p_3, p_4) as ferry path. The base

Fig. 5 More general paths for the ferry with two base stations



stations are placed once again at both ends of the ferry path. One can estimate the expected waiting time in terms of the parameter $\mathbf{p} := (p_1, p_2, p_3, p_4)$ in a similar way and choose \mathbf{p}^* optimally. This \mathbf{p}^* places the base stations optimally and designs the ferry routes optimally in more generality. If the breadth of $\Delta, 2d$, is large it might be better to consider zig zag ferry paths (Fig. 5) with two base stations placed at (p_1, p_2) and (p_3, p_4) .

5 Conclusions

We study continuous polling systems that cater to non-standard variants of gating disciplines like, globally gated service, gated service, mixed service (a mixture of globally gated and gated service disciplines) and elevator service disciplines. We obtain an expression for the expected stationary workload. We obtain this result under more general conditions than the usual symmetric conditions. We come up with a way of discretization such that the available Pseudo conservation laws of discrete polling systems can be utilized for obtaining the results for the continuous counterparts. We expressed the expected workload as a parametrized function of moment fixed points. The later are some stationary moments obtained as fixed points of a function defined on spaces of left continuous functions with right limits equipped with supremum norm and which are further parametrized by the number of discretization levels. We show the required convergence via the continuous dependence of the fixed points on the parameter. This way we obtained a common expression, which represents the expected virtual workload for continuous as well as the discretized polling systems, at different values of the parameter. We then showed the continuity of the expected virtual workload with respect to the parameter and hence obtained the expected virtual workload for the continuous system as the limit of the expected virtual workloads of the discrete systems, when the levels of discretization tend to infinity. We applied these results to a wireless LAN in which a ferry assists data transfer between the users of the network and a base station which acts as a gateway to the external world. We provided the analytical expressions for the expected virtual workload in a system of FWLAN, taking into account the radio channel considerations, which allowed us to optimize the trajectories of the Ferry. We covered both uplink, downlink as well as their combination. Minimizing the virtual workload results in a Pareto minimization of the expected waiting times in the system. We have further presented some extensions to the case of several base stations and to routing of several ferries. We observe that by using two base stations, one can achieve fairness : the expected waiting times are independent of the positions of arrivals.

In Kavitha (2011), using similar approach we analyzed an FWLAN, in which the ferry also supports local communication. This FWLAN is analyzed using a continuous polling

system with rerouting, whose performance is once again obtained using the discretization approach. One can further explore this discretization approach.

Acknowledgements The reviewers have pointed out a gap in the proof of Theorem 3 (of the revision 1) and we corrected the same in the final manuscript. We are thankful to the reviewers for pointing out the gap and allowing us to correct the same.

Appendix A: Limit of the virtual workload, V_{mix}^σ (9)

Proof of Lemma 1 Some terms of (9) are independent of σ :

$$\begin{aligned}\rho &= \sum_{i=1}^{2\sigma} \lambda_i b_i = \lambda \int_0^{|\mathcal{C}|} b(q) P_Q(dq) = \lambda \bar{b}, \\ \sum_{i=1}^{2\sigma} \lambda_i b_i^{(2)} &= \lambda \bar{b}^{(2)} \quad \text{and} \\ \sum_{i=1}^{\sigma} \rho_{2i} &= \lambda p_{gg} E[B|\text{globally gated}] = \lambda p_{gg} \bar{b}_{gg}.\end{aligned}\tag{21}$$

Recall $i^\sigma = (i-1)|\mathcal{C}|/\sigma$, note $\delta^\sigma(q) = i^\sigma$ when $q \in I_i$ and with these,

$$\begin{aligned}\sum_i^{\sigma} \rho_{2i} \frac{(i-1)|\mathcal{C}| \alpha^{-1}}{\sigma} &= \lambda p_{gg} \alpha^{-1} \sum_i^{\sigma} E[b_{gg}(Q) 1_{\{Q \in I_i\}}] i^\sigma \\ &= \lambda p_{gg} \alpha^{-1} \sum_i^{\sigma} b_{gg}(i^\sigma) f_Q(i^\sigma) \frac{|\mathcal{C}|}{\sigma} i^\sigma\end{aligned}\tag{22}$$

$$+ \lambda p_{gg} \alpha^{-1} \int_0^{|\mathcal{C}|} (b_{gg}(q) f_Q(q) - b_{gg}(\delta^\sigma(q)) f_Q(\delta^\sigma(q))) \delta^\sigma(q) dq\tag{23}$$

$$\xrightarrow{\sigma \rightarrow \infty} \lambda p_{gg} \alpha^{-1} \int_0^{\sigma} q b_{gg}(q) f_Q(q) dq = \lambda p_{gg} \alpha^{-1} E[Q b_{gg}(Q)].\tag{24}$$

The term (22) is a Riemann sum and hence converges to the integral (24) while the second term (23) converges to zero by continuity of the functions b_{gg} , f_Q via Bounded Convergence Theorem (BCT). Similarly, excluding the terms which converge to zero by BCT as in (23),

$$\begin{aligned}\sum_{i=1}^{2\sigma} \rho_i^2 &\approx \frac{\lambda |\mathcal{C}|}{\sigma} \sum_{i=1}^{\sigma} (p_g b_g^2(i^\sigma) + p_{gg} b_{gg}^2(i^\sigma)) f_Q^2(i^\sigma) \frac{|\mathcal{C}|}{\sigma} \\ &\rightarrow 0 \int_0^{\sigma} (p_{gg} b_{gg}^2(q) + p_g b_g^2(q)) f_Q^2(q) dq = 0.\end{aligned}\tag{25}$$

Following similar logic, with $\widehat{b}_g(q) := E[b_g(Q)1_{\{Q \leq q\}}]$ and from (21):

$$\begin{aligned} \sum_i^\sigma \rho_{2i} \sum_{j \leq 2i} \rho_j &= \sum_{i=1}^\sigma \rho_{2i} \sum_{j \leq i} \rho_{2j} + \sum_{i=1}^\sigma \rho_{2i} \sum_{j \leq i} \rho_{2j-1} \\ &= \frac{1}{2} \left(\sum_{i=1}^\sigma \rho_{2i} \right)^2 + \frac{1}{2} \sum_i^\sigma \rho_{2i}^2 + \lambda^2 p_{gg} p_g \sum_{i=1}^\sigma b_{2i} l_i \sum_{j \leq i} b_{2j-1} l_j \\ &\rightarrow \frac{\lambda^2}{2} \left((p_{gg} \bar{b}_{gg})^2 + p_{gg} p_g E[b_{gg}(Q) \widehat{b}_g(Q)] \right). \end{aligned} \tag{26}$$

By independence $E[b_{gg}(Q) \widehat{b}_g(Q)] = \bar{b}_{gg} \widehat{\bar{b}}_g$ and hence from limits (21)–(26), the limit of V_{mix}^σ (9) matches with the formula (7) of the Theorem 1. \square

Appendix B: Proofs

Proof of Lemma 2 Let $T(q) \equiv T$, with $E[T] = \tau$. In this case, $\mathcal{T}([a, c], T)$ is sum of the service times of all the arrivals in the interval $[0, T]$ that occurred in the segment $[a, c] \subset [0, |C|]$. Let \mathcal{N} represent the number of those Poisson arrivals that belong to segment $[a, c]$ and that arrived during the $[0, T]$ interval. In this case,

$$E[\mathcal{N}] = E[E[\mathcal{N}|T]] = \lambda P_Q([a, c]) E[T] = \lambda \tau \int_a^c f_Q(q) dq.$$

Let $\{B_i\}_{i \leq \mathcal{N}}$ represent the service times of these \mathcal{N} arrivals. Any arrival can fall in point q according to distribution $P_Q(dq) = f_Q(q) dq$ and the conditional expected service time of the arrival conditioned on the point of arrival is $b(q)$ and so,

$$E[B_i | \text{Arrival in } [a, c]] = \frac{1}{P_Q([a, c])} \int_a^c b(q) f_Q(q) dq.$$

Thus by Wald’s lemma,¹¹

$$E[\mathcal{T}([a, c], T)] = E \left[\sum_{i=1}^{\mathcal{N}} B_i \right] = E[\mathcal{N}] E[B_i] = \lambda \tau \int_a^c b(q) f_Q(q) dq. \tag{27}$$

Now we consider the other case, i.e., let every sample path of the random function T be a monotone increasing function of $q \in [a, c]$. The proof for T , monotone decreasing, follows

¹¹ For applying Wald’s lemma we need certain independence assumptions. These assumptions are satisfied in this proof (whenever Wald’s lemma is applied) because of the following: the service times (and the number) of incoming arrivals are independent of the system evolution till the time stamp at which they are served. For example, the number of arrivals (or those tagged in case of globally gated service) at around a point say q (i.e., the users awaiting service in segment $[q, q + dq]$) does not influence the system evolution till the server reaches point $\delta^\sigma(q)$, standing at which it would serve the awaiting users, i.e., till time $T_n^\sigma(q)$. To put it the other way, for example, $T_n^\sigma(q)$ is a stopping time for that sampled arrival process (sampled from the original gated Poisson arrival process), in which the users land in segment $[q, q + dq]$ and which originated during the time interval $C_n^\sigma(q)$.

in similar lines. For every M , define step function,

$$\delta^M(q) = a + \frac{i}{M} \text{ if } q \in a + \left[\frac{i}{M}, \frac{i+1}{M} \right).$$

Note that by monotonicity of T ,

$$\begin{aligned} \underline{T}^M(q) &:= T(\delta^M(q)) \leq T(q) \quad \text{and} \\ \bar{T}^M(q) &:= T\left(\delta^M(q) + \frac{1}{M}\right) \geq T(q). \end{aligned}$$

In the following the arrivals and their service times are independent¹² of T , \bar{T}^M and \underline{T}^M . More arrivals can occur during more time and hence workload during larger time will be larger. Hence and further using (27)

$$\begin{aligned} E[\mathcal{T}([a, c], T)] &\leq E[\mathcal{T}([a, c], \bar{T}^M)] \\ &= \lambda \sum_{i=0}^{M-1} \tau\left(a + \frac{i}{M}\right) \int_{a+i/M}^{a+(i+1)/M} f_Q(q)b(q)dq \\ &\leq \lambda \sum_{i=0}^{M-1} f_Q\left(a + \frac{i}{M}\right) b\left(a + \frac{i}{M}\right) \tau\left(a + \frac{i}{M}\right) \frac{1}{M} \\ &\quad + \lambda \int_a^c \tau(\delta^M(q)) (f_Q(q)b(q) - b(\delta^M(q)) f_Q(\delta^M(q))) dq \\ &\rightarrow \lambda \int_a^c f_Q(q)\tau(q)b(q)dq \quad \text{as } M \rightarrow \infty. \end{aligned}$$

The last limit is obtained because the partial sums define Riemann sums and they converge to the integral and because the second integral goes to zero by uniform continuity of b , f_Q functions (continuous function on compact intervals are uniformly continuous). Similarly using the lower bound on the time T , \underline{T}^M , we get,

$$E[\mathcal{T}([a, c], T)] \geq \lim_{M \rightarrow \infty} E[\mathcal{T}([a, c], \underline{T}^M)] = \lambda \int_a^c f_Q(q)\tau(q)b(q)dq. \quad (28)$$

Hence,

$$E[\mathcal{T}([a, c], T)] = \lambda \int_a^c f_Q(q)\tau(q)b(q)dq. \quad \square$$

Proof of Theorem 2 The operator \mathcal{F} given by (14) is reproduced here.

$$\mathcal{F}(\tau; \delta)(q) := \delta(q)\alpha^{-1} + \lambda\tau(|C|) \int_0^{\delta(q) \wedge |C|} b(y)f_Q(y)dy.$$

¹² The independence assumption holds for lower bound \underline{T}^M and the actual time T readily because of the reasons given in footnote 11. An upper bound is obtained by calculating the independent arrivals that would have arrived in a bigger interval i.e., \bar{T}^M .

We are interested in its fixed point. The function is clearly jointly continuous in τ, δ . Further,

$$\mathcal{F}(\tau_1; \delta)(q) - \mathcal{F}(\tau_2; \delta)(q) = \lambda (\tau_1(|\mathcal{C}|) - \tau_2(|\mathcal{C}|)) \int_0^{\delta(q) \wedge |\mathcal{C}|} b(y) f_Q(y) dy.$$

Thus for all δ

$$\|\mathcal{F}(\tau_1; \delta) - \mathcal{F}(\tau_2; \delta)\|_\infty \leq \rho \|\tau_1 - \tau_2\|_\infty.$$

Thus whenever $\rho < 1$, we have a fixed point of the operator \mathcal{F} for all δ and hence the existence of unique stationary first moments. Under this condition, we further see by parametrized contraction mapping theorem (Berger 1977) that the unique fixed point τ_*^δ is continuous in δ .

In fact from (14), when evaluated at $q = |\mathcal{C}|$, we find that:

$$\tau_*^\delta(|\mathcal{C}|) = \frac{\delta(q)\alpha^{-1}}{1 - \lambda \int_0^{\delta(|\mathcal{C}|) \wedge |\mathcal{C}|} b(y) f_Q(y) dy}.$$

Note that

$$\sup_{\delta \in \mathcal{D}} \lambda \int_0^{\delta(|\mathcal{C}|) \wedge |\mathcal{C}|} b(y) f_Q(y) dy = \lambda \int_0^{|\mathcal{C}|} b(y) f_Q(y) dy = \rho$$

and thus *the stationary moments exist if and only if* $\rho < 1$. Under this condition, one can solve the fixed point equation and complete the proof of Theorem 2. □

Proof of Theorem 3 The cycle time C_n^σ defined in (15) can be rewritten in the following fashion and via this expression we obtain the convergence of the required second moments (note $C_n^\sigma(0) = C_{n-1}^\sigma(|\mathcal{C}|)$ almost surely by continuity arguments):

$$C_n^\sigma(q) = |\mathcal{C}| \alpha^{-1} + U_n^\sigma(q) + \mathcal{I}_g([0, \delta^\sigma(q)], C_n^\sigma) + \mathcal{I}_{gg}([0, \delta^\sigma(q)], C_n^\sigma(0)), \tag{29}$$

$$U_n^\sigma(q) = \mathcal{I}_{gg}([\delta^\sigma(q), |\mathcal{C}|], C_{n-1}^\sigma(0)) + \mathcal{I}_g([\delta^\sigma(q), |\mathcal{C}|], C_{n-1}^\sigma). \tag{30}$$

Required Spaces: We first identify appropriate complete metric spaces in which the above quantities reside and whose norm gives the required convergence. We start with some definitions/notations.

Let (Ω, \mathcal{F}, P) be any probability space.

Let $\|\cdot\|_2$ be the L^2 norm, i.e., $\|\psi\|_2 := \sqrt{E[\psi^2]}$ for any random variable ψ .

Let \mathcal{D} represent the space of left continuous functions on $[0, |\mathcal{C}|]$ with right limits and equipped with supremum norm, $\|f\|_\infty := \sup_{q \in [0, |\mathcal{C}|]} |f(q)|$.

Let

$$\Psi := \{ \text{measurable functions } \psi : [0, |\mathcal{C}|] \times \Omega \rightarrow \mathcal{R} \text{ such that}$$

$$\text{the map } q \mapsto \|\psi(q, \cdot)\|_2 \in \mathcal{D} \}$$

$$\text{where } \|\psi(q, \cdot)\|_2 := \sqrt{\int \psi(q, \omega)^2 dP(\omega)} = \sqrt{E[\psi(q, \cdot)^2]}.$$

Define the following norm on this space:

$$\|\psi\|_{\infty,2} := \sup_{q \in [0, |C|]} \|\psi(q, \cdot)\|_2 \quad \text{for any } \psi \in \Psi.$$

By Lemma 3 ($\Psi, \|\cdot\|_{\infty,2}$) is a Banach space.

Contraction Mappings and fixed points: It is easy to see that (C_n^σ, U_n^σ) defined in (29)–(30) reside¹³ in Ψ for all σ . It is also easy to see that these under stationarity form the fixed point of the following function, defined point-wise (i.e., for every q) by:

$$\begin{aligned} \Theta_C^\sigma(\psi_c, \psi_u; \sigma)(q) &= |C|\alpha^{-1} + \sqrt{\tilde{B}}\psi_u(q) + \mathcal{T}_g([0, \delta^\sigma(q)], \psi_c) \\ &\quad + \mathcal{T}_{gg}([0, \delta^\sigma(q)], \psi_c(|C|)), \end{aligned} \quad (31)$$

$$\Theta_U^\sigma(\psi_c, \psi_u; \sigma)(q) = \frac{1}{\sqrt{\tilde{B}}}\mathcal{T}_{gg}([\delta^\sigma(q), |C|], \psi_c(|C|)) + \frac{1}{\sqrt{\tilde{B}}}\mathcal{T}_g([\delta^\sigma(q), |C|], \psi_c). \quad (32)$$

In the above, $\mathcal{T}_m([a, c], \psi)$ (with $m = g$ or gg) represents the sum of the modified service times, \tilde{B} , of the users that arrived in the interval $[a, c]$: 1) arrivals at around q (independent arrivals as explained in footnote 11) occur for time period $|\psi(q, \cdot)|$ as in Sect. 2.4.2 and 2) if $\psi(q, \cdot)$ is negative then the quantities added, \tilde{B} , are negative of the service times, else \tilde{B} are the service times themselves.

Let $\Theta := (\Theta_C, \Theta_U)$. The function $\Theta : \Psi^2 \rightarrow \Psi^2$ is affine linear (linear but for $|C|\alpha^{-1}$ term). Let Θ_{Lin} represent the linear part. With this,

$$\begin{aligned} \|\Theta(\psi_c, \psi_u; \sigma) - \Theta(\psi'_c, \psi'_u; \sigma)\|_{\infty,2} &= \|\Theta_{Lin}(\psi_c - \psi'_c, \psi_u - \psi'_u)\|_{\infty,2} \\ &\leq \sqrt{\tilde{B}}\|\psi_u - \psi'_u\|_{\infty,2} + \|\mathcal{T}_g([0, \delta^\sigma(q)], \psi_c - \psi'_c)\|_{\infty,2} \\ &\quad + \|\mathcal{T}_{gg}([0, \delta^\sigma(q)], \psi_c(|C|) - \psi'_c(|C|))\|_{\infty,2} \\ &\quad + \frac{1}{\sqrt{\tilde{B}}}(\|\mathcal{T}_g([\delta^\sigma(q), |C|], \psi_c - \psi'_c)\|_{\infty,2} \\ &\quad + \|\mathcal{T}_{gg}([\delta^\sigma(q), |C|], \psi_c(|C|) - \psi'_c(|C|))\|_{\infty,2}). \end{aligned}$$

In the above, terms like $\|\mathcal{T}_g([0, \delta^\sigma(q)], \psi_c)\|_{\infty,2}$ refer to the $\|\cdot\|_{\infty,2}$ norm of a $\psi \in \Psi$, where $\psi(q, \cdot)$ represents the random workload, $\mathcal{T}_g([0, \delta^\sigma(q)], \psi_c)$. Note that for any type of arrivals (i.e., with $m = g$ or $m = gg$) and for any interval I and time function T ,

$$|\mathcal{T}_m(I, T)| \leq \sum_i B_i$$

where B_i are the service times of the arrivals in the chosen time period $|T|$ and so it is clear that these type of terms converges (in monotone fashion) to zero as $\hat{B} \rightarrow 0$. In fact even the terms like $1/\sqrt{\hat{B}}|\mathcal{T}_m(\cdot, \cdot)|$ drop to zero as $\hat{B} \rightarrow 0$. Thus there exists a $\hat{\mathbf{B}} < \infty$ such that

$$\|\Theta_{Lin}\|_{\infty,2} < 1 \quad \text{for all } \hat{B} < \hat{\mathbf{B}} \text{ and for all } \sigma.$$

¹³ C_n^σ is a sum of two monotone functions (recall from (15), $C_n(q) = T_n(q) + [T_{n-1}|C| - T_{n-1}(q)]$) while U_n^σ is a monotone function of q for all realizations, hence their moments will also have similar properties and hence both the moments are in \mathcal{D} .

Hence Θ is a contraction for all σ and the contraction coefficient can also be chosen independent of σ , i.e., there exists $\hat{\mathbf{B}}$ and $L < 1$ such that (for all $\hat{B} < \hat{\mathbf{B}}$),

$$\|\Theta(\psi_c, \psi_u; \sigma) - \Theta(\psi'_c, \psi'_u; \sigma)\|_{\infty,2} \leq L\|(\psi_c, \psi_u) - (\psi'_c, \psi'_u)\|_{\infty,2} \quad \text{for all } \sigma.$$

Thus, from contraction mapping theorem, we have a unique fixed point (C_*^σ, U_*^σ) for every σ and these by construction of the map Θ , given by (29)–(30), are the stationary quantities corresponding to (C_n, U_n) with discretization levels σ .

Continuity of Θ with respect to σ : From (31)–(32), difference $\Theta(C_*^{\sigma'}, U_*^{\sigma'}; \sigma) - \Theta(C_*^{\sigma'}, U_*^{\sigma'}; \infty)$ is made up of terms like ϖ , the workload process defined for any $m = g$ or gg and any $\psi \in \Psi$ by,

$$\varpi_m^\sigma(q, \cdot; \psi) := \mathcal{T}_m((\delta^\sigma(q), q], \psi).$$

Clearly $\varpi_m^\sigma(\cdot, \cdot; C_*^{\sigma'}) \in \Psi$ for every $\sigma \leq \infty$ and for every σ' and using¹⁴ (33)

$$\|\varpi_m^\sigma(q, \cdot; C_*^{\sigma'})\|_2 \leq 2\hat{B}\|\mathcal{N}^\sigma(q, \cdot)\|_2,$$

where $\mathcal{N}^\sigma(q, \cdot)$ represents the number of customers that arrived in segment $[\delta^\sigma(q), q]$ for a time duration of $C_*^{\sigma'}(|\mathcal{C}|)$. Clearly¹⁵ using the properties of Poisson process,

$$\begin{aligned} \|\mathcal{N}^\sigma(q, \cdot)\|_2^2 &= E[\mathcal{N}^\sigma(q, \cdot)^2] \\ &= \left(\lambda P_Q([\delta^\sigma(q), q])\|C_*^{\sigma'}(|\mathcal{C}|) + \left(\lambda P_Q((\delta^\sigma(q), q])\|C_*^{\sigma'}(|\mathcal{C}|)\right)^2\right) \\ &\leq \left(\lambda \frac{\|f_Q\|_\infty}{\sigma} C_*^{\sigma'}(|\mathcal{C}|) + \left(\lambda \frac{\|f_Q\|_\infty}{\sigma} C_*^{\sigma'}(|\mathcal{C}|)\right)^2\right) \quad \text{for any } q. \end{aligned}$$

Hence for any given σ' and m ,

$$\lim_{\sigma \rightarrow \infty} \|\varpi_m^\sigma(\cdot, \cdot; C_*^{\sigma'})\|_{\infty,2} = 0 \quad \text{and similarly} \quad \lim_{\sigma \rightarrow \infty} \|\varpi_m^\sigma(\cdot, \cdot; C_*^{\sigma'}(|\mathcal{C}|))\|_{\infty,2} = 0.$$

From (31)–(32), difference

$$\begin{aligned} \Theta_C(C_*^{\sigma'}, U_*^{\sigma'}; \sigma) - \Theta_C(C_*^{\sigma'}, U_*^{\sigma'}; \infty) &= -\varpi_g^\sigma(\cdot, \cdot; C_*^{\sigma'}) - \varpi_{gg}^\sigma(\cdot, \cdot; C_*^{\sigma'}(|\mathcal{C}|)), \\ \Theta_U(C_*^{\sigma'}, U_*^{\sigma'}; \sigma) - \Theta_U(C_*^{\sigma'}, U_*^{\sigma'}; \infty) &= \frac{1}{\sqrt{\hat{B}}}\varpi_g^\sigma(\cdot, \cdot; C_*^{\sigma'}) + \frac{1}{\sqrt{\hat{B}}}\varpi_{gg}^\sigma(\cdot, \cdot; C_*^{\sigma'}(|\mathcal{C}|)) \end{aligned}$$

and thus for any $\sigma' \leq \infty$,

$$\|\Theta(C_*^{\sigma'}, U_*^{\sigma'}; \sigma) - \Theta(C_*^{\sigma'}, U_*^{\sigma'}; \infty)\|_{\infty,2} \rightarrow 0 \quad \text{as } \sigma \rightarrow \infty. \tag{34}$$

¹⁴From (15), $C_*^\sigma(q)$ equals sum of stationary quantities corresponding to $T_n(q)$ and $T_{n-1}(|\mathcal{C}| - T_{n-1}(q))$ and hence will be less than or equal to $2C_*^\sigma(|\mathcal{C}|)$ in distribution and this is true for all q , i.e.,

$$C_*^\sigma(q) \stackrel{\text{distribution}}{\leq} 2C_*^\sigma(|\mathcal{C}|) \quad \text{for all } q. \tag{33}$$

Also, recall service times $B \leq \hat{B}$ with probability one.

¹⁵Because of independent arrivals as reasoned in footnotes 11 and 12.

Continuity of fixed points with respect to σ : By the definition of fixed points,

$$\begin{aligned} & \| (C_*^\sigma, U_*^\sigma) - (C_*^\infty, U_*^\infty) \|_{\infty,2} \\ &= \| \Theta(C_*^\sigma, U_*^\sigma; \sigma) - \Theta(C_*^\infty, U_*^\infty; \infty) \|_{\infty,2} \\ &\leq \| \Theta(C_*^\infty, U_*^\infty; \sigma) - \Theta(C_*^\infty, U_*^\infty; \infty) \|_{\infty,2} + \| \Theta(C_*^\sigma, U_*^\sigma; \sigma) - \Theta(C_*^\infty, U_*^\infty; \sigma) \|_{\infty,2} \\ &\leq \| \Theta(C_*^\infty, U_*^\infty; \sigma) - \Theta(C_*^\infty, U_*^\infty; \infty) \|_{\infty,2} + L \| (C_*^\sigma, U_*^\sigma) - (C_*^\infty, U_*^\infty) \|_{\infty,2}. \end{aligned}$$

And hence

$$\| (C_*^\sigma, U_*^\sigma) - (C_*^\infty, U_*^\infty) \|_{\infty,2} \leq \frac{1}{1-L} \| \Theta(C_*^\infty, U_*^\infty; \sigma) - \Theta(C_*^\infty, U_*^\infty; \infty) \|_{\infty,2}.$$

Thus, using (34), as $\sigma \rightarrow \infty$

$$\| (C_*^\sigma, U_*^\sigma) - (C_*^\infty, U_*^\infty) \|_{\infty,2} \rightarrow 0.$$

It is easy to check that $\| C_*^\sigma(q, \cdot) \|_2 = c_*^{2\sigma}$. Thus, by the definition of $\| \cdot \|_{\infty,2}$ the theorem on second moments follows. \square

Lemma 3 $(\Psi, \| \cdot \|_{\infty,2})$ is a complete space.

Proof Say ψ_n is a Cauchy sequence. Then for any q , $\psi_n(q, \cdot)$ is Cauchy in L^2 space and hence there exist an L^2 -limit say call it $\psi(q, \cdot)$. Now for any q ,

$$\| \psi_n(q, \cdot) - \psi(q, \cdot) \|_2 = \lim_{m \rightarrow \infty} \| \psi_n(q, \cdot) - \psi_m(q, \cdot) \|_2.$$

For any $\epsilon > 0$ there exists N such that

$$\sup_q \| \psi_m(q, \cdot) - \psi_n(q, \cdot) \|_2 \leq \epsilon \quad \text{for all } n, m > N.$$

Combining the above two equations, whenever $n > N$

$$\sup_q \| \psi_n(q, \cdot) - \psi(q, \cdot) \|_2 \leq \epsilon$$

and hence $\| \psi_n - \psi \|_{\infty,2} \rightarrow 0$ as $n \rightarrow \infty$.

By completeness of \mathcal{D} the map (of limits of L^2 norms) $q \rightarrow \| \psi(q, \cdot) \|_2$ resides in \mathcal{D} . \square

Appendix C: Pareto optimality

Let W be the set of vectors of expected waiting times achieved by a class of design policies. A vector $w_1 \in W$ obtained by some policy dominates another vector w_2 if all entries of w_1 are smaller than or equal to those of w_2 with at least one entry being strictly smaller. This definition is given when w_1 has finite number of entries. We now extend the definition in an obvious way to our case, i.e., when w_1 is a function on $[0, |C|]$ that is integrable with respect to Lebesgue measure, \mathcal{L} . A function w_1 is said to dominate another function w_2 if $w_1 \leq w_2$ almost surely (w.r.t. Lebesgue measure) and if there exists a set \mathcal{B} such that $w_1 < w_2$ on \mathcal{B} and $\mathcal{L}(\mathcal{B}) > 0$. A minimum vector in W in the Pareto sense is one that is not dominated by any other vector in W .

Lemma 4 Let ζ be an integrable function on $[0, |C|]$ with $\zeta > c$ almost surely, with $c \geq 0$. Let,

$$w^* = \arg \min_{w \in W} \int_0^{|C|} w(q)\zeta(q) dq,$$

be a minimizer of the integral. Then w^* is a Pareto optimal solution.

Proof If it is not and there exists a vector w_1 which dominates w^* . Say $w_1 < w^*$ on \mathcal{B} with $\mathcal{L}(\mathcal{B}) > 0$ while $w_1 = w^*$ on the complement, \mathcal{B}^c . Then

$$\begin{aligned} \int_0^{|C|} (w_1(q) - w^*(q))\zeta(q) dq &= \int_{\mathcal{B}} (w_1(q) - w^*(q))\zeta(q) dq \\ &\leq -c \int_{\mathcal{B}} (w^*(q) - w_1(q)) dq < 0. \end{aligned}$$

and so the optimality of w^* is contradicted when $c > 0$. When $c = 0$, by definition of infimum one can find a subset of \mathcal{B} on which $\zeta > \tilde{c} > 0$ and the proof follows. \square

References

- Altman, E., & Levy, H. (1994). Queueing in space. *Advances of Applied Probability*, 26, 1095–1116.
- Altman, E., Khamisy, A., & Yechiali, U. (1992). On elevator polling with globally gated regime. *Queueing Systems*, 11, 85–90 (special issue on “Polling Models”, Eds. H. Takagi and O. Boxma).
- Altman, E., Konstantopoulos, P., & Liu, Z. (1992). Stability, monotonicity and invariant quantities in general polling systems. *Queueing Systems*, 11(1–2), 35–57.
- Baccelli, F., & Brémaud, P. (1991). *Applications of Mathematics: Vol. 26. Elements of queueing theory*. Berlin: Springer.
- Berger, M. S. (1977). *Nonlinearity and functional analysis*. New York: Academic Press.
- Boona, M. A. A., van der Meib, R. D., & Winands, E. M. M. (2011). Applications of polling systems. *Surveys in Operations Research and Management Science*, 16(2), 67–82.
- Boxma, O. (1989). Workloads and waiting times in single-server systems with multiple user classes. *Queueing Systems*, 5, 185–214.
- Boxma, O. J., & Groenendijk, W. P. (1987). Pseudo-conservation laws in cyclic-service systems. *Journal of Applied Probability*, 24(4), 949–964.
- Boxma, O.J., Levy, H., & Yechiali, U. (1992). Cyclic reservation schemes for efficient operation of multiple-queue single-server systems. *Annals of Operations Research*, 187–208.
- Boxma, O. J., Bruin, J., & Fralix, B. (2008). Waiting times in polling systems with various service disciplines. *EURANDOM report*, June 2008.
- Coffman, J. E. G., & Gilbert, E. (1986). A continuous polling system with constant service times. *IEEE Transactions on Information Theory*, 32, 584–591.
- Coffman, J. E. G., & Gilbert, E. (1987). Polling and greedy servers on the line. *Queueing Systems*, 2115–2145.
- Eliazar, I. (2003). The snowblower problem. *Queueing Systems*, 45, 357–380.
- Eliazar, I. (2005). From polling to snowplowing. *Queueing Systems*, 51(1–2), 115–133.
- Foss, S. G., & Chernova, N. I. (1996). Comparison theorems and ergodic properties of the polling systems. *Problemy Peredači Informacii*, 4, 46–71.
- Fuhrmann, S., & Copper, R. (1985). Applications of the decomposition principle in m/g/1 vacation models to two continuum cyclic queueing models. *AT&T Technical Journal*, 64, 1091–1098.
- Georgiadis, L., & Szpankowski, W. (1992). Stability of token passing ring. *Queueing Systems*, 11(1–2), 7–33.
- Kavitha, V. (2011). Continuous polling with rerouting and applications to ferry assisted wireless LANs. *ValueTools 2011*.
- Kavitha, V., & Altman, E. (2009). Queueing in space: design of message ferry routes in static adhoc networks. In *21st international teletraffic congress (ITC 21)*, Paris, France, 15–17 September 2009.
- Kavitha, V., & Altman, E. (2010a). Analysis and design of message ferry routes in sensor networks using polling models. In *WiOpt*, Avignon, France, May 31–Jun 04.

- Kavitha, V., & Altman, E. (2010b). Opportunistic scheduling of a message ferry in sensor networks. In *MobiOpp 2010*, Pisa, Italy, February.
- Khamisy, A., Altman, E., & Sidi, M. (1992). Polling systems with synchronization constraints. *Annals of Operations Research*, 35, 231–267 (special issue on Stochastic Modeling of Telecommunication Systems, Eds. P. Nain and K. W. Ross).
- Kroese, D., & Schmidt, V. (1992). A continuous polling system with general service times. *The Annals of Applied Probability*, 2(4), 906–927.
- Kroese, D., & Schmidt, V. (1993). Queueing systems on the circle. *ZOR. Zeitschrift Für Operations-Research*, 37(3), 303–331.
- Kroese, D., & Schmidt, V. (1994). Single-server queues with spatially distributed arrivals. *Queueing Systems*, 17, 317–345.
- Kroese, D., & Schmidt, V. (1996). Light-traffic analysis for queues with spatially distributed arrivals. *Mathematics of Operations Research*, 21, 135–157.
- Leskela, L., & Unger, F. (2010). Stability of a spatial polling system with greedy myopic service. *Annals of Operations Research*.
- Robert, P. (2010). The evolution of a spatial stochastic network. *Stochastic Processes and their Applications*.
- Rojas-Nandayapa, L., Foss, S., & Kroese, D. (2011). Stability and performance of greedy server systems: a review and open problems. *Queueing Systems*.
- Saad, W., Han, Z., Basar, T., Debbah, M., & Hjørungnes, A. (2009). A selfish approach to coalition formation among unmanned air vehicles in wireless networks. *Gamenets, Istanbul, Turkey*.
- Shi, Y., & Hou, Y. T. (2008). Theoretical results on base station movement problem for sensor network. In *IEEE INFOCOM 08*.
- Sidi, M., Levy, H., & Fuhrmann, S. W. (1992). A queueing network with a single cyclically roving server. *Queueing Systems (special issue on Polling Models, Eds. H. Takagi and O. Boxma)* (pp. 121–144).
- Takagi, H. (1986). *Analysis of polling systems*. Cambridge: The MIT Press.
- Tariq, M. M. B., Ammar, M., & Zegura, E. (2006). Message ferry route design for sparse ad hoc networks with mobile nodes. In *Proc. of ACM MobiHoc*, Florence, Italy, 22–25 May 2006 (pp. 37–48).
- Whitt, W. (1991). A review of $l = \lambda w$ and extensions. *Queueing Systems: Theory and Applications*, 9(3), 235–268.
- Yousefi, S., Altman, E., El-Azouzi, R., & Fathy, M. (2007). Connectivity in vehicular ad hoc networks in presence of wireless mobile base-stations. In *The proceedings of the 7th international conference on ITS telecommunications*, Sophia-Antipolis, France, June 6–8.