Ergodicity of a continuous polling model with a random walk routing

Eugene Spodarev

Friedrich–Schiller Universität, Institut für Stochastik,
Ernst Abbe Platz 1-4, 07743 Jena, Deutschland
E-mail: seu@minet.uni-jena.de

HISTORY:

Received 12 April 2001

In the paper the necessary and sufficient conditions of ergodicity of a continuous polling system are given. One server on a cirlce serves the customers that arrive to the system independently and are distributed along the circle according to some continuous probability distribution. The server's routing is a general random walk on the circle. While moving the server pays attention to all customers encoutered, then the served requests leave the system. The principal tools used in the proof of the main result are majorizing and ergodicity theorems for regenerative processes.
Keywords: Markov chain, ergodicity, majorizing, regenerative process, polling system, greedy server, stationary metrically transitive sequence, random walk, queueing theory

AMS Subject classification: Primary 60K25, 90B22; Secondary 60G50, 60J10

1. Introduction

Polling systems are the subject of studying for more then twenty years, their applications range from information and transport networks to multiprocessor computers and optimization of the hard drives' perfomance. The aim of this paper is to generalize the results of [2]–[6] to the case of a continuous polling system on a circle with a general random walk as server routing and to approach by this means the studying of circular continuous polling systems with a "gready" server. The ergodic theorems for the latter have not been proved yet. A system similar to the one considered below was first studied in [4], while the control sequences therein were either Poisson or determined, and the server was able to move only in one direction. The proof of ergodicity was not given there. The random measures' approach to a similar polling model (with general service times) was demonstrated in [8]. In [9] a polling model with the Brownian motion of zero drift as a routing procedure was examined and stability results for its configuration measure were proved. See [10] for further developments. The comparison of circular continuous polling systems with a "greedy" server and with the server moving in one direction was made in [5]. The heavy traffic behaviour of them was studied in [11]. Discrete polling systems (with a finite number of queues for customers) appeared to be much simpler, and the necessary conditions of ergodicity of a broad class of such systems were obtained in [12] (cf. [6] for sufficient conditions). The latter paper finished the description of a class of such systems that meet some requirements of monotonicity and additivity. This class comprises also discrete polling systems with one-directional service and with a "gready" server. Finally, the first attempts to consider continuous polling systems via some limiting procedures are to be seen in [3], [7].

The below paper is organized as follows: in section 2 the exact mathematical setting of the problem is given; the main ergodic theorem is proved in section 3. Section 4 contains the discussion of possible generalizations and open problems.

2. Polling systems

A server moves along the unit circle serving customers that are scattered on it. Service requests arrive to the system after the periods of time τ_i, where $\{\tau_i\}_{i=1}^{\infty}$ are i. i. d. random variables with finite mean $E\tau_i = b$. Let us call $\lambda = 1/b$ the *input intensity*. When the i–th customer enters the system it is scattered instantly on the circle with coordinate $y_i \in [0,1)$; random variables $\{y_i\}_{i=1}^{\infty}$ are i. i. d. on $[0,1]$ with distribution density $p(x)$. The velocity of the server is a random variable $v > 0$, $Ev = 1/\alpha > 0$. The distance ρ_n covered by the server up to the n–th change of direction of moving is a random walk: at this instant n the server chooses the future direction clockwise with probability p and counterclockwise with probability q, $p \neq q$, $p + q = 1$. Then it walks in this chosen direction a distance γ_n. In other words,

$$\rho_{n+1} = \rho_n + \gamma_{n+1} \left(I\{\delta_{n+1} = 1\} - I\{\delta_{n+1} = -1\} \right),$$

$$\delta_n = \begin{cases} 1, \text{with probability } p \\ -1, \text{with probability } q \end{cases},$$

where $f_n = (\gamma_n, \delta_n)$ is an i. i. d. sequence of random elements. Distances γ_n are distributed continuously on $[0, 1/2]$. When the server reaches a point y_i on the circle where the i–th customer is waiting it stops to perform service of duration a_i; $\{a_i\}_{i=1}^{\infty}$ are i. i. d. random variables with finite $Ea_i = a > 0$. Being served the customer leaves the system.

Let us mention several definitions that will be used in the sequel.

Definition 2.1. Let two stochastic processes $X(t)$ and $X^*(t)$, $t \geqslant 0$ with the values from X be defined on the same probability space. We say that $X(t) \leq_{st} X^*(t)$ iff for all $x \in X$ $P\{X(t) \geq x\} \leq P\{X^*(t) \geq x\} \quad \forall t \geq 0$.

Definition 2.2. Stochastic process $X(t)$, $t \geqslant 0$ defined on the measurable space $\Big(\mathbf{X}, \mathcal{B}(\mathbf{X})\Big)$ is ergodic iff for all $B \in \mathcal{B}(\mathbf{X})$ and all its initial values $X(0)$ the following is true: $\lim\limits_{t\to\infty} P\{X(t) \in B\} = \mu(B)$, where $\mu(\cdot)$ is a probability measure.

Definition 2.3. Probability measure $\mu(\cdot)$ on $\mathbb{R}_+$ is called a lattice distribution iff it it concentrated in the nodes of a lattice with step $h > 0$, i. e. in the points $x_j = x_0 + jh, \quad x_0 \geq 0, \quad \forall j \in \mathbb{N}$.

Later on suppose that the random variables τ_i are not distributed on a lattice for all $i \in \mathbb{N}$. Let us denote by $\xrightarrow{D}$ the convergence in distribution (weak convergence). Let ζ_n be the instant of time when the n–th customer leaves the system, $D(t) \stackrel{def}{=} \max\{n : \zeta_n \leq t\}$. Denote by $q(t)$ the number of customers present in the system at time t. Let

$$N(t) = \max\{n : \tau_1 + \ldots + \tau_n < t\}.$$

Suppose $q(0) \stackrel{a.s.}{=} 0$. We shall say that the server completed the *cycle* of moving if he visited (at least once) all points of the circle and then returned to 0. Thus the duration of the n–th cycle is a random variable T_n, $T_1 = \inf\{k : |\rho_k| > 1\}$. All T_n are independent and identically distributed. Let us denote by $\rho(t)$ the whole distance covered by the server up to a moment t.

3. Necessary and sufficient conditions of ergodicity

Here we shall find the conditions under which the process $q(t)$ is ergodic. First let us show that the mean cycle duration is finite:

Lemma 3.1. For all $p > 0$, $q > 0$ ($p \neq q$, $p + q = 1$), and $n \in \mathbb{N}$ the following is true:

$$E\,T_n < \infty.$$

Proof. We suppose $p > q$ without loss of generality. Denote $\Delta\rho_n = \rho_n - \rho_{n-1}$. Then one gets from Wald's equality for random walks that $E\,\Delta\rho_1\,E\,T_n = E\,\rho(T_n)\ \forall n \in \mathbb{N}$. If $p > q$ then $E\,\Delta\rho_1 > 0$, which implies

$$E\,T_n \leq 1 + \frac{1}{|E\,\Delta\rho_1|} < \infty.$$

□

Lemma 3.2. If process $q(t)$ is ergodic then $\lambda a < 1$.

Proof. Suppose the contrary takes place: $q(t)$ is ergodic and $\lambda a \geqslant 1$. Then by the existence of the limiting stationary distribution

$$\frac{D(t)}{N(t)} \xrightarrow{a.s.} 1 \quad \text{as} \quad t \to \infty.$$

The total distance up to the moment ζ_n is $\rho(\zeta_n) = (\zeta_n - \sum\limits_{i=1}^{n} a_i)v$. Since $D(\zeta_n) \sim N(\zeta_n)$ as $n \to \infty$, $D(\zeta_n) = n$ then by the law of large numbers the distance per unit time is

$$\frac{\rho(\zeta_n)}{\zeta_n} \sim \left(1 - \frac{N(\zeta_n)}{\zeta_n} \frac{\sum\limits_{i=1}^{n} a_i}{n} \right) v \xrightarrow{a.s.} (1 - \lambda a)\, v \quad \text{as} \quad n \to \infty.$$

Then by Lebesgue theorem

$$E\,\frac{\rho(\zeta_n)}{\zeta_n} \longrightarrow (1 - \lambda a)\alpha^{-1} \leq 0 \quad \text{as} \quad n \to \infty,$$

thus we have arrived at the contradiction, since in ergodic case the mean velocity of the server $\rho(\Delta t)/\Delta t$ in an interval Δt should be positive. □

Theorem 3.3. Let $\{\tau_i\}_{i=1}^{\infty}$, $\{a_i\}_{i=1}^{\infty}$ be the sequences of independent identically distributed random variables with $Ea_i = a$, $E\tau_i = 1/\lambda$. Then the process $q(t)$ is ergodic iff $\lambda a < 1$.

Proof. By lemma 3.2 the necessity is proved. Let us prove the sufficiency. Suppose $q(0) \overset{a.s.}{=} 0$. First we show that $q(t)$ is a regenerative process (see [1], p. 87–90). Indeed, when the next customer enters the empty system, the past of the process is erased, because one can suppose without loss of generality that at this instant the server is placed at the point 0. Let $c_n = \sum\limits_{i=1}^{n} \tau_i$ be the moment of arrival of the n–th service request to the system. Let us fix regeneration moments

$$\chi_1 = c_1, \quad \chi_n = \inf\{c_i > \chi_{n-1},\ i \in \mathbb{N}:\ q(c_i - 0) = 0, \quad q(c_i) > 0\}.$$

Then the n–th regeneration cycle is $\theta_n = \chi_{n+1} - \chi_n,\ \ n \geq 1$. Suppose $\lambda a < 1$. Let us prove that $q(t)$ is ergodic. By means of Smith's theorem (cf. [1], p. 92) $q(t)$ is ergodic if it is aperiodic, i.e. the distribution of θ_i is not a lattice distribution for all $i \in \mathbb{N}$, and $E\theta_i < \infty$. The aperiodicity of $q(t)$ follows from the fact that the input stream $\{\tau_i\}_{i=1}^{\infty}$ is not concetrated on a lattice. Let us show then that $E\theta_i < \infty$ for every $i \in \mathbb{N}$. Since all θ_i have the same distribution it suffices to show that for $i = 1$.

Besides the system $\mathbf{S}$ described in section 2 introduce another system $\mathbf{S}^*$ that differs from $\mathbf{S}$ only by the fact that the sever in it serves only those customers that arrived to $\mathbf{S}^*$ during the last cycle . It is supposed that the server does not perform service during its first cycle in $\mathbf{S}^*$. Let X_n be the number of customers in $\mathbf{S}^*$ at the end of the n–th cycle. Then $\{X_n\}_{n=1}^{\infty}$ is an aperiodic homogeneous Markov chain. This means, in particular, that $\{X_n\}_{n=1}^{\infty}$ is regenerative, while the moments when it equals to 0 are its regeneration points. Let $q^*(t)$ be the number of customers in $\mathbf{S}^*$ at the moment t, θ_1^* — the first regeneration cycle of $\{X_n\}_{n=1}^{\infty}$. If one can prove that

$$\theta_1 \leq_{st} \theta_1^*, \tag{3.1}$$

$$E\theta_1^* < \infty \tag{3.2}$$

then we are done. Let us prove that (3.1) holds. We shall show that the chain $\{X_n\}_{n=1}^{\infty}$ is ergodic, i.e. that there exists

$$\lim_{n\to\infty} P\{X_n = i\,|X_0 = 0\} = \frac{1}{\pi_i} > 0 \qquad \forall i \in \mathbb{Z}_+.$$

Suppose the contrary is true: $\lim_{n\to\infty} P\{X_n = i \,|\, X_0 = 0\} = 0$ (only these two cases are possible — cf. [13], p. 47). This means that

$$P\{X_n > y\} \to 1 \quad \text{as} \quad n \to \infty \quad \forall y \in \mathbb{N}, \tag{3.3}$$

or in other terms, $X_n \xrightarrow{P} \infty$ as $n \to \infty$. Evidently, $X_{n+1} = \xi(X_n) + \eta_n$ where $\xi(X_n)$ is the number of customers that entered $\mathbf{S}^*$ during the n–th cycle. We have $E\eta_n = \lambda\alpha$, $E\left(\xi(X_n) \,|\, X_n\right) = \lambda a X_n$ for sufficiently large n. Then

$$E(X_{n+1} - X_n \,|\, X_n) = (\lambda a - 1)X_n + \lambda\alpha,$$

hence $E(X_{n+1} - X_n) < 0$ if X_n is large enough, which contradicts $X_n \xrightarrow{P} \infty$. Then the Markov chain $\{X_n\}_{n=1}^{\infty}$ is ergodic, and (3.2) is proved.

Let us show now that (3.1) holds. Let t_n (t_n^*) be the moment of completion of the n–th cycle in $\mathbf{S}$ ($\mathbf{S}^*$),

$$T_n = t_n - t_{n-1}, \quad t_n^* = t_n^* - t_{n-1}^* \quad \forall n \in \mathbb{N},$$

let $D^*(t)$ be the number of customers that left $\mathbf{S}^*$ up to the moment t. Let the service times in both systems be attributed to the customers according to the order of their service in $\mathbf{S}$ (without loss of generality). Then one can show that

$$D^*(t_n^*) - D^*(t_{n-1}^*) \leq_{st} D(t_n) - D(t_{n-1}) \quad \forall n \in \mathbb{N},$$

that follows from

$$T_n^* \leq_{st} T_n \qquad \forall n \in \mathbb{N} \tag{3.4}$$

(it can be verified by induction on n). If Q_n (Q_n^*) is the number of cycles in $\mathbf{S}$ ($\mathbf{S}^*$) that are necessary to serve n customers in one regeneration period then it folows from (3.4) that $Q_n \leq_{st} Q_n^* \;\; \forall n \in \mathbb{N}$, and thus (3.1) is proved. $\square$

4. Generalizations and open questions

Remark 4.1. The method of majorizing used in the proof of theorem 3.3 is not the only one possible. In similar problems the Foster – Lyapunov criterion was often used, although it imposes restrictions on the control sequences and yields sometimes superflously strong sufficient conditions of ergodicity. At the same time the stochastic majorizing allows us to broaden the choice of possible control sequences and to get the following result with minimal modifications in the above proof:

Let $\{\tau_i\}_{i=1}^{\infty}$, $\{a_i\}_{i=1}^{\infty}$ be stationary metrically transitive sequences independent from each other, let $Ea_i = a$, $E\tau_i = 1/\lambda$ $\forall i \in \mathbb{N}$. Then the process $q(t)$, $t \geq 0$ is stochastically bounded iff $\lambda a < 1$.

Remark 4.2. There could be further generalizations of the above problem in order to approach the question of ergodicity of the system with "greedy" server, i.e. when the server goes towards the nearest customer on the circle. Thus, one can make the routing random walk on the circle dependent on the present state of the system, i.e. on the number and position of the customers on the circle. Then one can suppose the server's routing to be a semimartingale. The only difficulty that arises here will lie in the proof that the mean cycle duration in finite. Once this obstacle being overcome, the rest of the reasoning in section 3 does not change.

Acknowledgements

The above research is a part of the author's PhD project that was made during his stay at the Moscow State University, Faculty of Mathematics and Mechanics, Chair of Probability Theory under the supervision of Prof. L. G. Afanasieva. The author would like to thank Prof. L. G. Afanasieva for the statement of problems, constant discussions and many helpful ideas in the course of the research.

References

[1] Afanasieva, L. and Bulinskaya, E. (1980). *Stochastic processes in queueing and inventory theory.* Moscow State University, Moscow. (in russian).

[2] Altman, E. and Foss, S. (1993). Polling on a graph with general arrival and service time distribution. Rapport de recherche. INRIA Sophia–Antipolis.

[3] Altman, E. and Foss, S. (1997). Polling on a space with general arrival and service time distributions. *Operations Research Letters* **20,** 187–194.

[4] Coffman, E. G. and Gilbert, E. N. (1986). A continuous polling system with constant service times. *IEEE Tran. Inform. Theory* **32,** 584–591.

[5] Coffman, E. G. and Gilbert, E. N. (1987). Polling and greedy servers on a line. *Queueing Systems* **2,** 115–145.

[6] Foss, S. G. and Chernova, N. I. (1995). Ergodic properties of polling systems. Preprint 6. Institute of Mathematics SO RAN Novosibirsk. (in russian).

[7] Foss, S. G. and Chernova, N. I. (1996). On stability of polling models with an infinite number of queues. *Siberian Math. J.* **37,** 940–956.

[8] Kroese, D. and Schmidt, V. (1992). A continuous polling system with general service times. *Ann. Appl. Probab.* **2,** 906–927.

[9] Kroese, D. and Schmidt, V. (1993). Queueing systems on a circle. *Z. Oper. Res.* **37,** 303–331.

[10] Kroese, D. and Schmidt, V. (1994). Single server queues with spatially distributed arivals. *Queueing Systems* **17,** 317–345.

[11] Kroese, D. P. (1997). Heavy traffic analysis for continuous polling models. *J. Appl. Probab.* **34,** 720–732.

[12] Schassberger, R. (1995). Stability of polling networks with state-dependent server routing. *Prob. in Eng. and Inf. Sci.* **9,** 539–550.

[13] Shiryaev, A. N. (1989). *Probability theory.* Nauka, Moscow. (in russian).