On Polling Systems where Servers wait for Customers

Larissa G. Afanassieva, Franck Delcoigne, Guy Fayolle

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

On Polling Systems where Servers wait for Customers

Larissa G. Afanassieva , Franck Delcoigne , Guy Fayolle

N° 3058

December 1996

THÈME 1

Rapport de recherche

ISSN 0249-6399

On Polling Systems where Servers wait for Customers

Larissa G. Afanassieva *, Franck Delcoigne †, Guy Fayolle †

Thème 1 — Réseaux et systèmes
Projet MEVAL

Rapport de recherche n° 3058 — December 1996 — 21 pages

Abstract: In this paper, a particular polling system with N queues and V servers is analyzed. Whenever a server visits an empty queue, it waits for the next customer to come to this queue. A customer chooses his destination according to a routing matrix P. The model originates from specific problems arising in transportation networks. A global classification of the process describing the system is given under general assumptions. It is shown that this process can only be *transient* or *null recurrent*. In addition, a detailed classification of each node, together with limit laws (after proper time-scaling) are obtained. The method of analysis relies on the central limit theorem and a coupling with a reference system in which transportation times are identically zero.

Key-words: Network, Polling, Random walk, Recurrence, Transience, Central Limit theorem.

(Résumé : tsvp)

* Postal address: Moscow State University, Mechanico-Mathematical Faculty, Chair of probability, Laboratory of Large Random Systems, Vorobyevy Gori, 119899 Moscow, RUSSIA.

† Postal address: INRIA - Domaine de Voluceau, Rocquencourt - BP 105 - 78153 Le Chesnay Cedex - FRANCE.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : (33 1) 39 63 55 11 – Télécopie : (33 1) 39 63 53 30

Sur des systèmes à scrutin où les serveurs attendent les clients

Résumé : Dans cet article, on analyse un système de *polling* particulier, comportant N stations et V serveurs. Quand un serveur arrive à une station vide, il attend la prochaine arrivée à cette station. Chaque client choisit sa destination selon une matrice de routage P. Le contexte du modèle est surtout celui des réseaux de transport. On obtient une classification globale du système, sous des hypothèses assez larges, en montrant notamment qu'il ne peut être que transient ou récurrent. On donne aussi une classification détaillée pour chaque file, ainsi que des lois limites après changement d'échelle temporelle convenable. La démarche proposée s'appuie sur le théorème de la limite centrale et sur un couplage avec un processus de référence, où les temps de transport sont identiquement nuls.

Mots-clé : Marche aléatoire, Réseau, Récurrence, Scrutin, Transience, théorème Central Limite.

1 Description of the model

Consider an open network, consisting of N stations (nodes, parking lots) and V cars, which circulate among the stations. Let $\mathcal{S} \stackrel{\text{def}}{=} \{1, \ldots, N\}$. The arrivals of customers form a simple point process, defined by a metrically transitive sequence of interarrival times $\{a_n, n \geq 1\}$. Assuming the first customer arrives at $t_1 = 0$, then the $(n+1)$-th customer enters the system at time $t_{n+1} = \sum_{j=1}^{n} a_j$ and is directed to some node i with probability $\gamma_i,\ i \in \mathcal{S}, \gamma_i > 0$ and $\sum_{i=1}^{N} \gamma_i = 1$. All customers choose their destination via some ergodic routing matrix $\mathbf{P} = (p_{ij})_{i,j \in \mathcal{S}}$. Let $\pi = (\pi_1, \ldots, \pi_N)$ be the invariant measure associated to $\mathbf{P}$. A car arriving at a station where there are waiting customers takes one of them to some destination. Whenever the car finds a station empty, it stops and wait for the next arrival at this node. After having reached their destination, customers leave the network. Also, a customer who, upon arrival, does not find an available car waits in a queue (no impatiance phenomenon is assumed). Capacities of waiting rooms for clients are supposed unlimited and there are at least V available parking lots for cars at each station, i.e. *empty displacements* of cars are not allowed in this model.

We introduce the following quantities, for all $i, j \in \mathcal{S}, n \geq 1$:

- τ_{ij}, the time to go from node i to node j, $\forall 1 \leq i, j \leq N$. The N^2 random variables do not depend on the arrival process, but can be correlated between each other.
- $q_j(n)$, the number of customers at node j at time $t_n - 0$;
- $x_j(n)$, the number of cars at node j at time $t_n - 0$;
- $S(n)$, the position of the server at time $t_n - 0$ when $V = 1$.
- v_n, the set of nodes at which there is at least one vehicle at time t_n;

Assume the n-th customer arrives at node $i(n)$ and intends to go to node $j(n)$. The pairs

$$\{(i(n), j(n)), n \geq 1\},$$

are supposed to form a sequence of independent and identically distributed ($i.i.d.$) random variables. This means concretely that the choice of a destination by a customer depends only on his arrival node and on nothing else. Let also

$$\lambda = (Ea_n)^{-1}, \ \tau_i = \sum_{j=1}^{N} p_{ij} E(\tau_{ij}), \ \tau = \sum_{i=1}^{N} \pi_i \tau_i \,. \tag{1.1}$$

The main goal of the study is to analyze the behaviour of the process

$$\begin{aligned} Q \stackrel{\text{def}}{=} \{Q(n), n \geq 1\} &\stackrel{\text{def}}{=} \{(q_j(n), \ x_j(n); \ j \in \mathcal{S}), n \geq 1\} \\ &\stackrel{\text{def}}{=} \{\vec{q}(n), \vec{x}(n), n \geq 1\}. \end{aligned} \tag{1.2}$$

Definition 1 *The network is said to be ergodic if, and only if, there exists a stationary and a.s. finite sequence*

$$\widehat{Q} \stackrel{def}{=} \{\widehat{Q}(n), n \geq 1\} \stackrel{def}{=} \{(\widehat{q}_j(n), \ \widehat{x}_j(n); \ j \in \mathcal{S}), n \geq 1\},$$

such that $Q(n+k)$ converges weakly to $\widehat{Q}(k)$ as $n \to \infty, \forall k \geq 1$. Accordingly, a node i is ergodic if, and only if, $q_i(n+k)$ weakly converges to a stationary sequence $\widehat{q}_i(k)$ as $n \to \infty$.

Definition 2 *The network is transient if*

$$\lim_{n \to \infty} \sum_{j=1}^{N} q_j(n) = \infty, \quad a.s.$$

A node i is transient if

$$\lim_{n \to \infty} q_i(n) = \infty, \quad a.s.$$

Definition 3 *The process Q is said to be null recurrent if it is non-ergodic and there exists a state, say $((k_j, \ r_j); \ j \in \mathcal{S})$, such that*

$$P(\{q_j(n) = k_j, \ x_j(n) = r_j, \ j \in \mathcal{S}; \text{ for infinitely many } n\}) = 1.$$

Similar definitions hold for each separate node.

General notation

- Occasionally, it will be convenient to consider the system in continuous time and the same symbols will be used, e.g. $Q(t)$, whenever no ambiguity arises.

- Taboo probabilities: If $A \subset \mathcal{S}, i, j \in \mathcal{S}$, then ${}_AN_{ij}$ represents the number of times the Markov chain associated to $\mathbf{P}$ visits node j, starting from i, under the restriction that none of the nodes in the set A is entered in-between.

- $\{\mathcal{F}_n, n \geq 0\}$ will denote the increasing sequence of σ-algebras generated by $\{Q(t),\ t \leq n\}$.

- $A_i(t)$ stands for the event that an external arrival takes place at node i at time t.

- $\mathcal{Z}$ [resp. $\mathcal{R}$] is the set of integers [resp. of real numbers]. The positive parts of these sets are $\mathcal{Z}_+$ and $\mathcal{R}_+$ respectively. The components of a vector $\vec{X} \in \mathcal{R}^d$ will be denoted by X_i, $i = 1, \ldots, d$.

- The complementary $\mathcal{S} - A$ of $A \in \mathcal{S}$ will be written $\overline{A}$ and the indicator function of an arbitrary event $\mathcal{B}$ will be denoted by $\mathbb{1}_{\{\mathcal{B}\}}$.

2 Classification of the network

The purpose of this section is to classify the process $Q(n)$, viewed as a random walk on $\mathcal{Z}_+^{2N}$, in terms of ergodicity, transience and recurrence. Throughout this paper, we suppose $N > 1, \lambda > 0$. One of the basic results is that the process Q is never ergodic.

Theorem 2.1 *If $\gamma \neq \pi$, then the network is transient. Moreover, for any k such that*

$$\gamma_k > \sum_{i=1}^{N} \gamma_i p_{ik}, \tag{2.1}$$

node k is transient.

Proof When $\gamma \neq \pi$, there exists at least one node, say k, such that condition (2.1) holds. Define then, for all $m \geq 0$, the quantity

$$\alpha_k(n;m) \stackrel{\text{def}}{=} P(q_k(n) \leq m \,/\, q_k(1) = m).$$

Out of the first n arrivals in the system, let $\nu_{ij}(n)$ be the number of clients who arrived at node i and wanted to reach node j. Let

$$\nu_i(n) \stackrel{\text{def}}{=} \sum_{j=1}^{N} \nu_{ij}(n), \quad \widetilde{\nu}_i(n) \stackrel{\text{def}}{=} \sum_{j=1}^{N} \nu_{ji}(n). \tag{2.2}$$

In (2.2), $\nu_i(n)$ represents the total number of *external* customers arriving at node i, while $\widetilde{\nu}_i(n)$ stands for the number of *internal* customers intending to go to station i. Now the following inequality is obvious:

$$\alpha_k(n;m) \leq P\{\nu_k(n) \leq \widetilde{\nu}_k(n)\} = \beta_k(n). \tag{2.3}$$

To study the asymptotic behaviour of $\beta_k(n)$, as $n \to \infty$, let $A_{ij}(s)$ be the event that the s-th customer arrives at node i and wants to go to node j. One can write

$$\nu_k(n) = \sum_{s=1}^{n}\sum_{j=1}^{N} \mathbb{1}_{\{A_{kj}(s)\}}, \quad \widetilde{\nu}_k(n) = \sum_{s=1}^{n}\sum_{j=1}^{N} \mathbb{1}_{\{A_{jk}(s)\}}, \tag{2.4}$$

$$\nu_k(n) - \widetilde{\nu}_k(n) = \sum_{s=1}^{n} \delta_k(s)$$

where

$$\delta_k(s) = \sum_{j=1}^{N} (\mathbb{1}_{\{A_{kj}(s)\}} - \mathbb{1}_{\{A_{jk}(s)\}}).$$

From our general assumptions, it follows immediately that the random vectors $\delta(s) = (\delta_1(s), \ldots, \delta_N(s)), s \geq 1$, are *i.i.d.* Furthermore

$$E\delta_k(s) = \gamma_k - \sum_{j=1}^{N} \gamma_j p_{jk} = \rho_k > 0.$$

Using well-known large deviation upper bounds (see e.g. [5]), one obtains the estimate

$$\beta_k(n) \leq c_1 e^{-c_2\sqrt{n}} \tag{2.5}$$

where c_1 and c_2 are fixed constants. It follows from (2.3) and (2.5) that

$$\sum_{n=1}^{\infty} \alpha_k(n,m) < \infty.$$

The Borel-Cantelli Lemma ensures that the process $\{q_k(n), n \geq 1\}$ takes its values within an arbitrary bounded interval $[0, A]$ a finite number of times only: this means that $q_k(n)$ is transient, and so is the network. The proof of the theorem is concluded. ■

Theorem 2.2 *If*

$$\gamma = \pi, \tag{2.6}$$

then the network is always transient for $N \geq 4$ and any arbitrary τ. It is null recurrent for $N = 2, 3$ and $\tau = 0$.

Proof

Case $N \geq 4$. It suffices to prove the transience for $\tau = 0$, since one can easily show that a system with $\tau = 0$ is pathwise dominated by a system having the same arrival and routing processes, but an arbitrary $\tau > 0$.

For the sake of shortness, let us denote by $\mathcal{V}$ the state of $Q(n)$ when there are no customers in the network and all cars are waiting at node 1. Since $\{i(n), j(n)\}$ is a sequence of *i.i.d.* random vectors and $\tau = 0$, the state $\mathcal{V}$ is a regeneration point for the process Q. Setting

$$\alpha(n) = P(Q(n) = \mathcal{V}),$$

the following standard classification holds (see e.g. [7]):

(i) Q is transient if

$$\sum_{n=1}^{\infty} \alpha(n) < \infty\,; \tag{2.7}$$

(ii) Q is ergodic if

$$\liminf_{n\to\infty} \alpha(n) = \alpha > 0\,; \tag{2.8}$$

(iii) Q is null recurrent if

$$\sum_{n=1}^{\infty} \alpha(n) = \infty \quad \text{and} \quad \lim_{n\to\infty} \alpha(n) = 0. \tag{2.9}$$

Assuming $Q(0) = \mathcal{V}$ with probability one, we have the inequality

$$\alpha(n) \leq P\left(\nu_i(n) = \tilde{\nu}_i(n), \forall i \in \mathcal{S}\right) \stackrel{\text{def}}{=} \beta(n). \tag{2.10}$$

Expression (2.4), giving $\nu_i(n) - \tilde{\nu}_i(n)$ in terms of the vectors $\delta(s), s \leq n$, together with condition (2.6), which yields

$$E\,\delta_i(n) = 0, \quad i \in \mathcal{S},$$

permit to use the *Local Central Limit Theorem*, since the vectors $\delta(n), n \geq 1$, are *i.i.d.*. Consequently,

$$\beta(n) \sim C/n^{\frac{N-1}{2}}, \tag{2.11}$$

so that, with (2.10), condition (2.7) holds for $N > 3$.

Case $N = 2, 3$. Showing the property for $V = 1$ will ensure it holds for any V. Thus take $V = 1$ and let $\vec{\phi}(n)$ be the vector with components

$$\phi_i(n) = \nu_i(n) - \tilde{\nu}_i(n), \quad \forall i \in \mathcal{S},\ \forall n > 0.$$

Introduce the events

$$C_n \stackrel{\text{def}}{=} \left\{\vec{\phi}(n) = \vec{0}\right\}$$

and

$$L_{i1}(n) \stackrel{\text{def}}{=} \bigcup_{s=1}^{n} \left\{\mathbb{1}_{\{A_i(s)\}} = \mathbb{1}_{\{A_{i1}(s)\}} = 1\right\} \left\{\mathbb{1}_{\{A_i(k), k>s\}} = 0\right\}.$$

$L_{i1}(n)$ says that, among the n first arrivals in the network, the destination of the last client arrived at node i was node 1.
Clearly $C_n \subset \{S(n) = 1\}$. Moreover the event $\{S(n) = 1; \phi_1(n) = 0\}$ says, in particular, that among the n first arrivals, all clients whose destination was 1 have been served right after the n^{th} external arrival. Now the result follows from the Central limit theorem.

(i) $N = 2$. We have, for all n sufficiently large,

$$\frac{D_1}{\sqrt{n}} \sim P(C_n \cap L_{21}(n)) \le P(Q(n) = \mathcal{V}) \le P(C_n) \sim \frac{D_2}{\sqrt{n}},$$

for some constants D_1 and D_2, which proves the null recurrence according to (2.9).

(ii) $N = 3$. First, the irreducibility of the routing matrix P allows to assume *ad libitum*, for instance $p_{23}p_{31} \neq 0$. Then, for all n sufficiently large, there exit constants D_3 and D_4 such that

$$\frac{D_3}{n} \sim P(C_n \cap L_{21}(n) \cap L_{31}(n)) \le P(Q(n) = \mathcal{V}) \le P(C_n) \sim \frac{D_4}{n}.$$

Again (2.9) is satisfied and the theorem is completely proved. ■

3 Local behaviour

A detailed classification of the nodes will be presented in the next 5 theorems. For many quantities of interest, the reader is referred to definitions given in section 1, in particular in (1.1) and (1.2).

Theorem 3.1 *If*

$$\gamma = \pi,$$

then all the nodes are non ergodic.

Proof For the non ergodicity, it suffices to carry out the analysis in the case $\tau = 0$ and then to prove that, for any k, the sequence $\{q_k(n), n \ge 1\}$ forms a submartingale with respect to the family of σ-algebras $\{\mathcal{F}_n, n \ge 1\}$. We have

$$q_k(n+1) = q_k(n) = 0, \quad \forall k \in v_n, \tag{3.1}$$

and

$$q_k(n+1) - q_k(n) \ge \mathbb{1}_{\{A_k(n)\}} - \sum_{i \in v_n} {}_{v_n}N_{ik}\, \mathbb{1}_{\{A_i(n)\}}, \quad \forall k \in \overline{v}_n. \tag{3.2}$$

It follows from (3.2) that, for all $k \in \overline{v}_n$,

$$E(q_k(n+1) - q_k(n)/\mathcal{F}_n) \geq \pi_k - \sum_{i \in v_n} \pi_i \; E({}_{v_n}N_{ik}).$$

Setting

$$\begin{aligned}
\mathbf{P}_{\overline{v}_n\overline{v}_n} &= (p_{ij}, \; i \in \overline{v}_n, \; j \in \overline{v}_n), \\
\mathbf{P}_{v_n\overline{v}_n} &= (p_{ij}, \; i \in v_n, \; j \in \overline{v}_n), \\
\pi_{v_n} &= (\pi_i, \; i \in v_n),
\end{aligned}$$

the above relation yields, for all $k \in \overline{v}_n$,

$$\begin{aligned}
E(q_k(n+1) - q_k(n)/\mathcal{F}_n) &\geq \pi_k - \sum_{i \in v_n} \pi_i \sum_{j \in \overline{v}_n} p_{ij} E({}_{v_n}N_{jk}) \\
&= \pi_k - \left({}^t\pi_{v_n} \mathbf{P}_{v_n\overline{v}_n} \sum_{q \geq 0} (\mathbf{P}_{\overline{v}_n\overline{v}_n})^q \right)_k = 0,
\end{aligned}$$

which, together with (3.1), shows that $\{q_k(n), n \geq 1\}$ is a submartingale. Consequently q_k will no reach a compact set containing 0 in an integrable time and the nodes are non ergodic. The proof is concluded. ∎

In fact, there is a more precise result.

Theorem 3.2 *If*

$$\gamma = \pi \quad \text{and} \quad \tau = 0,$$

then each node is null recurrent and has always a positive probability of being empty.

Proof Let $\alpha_i(n)$ be the probability that node i is empty at the instant of the $(n+1)$-th arrival in the system. We shall prove that

$$\liminf_{n \to \infty} \alpha_i(n) = \alpha_i > 0,$$

and this will be sufficient for node i to be null recurrent. Here, there is no loss of generality in assuming $V = 1$ and $q_i(0) = 0, \forall i \in \mathcal{S}$. Then, with the notation of theorem 2.1, we have

$$\bigcap_{k \neq i} \{\nu_k(n) > \tilde{\nu}_k(n)\} \subset \{S(n) = i\} \subset \{q_i(n) = 0\}.$$

Thus

$$P(q_i(n)=0) \geq P\left(\bigcap_{k\neq i}\left\{\frac{\sum_{t=1}^{n}\delta_k(t)}{\sqrt{n}}>0\right\}\right).$$

Hence, using $\gamma=\pi$ and the *Central Limit Theorem*, we get

$$\liminf_{n\to\infty}\alpha_i(n)=\alpha_i>0.$$

■

Theorem 3.3 *When $\gamma=\pi$, the nodes are*

(a) *transient if $\lambda\tau>V$;*

(b) *recurrent if $\lambda\tau<V$, under the additional assumption*

$$E(\tau_{ij}\tau_{kl})<\infty, \quad \forall i,j,k,l\in\mathcal{S}, \tag{3.3}$$

and in this case all conclusions of theorem 2.2 hold.

Proof

Case (a) One can use standard queueing theory arguments. Consider an arbitrary node, i say. It can be viewed as a single queue with V servers working in parallel and an equivalent average service time which is larger than

$$\sum_{k\neq i}\frac{\pi_k\tau_k}{\pi_i}+\tau_i=\frac{\tau}{\pi_i}.$$

The corresponding traffic intensity ρ_i satisfies the inequalities

$$\rho_i\geq\frac{\lambda\tau\gamma_i}{\pi_i}=\lambda\tau>V,$$

which means that each node is transient.

Case (b) Introduce the system $\widetilde{Q}$, obtained from Q by taking $\tau=0$. This means in particular that Q and $\widetilde{Q}$ work under the same arrival and routing processes. Assuming $Q(0)=\widetilde{Q}(0)$, the idea is to couple Q and $\widetilde{Q}$ pathwise, at properly chosen instants.

Step 1. Let us show first that, if all cars are blocked at a given node i at some time T in system Q, then $Q(T) = \widetilde{Q}(T)$. Let $U_{ij}(t)$ [resp. $\widetilde{U}_{ij}(t)$] be the number of clients who arrived at node i and have been transported to node j on the time interval [0,t] in Q [resp. $\widetilde{Q}$]. The machinery of the two systems ensures that

$$U_{ij}(t) \le \widetilde{U}_{ij}(t), \quad \forall i, j \in \mathcal{S}, \forall t \ge 0. \tag{3.4}$$

Now the definition of T yields

$$U_{ii}(T) = \widetilde{U}_{ii}(T) \quad \text{and} \quad \sum_{j=1}^{N} U_{ji}(T) = \sum_{j=1}^{N} \widetilde{U}_{ji}(T),$$

whence, by (3.4),

$$U_{ji}(T) = \widetilde{U}_{ji}(T), \quad \forall j \in \mathcal{S}. \tag{3.5}$$

Suppose there $\exists j \neq i$, such that $q_j(T) > \tilde{q}_j(T)$ and let r be the destination node of the last customer served at node j in system $\widetilde{Q}$, before time T. But then $U_{jr}(T) < \widetilde{U}_{jr}(T)$. For $r = i$, this immediately contradicts (3.5). Otherwise, $q_r(T) > \tilde{q}_r(T)$, since there were more visits of cars at node i, before time T, in system $\widetilde{Q}$ than in system Q, remembering that at time T all cars are blocked at node i. Now, by induction, there would exist s, such that $U_{si}(T) < \widetilde{U}_{si}(T)$, contradicting again (3.5).

Step 2. Now we will construct an increasing sequence of stopping times $\{T_k, k \ge 1\}$ satisfying the following properties:

$$E(T_{k+1} - T_k / \mathcal{F}_{T_k}) < C < \infty \quad \text{and} \quad \exists\ i(k) \text{ with}$$

$$q_{i(k)}(T_k) = 0, \quad x_{i(k)}(T_k) = V, \ \forall k \ge 0.$$

Take, using the notation (1.2),

$$\vec{0} < \vec{q}(0) < \vec{K},$$

with strict inequalities, $\vec{K}$ being a fixed positive vector. The point is to prove that at least one arbitrary component of $\vec{q}(t)$ will reach 0 in an integrable time. This will be done by means of a supermartingale argument, focusing on the evolution of a fixed component, say q_i, at arrival times $u_1, u_2, \ldots,$ at node i. Clearly, the time necessary to empty one arbitrary component is not greater than the time to empty q_i, conditioned on the fact that all other components remain strictly positive in the meanwhile.

This reads more precisely, setting $\mathcal{G}(t) = \bigcap_{l=1}^{N} \{q_l(r) \neq 0, \forall r < t\}$,

$$\inf\{t \geq 0/\ \exists j, q_j(t) = 0, \forall j \in \mathcal{S}\} \leq \inf\{u_s, s \geq 1, 1\!\!1_{\{\mathcal{G}(u_s)\}} 1\!\!1_{\{q_i(u_s)=0\}} = 1\}.$$

Let $M_{vk}((a,b]), 1 \leq v \leq V$, be the number of visits made by an arbitrary vehicle v going from node k to node i, on the time interval $(a,b]$. For the sake of shortness, the position of v at $t = 0$ is omitted. Then the following global equation holds:

$$\frac{1}{L} E[q_i(u_L) - q_i(0)/\ \mathcal{F}_0] = 1 - \frac{1}{L} \sum_{s=1}^{L} \sum_{v=1}^{V} \sum_{k=1}^{N} p_{ki} E[M_{vk}((u_{s-1}, u_s])/\mathcal{F}_0], \text{on } \mathcal{G}(u_L). \tag{3.6}$$

Choose $\epsilon > 0$ arbitrarily small. From the ergodic renewal theorem, together with additive properties of the M_{vk}'s and (3.3), there exists L (depending on ϵ) in (3.6), such that

$$\frac{1}{L} E[q_i(u_L) - q_i(0)/\ \mathcal{F}_0] \leq 1 - V \sum_{k=1}^{N} \frac{p_{ki}\pi_k}{\lambda\tau\gamma_i} + \epsilon = 1 - \frac{V}{\lambda\tau} + \epsilon < 0, \text{ a.s on } \mathcal{G}(u_L). \tag{3.7}$$

The supermartingale constructed in (3.7) will reach 0 (see e.g. [6]) after an integrable time and remains at 0 with a positive probability. Moreover when a queue, say m, is empty, one can always force all cars to go to node m, before the next arrival at m occurs: this takes a (residual) time R which, by (3.3), is uniformly integrable with respect to the positions of cars at time u_L in equation (3.7).

Hence, $T_1 \leq U_L + R$. Then letting

$$A(T_1) = \inf\{t \,/\, t \geq T_1,\ q_j(t) > 0, \forall j \in \mathcal{S}\},$$

one can simply choose

$$T_2 - T_1 = \begin{cases} \Delta + \widetilde{R} & \text{if } A(T_1) > \Delta, \\ A(T_1) + \widetilde{U} + \widetilde{R} & \text{otherwise,} \end{cases}$$

where Δ is a fixed constant and $\widetilde{U}, \widetilde{R}$ are respectively obtained in the same way as U, R above. The introduction of Δ is just in case the second moment of the residual interarrival time would not exist.

This procedure can be repeated, remarking that, once a coordinate has reached the value 0, its next strictly positive value at some arrival instant will be exactly 1.
Thus the existence of a sequence $\{T_n, n \geq 1\}$, with $\sup\limits_n E(T_{n+1} - T_n) < \infty$, is established.

Step 3. Since

$$Q(T_n) = \widetilde{Q}(T_n),\ n \geq 1,$$

the conclusion of part (b) is immediate from theorems 2.2 and 3.2, using the uniform boundedness of $E(T_{n+1} - T_n)$. Theorem 3.3 is completely proved. ∎

In the next two theorems, pathwise comparisons are needed between the process $\widetilde{Q}$, introduced in theorem 3.3, and its restriction to a polling system where only a subset of the nodes is visited.

Set $\Lambda \stackrel{\text{def}}{=} \{i_1, \ldots, i_k\}$, for all k-tuple of integers $1 \leq i_1 < \ldots < i_k \ldots \leq N$, and let $\widetilde{\mathbf{P}}^\Lambda$ denote the transition matrix of the restriction to Λ of a Markov chain with state space $\mathcal{S}$ and transition matrix $\mathbf{P}$. Introducing the two vectors $\widetilde{\pi}^\Lambda$ and $\widetilde{\gamma}^\Lambda$, with components

$$\widetilde{\pi}_j^\Lambda = \frac{\pi_j}{\sum_{i \in \Lambda} \pi_i}, \quad \widetilde{\gamma}_j^\Lambda = \frac{\gamma_j}{\sum_{i \in \Lambda} \gamma_i}, \quad \forall j \in \Lambda, \tag{3.8}$$

it is well known (see e.g. [7]) that the invariant measure of $\widetilde{\mathbf{P}}^\Lambda$ is indeed given by $\widetilde{\pi}^\Lambda$. The exact form of $\widetilde{\mathbf{P}}^\Lambda$ can be found in [7] but is not needed there.

To avoid uninteresting technicalities, we take $x_i(0) = 0,\ \forall i \in \overline{\Lambda}$. Let $\tau = 0$ and consider the process $\widetilde{Q}^\Lambda$ on $\mathcal{Z}_+^\Lambda \times \mathcal{Z}^{\overline{\Lambda}}$ constructed as follows:

- it corresponds to a new polling system with zero transfer times;
- its sample paths are obtained by assuming the server never stops at the nodes in $\overline{\Lambda}$ and that the arrival and routing processes are the same as for Q.

Then, writing $\tilde{q}_k^\Lambda(n)$ for the number of customers at the n-th arrival instant, waiting at node $k, \forall k \in \Lambda$, in the system $\tilde{Q}^\Lambda$, we have

$$\tilde{q}_i^\Lambda(n) \leq q_i(n), \quad \forall i \in \Lambda, \quad \forall n \geq 0.$$

Theorem 3.4 *Let $\partial B = \{(x_1, \ldots, x_N) \in \mathcal{Z}_+^N \,; x_i = 0, \forall i \in \Lambda\}$, for any arbitrary Λ, and assume $\gamma = \pi$. Then*

$$P\Big((q_1(n), \ldots, q_N(n)) \in \partial B, \text{ for infinitely many } n\Big) = 0, \quad \forall\, |\Lambda| \geq 4.$$

In other words, the boundaries of codimension ≥ 4 are transient.

Proof As already remarked, it suffices to prove the statement for $\tau = 0$, since the case $\tau \neq 0$ follows by direct sample path comparison. Consider $\tilde{Q}^\Lambda$, with $\Lambda = \{i_1, \ldots, i_k\}$. Here $\tilde{\gamma}^\Lambda = \tilde{\pi}^\Lambda$ (see (3.8)) and theorem 2.2 implies

$$P(\tilde{q}_{i_1}^\Lambda(n) = \ldots = \tilde{q}_{i_k}^\Lambda(n) = 0) \sim \frac{C^\Lambda}{n^{\frac{k-1}{2}}},$$

for some constant C^Λ. Since $|\Lambda| \geq 4$, it follows that

$$\sum_{n \geq 0} P(q_{i_1}(n) = \ldots = q_{i_k}(n) = 0) < \infty$$

and the proof is concluded by using the Borel-Cantelli lemma. ∎

The next theorem deals with the complete classification when $\gamma \neq \pi$, in which case we introduce a numbering of the nodes, according to the following inequalities:

$$\frac{\gamma_1}{\pi_1} = \cdots = \frac{\gamma_{l-1}}{\pi_{l-1}} < \frac{\gamma_l}{\pi_l} \leq \cdots \leq \frac{\gamma_N}{\pi_N}. \tag{3.9}$$

Note that, since $l \leq N$, (3.9) excludes $\gamma = \pi$.

Theorem 3.5 *Assume (3.9) holds.*

(a) *For any $i \geq l$,* $\lim_{n \to \infty} q_i(n) = \infty$ a.s.

(b) *Assume $l = 2$ and put $\rho_1 = \dfrac{\lambda \tau \gamma_1}{\pi_1}$. Then,* (b1) *for* $\rho_1 < V$, *node* 1 *is ergodic;*

(b2) *for* $\rho_1 > V$, *node* 1 *is transient.*

(c) *Assume* $l \geq 3$. *Then the set of nodes* $\Lambda \stackrel{def}{=} \{1, \ldots, l-1\}$ *behaves as a polling system, with parameters* γ^Λ, π^Λ *and* $\mathbf{P}^\Lambda$, *which have been formally defined in the preamble before theorem 3.4.* mutatis mutandis.

This theorem claims, among other things, that the original polling system has at most one ergodic node when $\gamma \neq \pi$.

Proof

(a) Choosing a node i with $i \geq l$, we consider the system $\widetilde{Q}^\Lambda$ corresponding to $\Lambda = \{1, i\}$. Here $\widetilde{\gamma}^\Lambda \neq \widetilde{\pi}^\Lambda$ and, more precisely, $\widetilde{\gamma}_i^\Lambda > \sum_{j \in \Lambda} \widetilde{\gamma}_j^\Lambda \widetilde{p}_{ji}^\Lambda$. Using now theorem 2.1, it follows that $\widetilde{q}_i^\Lambda$ is transient and so is q_i, since $\widetilde{q}_i^\Lambda(n) \leq q_i(n)$, $\forall n \geq 1$.

(b) It follows from part (a) that

$$\lim_{n \to \infty} q_i(n) = \infty, \quad \forall i \geq 2, \text{ a.s.}$$

Hence, the traffic intensity at node 1 is exactly

$$\rho_1 = \frac{\lambda \tau \gamma_1}{\pi_1}$$

and the result follows easily.

(c) Using again part (a),

$$\lim_{n \to \infty} q_i(n) = \infty, \quad \forall i \geq l, \text{ a.s.}$$

So, for n sufficiently large, the cars will not stop anymore at the nodes belonging to $\overline{\Lambda}$. In fact, omitting the details, one sees that the original network decomposes into two subsystems:

- a system of $N - l + 1$ transient nodes.
- a polling system Q^Λ of $l - 1$ nodes.

The proof of the theorem is concluded. ∎

4 Time scaling and limit laws

When $\gamma = \pi$, it has been shown above that each node is null-recurrent and the system is either null recurrent or transient. In this section, limit laws are obtained for the joint distribution of the position of the server and the number in the queues.

Notation Here vectors will often be written with arrows (e.g. $\vec{x}$, $\vec{q}(n)$, $\vec{k}$, etc.). Setting

$$ {}_i\overrightarrow{N} \stackrel{\text{def}}{=} ({}_iN_{i1}, \ldots, {}_iN_{iN}), \quad \forall i \in \mathcal{S}, $$

the quantities $\left\{\overrightarrow{Y}(s;i), s \geq 1\right\}$ will stand for a sequence of vectors which, for each fixed i, are i.i.d., independent of the arrival process and distributed as ${}_i\overrightarrow{N}$. Let

$$ {}_i\overrightarrow{H}_j \stackrel{\text{def}}{=} ({}_iH_{j1}, \ldots, {}_iH_{jN}), \quad \forall i \in \mathcal{S}, $$

where ${}_iH_{jk}$, in the Markov chain with matrix $\mathbf{P}$, is the number of visits to k, starting from j, without hitting i. In addition, define

$$ \overrightarrow{Z}(n;i) \stackrel{\text{def}}{=} \sum_{s=1}^{n} \left(\mathbb{1}_{\{A_i(s)\}} \overrightarrow{Y}(s;i) - \sum_{j=1}^{N} \mathbb{1}_{\{A_j(s)\}} \vec{e}_j \right), \quad \forall i \in \mathcal{S}, $$

where $\vec{e}_j$ is the j-th unit vector of R^N and, by convention, ${}_iN_{ii} = 1, \forall i \in \mathcal{S}$. Letting $\Delta\overrightarrow{Z}(n;i) \stackrel{\text{def}}{=} \overrightarrow{Z}(n+1;i) - \overrightarrow{Z}(n;i)$, it follows (since $\gamma = \pi$) that

$$ E[\Delta\overrightarrow{Z}(n;i)] = \vec{0}, \quad \forall i \in \mathcal{S}. $$

Introduce also the following covariance matrices, which do not depend on n:

$$ \Gamma_i \stackrel{\text{def}}{=} \Big(E(\Delta Z_j(n;i)\ \Delta Z_k(n;i)) \Big)_{j,k \in \mathcal{S}}, \quad \forall i \in \mathcal{S}. $$

In the sequel $\overrightarrow{\mathcal{N}}(\Gamma_i)$ will represent a random vector in R^N, normally distibuted, centered at the origin, having its i-th coordinate identically 0 and Γ_i as covariance matrix. Using the symbols " (*) " for convolution, we are in a position to state the main theorem.

Theorem 4.1 *If $\gamma = \pi$, $\tau = 0$, $\vec{x}(0) = \vec{x}$, $\vec{q}(0) = \vec{0}$, $\forall j \in \mathcal{S}$, then*

$$\begin{aligned}
\lim_{n\to\infty} P(\exists i \neq j, \{i,j\} \subset v_n) &= 0, & (4.1)\\
P\left(v_n = \{i\}; \frac{\vec{q}(n)}{\sqrt{n}} \leq \vec{k}\right) &= P\left(\vec{0} \leq \frac{\overrightarrow{Z}(n;i)}{\sqrt{n}} \leq \vec{k} + \sum_{j\neq i} \frac{{}_i\vec{H}_j^{(*)x_j}}{\sqrt{n}}\right), & (4.2)\\
\lim_{n\to\infty} P\left(v_n = \{i\};\ \frac{\vec{q}(n)}{\sqrt{n}} \leq \vec{k}\right) &= P\left(\overrightarrow{\mathcal{N}}(\Gamma_i) \in [\vec{0}, \vec{k}]\right), & (4.3)
\end{aligned}$$

$\forall \vec{k} \in \mathcal{R}_+^N,\ \forall i, j \in \mathcal{S}.$

Proof Equation (4.1) is a staightforward consequence of theorem 3.4 and the limit law in (4.3) is obtained by letting $n \to \infty$ in (4.2). It is also worth noting the choice $\vec{k} = \vec{\infty}$ in (4.2), which gives the time-dependent probability of having at least one vehicle at a given node. Thus we are left with the proof of (4.2).

The argument will first be developed for $V = 1$ and relies on a counting (and in some sense combinatorial) argument, where the time does not play any role. Consider the system at the n-th arrival instant and condition its state on

$$\vec{\nu}(n) = \vec{b} \stackrel{\text{def}}{=} (b_1, b_2, \ldots, b_N),$$

where $\vec{\nu}(n)$ is the vector of the number of arrivals at each queue, introduced in section 2. Let $X(t;n)$ denote the position of the server after t effective visits (or services) up to time n. Clearly, $t \leq n$ and the evolution of $X(t;n)$ is Markovian, with transition matrix $\mathbf{P}$. Thus one can write, for some fixed i, with the notation $S(n)$ instead of v_n since $V = 1$,

$$P(S(n) = i\,/\vec{\nu}(n); S(0) = i) = P\left({}_iN_j^{(*)\nu_i(n)} \leq \nu_j(n),\ \forall j \neq i\,/\vec{\nu}(n); S(0) = i\right),$$

which in turn yields

$$P(S(n) = i\,/\ S(0) = i) = P\left(\frac{\overrightarrow{Z}(n;i)}{\sqrt{n}} \geq \vec{0}\right).$$

Similarly, remarking that the event $\{\vec{q}(n) = \vec{k}\}$ can occur only if $X(t; n)$ takes the value i exactly b_i times and the value j exactly $b_j - k_j$ times, for all $j \neq i$, we get

$$P\Big(S(n) = i; \frac{\vec{q}(n)}{\sqrt{n}} \leq \vec{k} \,/\, S(0) = i\Big) = P\left(\vec{0} \leq \frac{\overrightarrow{Z}(n; i)}{\sqrt{n}} \leq \vec{k}\right),$$

and the result follows from the central limit theorem.

The extension to $V > 1$ is not difficult, noting simply that when all cars are at some node i, the procedure for $V = 1$ can be reproduced. Details are omitted.

■

The last theorem given hereafter does in some sense justify, beyond its theoretical interest, the detailed analysis made for the system $\tilde{Q}$ in which $\tau = 0$.

Theorem 4.2 *When $\gamma = \pi$ and $\lambda\tau < V$, the distribution of $Q(t)$, as $t \to \infty$, satisfies*

$$\lim_{t\to\infty} P\left(\frac{\vec{Q}(t)}{\sqrt{t}} \leq \vec{k}\right) = \sum_{i=1}^{N} P\left(\overrightarrow{\mathcal{N}}(\Gamma_i) \in [\vec{0}, \lambda^{1/2}\vec{k}]\right), \quad \forall \vec{k} \in \mathcal{R}_+^N. \tag{4.4}$$

The remarkable fact is that, after a scaling in $1/\sqrt{t}$, neither transfer times nor the number of vehicules do appear in the explicit form of the limiting distribution. This is indeed a phase transition phenomenon. *Obviously, this is not the case as far as speed of convergence is concerned.*

Proof Let $\nu(a, b) = \sum_{i=1}^{N} \nu_i(b) - \nu_i(a)$ be the total number of customers arrived in the time interval (a, b). Pathwise, we have

$$\tilde{Q}(t) \leq Q(t) \leq \tilde{Q}(t) + \nu(t - T_{l(t)}), \tag{4.5}$$

where $\{T_n, n \geq 1\}$ is the sequence introduced in theorem 3.3 and

$$l(t) = \inf\{n \,/\, T_n \leq t < T_{n+1}\}.$$

But the constructive procedure of this sequence shows that the random variables $T_{n+1} - T_n, n \geq 1$ can be bounded by the increments of a renewal point process, having a proper distribution for its stationary residual time, whence

$$\lim_{t\to\infty} P\left(\frac{\nu(t - T_{l(t)})}{\sqrt{t}} > a\right) = 0, \ \forall a \geq 0,$$

and the relation (4.4) follows directly. The theorem is proved. ■

Remark *The fact that τ does not take place in (4.4) can be explained by coupling. The non-influence of V follows from theorem 3.4 and equation (4.4): asymptotically, all cars are blocked at some node (not always the same) and only one customer can be transported, no matter the value of V may be.*

Acknowledgement

The authors are grateful to M. Parent (program director in the PRAXITELE project now developed at INRIA) for many fruitful discussions, which helped to the definition of the model analyzed in this paper.

References

[1] A.A. Borovkov and R. Schassberger. Ergodicity of a polling network, Stochastic Processes & Appl., 50, (1994), pp. 253-262.

[2] Ch. Fricker and M.R. Jaibi. Stability of a polling model with a Markovian scheme, Rapport de Recherche INRIA N^0 2278, May 1994.

[3] G. Fayolle and Jean-Marc Lasgouttes. A State-dependent polling model with Markovian routing. *IMA Volumes in Math. & its Applications*, 71, 1995, pp.283-301.

[4] L.G. Afanassieva, G. Fayolle, S.Yu. Popov. Models for transportation networks, Rapport de recherche INRIA, N^0 2837, March 1996.

[5] I.A. Ibragimov and Yu.V. Linnik. *Independent and stationary sequences of random variables*, Wolters Noordhoff Publishing, Gronigen.

[6] G. Fayolle, V.A. Malyshev, M.V. Menshikov. *Topics in the Constructive Theory of Markov Chains*, Cambridge Univ. Press, 1995

[7] W. Feller. *An Introduction to Probability Theory and Applications*, Second Edition.

[8] J. Jacod, A.N. Shiryaev. *Limit Theorems for Stochastic Processes*, Springer 1987.

[9] Kai Lai Chung. *Markov chains with stationary transition probabilities*, Springer, Berlin 1967.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l' Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399