Journal of Global Optimization **15:** 299–314, 1999.

On optimization of continuous-time Markov networks in distributed computing

ALAIN BUI
LaRIA, Université de Picardie Jules Verne, France

(Accepted in revised form 28 July 1999)

Abstract. The paper presents a new stochastic model for studying the optimization of functioning rules in distributed computing. In this model a network is represented by a finite number of continuous-time homogeneous Markov processes which are connected by relations between entries of their intensity matrices. Good functioning rules are those optimizing a guide function defined according to the context. Two specific optimization problems are studied: a problem of resource allocation with conflicts between processes, and a problem of access to shared resources. The latter is a linearly constrained nonconvex problem with an objective function which is a sum of ratios of linear functions of special form.

Key words: Modeling, Markov process, Distributed computing

1. Introduction

The aim of this paper is to develop a new continuous-time stochastic model for optimizing the functioning of distributed computing networks. Although this model is technically more difficult to handle than the discrete-time model developed in our previous papers [2, 4, 5] it is closer to reality, and at the same time, more flexible for the applications. Taking into account the importance of local considerations, the model involves local states for each site, transitions between these states and relations characterizing the network, along with a guide function which defines the criterion of good functioning. More specifically, a distributed system is represented by a finite number of interconnected homogeneous continuous-time Markov processes. The interconnection into network is ensured by relations between the entries of the intensity matrices of these Markov processes. The entries of the intensity matrices are the ‘variables’ of the model, the relations describing the interconnection constitute the ‘constraints’, while the guide function defining the criterion of functioning is the objective function to be optimized. An optimal functioning rule of the system is then an optimal solution of this constrained optimization problem. As illustrations, we shall discuss two optimization problems arising from this model: the problem of resource allocation with conflicts between processes, formulated through the model of the ‘dining philosophers problem’, and the problem of access to shared resources with mutual exclusion. It turns out that in each considered case, by exploiting the particular structure of the problem, an optimal solution can be found in closed form, even though the objective function may be highly nonconvex (sum of ratios of linear functions). The emphasis on closed forms for optimal solutions is motivated by our primary interest in clarifying general optimal functioning rules rather than in solving specific numerical problems.

We refer the reader to [1, 6], [14, 13] for the general concepts, models and applications of distributed computing; to [2–5], for studies on Optimization and Performance Evaluation in Distributed Computing, especially via a discrete-time stochastic model; to [15] for basic concepts and results in global optimization; and finally to [10, 11] for basic concepts and results on Probability Theory and Stochastic Processes.

2. Continuous-time Markov networks

Let $({}^kX_t)_{t\geqslant 0},\ k \in \{1, \dots , N\}$ be N homogeneous continuous-time Markov processes defined on a probability space $(\Omega, \mathcal{A}, \Pr)$, having the same finite state space S.

Suppose that their stochastic matrices

$$ {}^kP(t) = \left({}^kp(t, i, j)\right)_{j,j\in S}, \qquad t > 0,\ k \in \{1, \dots , N\}, $$

satisfy the condition $\lim_{t\to 0} {}^kP(t) = I$, I being the identity matrix, viz. $\lim_{t\to 0} {}^k P(t, i, j) = 1$ if $i = j$ and = 0 if $i \neq j$.

We recall that then the limit

$$ {}^kP(t) = \lim_{t>0,t\to 0} \frac{{}^kP(t) - I}{t} = {}^kQ \quad \text{exists,} $$

where kQ is a matrix whose entries satisfy

$$ {}^kq_{ij} \geqslant 0 \quad \text{if} \quad i \neq j, \quad \text{and} \quad {}^kq_{ii} \leqslant 0, \quad \text{with} $$
$$ \sum_{j\in S} {}^kq_{ij} = 0, \quad i \in S, \quad k \in \{1, \dots , N\}. $$

For convenience, denote ${}^kq_i = -({}^kq_{ii}) \geqslant 0, i \in S$. The following results will be used in the next section:

a) ${}^kP(t) = e^{{}^kQt} = I + \sum_{n\geqslant 1} ({}^kQ)^n \frac{t^n}{n!}, \quad t \geqslant 0,$ ${}^kQ$ is called the *intensity matrix* of ${}^kP(t)$. (cf. [10], p. 238)
b) The limit $\lim_{t\to\infty} {}^kP(t, i, j) = {}^k\pi(i, j), i, j \in S$, exist for any k and their convergence is exponentially fast. In matrix notation,

$$ \lim_{t\to\infty} {}^kP = {}^k\Pi \text{ and } {}^k\Pi.{}^kQ = 0 $$

In the present case, we suppose that $\forall k \in 1, \dots, N$ there is only one class of (recurrent) states. The ${}^k\pi(i, j)$'s are then independent of i and we shall denote them by ${}^k\pi_j$. In this case, ${}^k\Pi$ is a matrix with identical rows, and a row of ${}^k\Pi$ is a stationary probability distribution for the k^{th} process.

c) Let now ${}^kW = ({}^kw(i, j))_{i,j\in S}$ be the matrix whose entries ${}^kw(i, j)$ are the means $E_i({}^k\tau_j)$ of the first passage times ${}^k\tau_j$ (of the k^{th} process) starting from $i, {}^k\tau_j$ being defined as ${}^k\tau_j = \min\{t \geqslant 0 : \ {}^kX_t = j\}$. Then

$$ {}^kW = (\hat{E}.{}^kZ_{dg} - {}^kZ).({}^k\Pi_{dg})^{-1} \tag{1} $$

where $\hat{E}$ is the matrix whose entries are all equal to 1, ${}^kZ = ({}^k\Pi - {}^kQ)^{-1}$ and ${}^kZ_{dg}$ denotes the matrix resulting from setting all the entries of ${}^kZ$ off the main diagonal equal to 0, and similarly for ${}^k\Pi_{dg}$. (cf. [10], Proposition 8.9, p. 255).

d) Let ${}^k\mu(i)$ be the mean recurrence time of state i, i.e., the mean time elapsed between an entry into state i and the next return to i (of the k^{th} process), consisting of a sojourn time in i ending in a jump to some $j \in S, j \neq i$, followed by a first passage from j to i. Then

$$ {}^k\mu(i) = \frac{1}{{}^k\pi j.{}^kq_i}, \qquad i \in S \tag{2} $$

Let $k\xi(i)$ be the Smoluchowskian mean recurrence time (cf.[10], p. 256), i.e. the mean time elapsed between an exit from i and the next return to i. Then

$$ {}^k\xi(i) = \frac{1 - {}^k\pi_i}{{}^k\pi_i.{}^kq_i} \tag{3} $$

In our model, the N continuous time Markov processes represent N processors, connected in network by relations $\mathcal{R}_h, h \in H$, between entries of their intensity matrices ${}^kQ, k \in \{1, \dots, N\}$. According to the context, a guide function F is defined on the entries of ${}^kQ, \ k \in \{1, \dots, N\}$. A functioning rule for such a network is represented by a vector whose components are the entries ${}^kq(i, j), \ i, j \in S, k \in \{1, \dots, N\}$, of intensity matrices satisfying $\mathcal{R}_h, h \in H$. Choices for good functionings are based on criteria which optimize F under constraints $\mathcal{R}_h, h \in H$.

In general many criteria are available. In this paper we focus on those criteria for which the optimal solution of the corresponding problem can be found in closed form, so that general functioning rules based on these criteria can be explicitly derived in terms of the values of the basic parameters of the problem.

Next we discuss the application of this approach to two typical problems: the so-called 'dining philosophers problem' which models a problem of resource allocation with conflicts, and the problem of access to shared resources, which is a mutual exclusion problem.

3. A problem of resource allocation with conflicts

In this problem of resource allocation with conflicts between processes in distributed computing (cf. [7,9]), the N processors are represented by N philosophers arranged in a circle around a spaghettis plate, with a fork between each pair of philosophers. In order to eat, each philosopher needs his two adjacent forks. As above-mentioned, we can consider several possibilities for using the model for a given problem. Here we propose to examine two typical cases.

a) In the first case, the state space S consists of four elements: 'waiting (request)' (state 1), 'with one fork (refusal)' (state 2), 'with 2 forks (acceptance)' (state 3), 'thinking (execution)' (state 4). The evolution of the k^{th} philosopher is represented by the intensity matrix

$$^{k}Q = \begin{pmatrix} -^{k}q_{12} & ^{k}q_{12} & 0 & 0 \\ 0 & -^{k}q_{23} & ^{k}q_{23} & \\ 0 & 0 & -^{k}q_{34} & -^{k}q_{34} \\ ^{k}q_{41} & 0 & 0 & -^{k}q_{41} \end{pmatrix}$$

For a small lapse of time Δt, the corresponding stochastic matrix $^{k}P(\Delta t)$ is equivalent to

$$\begin{pmatrix} 1-{}^{k}q_{12}\Delta t & ^{k}q_{12}\Delta t & 0 & 0 \\ 0 & 1-{}^{k}q_{23}\Delta t & ^{k}q_{23}\Delta t & \\ 0 & 0 & 1-{}^{k}q_{34}\Delta t & -^{k}q_{34}\Delta t \\ ^{k}q_{41}\Delta t & 0 & 0 & 1-{}^{k}q_{41}\Delta t \end{pmatrix}.$$

An efficient functioning of the network should ensure that the time during which the philosophers keep themselves busy to eat is as long as possible. This suggests that the chance of a philosopher to pass from state 2 to state 3 should be linked with the chance of his neighbour to pass from state 1 to state 2, SO that the connection into network can be expressed as

$$\begin{cases} ^{k}q_{12} + {}^{k+1}q_{23} = c_k, \quad k \in \{1, \ldots, N\} \\ ^{N}q_{12} + {}^{1}q_{23} = c_N, \quad \text{where} \quad c_1, \ldots, c_N \ \text{are constant.} \end{cases}$$

PROPOSITION 3.1. *(i) A stationary probability for the k^{th} process is*

$$\left(^{k}\pi_1, \ ^{k}\pi_2, \ ^{k}\pi_3, \ ^{k}\pi_4\right)$$

$$\left(\frac{^{k}q_{23}\cdot{}^{k}q_{34}\cdot{}^{k}q_{41}}{^{k}\Delta}, \frac{^{k}q_{12}\cdot{}^{k}q_{34}\cdot{}^{k}q_{41}}{^{k}\Delta}, \frac{^{k}q_{12}\cdot{}^{k}q_{23}\cdot{}^{k}q_{41}}{^{k}\Delta}, \frac{^{k}q_{12}\cdot{}^{k}q_{23}\cdot{}^{k}q_{34}}{^{k}\Delta},\right)$$

where

$$^{k}\delta = {}^{k}q_{23}\cdot{}^{k}q_{34}\cdot{}^{k}q_{41} + {}^{k}q_{12}\cdot{}^{k}q_{34}\cdot{}^{k}q_{41} + {}^{k}q_{12}\cdot{}^{k}q_{23}\cdot{}^{k}q_{41}{}^{k}q_{12}\cdot{}^{k}q_{23}\cdot{}^{k}q_{34}$$

(ii) The mean time to reach state 3 from state 1 is $\frac{^{k}q_{12}+{}^{k}q_{23}}{^{k}q_{12}\cdot{}^{k}q_{23}}$.

Proof. (i) The Markov process having only a recurrent class of states, ${}^k\Pi$ is a matrix with identical rows $\left({}^k\pi_1,\ \ {}^k\pi_2,\ \ {}^k\pi_3,\ \ {}^k\pi_4\right)$. Solving the equation ${}^k\Pi.{}^kQ = 0$, we obtain the result announced. (ii) Using formula (1), we obtain ${}^kW$, with ${}^kw(1, 3)$ as indicated. □

If we adopt the criterion that the mean time to reach state 3 starting from state 1 should be, globally, as short as possible, then the guide function is

$$\sum_{k=1}^{N} \frac{{}^kq_{12} + {}^kq_{23}}{{}^kq_{12} \cdot {}^kq_{23}}$$

to be minimized under the specified constraints. Since the guide function and the constraints deals only with ${}^kq_{12}$ and ${}^kq_{23},\ k \in \{1, \dots, N\}$, we can assume that ${}^kq_{34}$ and ${}^kq_{41},\ k \in \{1, \dots, N\}$, are independent of k, and denote them by (q_{34} and q_{41}, respectively.

Now, to simplify the notation, let us write x_k to mean ${}^kq_{12}$, and y_k to mean ${}^kq_{23}$. Then, the search for optimal functioning rules amounts to solving the following problem:

Problem 1.

$$\text{Minimize } F(x_1, y_1, \dots, x_k, y_k, \dots, x_N, y_N) = \sum_{k=1}^{N} \frac{x_k + y_k}{x_k \cdot y_k}$$

subject to

$$\begin{cases} x_k + y_{k+1} = c_k,\ k \in \{1, \dots, N\} \\ x_N + y_1 = c_N \end{cases} \tag{4}$$

PROPOSITION 3.2. *There exists a unique optimal functioning rule:* $({}^1Q, \dots, {}^NQ)$, *where*

$$({}^1q_{12}, {}^1q_{23}) = \left(\frac{1}{2}c_1, \frac{1}{2}c_N\right),$$
$$({}^kq_{12}, {}^kq_{23}) = \left(\frac{1}{2}c_k, \frac{1}{2}c_{k-1}\right), \quad k \in \{2, \dots, N\}.$$

Proof. Let $\lambda_1, \dots, \lambda_N$ be the Lagrange multipliers. We have to solve the system of equations:

$$\frac{\partial}{\partial x_k}F + \frac{\partial}{\partial x_k}\sum_{j=1}^{N}\lambda_j L_j = -\frac{1}{(x_k)^2} + \lambda_k = 0, \quad k \in \{1, \dots, N\}.$$

$$\frac{\partial}{\partial y_k}F + \frac{\partial}{\partial y_k}\sum_{j=1}^{N}\lambda_j L_j = -\frac{1}{(y_1)^2} + \lambda_N = 0,$$

$$\frac{\partial}{\partial y_k}F + \frac{\partial}{\partial y_k}\sum_{j=1}^{N}\lambda_j L_j = -\frac{1}{(y_k)^2} + \lambda_k = 0, \quad k \in \{2, \dots, N\}.$$

This implies that $x_k = y_{k+1}$ for $k \in \{1, \dots, N-1\}$ and $x_N = y_1$. By constraints (4), we conclude that $x_k = \frac{1}{2}c_k$, $k \in \{1, \dots, N\}$, and $y_1 = \frac{1}{2}c_N$, $y_k = \frac{1}{2}c_{k-1}$, $k \in \{2, \dots, N\}$ Since F is a convex function, the solution

$$\begin{aligned}&\left(\left({}^1q_{12},\ {}^1q_{23},\right), \dots, \left({}^kq_{12},\ {}^kq_{23},\right), \dots, \left({}^Nq_{12},\ {}^Nq_{23},\right)\right)\\ &\quad = \left(\left(\frac{1}{2}c_1, \frac{1}{2}c_N\right), \dots, \left(\frac{1}{2}c_k, \frac{1}{2}c_{k-1}\right), \dots, \left(\frac{1}{2}c_N, \frac{1}{2}c_{N-1}\right)\right)\end{aligned}$$

is a minimum for F. □

b) Consider now the case when we lump state "waiting" and state "thinking". Then S contains only 3 states. If we allow each philosopher with one fork to abandon its single fork before reaching the state with two forks, then the evolution of the k^{th} process is represented by the intensity matrix

$${}^kQ = \begin{pmatrix} -{}^kq_{12} & {}^kq_{12} & 0 \\ {}^kq_{21} & -({}^kq_{21} + {}^kq_{23}) & {}^kq_{23} \\ {}^kq_{31} & 0 & -{}^kq_{31} \end{pmatrix}$$

PROPOSITION 3.3. *For the k^{th} process, (i) A stationary probability is*

$$\left({}^k\pi_1,\quad {}^k\pi_2,\quad {}^k\pi_3\right) = \left(\frac{{}^kq_{31}({}^kq_{21} + {}^kq_{23})}{{}^k\Delta}, \frac{{}^kq_{12}\cdot{}^kq_{31}}{{}^k\Delta}, \frac{{}^kq_{12}\cdot{}^kq_{23}}{{}^k\Delta}\right)$$

where

$${}^k\Delta = {}^kq_{31}({}^kq_{21} + {}^kq_{23}) + {}^kq_{12}\cdot{}^kq_{31} + {}^kq_{12}\cdot{}^kq_{23}.$$

(ii) The Smoluchowskian mean recurrence time of state 3 is

$${}^k\xi(3) = \frac{{}^kq_{12} + {}^kq_{23}}{{}^kq_{12}\cdot{}^kq_{23}} + \frac{1}{{}^kq_{23}}.$$

Proof.

(i) As in the preceding case, the Markov process has only a recurrent class of states, then ${}^k\Pi$ is a matrix with identical rows. Solving ${}^k\Pi.{}^kQ = 0$, we easily obtain the result announced.

(ii) The result follows from (i) and equation (3) written as ${}^k\xi(3) = \frac{1}{{}^k\pi_3.{}^kq_3} - \frac{1}{{}^kq_3}$. □

Suppose now that the rule for coming back to state 1 is the same for all philosophers, so that ${}^kq_{21}$ and ${}^kq_{31}$ are independent of k (we denote them q_{21} and q_{31}).

Moreover, suppose that we would like to impose on the ${}^kq_{12}$'s a global level, and in a similar manner, on the ${}^kq_{23}$'s. Then the connection into network is expressed by

$$\sum_{k=1}^{N} {}^kq_{12} = A \text{ and } \sum_{k=1}^{N} {}^kq_{23} = B \tag{5}$$

In order to maximize the time during which the philosophers keep themselves busy to eat, it suffices to (globally) minimize their Smoluchowskian mean recurrence times of state 3. This criterion leads to the problem:

$$\text{Minimize } \sum_{k=1}^{N} {}^k\xi(3) = \sum_{k=1}^{N} \left(\frac{{}^kq_{12} + {}^kq_{23}}{{}^kq_{12} \cdot {}^kq_{23}} + \frac{1}{{}^kq_{23}} \right) \text{ under constraints (5).}$$

To simplify the notation, let us write x_k to mean ${}^kq_{12}$, and y_k to mean ${}^kq_{23}$. Then the problem to be solved is

Problem 2.

$$\text{Minimize } F(x_1, y_1, \ldots, x_k, y_k, \ldots, x_N, y_N) = \sum_{k=1}^{N} \frac{q_{21} + (x_k + y_k)}{x_k . y_k}$$

subject to

$$\mathcal{L}_1 = \sum_{k=1}^{N} x_k - A = 0 \quad \text{and} \quad \mathcal{L}_2 = \sum_{k=1}^{N} y_k - B = 0.$$

PROPOSITION 3.4. *There exists a set of optimal functioning rules (corresponding to problem 2)* $({}^1Q, \ldots, {}^NQ)$, *where* ${}^kq_{23} = \left(\frac{B}{A}\right).{}^kq_{12}$.

Proof. Let λ_1 and λ_2 be the Lagrange multipliers. We have to solve the system of equations

$$\frac{\partial}{\partial x_k} F + \frac{\partial}{\partial x_k}(\lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2) = -\frac{q_{12}}{(x_k)^2 y_k} - \frac{1}{x_k} + \lambda_1 = 0.$$

$$\frac{\partial}{\partial y_k} F + \frac{\partial}{\partial y_k}(\lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2) = -\frac{q_1 2}{x_k (y_k)^2} - \frac{1}{y_k} + \lambda_2 = 0, \quad k \in \{1, \ldots, N\}.$$

This implies that $\frac{y_k}{x_k} = \frac{\lambda_1}{\lambda_2} = \left(\sum_{k=1}^{N} y - k\right) / \left(\sum_{k=1}^{N} x_k\right) = \frac{B}{A}$ and consequently, $y_k = \frac{B}{A} x_k,\ k \in \{1, \ldots, N\}$.

Since F is a convex function of $x_1, y_l, \ldots, x_k, y_k, \ldots, x_N, y_N$ (as sum of convex functions), the solutions $\left((x_1, \frac{B}{A}x_1), \ldots, (x_k, \frac{B}{A}x_k), \ldots, (x_N, \frac{B}{A}x_N)\right)$ are minima for F. $\square$

4. A problem of access to shared resources

Mutual exclusion must be ensured when many processes require an access to shared resources. Such a protocol consists of a policy which allows at most one process to work with the resource. Solutions are proposed in (cf. [8, 12] and solutions by discrete time Markov model are given in [4,5]). In the present case, the state space has four elements: 'request' (state 1), 'refusal' (state 2), 'acceptance' (state 3), 'execution' (state 4).

In our present continuous-time model, the evolution of the k^{th} process is represented by the intensity matrix

$$^k Q = \begin{pmatrix} -(^k q_{12} + {}^k q_{13}) & ^k q_{12} & ^k q_{13} & 0 \\ ^k q_{21} & -{}^k q_{21} & 0 & 0 \\ 0 & 0 & -{}^k q_{34} & -{}^k q_{34} \\ ^k q_{41} & 0 & 0 & -{}^k q_{41} \end{pmatrix}$$

where $^k q_{12}$ and $^k q_{13}$ are supposed to be elements of $]0, \chi]$ with χ a positive constant.

PROPOSITION 4.1. *(i) A stationary probability for the k^{th} process is*

$$\begin{aligned} &\left(^k\pi_1, {}^k\pi_2, {}^k\pi_3, {}^k\pi_4, \right) \\ &\quad = \left(\frac{^k q_{21} \cdot {}^k q_{34} \cdot {}^k q_{41}}{^k\Delta}, \frac{^k q_{12} \cdot {}^k q_{34} \cdot {}^k q_{41}}{^k\Delta}, \frac{^k q_{13} \cdot {}^k q_{21} \cdot {}^k q_{41}}{^k\Delta}, \frac{^k q_{12} \cdot {}^k q_{21} \cdot {}^k q_{34}}{^k\Delta} \right) \end{aligned}$$

where

$$^k\Delta = {}^k q_{21} \cdot {}^k q_{34} \cdot {}^k q_{41} + {}^k q_{12} \cdot {}^k q_{34} \cdot {}^k q_{41} + {}^k q_{13} \cdot {}^k q_{21} \cdot {}^k q_{41} + {}^k q_{13} \cdot {}^k q_{21} \cdot {}^k q_{34}$$

(ii) The mean recurrence times for state 2 and state 3 are $^k\mu(2) = \frac{^k\Delta}{^k q_{12} \cdot {}^k q_{21} \cdot {}^k q_{34} \cdot {}^k q_{41}}$ *and* $^k\mu(3) = \frac{^k\Delta}{^k q_{13} \cdot {}^k q_{21} \cdot {}^k q_{34} \cdot {}^k q_{41}}$*, respectively.*

Proof. (i) Here the Markov process has only a recurrent class, $^k\Pi$ is a matrix with identical rows $(^k\pi_1, {}^k\pi_2, {}^k\pi_3, {}^k\pi_4)$. Solving the equation $^k\Pi . {}^k Q = 0$, we have the result announced. (ii) We use the formula $^k\pi(i) = \frac{1}{^k\pi_i . {}^k q_i}$. □

One criterion of good functioning is to simultaneously minimize the mean recurrence time of state 2 and maximize the mean recurrence time of state 3 (i.e., to come back as often as possible to state 'refusal', and as late as possible to state 'acceptance'). These two simultaneous objectives (minimize $^k\mu(2)$ and maximize $^k\mu(3)$ can be combined into a single one by minimizing the function $\sum_{k=1}^{N} \frac{^k\mu(2)}{^k\mu(3)} = \sum_{k=1}^{N} \frac{^k q_{13}}{^k q_{12}}$ Such a function has been encountered in other contexts (cf.[4]). In the other hand, in view of the importance of state 2 and state 3, by imposing a level on the sum of all $^k q_{12}$ and the sum of all $^k q_{13}$ the connection into network can be expressed in either of the following ways:

$$\sum_{k=1}^{N} {}^{k}q_{12} \geqslant A \text{ and } \sum_{k=1}^{N^k} {}^{k}q_{13} \geqslant B \tag{6}$$

or

$$\sum_{k=1}^{N} {}^{k}q_{12} = A \text{ and } \sum_{k=1}^{N^k} {}^{k}q_{13} = B \tag{7}$$

The set of inequalities (6) (or (7)) are the constraints of the problem. To simplify the notation, let us write x_k instead of ${}^{k}q_{13}$, and y_k instead of ${}^{k}q_{12}$. Recall that x_k and y_k are supposed to lie in the interval $]0, \chi]$, where χ is a positive constant. Then A and B must satisfy $A \leqslant N\chi$ and $B \leqslant N\chi$. Furthermore, since the objective function (to be minimized) is $\sum_{k=1}^{N} \frac{x_k}{y_k}$ we will also assume that $x_k \geqslant \eta$, where $0 < \eta < \frac{A}{N}$.

a) With constraints (6) the problem we have to solve is

Problem 3.

$$\text{Minimize } \sum_{k=1}^{N} \frac{x_k}{y_k}$$

subject to

$$\begin{cases} \sum\limits_{k=1}^{N} x_k = A \quad \sum\limits_{k=1}^{N} y_k \geqslant B \\ \eta \leqslant x_k \leqslant \chi, \;\; 0 < y_k \leqslant \chi, k \in \{1, \ldots, N\} \end{cases}$$

Although the objective function in this problem is nonconvex, an optimal solution can be found in closed form. To see this observe that if a feasible solution $((x_l, y_1), \ldots, (x_k, y_k), \ldots (x_N, y_N))$ is such that $y_k < \chi$ for at least one $k \in \{1, \ldots, N\}$, then $\sum_{k=1}^{N} \frac{x_k}{y_k} \geqslant \sum_{k=1}^{N} \frac{x_k}{\chi}$ This implies that an optimal solution must satisfy $y_k = \chi$ for every $k \in \{1, \ldots, N\}$. Then the problem becomes:

$$\text{Minimize } \sum_{k=1}^{N} x_k \text{ subject to}$$

$$\begin{cases} \sum\limits_{k=1}^{N} x_k \geqslant A \\ \eta \leqslant x_k \leqslant \chi, \;\; k \in \{1, \ldots, N\}. \end{cases}$$

It is readily seen that the first inequality constraint in this linear program can be replaced by the equality:

$$\sum_{k=1}^{N} x_k = A$$

Indeed, if a feasible solution x to this linear program is such that $\sum_{k=1}^{N} x_k > A$ then, since $N_\eta < A$, there exists $h \in \{1, \dots, N\}$ such that $\eta < x_h$ and a better feasible solution can be obtained by slightly diminishing x_h while keeping every other x_k, $k \neq h$, unchanged. Thus, the above linear program can be written as

$$\text{Minimize } \sum_{k=1}^{N} x_k \text{ subject to}$$
$$\begin{cases} \sum_{k=1}^{N} x_k = A \\ \eta \leqslant x_k \leqslant \chi, \quad k \in \{1, \dots, N\}. \end{cases} \tag{8}$$

Of course, this linear program can be solved numerically by the simplex method. However, to find a closed form of the optimal solution, we observe that the extreme points of the constraint polytope of this linear program can be computed by simple formulas. In fact, for each integer $m \in \{1, \dots, N\}$, let $x^{(m)}$ be a vector with components

$$\begin{cases} x_k = \eta, & \text{if } 1 \leqslant k < m; \\ x_k = \chi, & \text{if } m \leqslant k < N; \\ x_N = A - (m-1)\eta - (N-m)\chi. \end{cases} \tag{9}$$

We say that a vector $x' \in R^N$ is obtained from $x \in R^N$ by a permutation of components if there exists a 1-1 mapping ϖ of the set $\{1, \dots, N\}$ onto itself such that $x'_i = x_{\varpi(i)}, \quad i \in \{1, \dots, N\}$.

LEMMA 4.1. *The set of extreme points of the polytope*

$$\sum_{k=1}^{N} x_k = A, \quad \eta \leqslant x_k \leqslant \chi, k \in \{1, \dots, N\} \tag{10}$$

consists of all points obtained from $x^{(m)}$ by a permutation of the components, where

$$\frac{N_\chi - A}{\chi - \eta} \leqslant m \leqslant 1 + \frac{N_\chi - A}{\chi - \eta} \tag{11}$$

Proof. An extreme point of the polytope defined by (10) corresponds to a basic solution of this linear inequalities system. Since the polytope lies in the linear manifold $\sum_{k=1}^{N} x_k = A$, each extreme point is determined by a set of $N-1$ inequalities among

$$\eta \leqslant x_k, x_k \leqslant \chi, k \in \{1, \dots, N\}$$

such that these $N-1$ inequalities are satisfied by the given extreme point as equalities. In other words, each extreme point is the unique solution of a linear system of the form (9), modulo a permutation of the index set $\{1, \dots, N\}$. Here, one must have $\eta \leqslant X_N^{(m)} \leqslant \chi$, hence (11). Furthermore, it can easily be verified that the latter condition implies that $m \in \{1, \dots, N\}$. □

Note that condition (11) completely determines m, namely:

$$m = 1 + \left\lfloor \frac{N\chi - A}{\chi - \eta} \right\rfloor, \tag{12}$$

where $\lfloor . \rfloor$ denotes the integer part. Now, it is obvious that the value of the objective function in (8) is unchanged by a permutation of the index set. Therefore this value is the same at every extreme point of the constraint polytope, and so an optimal solution of (8) is $x^{(m)}$ given by (9). We have thus proved the following

PROPOSITION 4.2. *The optimal functioning rule corresponding to problem 3 is defined by*

$$\begin{cases} {}^{k}q_{12} = \eta & \textit{if } 1 \leqslant k < m; \\ {}^{k}q_{12} = \chi & \textit{if } m \leqslant k < N; \\ {}^{N}q_{12} = A - (m-1)\eta - (N-m)\chi. \end{cases}$$
$${}^{k}q_{13} = \chi, \quad \textit{for every } k \in \{1, \dots, N\}$$

where m is the integer (12).

b) With constraints (7) the problem to be solved is:

Problem 4.

$$\text{Minimize } \sum_{k=1}^{N} \frac{x_k}{y_k} \text{ subject to}$$

$$\begin{cases} \sum_{k=1}^{N} x_k = A \\ \sum_{k=1}^{N} y_k = B \\ \eta \leqslant x_k \leqslant \chi, 0 \leqslant y_k \leqslant \chi, \quad k \in \{1, \dots, N\}. \end{cases}$$

(Recall that $\max(A, B) \leqslant N_\chi$*, and $N_\eta < A$).*

Because of the two equality constraints, this nonconvex problem cannot be reduced to a linear program as Problem 3. However, we shall show that, again due to the special structure of this problem, it is possible to find an optimal solution in almost closed form.

For fixed $x = (x_1, \dots, x_N)$ each function $y \mapsto \frac{x_k}{y_k}$ is convex on the interval $0 < y_k \leqslant \chi$, hence the function $\sum_{k=1}^{N} \frac{x_k}{y_k}$ is convex on the hyperrectangle $\{y | 0 < y_k \leqslant \chi, (k = 1, \dots, N)\}$. Let $\phi(x)$ be the optimal value of the convex program

$$\text{Minimize } \sum_{k=1}^{N} \frac{x_k}{y_k} \text{ s. t. } \quad \sum_{k=1}^{N} y_k = B, 0 < y_k \leqslant \chi, k \in \{1, \dots, N\}.$$

Since $\phi(x)$ is the pointwise minimum of a family of linear functions $x \mapsto \frac{x_k}{y_k}$ is a concave function (see e.g. [15]) and so Problem 4 is equivalent to the following linearly constrained concave minimization problem

$$\text{Minimize } \phi(x) \text{ s. t. } \quad \sum_{k=1}^{N} x_k = A, \eta < x_k \leqslant \chi, k \in \{1, \dots, N\}.$$

As is well known, an optimal solution of this concave minimization problem is achieved at an extreme point of the constraint polytope, i.e. the polytope (10). Let $\bar{x} = x^{(m)}$ for m given by (12) and let $\bar{y} = (\bar{y}, \dots, \bar{y}_N)$ be an optimal solution of the convex program

$$\text{Minimize } \sum_{k=1}^{N} \frac{\bar{x}_k}{y_k} \text{ s. t. } \quad \sum_{k=1}^{N} y_k = B, 0 < y_k \leqslant \chi, k \in \{1, \dots, N\}. \tag{13}$$

We can now state

PROPOSITION 4.3. *An optimal functioning rule corresponding to Problem 4 is given by*

$$\begin{cases} {}^k q_{12} = \eta & \textit{if } 1 \leqslant k < m; \\ {}^k q_{12} = \chi & \textit{if } m \leqslant k < N; \\ {}^N q_{12} = A - (m-1)\eta - (N-m)\chi. \end{cases} \tag{14}$$
$${}^k q_{13} = \chi, \quad \textit{for every } k \in \{1, \dots, N\}$$

where m is the integer (12).

Proof. Clearly, if a vector x' is obtained from x by a permutation of components, then $\phi(x') = \phi(x)$. Hence, $\phi(x)$ takes on the same value at every extreme point of the polytope (10). This implies that an optimal solution of Problem 4 is $(\bar{x}, \bar{y})$ where $\bar{x} = x^{(m)}$ with m defined by (12), while $\bar{y}$ is an optimal solution of the Program (13).

The convex Program (13) whose optimal solution yields $\bar{y}$ is easy to solve. We can also show that if the parameters B, χ and η satisfy suitable conditions then a closed form of $\bar{y}$ can be found. To this end, consider the convex program obtained from (13) by omitting the constraints $0 < y_k \leqslant \chi, k \in \{1, \dots, N\}$, i.e.

$$\text{Minimize } \sum_{k=1}^{N} \frac{\bar{x}_k}{y_k} \text{ s. t. } \quad \sum_{k=1}^{N} y_k = B. \tag{15}$$

Denote $f(\bar{x}, y) = \sum_{k=1}^{N} \frac{\bar{x}_k}{y_k}$. The Lagrange multipliers method gives

$$\frac{\partial}{\partial y_k} f + \lambda \frac{\partial}{\partial y_k}\left(\sum_{j=1}^{N} y_j - B\right) = -\frac{\bar{x}_k}{(y_k)^2} + \lambda = 0, k \in \{1, \dots, N\}$$

so that $y_k = \sqrt{\frac{\bar{x}_k}{\lambda}}$, $k \in \{1, \dots, N\}$ which implies that $\sqrt{\lambda} = \frac{1}{B}\sum_{k=1}^{N} \sqrt{\bar{x}_k}$. Consequently, an optimal solution $\hat{y}$ of Problem (15) is given by

$$\hat{y}_k = \frac{B\sqrt{\bar{x}_k}}{\sum_{j=1}^{N} \sqrt{\bar{x}_j}}, \quad k \in \{1, \dots, N\}$$

and the optimal value is

$$\sum_{k=1}^{N} \frac{\bar{x}_k}{\hat{y}_k} = \frac{1}{B}\left(\sum_{k=1}^{N} \sqrt{\bar{x}_k}\right)^2.$$

For $\hat{y}$ to be an optimal solution of (13), it suffices that it is feasible to (13), i.e.

$$\frac{B\sqrt{\bar{x}_k}}{\sum_{j=1}^{N} \sqrt{\bar{x}_j}} \leqslant \chi, \text{ for every } k \in \{1, \dots, N\} \tag{16}$$

Since $\bar{x} \in \{\eta, \chi\}$ for $k\{= 1, \dots, N-1\}$ and $\eta \leqslant \bar{x}_N \leqslant \chi$, the above condition holds when

$$\frac{B\sqrt{\chi}}{N\sqrt{\eta}} \leqslant \chi, \quad \text{i.e. } B \leqslant N\sqrt{\eta\chi}.$$

Thus if $B \leqslant N\sqrt{\eta\chi}$, then $\hat{y}$ is an optimal solution of Problem (13). □

A more refined result is the following:

PROPOSITION 4.4. *Assume*

$$B < \chi\left[1 + \sqrt{(N-1)\left(\frac{A}{\chi} - 1\right)}\right]. \tag{17}$$

Then $\hat{y}$ is an optimal solution of (13), i.e. in (14):

$\bar{y}_k = \hat{y}_k$, *for every* $k \in \{1, \dots, N\}$.

Proof. Denote

$$\begin{aligned}\nu(\eta) &= \sum_{k=1}^{N} \sqrt{\bar{x}_k} \\ &= (m-1)\sqrt{\eta} + (N-m)\sqrt{\chi} + \sqrt{A - (m-1)\eta - (N-m)\chi}.\end{aligned}$$

We have to prove that under condition (17):

$$\frac{B\sqrt{\chi}}{N\sqrt{\eta}} \leqslant \chi. \tag{18}$$

For $1 < m \leqslant N$, let η_m be the value of η for which

$$\frac{N_\chi - A}{\chi - \eta_m} = m - 1. \tag{19}$$

Then for $\eta_m < \eta \leqslant \eta_{m+1}$, we have

$$\left\lfloor \frac{N_\chi - A}{\chi - \eta} \right\rfloor = m - 1.$$

From (19):

$$\eta_m = \chi - \frac{N_\chi - A}{m - 1} \tag{20}$$

which shows that η_m increases as m increases. On the other hand, for $\eta_m \leqslant \eta < \eta_{m+1}$, the derivative of the function $\nu(\eta)$ is

$$\nu'(\eta) = (m-1) - \frac{1}{2\sqrt{\eta}} - \frac{m-1}{2\sqrt{A - (m-1)\eta - (N-m)\chi}} > 0.$$

because $\eta < \eta_{m+1}$ implies by (20): $A - (N-m)\chi - m\eta > 0$. So the function $\nu(\eta)$ is increasing in the interval $[\eta_m, \eta_{m+i})$. Furthermore,

$$\begin{aligned}
\nu(\eta_{m+1} - 0) =&(m-1)\sqrt{\chi - \frac{N_\chi - A}{m}} + (N-m)\sqrt{\chi} \\
&+ \sqrt{A - (m-1)(\chi - \frac{N_\chi - A}{m}) - (N-m)\chi} \\
=&(m-1)\sqrt{\frac{A - (N-m)\chi}{m}} + (N-m)\sqrt{\chi}\sqrt{\frac{A - (N-m)\chi}{m}} \\
=&m\sqrt{\frac{A - (N-m)\chi}{m}} + (N-m)\sqrt{\chi} \\
=&\nu(\eta_{m+1}).
\end{aligned}$$

Hence $\nu(\eta)$ is increasing in each interval $[\eta_m, \eta_{m+1}]$, $m \in \{2, \ldots, N-1\}$, i.e. in the whole interval $[\eta_2, \eta_N]$. For $0 < \eta < \eta_2$, we have $m = 1$, so

$$\nu(\eta) = (N-1)\sqrt{\chi} + \sqrt{A - (N-1)\chi}.$$

while

$$\nu(\eta_2) = \sqrt{\eta_2} + (N-2)\sqrt{\chi} + \sqrt{A - (\eta_2) - (N-2)\chi}.$$

This yields by substituting $\eta_2 = A - (N-1)\chi$,

$$\nu(\eta_2) = \sqrt{A - (N-1)\chi} + (N-1)\sqrt{\chi}.$$

i.e. $\nu(\eta) = \nu(\eta_2)$ for $0 < \eta \leqslant \eta_2$.

We have thus proved that the function $\nu(\eta)$ is increasing in the whole interval $0 < \eta \leqslant \eta_N$. (Note that since $N_\eta < A$, we always have $\frac{N\chi}{\chi-\eta}N$, so $\eta \leqslant \eta_N$). Since $\eta_N = \frac{A-\chi}{N-1}$ it then follows from the assumption (17) that

$$\frac{B\sqrt{\chi}}{\nu(\eta_N)} = \frac{B\sqrt{\chi}}{\sqrt{(N-1)(A-\chi)} + \sqrt{\chi}} \leqslant \chi.$$

Therefore, for all $\eta \in]0, \eta_N]$

$$\frac{B\sqrt{\chi}}{\nu(\eta)} \leqslant \frac{B\sqrt{\chi}}{\nu(\eta_N)} \leqslant \chi,$$

as desired. □

Finally, it is worth noticing that the ‘bounding’ method earlier introduced in [3] for the problem of routing distributed algorithms would give essentially the same result, though with a less satisfactory theoretical foundation concerning the global optimality of the solution.

Acknowledgments

The author is grateful to a referee for useful remarks and suggestions which have helped to improve the paper.

References

Brand, D. and Zafiropulo, P. (1983), On communicating finite-state machines, *Journal of the ACM* 30 (2): 323–342.

Bui, A. Bui, M. and Lavault, C. (1999), On the hierarchy of functioning rules in distributed computing, RAIRO Op. Res. 33(1): 15–27.

Bui, A. (1994), A Bounding problem concerning distributed routing Algorithms, *Revue R. Math. Pures et Appl*. 34: 1–15.

Bui, A. (1997), Optimization as an Instrument for Evaluating distributed network performance, Communication to the Workshop *from local to global Optimization* Linköping, August 1997.

Bui, M. (1992), Tuning distributed control algorithms for optimal functioning, *Journal of Global Optimization*, 2: 177–199.

Casavant, T. L. and Kuhl, J. G. (1990), Communicating Finite Automata Approach to Modeling Distributed Computation and its Application to Distributed Decision Making, *IEEE Trans. on Computers* 39 (5): 628–639.

Chandy, K. M. and Misra, J. (1994), The drinking philosophers problem, *ACM Trans. on Prog. Lang. and Syst.* 6 (4): 632–646.

Chandy, K. M. and Misra, J. (1988), *Parallel Program Design: A Foundation*, Addison-Wesley.

Dijkstra, E. W. (1971), Hierachical ordering of Sequential Processes, *Acta Informatica*, 1: 115–138.

Iosifescu, M. (1980), *Finite Markov Processes and their Applications*, New York: Wiley.

Kemeny, J. G. and Snell, J. L. (1983), *Finite Markov Chains*, Berlin/New York: Springer-Verlag.

Lamport, L. (1986), The mutual Exclusion Problem, Part I and Part II. *Journal of ACM*, 33 (2): 327–348.

Lynch, N. A. (1996), *Distributed Algorithms*, Morgan Kaufmann Pub.

Tel, G. (1994), *Introduction to distributed algorithms*, Cambridge University.

Tuy, H. (1998), *Convex Analysis and Global Optimization*, Dordrecht/Boston/London: Kluwer Academic Publishers.