Multiserver real-time system with shortage of maintenance teams

Edward Ianovsky 1*, Joseph Kreimer 2

1 Rothman School of Business Administration, University of Toronto, Canada
2 Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, P.O.Box 653, Beer-Sheva 84105, Israel
E-Mail: kremer@bgumail.bgu.ac.il

accepted March 15, 2009

Summary
We consider a real-time multiserver system with identical servers (e.g. unmanned air vehicles, machine controllers, etc.) that provide service for requests of real-time jobs arriving via several different channels (e.g. surveillance regions, assembly lines, etc.) working under maximum load regime. There is exactly one job in each channel at any instant. Each channel has its own specifications, and therefore different kinds of equipment and inventory are needed to serve different channels. There is limited number of identical maintenance teams (less than the total number of servers in the system). Our goal is to compute analytically steady-state probabilities of this system, its availability, loss penalty function and other performance characteristics, and to discuss optimality conditions.

Key words: Availability, optimization, performance, real-time system, steady-state.

1. INTRODUCTION

Real-time systems (RTS) are defined as those for which correctness depends not only on the logical properties of the computed results, but also on their temporal properties. In RTS a calculation that uses temporally invalid data, may be useless, and sometimes harmful – even if such a calculation is functionally correct. Examples include industrial automation, traffic control, aerospace, robotics, intelligence and defense systems, telecommunication and distributed process control, just to name a few.

* corresponding author
Several authors ([23], [21] and [25]) have proposed a number of priority algorithms for scheduling real-time tasks on a single processor. The problem of determining a minimum number of processors in multiprocessor computer system executing real-time tasks was studied in [2]. Different scientific communities are treating RTS problems. During last two decades several metaheuristic methods, such as Simulated Annealing [12], Tabu Search [5] and Greedy Randomized Adaptive Search Procedure [3] were developed. Good surveys on applications of Artificial Intelligence to real-time decision problems (RTDP) are given in [4], [26], [22] and [24].


We will concentrate our attention on RTS with a zero deadline for the beginning of job processing. In these RTS, jobs are processed immediately upon arrival, if there are available servers. That part of the job which is not processed immediately is lost forever, since queueing of jobs is not allowed. The particular interest in such RTS was aroused by military intelligence problems involving unmanned air vehicles (UAV), which demonstrated very high efficiency during the Gulf War. It was proved ([18], [20]) that the non-mix policy of never relieving an operative server maximizes the availability of a multiserver single-channel RTS involving preventive maintenance and working in general regime with any arrival pattern under consideration and constant service and maintenance times. This policy appears to be optimal for any finite time interval, and not only for infinite horizon. In [19] and [15] multiserver (identical servers) and multichannel (identical channels) RTS (with unrestricted and restricted number of maintenance facilities respectively), working under maximum load regime were treated as finite source queues [6]. In [14] and [16] the two-dimensional birth-and-death processes were applied in analysis of a multiserver RTS (with ample and restricted number of maintenance teams respectively) with two different channels operating under a maximum load regime, when both service and maintenance times are exponentially distributed. In [17] the results of [12] for different channels operating under a maximum load regime were extended. In [7] and [10] were obtained optimal assignment probabilities to maximize availability of RTS (with ample and restricted number of maintenance teams respectively) with large number of servers and two channels. In [13] a multiserver and multichannel RTS with identical servers and channels and ample maintenance facilities, working in general regime with exponentially distributed service, maintenance, jobs inter-arrival and duration times were treated as Markov chains in order to obtain various performance characteristics. In [8] and [9] we obtained moments of RTS with ample and limited maintenance facilities respectively. In [1] RTS with preemptive priorities policy were considered.

This work extends the model developed in [17] and the results of [7] as follows. The RTS under consideration assumes immediate on-line data acquisition. We compute analytically steady-state probabilities of this system, its availability, loss penalty function and other performance characteristics, and provide limiting values of performance measurements for RTS with large number of servers and shortage of maintenance teams. Then we obtain optimal assignment probabilities which maximize system availability and minimize its loss penalty function. Finally we discuss the rate of convergence.

The paper is organized as follows: In Section 2, the description of the model is presented. Section 3 provides optimal values of assignment probabilities and rate of convergence. In Section 4 some numerical results are presented. Finally, in Appendix we provide some auxiliary Lemmas and Corollaries on which the proof of our main results is based.
2. DESCRIPTION OF THE MODEL

We consider a multiserver RTS consisting of \( N \) identical servers that provide service for requests of real-time jobs arriving via \( r \) different channels required to be under nonstop surveillance. There is exactly one job in \( i \)-th channel at any instant (there are no additional job arrivals to the busy channel), and therefore one server at most is used to serve the \( i \)-th channel (with others being on stand-by or providing the service to another channel or in maintenance or waiting for maintenance) at any given time.

Each channel has its own specifications and conditions, etc., and therefore different kinds of equipment and inventory are needed to serve different channels.

A server is operative for a period of time \( S_i \) before requiring \( R \) hours of maintenance. \( S_i \) and \( R \) are independent exponentially distributed random values with parameters \( \mu_i \) \( (i=1,...,r) \) and \( \lambda \) respectively.

It is assumed that there are \( K \) \( (K < N) \) maintenance teams available to repair (with repair times \( R_i \) being i.i.d.r.v.) the servers. Thus, a shortage of maintenance facilities is possible. In that case the server waits for maintenance. This server is assigned to the \( i \)-th region with probability \( p_i \) \( (i=1,...,r) \). It receives the appropriate kind of maintenance (equipment, programming, etc.), and therefore cannot be sent to another channel. Assignment probabilities \( p_i \) may depend upon inventory conditions. They also can be used as control parameters. The duration \( R_i \) of repair is exponentially distributed with parameter \( \lambda_i \), and does not depend on the channel. After maintenance, the server will either be on stand-by or serving the region it was assigned to.

The system works under a maximum load (worst case) of nonstop data arrival to each one of \( r \) channels. This kind of operation is typical in high performance data acquisition and control systems, such as self-guided missiles, space stations, satellites, etc.

If, during some period of time of length \( T \), there is no available server to serve one of the jobs, we will say that the part of the job of length \( T \) is lost.

Queues of jobs cannot exist in RTS, nevertheless they can be fitted into a framework of finite source queues, while using a dual approach of changing the roles between jobs and servers.

Denote: \( \lambda_i = \lambda p_i \) and \( \rho_i = \lambda_i / \mu_i \), \( (n_1, n_2, ..., n_i, ..., n_r) \) the state of the system, where \( n_i \) \( (i=1,...,r) \) is a number of fixed servers assigned to the \( i \)-th channel (obviously \( \sum_{i=1}^{r} n_i \leq N \)), and \( p_{n_1,n_2,...,n_r} \) the corresponding steady state probability. There are \( \left( \begin{array}{c} N + r \\ r \end{array} \right) \) states in total.

The above RTS can be presented as a closed queuing network consisting of \( N \) customers (\( N \) servers of the RTS), \( r+1 \) stations (\( r \) channels and maintenance station of the RTS) with one server at \( i \)-th station (one job in the \( i \)-th channel of the RTS) \( (i=1,...,r) \) and \( K \) servers at \( r+1 \)-th station (\( K \) maintenance teams of the RTS). The network has transition probabilities \( p_i \) (assignment probabilities of the RTS) from \( r+1 \)-th station to \( i \)-th one \( (i=1,...,r) \), and transition probabilities from \( i \)-th station to \( r+1 \)-th are equal to 1. Other transition probabilities are equal to 0. Service times at the network stations (operating and maintenance times of the RTS) are exponentially distributed. The customers of the network cannot leave it and cannot come to it from outside. Thus, the network is a closed Jackson network by definition.
From the description of the RTS as a closed Jackson network we obtain that the steady-state probabilities $p_{n_1, n_2, \ldots, n_r}$ are given by the following

**Theorem 1:** A real-time system with $N$ servers, $K$ ($K < N$) maintenance crews, $r$ ($r \geq 2$) different channels operating under a maximum load regime with one job in each one of the channels at any instant, and exponentially distributed operating and maintenance times (with parameters $\mu_i$ (for the $i$-th channel ($i = 1, r$)) and $\lambda$ respectively) has following values of the steady-state probabilities

$$p_{n_1, n_2, \ldots, n_r} = \frac{\min\left(\sum_{i=1}^{r} n_i - N - K\right)}{K!} \prod_{j=1}^{r} \rho_j^{n_j} p_{0, \ldots, 0},$$

$$p_{0, \ldots, 0} = \left[ \sum_{n_1, n_2, \ldots, n_r = 0}^{K} \frac{\min\left(\sum_{i=1}^{r} n_i - N - K\right)}{K!} \prod_{j=1}^{r} \rho_j^{n_j} \right]^{-1},$$

where $(n_1, n_2, \ldots, n_r)$ is a state of the RTS, $n_i$ is number of fixed servers of the RTS in the $i$-th channel and $\rho_i = \frac{\lambda p_i}{\mu_i}$ ($i = 1, r$).

### 3. OPTIMIZATION

In this Section we will obtain the optimal values of assignment probabilities for large values of $N$, and the rate of convergence, when $N \to \infty$.

We will use the following notations:

$$S_N^r(\rho_1, \ldots, \rho_r) = \sum_{n_1, n_2, \ldots, n_r = 0}^{K} \frac{\min\left(\sum_{i=1}^{r} n_i - N - K\right)}{K!} \prod_{j=1}^{r} \rho_j^{n_j},$$

$$P_N^k(\rho_1, \ldots, \rho_r)$$ the probability that channel $k$ ($k = 1, r$) is not served,

$$Av_N(\rho_1, \ldots, \rho_r) = \sum_{k=1}^{r} \left(1 - P_N^k(\rho_1, \ldots, \rho_r)\right) / r$$ system’s availability and

$$TC_N(\rho_1, \ldots, \rho_r) = \sum_{k=1}^{r} C_k P_N^k(\rho_1, \ldots, \rho_r)$$ an average loss penalty cost, where $C_k$ ($k = 1, r$) the cost of the time unit during which the $k$-th channel is not served.
Using these notations we obtain:

\[
\frac{S^{-1}_N(\rho_1, \ldots, \rho_{k-1}, \rho_{k+1}, \ldots, \rho_r)}{S^-_N(\rho_1, \ldots, \rho_r)} = \sum_{n_1, \ldots, n_{r-1}, n_{r+1}, \ldots, n_r \geq 0} \frac{\min \left\{ \frac{\sum n_i - n_i K}{K}, \prod \rho_j^{ \frac{n_j}{r} } \right\}}{K!} \prod_{j=1}^r \rho_j^{ \frac{n_j}{r} } = \\
\sum_{n_1, \ldots, n_{r-1}, n_{r+1}, \ldots, n_r \geq 0} \frac{\min \left\{ \frac{\sum n_i - n_i K}{K}, \prod \rho_j^{ \frac{n_j}{r} } \right\}}{K!} \prod_{j=1}^r \rho_j^{ \frac{n_j}{r} } 
\]

\[
TC_N(\rho_1, \ldots, \rho_r) = \sum_{k=1}^r C_k \frac{S^{-1}_N(\rho_1, \ldots, \rho_{k-1}, \rho_{k+1}, \ldots, \rho_r)}{S^-_N(\rho_1, \ldots, \rho_r)} 
\]

and for \( C_1 = C_2 = \ldots = C_r = 1 \) we have

\[
Av_N(\rho_1, \ldots, \rho_r) = 1 - TC_N(\rho_1, \ldots, \rho_r)/r.
\]

Let \( 0 \leq \rho_i \leq 1 \) (\( i = 1, \ldots, r \)), \( \sum_{i=1}^r \rho_i = 1 \), \( \lambda > 0 \), \( \mu_i > 0 \) and \( \rho_i = \frac{\lambda \rho_i}{\mu_i} \) (\( i = 1, \ldots, r \)), \( \rho = \max_{i=1}^r \rho_i \). Then \( \rho > 0 \) and we can formulate the following theorems:

**Theorem 2:** A real-time system with \( N \) servers, \( K \) (\( K < N \)) maintenance crews, \( r \) (\( r \geq 2 \)) different channels operating under a maximum load regime, and exponentially distributed operating and maintenance times (with parameters \( \mu_i \) (\( i = 1, \ldots, r \)) and \( \lambda \) respectively) has following limiting values (as \( N \to \infty \)) of probability that the \( i \)-th channel is not served

\[
P^{(i)}(\rho_1, \ldots, \rho_r) = \lim_{N \to \infty} P_N^{(i)}(\rho_1, \ldots, \rho_r) = 1 - \rho_i \min \left\{ K, \frac{1}{\rho} \right\}, \quad i = 1, \ldots, r.
\]

The proof follows directly from the sequence of Lemmas and Corollaries given in Appendix.

Now we immediately obtain from the Theorem 1, the limiting values of the loss penalty cost function and system’s availability

**Corollary 2.1:** \( TC(\rho_1, \ldots, \rho_r) = \lim_{N \to \infty} TC_N(\rho_1, \ldots, \rho_r) = \sum_{i=1}^r C_i \left( 1 - \rho_i \min \left\{ K, \frac{1}{\rho} \right\} \right) \),

\[
Av(\rho_1, \ldots, \rho_r) = \lim_{N \to \infty} Av_N(\rho_1, \ldots, \rho_r) = \frac{\min \left\{ K, \frac{1}{\rho} \right\}}{r} \sum_{i=1}^r \rho_i.
\]
Corollaries 2.2 and 2.3 provide the rate of convergence.

**Corollary 2.2:** The rate of convergence $P_N^{(i)}(\rho_1, \ldots, \rho_r) \xrightarrow{N \to \infty} P^{(i)}(\rho_1, \ldots, \rho_r)$, $i = 1, r$, is
\[
\begin{align*}
&O(\alpha_1^N), \text{ if } K \rho > 1 & &\text{if } m = 1 \\
&O(1/N), \text{ if } K \rho > 1 & &\text{if } m \in \{2, 3, \ldots, r\}, \text{ where } \max\left\{\frac{1}{K \rho}, \max_{\rho_i < \rho} \frac{\rho_i}{\rho}\right\} < \alpha_i < 1, \quad K \rho < \alpha_2 < 1, \text{ and } m \text{ is number of } \rho_i = \rho \quad (i = 1, r).
\end{align*}
\]

**Corollary 2.3:** The rate of convergence of $TC_N(\rho_1, \ldots, \rho_r) \xrightarrow{N \to \infty} TC(\rho_1, \ldots, \rho_r)$ and $Av_N(\rho_1, \ldots, \rho_r) \xrightarrow{N \to \infty} Av(\rho_1, \ldots, \rho_r)$ is
\[
\begin{align*}
&O(\alpha_1^N), \text{ if } K \rho > 1 & &\text{if } m = 1 \\
&O(1/N), \text{ if } K \rho > 1 & &\text{if } m \in \{2, 3, \ldots, r\}, \text{ where } \max\left\{\frac{1}{K \rho}, \max_{\rho_i < \rho} \frac{\rho_i}{\rho}\right\} < \alpha_i < 1, \quad K \rho < \alpha_2 < 1, \text{ and } m \text{ is number of } \rho_i = \rho \quad (i = 1, r).
\end{align*}
\]

Finally, the next Theorem provides the optimal values of assignment probabilities.

**Theorem 3:** In real-time system with $N$ servers, $K$ ($K < N$) maintenance crews, $r$ ($r \geq 2$) different channels operating under a maximum load regime, and exponentially distributed operating and maintenance times (with parameters $\mu_i$ ($i = 1, r$) and $\lambda$ respectively) the optimal assignment probabilities $p_i^*$ ($i = 1, r$), which minimize system loss penalty function when $N \to \infty$ are determined as follows:

(a) If $\lambda K \geq \sum_{i=1}^r \mu_i$, then $p_j^* = \mu_j / \sum_{i=1}^r \mu_i$ for $j = 1, r$.

(b) If $\lambda K < \sum_{i=1}^r \mu_i$, then
\[
p_j^* = \begin{cases} 
\mu_j / \lambda K, & \text{if } \sum_{i=1}^r \mu_i / \lambda K \leq 1, \\
1 - p_1^* - \ldots - p_{j-1}^*, & \text{if } \sum_{i=1}^r \mu_i / \lambda K \leq 1 & \text{and } \sum_{i=1}^r \mu_i / \lambda K > 1, \\
0, & \text{if } \sum_{i=1}^r \mu_i / \lambda K > 1,
\end{cases}
\]

where the channels are numbered in the following manner: $\frac{C_1}{\mu_1} \geq \frac{C_2}{\mu_2} \geq \ldots \geq \frac{C_r}{\mu_r}$.

Corresponding optimal value of loss penalty function is

(a) $TC(\rho_1^*, \ldots, \rho_r^*) = 0$, if $\lambda K \geq \sum_{i=1}^r \mu_i$.
(b) \[ TC(p_1^*, \ldots, p_r^*) = \sum_{i=1}^{r} C_i (1 - Kp_i^*), \] if \( \lambda K < \sum_{i=1}^{r} \mu_i \), where \( p_i^* = \frac{\lambda p_i^*}{\mu_i}, i = 1, r. \)

Similar results for limiting values (as \( N \to \infty \)) of system availability \( Av(p_1, \ldots, p_r) \) can be obtained from (3).

**Proof:**

In this proof we shall use the expression for limiting values of system loss penalty function given by Corollary 2.1.

(a) If there is a set \((p_1, \ldots, p_r)\) such that \( K \rho < 1 \) then for any \( i = 1, r \) we get \( K \rho < 1 \) or equivalently \( \lambda K p_i < \mu_i \), hence \( \lambda K < \sum_{i=1}^{r} \mu_i \). Thus, if \( \lambda K \geq \sum_{i=1}^{r} \mu_i \) then \( K \rho \geq 1 \) for any set \((p_1, \ldots, p_r)\). Therefore, if \( \lambda K \geq \sum_{i=1}^{r} \mu_i \) then \( TC(p_1, \ldots, p_r) = \sum_{n=1}^{r} C_n (1 - \rho_n / \rho) \) for any set \((p_1, \ldots, p_r)\) and the unique set of probabilities, which minimize \( TC(p_1, \ldots, p_r) \), is \( p_n = \mu_n \sum_{i=1}^{r} \mu_i, n = 1, r. \)

(b) If \( \lambda K < \sum_{i=1}^{r} \mu_i \) then the point \( p_n = \mu_n \sum_{i=1}^{r} \mu_i, n = 1, r \) is in the set of points satisfying \( K \rho < 1 \). Hence, if \( K \rho > 1 \) then decreasing of \( \rho \) results in equation \( K \rho = 1 \) before than \( \rho_1 = \rho_2 = \ldots = \rho_r = \rho \) happens. But this decreasing of \( \rho \) results in decreasing of \( TC(p_1, \ldots, p_r) = \sum_{n=1}^{r} C_n (1 - \rho_n / \rho) \). Therefore, for any set \((p_1^{(1)}, \ldots, p_r^{(1)})\): \( K \rho^{(1)} > 1 \) there is a set \((p_1^{(2)}, \ldots, p_r^{(2)})\): \( K \rho^{(2)} = 1 \) such that \( TC(p_1^{(1)}, \ldots, p_r^{(1)}) > TC(p_1^{(2)}, \ldots, p_r^{(2)}) \). Thus, it is clear that \( \min_{K \rho < 1} TC(p_1, \ldots, p_r) = \min_{(p_1, \ldots, p_r)} TC(p_1, \ldots, p_r) \), since \( TC(p_1, \ldots, p_r) \) is continuous in the set of points \((p_1, \ldots, p_r)\) satisfying \( K \rho = 1 \), and for \( K \rho < 1 \) \( TC(p_1, \ldots, p_r) = \sum_{n=1}^{r} C_n (1 - K \rho_n) \) is a linear function of \( p_1, \ldots, p_r. \)

Consequently, we obtain the following LP (lineal programming) problem for determining the optimal assignment/routing probabilities:

\[
\min \sum_{n=1}^{r} C_n \left( 1 - \frac{\lambda K}{\mu_n} p_n \right) \text{ or equivalently } \max \sum_{n=1}^{r} C_n \frac{p_n}{\mu_n} \\
\text{subject to:}
\]

\[
0 \leq p_n < \frac{\mu_n}{\lambda K}, n = 1, r; \
\sum_{n=1}^{r} p_n = 1.
\]

The proof follows directly from the solution of this problem for the case \( \lambda K < \sum_{i=1}^{r} \mu_i \).

Q.E.D.
Note: From the Corollary 2.2 we can conclude that if the convergence rate is exponential, then the convergence is faster when the corresponding values of \( \max \left\{ \frac{1}{Kp}, \max_{\rho_i \in \mathbb{R}} \rho_i \right\} \) or \( Kp \) are smaller.

5. NUMERICAL RESULTS

In this Section we provide some numerical results, supporting the assertions of Theorems 2 and 3.

Table 1. The numerical results

<table>
<thead>
<tr>
<th>( \frac{N}{\Delta} )</th>
<th>0.1</th>
<th>0.01</th>
<th>0.001</th>
<th>0.0001</th>
<th>0.00001</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5.169974</td>
<td>4.803308</td>
<td>4.781544</td>
<td>4.779519</td>
<td>4.779318</td>
</tr>
<tr>
<td>10</td>
<td>3.432746</td>
<td>2.637664</td>
<td>2.614681</td>
<td>2.613104</td>
<td>2.612953</td>
</tr>
<tr>
<td>50</td>
<td>2.715281</td>
<td>0.818982</td>
<td>0.740710</td>
<td>0.739046</td>
<td>0.738952</td>
</tr>
<tr>
<td>100</td>
<td>2.714286</td>
<td>0.662792</td>
<td>0.534120</td>
<td>0.531562</td>
<td>0.531461</td>
</tr>
<tr>
<td>500</td>
<td>2.714286</td>
<td>0.606613</td>
<td>0.381024</td>
<td>0.372379</td>
<td>0.372203</td>
</tr>
<tr>
<td>1000</td>
<td>2.714286</td>
<td>0.606557</td>
<td>0.361065</td>
<td>0.336498</td>
<td>0.335295</td>
</tr>
<tr>
<td>( \infty )</td>
<td>2.714286</td>
<td>0.606557</td>
<td>0.361065</td>
<td>0.336498</td>
<td>0.335295</td>
</tr>
</tbody>
</table>

The table 1 contains the values of \( TC_N(\rho_1, \ldots, \rho_r) \), when \( \lambda K \geq \sum_{i=1}^{r} \mu_i \), where \( r = 3, \ K = 5, \ \lambda = 2, \mu_i = 2, \mu_2 = 6, \mu_3 = 3, \ C_1 = 4, \ C_2 = 5, \ C_3 = 1 \) and \( p_1 = p_1^* - \Delta, \ p_2 = p_2^* + \Delta, \ p_3 = p_3^* \), \( p_1^* = 0.2, \ p_2^* = 0.6, \ p_3^* = 0.2 \).

Table 2. The numerical results

<table>
<thead>
<tr>
<th>( \frac{N}{\Delta} )</th>
<th>0.1</th>
<th>0.01</th>
<th>0.001</th>
<th>0.0001</th>
<th>0.00001</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>4.46625</td>
<td>3.915946</td>
<td>3.888256</td>
<td>3.885792</td>
<td>3.885549</td>
</tr>
<tr>
<td>10</td>
<td>3.556086</td>
<td>2.740437</td>
<td>2.712353</td>
<td>2.710231</td>
<td>2.710025</td>
</tr>
<tr>
<td>50</td>
<td>2.634585</td>
<td>0.531772</td>
<td>0.443593</td>
<td>0.441945</td>
<td>0.441866</td>
</tr>
<tr>
<td>100</td>
<td>2.633803</td>
<td>0.361071</td>
<td>0.213223</td>
<td>0.210545</td>
<td>0.210465</td>
</tr>
<tr>
<td>1000</td>
<td>2.633803</td>
<td>0.306056</td>
<td>0.036109</td>
<td>0.020397</td>
<td>0.020112</td>
</tr>
<tr>
<td>10000</td>
<td>2.633803</td>
<td>0.306056</td>
<td>0.031111</td>
<td>0.003611</td>
<td>0.002031</td>
</tr>
<tr>
<td>( \infty )</td>
<td>2.633803</td>
<td>0.306056</td>
<td>0.031111</td>
<td>0.003611</td>
<td>0.000312</td>
</tr>
</tbody>
</table>

The table 2 contains the values of \( TC_N(\rho_1, \ldots, \rho_r) \), when \( \lambda K \geq \sum_{i=1}^{r} \mu_i \), where \( r = 3, \ K = 7, \ \lambda = 2, \mu_i = 2, \mu_2 = 6, \mu_3 = 3, \ C_1 = 4, \ C_2 = 5, \ C_3 = 1 \) and \( p_1 = p_1^* - \Delta, \ p_2 = p_2^* + \Delta, \ p_3 = p_3^* \), \( p_1^* = 2/11, \ p_2^* = 6/11, \ p_3^* = 3/11 \).
### Table 3. The numerical results

| $N$ | $p_1^{\text{opt}}$ | $p_2^{\text{opt}}$ | $T_{C_N}^{\text{opt}}$ | $T_{C_{\infty}}^{\text{opt}}$ | $|p_1^{\text{opt}} - p_1^*|$ | $|p_2^{\text{opt}} - p_2^*|$ | $|T_{C_N}^{\text{opt}} - T_{C_N}^*|$ |
|-----|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 5   | 0.289966         | 0.710034         | 4.207115         | 1.918872         | 0.089966         | 0.110034         | 0.577612         |
| 10  | 0.236364         | 0.711684         | 2.42267          | 0.868571         | 0.036364         | 0.111684         | 0.191594         |
| 50  | 0.207877         | 0.600393         | 0.73624          | 0.356203         | 0.000787         | 0.000393         | 0.000356         |
| 100 | 0.20089          | 0.599041         | 0.531282         | 0.340941         | 0.000089         | 0.000059         | 0.000179         |
| 500 | 0.199955         | 0.599736         | 0.372146         | 0.372193         | 0.000004         | 0.000026         | 0.000046         |
| 1000| 0.199995         | 0.599863         | 0.352692         | 0.352715         | 0.000005         | 0.000137         | 0.000023         |

The table 3 of $p_i^{\text{opt}}$ ($i = 1, r - 1$) and $T_{C_N}^{\text{opt}} = \min_{0 \leq p_i \leq 1, i=1,r} TC_N(p_1, ..., p_r)$ values, when $\lambda K \leq \sum_{i=1}^r \mu_i$, for $r = 3$, $K = 5$, $\lambda = 2$, $\mu_1 = 2$, $\mu_2 = 6$, $\mu_3 = 3$, $C_1 = 4$, $C_2 = 5$, $C_3 = 1$, where

$T_{C_N}^* = TC_N(p_1^*, p_2^*, p_3^*)$, $p_i^* = \frac{\lambda p_i^*}{\mu_i} (i = 1, 3)$, $p_1^* = 0.2$, $p_2^* = 0.6$, $p_3^* = 0.2$.

$T_{C_{\infty}}^{\text{opt}} = TC(p_1^{\text{opt}}, p_2^{\text{opt}}, p_3^{\text{opt}})$, $p_i^{\text{opt}} = \frac{\lambda p_i^{\text{opt}}}{\mu_i} (i = 1, 3)$.

### Table 4. The numerical results

| $N$ | $p_1^{\text{opt}}$ | $p_2^{\text{opt}}$ | $T_{C_N}^{\text{opt}}$ | $T_{C_{\infty}}^{\text{opt}}$ | $|p_1^{\text{opt}} - p_1^*|$ | $|p_2^{\text{opt}} - p_2^*|$ | $|T_{C_N}^{\text{opt}} - T_{C_N}^*|$ |
|-----|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 7   | 0.250339         | 0.749661         | 3.117187         | 1.009033         | 0.068521         | 0.204206         | 0.770553         |
| 10  | 0.216506         | 0.66903          | 2.217302         | 0.774478         | 0.034688         | 0.123576         | 0.493281         |
| 50  | 0.184421         | 0.558288         | 0.369965         | 0.114297         | 0.002602         | 0.012834         | 0.071898         |
| 100 | 0.182986         | 0.551354         | 0.176506         | 0.053715         | 0.001168         | 0.005899         | 0.033954         |
| 1000| 0.181923         | 0.546003         | 0.016883         | 0.005108         | 0.000105         | 0.000549         | 0.003221         |

The table 4 of $p_i^{\text{opt}}$ ($i = 1, r - 1$) and $T_{C_N}^{\text{opt}} = \min_{0 \leq p_i \leq 1, i=1,r} TC_N(p_1, ..., p_r)$ values, when $\lambda K \geq \sum_{i=1}^r \mu_i$, for $r = 3$, $K = 7$, $\lambda = 2$, $\mu_1 = 2$, $\mu_2 = 6$, $\mu_3 = 3$, $C_1 = 4$, $C_2 = 5$, $C_3 = 1$, where

$T_{C_N}^* = TC_N(p_1^*, p_2^*, p_3^*)$, $p_i^* = \frac{\lambda p_i^*}{\mu_i} (i = 1, 3)$, $p_1^* = 2/11$, $p_2^* = 6/11$, $p_3^* = 3/11$.

$T_{C_{\infty}}^{\text{opt}} = TC(p_1^{\text{opt}}, p_2^{\text{opt}}, p_3^{\text{opt}})$, $p_i^{\text{opt}} = \frac{\lambda p_i^{\text{opt}}}{\mu_i} (i = 1, 3)$.

**Note:** It can be easily seen that in most of numerical examples we obtained results very closed to optimal and to limiting values already for $N = 50$.

**Acknowledgement**

This research was supported by the Paul Ivanier Center for Robotics Research and Production Management at Ben-Gurion University.
REFERENCES


APPENDIX

Here we provide several Lemmas and Corollaries, which we need in order to prove our main results in Section 3.

**Lemma 1:**

\[ S_N^r (\rho_1, ..., \rho_r) = \frac{1}{\rho_i - \rho_j} \left[ \rho_i S_N^{r-1}(\rho_1, ..., \rho_{j-1}, \rho_{j+1}, ..., \rho_r) - \rho_j S_N^{r-1}(\rho_1, ..., \rho_{j-1}, \rho_{j+1}, ..., \rho_r) \right], \]

where:

\[ \rho_1, ..., \rho_r > 0 \quad \text{and} \quad \rho_i \neq \rho_j \quad \text{for} \quad i, j = 1, r, \quad i \neq j, \quad r \geq 2, \quad N \geq 1. \]

**Lemma 2:**

\[ S_N^r (\rho_1, ..., \rho_r) = \sum_{i=1}^{r} \frac{\rho_i^{r-1} S_N^{i-1}(\rho_i)}{\prod_{k=1, k \neq i}^{r} (\rho_i - \rho_k)}, \]

where:

\[ \rho_1, ..., \rho_r > 0 \quad \text{and} \quad \rho_i \neq \rho_j \quad \text{for} \quad i, j = 1, r, \quad i \neq j, \quad r \geq 2, \quad N \geq 1. \]
Lemma 3:
\[
\sum_{i=1}^{r} \rho_i^{r-2} \prod_{k=2}^{r} (\rho_i - \rho_k) = 0,
\]
where:
\[
\rho_1, ..., \rho_r > 0 \text{ and } \rho_i \neq \rho_j \text{ for } i, j = 1, r, i \neq j, r \geq 2.
\]

Corollary A1:
\[
\sum_{n=2}^{r} \frac{\rho_n^{r-3}}{1 - \frac{\rho_1}{\rho_n} \prod_{k=2}^{r} (\rho_n - \rho_k)} \prod_{k=2}^{r} (\rho_1 - \rho_k) = -\rho_1^{r-2},
\]
where:
\[
\rho_1, ..., \rho_r > 0 \text{ and } \rho_i \neq \rho_j \text{ for } i, j = 1, r, i \neq j, r \geq 2, \prod_{k=2}^{r} (\rho_n - \rho_k) = 1.
\]

Lemma 4:
\[
\sum_{n_1}^{N-n_2} \sum_{n_2=0}^{n_1} \sum_{n_3=0}^{n_2} \ldots \sum_{n_l=0}^{n_{l-1}} = \binom{N+i}{i}, \quad i \geq 1.
\]

Let \( \rho_i = \ldots = \rho_m = \rho \quad (m \in \{0, 1, \ldots, t\}), \quad 0 < \rho_i < \rho \quad \text{and} \quad \rho_i \neq \rho_j \quad (i, j = m+1, t, \quad i \neq j), \quad S^0_n(\cdot) = 1 \quad (n = 0, N).

Lemma 5:
For \( t \geq 1 \) it is true:
\[
S^t_N(\rho_1, \ldots, \rho_m, \rho_{m+1}, \ldots, \rho_t) = \left\{ \begin{aligned}
&\sum_{i=0}^{N} \binom{i+m-1}{i} K_{\min(i,N-k)} K! \rho_j^{i} \min(N-i, K)! , & \text{if } m = t, \\
&\sum_{i=0}^{N} \binom{i+m-1}{i} (Kp)^i S_{N-i}(\rho_i) , & \text{if } m = t-1, \\
&\sum_{i=0}^{N} \binom{i+m-1}{i} (Kp)^i \rho_j^{i-m+1} S_{N-i}(\rho_j) \prod_{n=m+1}^{t} (\rho_j - \rho_n) , & \text{if } m = 1, t-2.
\end{aligned} \right.
\]

Lemma 6:
\[
F(s, i, q) = \sum_{m=0}^{s} (m+1)(m+2) \ldots (m+i)q^m =
\]
\[
= i!q^{s+1} - q^{s+1} \sum_{k=1}^{s+1} \frac{1}{(1-q)^k} \left( s + i + k + 1 \right), \quad q \neq 1, \quad i \geq 1.
\]
\[
F(s, i, 1) = \sum_{m=0}^{s} (m+1)(m+2) \ldots (m+i) = \frac{(s+1) \ldots (s+i+1)}{i+1}, \quad i \geq 1.
\]
Lemma 7: \[ \sum_{i=0}^{N} \binom{i+m-1}{i} (K\rho)_i S_{N-i}^1 (\rho_j) = \]

\[ = \frac{S_j^1 (\rho_j)}{(1 - \frac{\rho}{\rho_j})^m} \sum_{i=0}^{m} \frac{1}{\rho_j} \sum_{e=0}^{\min(N-s, K)} K^{\min(s, N-K)} K! (s + m - e) \rho^e, \text{ where:} \]

\[ 0 < \rho_j < \rho, \ m \geq 2. \]

Lemma 8: \[ S_N^t (\rho_1, \ldots, \rho_m, \rho_{m+1}, \ldots, \rho_t) = \sum_{n=m+1}^{t} \frac{\rho_n^{t-1} S_n^1 (\rho_n)}{(\rho_n - \rho)^m \prod_{i=m+1}^{t} (\rho_n - \rho_i)} - \]

\[ - \rho \sum_{e=1}^{m} \left( \sum_{s=0}^{\min(N-s, K)} K^{\min(s, N-K)} K! (s + m - e) \rho^e \right) \left( \sum_{n=m+1}^{t} \frac{\rho_n^{t-m-2}}{(1 - \frac{\rho}{\rho_n}) \prod_{i=m+1}^{t} (\rho_n - \rho_i)} \right), \ m = 1, t-2. \]

Corollary A2:

Let \( \rho_1 = \ldots = \rho_m = \rho \ (m \in \{0, 1, \ldots, t\}), \) \( 0 < \rho_j < \rho \) and \( \rho_i \neq \rho_j \) \((i, j = m+1, t, i \neq j)\), \( \sum_{k=1}^{t} (\rho_i - \rho_i) = 1. \) Then, for \( t \geq 1, \)

\[ S_N^t (\rho_1, \ldots, \rho_m, \rho_{m+1}, \ldots, \rho_t) = \]

\[ \begin{cases} 
\sum_{s=0}^{\min(N-s, K)} K^{\min(s, N-K)} K! (s + m - 1) \rho^s, & \text{if } m = t, \\
\sum_{n=m+1}^{t} \frac{\rho_n^{t-1} S_n^1 (\rho_n)}{(\rho_n - \rho)^m \prod_{i=m+1}^{t} (\rho_n - \rho_i)} - \\
- \rho \sum_{e=1}^{m} \left( \sum_{s=0}^{\min(N-s, K)} K^{\min(s, N-K)} K! (s + m - e) \rho^e \right) \left( \sum_{n=m+1}^{t} \frac{\rho_n^{t-m-2}}{(1 - \frac{\rho}{\rho_n}) \prod_{i=m+1}^{t} (\rho_n - \rho_i)} \right), & \text{if } m = 0, t-1. 
\end{cases} \]
Lemma 9:
\[
\sum_{s=0}^{N} \frac{K^{\min(s, N-K)} K!}{\min(N-s, K)!} \left( \frac{s+i}{s} \right) \rho^s = \frac{1}{(1-K\rho)^{i+1}} + O(\alpha^r), \text{ for } 0 < K\rho < \alpha < 1 \text{ and } i \geq 0.
\]

Corollary A3:
Lemma 9, for \( K\rho \geq 1 \), results in:
\[
\sum_{s=0}^{N} \frac{K^{\min(s, N-K)} K!}{\min(N-s, K)!} \left( \frac{s+i}{s} \right) \rho^s = O\left( \frac{1}{N} \right), \text{ for } K\rho \geq 1 \text{ and } i \geq 0.
\]

Corollary A4:
Let \( \rho_1 = ... = \rho_m = \rho \ (m \in \{0, 1, \ldots, t\}, \ t \in \{0, 1, \ldots, r\}) \), \( 0 < \rho_i < \rho \ (i = m+1, t) \), \( \rho_i = 0 \ (i = t+1, r) \),
\[
\sum_{k=1}^{0} = 0 \text{ and } \prod_{k \neq t} (\rho_i - \rho_k) = 1. \text{ Then, for } K \max_{i=1,t}(\rho_i) < \alpha < 1 \text{ and } r \geq 1, \text{ we have }
\]
\[
S_N^{\tau}(\rho_1, \ldots, \rho_m, \rho_{m+1}, \ldots, \rho_r) = \frac{1}{\prod_{j=1}^{r} (1-K\rho_j)} + O(\alpha^r).
\]