Monitoring of Stochastic Particle Systems: Analysis and Optimization

Iddo Eliazar∗† 
eldar@post.tau.ac.il

Uri Yechiali†
uriy@post.tau.ac.il

March 11, 2006

Abstract

Consider a system to which particles of random lifetimes flow stochastically. The system is monitored at discrete time epochs following a renewal process. When the system is detected non-empty (upon monitoring), a service procedure is initiated – clearing the system off particles. Once the service procedure is concluded, the system’s evolution regenerates.

Particles may represent customers or jobs in a queueing system, contaminants in a physical system, hazardous chemical or biological agents in an environmental system, standing buy/sell orders at a brokerage center, etc.

This class of stochastic systems is modeled and analyzed. We (i) derive the joint transform of the time-to-first-detection and the number of particles present in the system at that epoch, and compute their statistics; (ii) define and calculate various path-functionals and performance measures of the system; and, (iii) study the issue of optimal monitoring schemes.

Keywords: stochastic particle systems; queueing theory; discrete-time monitoring; optimal monitoring.

1 Introduction

A multitude of physical or ‘real world’ systems can be characterized, schematically, as follows:

Independent particles flow stochastically into a system. The particles remain in the system for a random duration of time – their lifetime – and then exit (or vanish). While in the system, the particles need to be attended and processed. Examples of such systems include: impatient customers arriving to

∗School of Chemistry, Raymond & Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
†Department of Statistics & Operations Research, School of Mathematical Sciences, Raymond & Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
a service center; pollutants flowing into a physical system; hazardous chemical or biological agents contaminating an environmental system; standing buy/sell orders arriving to a brokerage center; etc. Often, service in such systems is not provided continuously. Rather, the systems are monitored discretely (i.e., at discrete time epochs), and when detected non-empty a service procedure is initiated.

The case of impatient customers, arriving to a service center with vacationing server(s), was studied in the realm of queueing theory. Queueing systems in which the server(s) is (are) unavailable while on vacation are treated in Levy & Yechiali [4], Takagi [6], Yechiali [8], and Levy & Yechiali [5]. Impatient customers, on the other hand, who abandon the queue if the server does not return from its vacation before their impatience time runs out, are considered in Altman & Yechiali [1]. In such systems the server’s vacationing regime can be looked upon as a stochastic monitoring process.

Our aim in this work is to introduce and study a general model for the type of discretely-monitored stochastic particle systems described above. We consider a generic particle system with the following features: (i) particle lifetimes are governed by a common arbitrary probability law; (ii) monitoring takes place following an arbitrary renewal process; (iii) the service procedure is general (e.g., sequential, parallel, etc.). Upon the completion of a service procedure the system is set back to its empty state (i.e., with no particle present), and the monitoring process resumes anew. We study this model focusing on its analysis and optimization.

The paper is organized as follows. In section 2 we introduce the generic system-model and describe, in detail, the system’s three underlying processes: inflow, monitoring, and service. In section 3 we conduct an analysis of the system, computing the joint statistics of the following pair of random variables: the time-to-first-detection and the number of particles present in the system at that time epoch. In section 4 we (i) introduce various path-functionals of the systems and study their statistics; and, (ii) study the behavior of various service procedures. Last, in section 5, we define several performance measures (i.e., quantitative indicators measuring the systems’ performance), and use them in order to devise optimal monitoring schemes.

2 The model

Consider a system, initiated at time $t = 0$, to which there is a random inflow of particles. The system is monitored at discrete time epochs. If the monitor finds the system clear of particles, no action is taken. However, when the monitor finds the system non-empty a service procedure is initiated. The service procedure clears off particles from the system, at the end of which the system is set back to its ‘empty state’ (with no particles in it).

Let us specify in further detail the system’s three underlying processes: inflow, monitoring, and service:
The inflow process

Particles arrive to the system according to a Poisson process with rate $\lambda$. A particle arriving to the system, while not in service, remains in it for a random duration of time – the particle’s ‘lifetime’ – and then departs. The particles are assumed independent and identically distributed (IID), and their common lifetime distribution function is denoted by $F(s), s \geq 0$.

The monitoring process

The system is monitored at a random sequence of discrete time epochs denoted $0 < T_1 < T_2 < T_3 < \cdots$. The sequence of monitoring epochs $\{T_k\}_{k=1}^{\infty}$ is assumed a renewal process. That is, $T_k = \Delta_1 + \cdots + \Delta_k$ where $\{\Delta_k\}_{k=1}^{\infty}$ is an IID sequence of positive-valued random variables, all distributed like the generic random variable $\Delta$ – the ‘inter-monitoring time’.

The service process

Once the monitor finds the system non-empty, a service procedure commences. We denote by $S_n$ ($n = 1, 2, 3, \cdots$) the random duration of time it takes the service procedure to clear the system and restore it to the ‘empty state’ – given that $n$ particles were found in the system when monitored.

Particles arriving to the system while it is in service can be either rejected, or accepted and treated within the service procedure in commence – pending on the specific features of the service regime.

The three processes – inflow, monitoring, and service – are assumed independent. Once the system is restored to its ‘empty state’ the entire process regenerates: the system’s clock is reset to $t = 0$, and a new monitoring sequence (independent of the past) is initiated.

3 Analysis

In this section we conduct the basic analysis of the system described in section 2 above. We begin with an analysis of the inflow process, and then turn to study issues regarding the time-to-first-detection: (i) how long does it take till a system is first detected non-empty? and, (ii) how many particles are in the system at that time epoch?

3.1 The distribution of the inflow process

Consider a system with inflow alone – that is, with no monitoring and service taking place. Let $N(t)$ $(t \geq 0)$ denote the number of particles in the system at time $t$, and let $D(t)$ $(t \geq 0)$ denote the number of particles that have departed the system up to time $t$. The following result, which is a generalization of Bartlett’s theorem for Poisson Processes (see, for example, Kingman [3]), will serve us in the sequel:
**Proposition 1**\( N(t) \) and \( D(t) \) are independent and Poisson-distributed random variables. The mean of \( N(t) \) is \( \lambda \int_0^t (1 - F(s))ds \), and the mean of \( D(t) \) is \( \lambda \int_0^t F(s)ds \).

Readers familiar with the theory of queueing systems should identify \( N(t) \) as the queue-size (at time \( t \)) of the \( M/G/\infty \) queueing model (where jobs arrive at rate \( \lambda \) and require service times governed by the distribution \( F \)). The proof of proposition 1 can be found in Takacs [7].

The following shorthand notation will be henceforth used
\[
\Lambda(t) = \lambda \int_0^t (1 - F(s))ds.
\]

For example, if the particles’ lifetime is exponentially-distributed with parameter \( \eta \) (that is, \( 1 - F(s) = \exp\{-\eta s\} \)) then
\[
\Lambda(t) = \frac{\lambda}{\eta}(1 - \exp\{-\eta t\})).
\]

And, if the particles’ lifetime is Pareto-distributed with exponent \( p \) – namely, \( 1 - F(s) = (1 + s)^{-p} \) – then
\[
\Lambda(t) = \begin{cases} 
\frac{\lambda}{1-p} \left\{(1+t)^{1-p} - 1\right\} & 0 < p < 1, \\
\lambda \ln(1+t) & p = 1, \\
\frac{\lambda}{1-p} \left\{1 - (1+t)^{-(p-1)}\right\} & p > 1.
\end{cases}
\]

### 3.2 The time-to-first-detection

Let \( K \in \{1, 2, 3, \ldots \} \) be the number of the first monitoring epoch at which the system was found non-empty. Namely
\[
K = \inf\{k \geq 1 \mid N(T_k) > 0\}.
\]

The random variable \( K \) is geometrically-distributed with parameter \( 1 - E[\exp\{-\Lambda(\Delta)\}] \), that is
\[
P(K = k) = E[\exp\{-\Lambda(\Delta)\}]^{k-1} (1 - E[\exp\{-\Lambda(\Delta)\}]).
\]

We explain:

Observe the system at the first monitoring epoch \( T_1 \triangleq \Delta \). If there are no particles in the system at time \( T_1 \) then the entire system process regenerates, and otherwise – a service procedure initiates. Hence, \( K \) is geometrically-distributed with parameter \( P(N(T_1) > 0) \). However
\[
P(N(T_1) = 0) = P(N(\Delta) = 0) = E[P(N(\Delta) = 0 \mid \Delta)] = E[\exp\{-\Lambda(\Delta)\}],
\]

\[
(3)
\]
and hence $P(N(T_1) > 0)$ equals $1 - E[\exp\{-\Lambda(\Delta)\}]$.

We now turn to analyze the random pair $(\tau, N(\tau))$ where: $\tau := T_K$ is the ‘time-to-first-detection’ – the time till the first monitoring epoch at which a non-empty system was detected (and a service procedure was called upon); and, $N(\tau)$ is the number of particles present at the system at the time $\tau$. It should be emphasized that while the inflow process and the monitoring epochs are independent, the random time $\tau$ and the inflow process are, on the other hand, highly correlated.

The random pair $(\tau, N(\tau))$ satisfies the following equation (in law):

$$(\tau, N(\tau)) \overset{d}{=} (T_1, N(T_1)) + (\tau', N'(\tau')) \cdot I_{\{N(T_1) = 0\}}, \quad (4)$$

where $(\tau', N'(\tau'))$ is an IID copy of $(\tau, N(\tau))$, which is independent of $(T_1, N(T_1))$. The explanation of equation (4) is identical to the explanation regarding the distribution of random variable $K$.

Observe the system at the first monitoring epoch $T_1$. If particles are present in the system at time $T_1$ then $(\tau, N(\tau)) = (T_1, N(T_1))$. However, if the system is empty at time $T_1$ (i.e., if $N(T_1) = 0$) then $\tau = T_1 + \tau'$ and $N(\tau) = N'(\tau')$.

Writing these two scenarios in a single equation yields (4).

Equation (4), in turn, leads to:

**Proposition 2** The joint transform of the random pair $(\tau, N(\tau))$ is

$$E\left[\exp\{-\omega \tau\} z^{N(\tau)}\right] = \frac{E[\exp\{-\omega \Delta - (1 - z)\Lambda(\Delta)\}] - E[\exp\{-\omega \Delta - \Lambda(\Delta)\}]}{1 - E[\exp\{-\omega \Delta - \Lambda(\Delta)\}]}. \quad (5)$$

where $\omega \geq 0$ and $|z| \leq 1$.

The proof of proposition 2 is given in the appendix. Since $\tau$ is positive valued, and since $N(\tau)$ is integer valued, we used a ‘hybrid’ transform (Laplace transform for the positive-valued variable $\tau$, and $z$-transform for the integer-valued random variable $N(\tau)$).

In particular, proposition 2 implies the following corollaries:

- The Laplace transform of the time-to-first-detection $\tau$ is

$$E[\exp\{-\omega \tau\}] = \frac{E[\exp\{-\omega \Delta\} - E[\exp\{-\omega \Delta - \Lambda(\Delta)\}]]}{1 - E[\exp\{-\omega \Delta - \Lambda(\Delta)\}]]. \quad (6)$$

- The $z$-transform of $N(\tau)$ is

$$E\left[z^{N(\tau)}\right] = \frac{E[\exp\{-z(1 - z)\Lambda(\Delta)\}] - E[\exp\{-z\Lambda(\Delta)\}]}{1 - E[\exp\{-z\Lambda(\Delta)\}]}. \quad (7)$$

This, in turn, yields the probability distribution of $N(\tau)$:

$$P(N(\tau) = n) = \frac{1}{n!} E[\exp\{-\Lambda(\Delta)\} \Lambda(\Delta)^n] \frac{1}{1 - E[\exp\{-\Lambda(\Delta)\}]} \quad \left(= \frac{P(N(\Delta) = n)}{P(N(\Delta) > 0)}\right)$$

$(n = 1, 2, 3, \cdots)$.
• The covariance $\text{Cov}(\tau, N(\tau))$ of the pair $(\tau, N(\tau))$ is

$$
\text{Cov}(\tau, N(\tau)) = \frac{\mathbb{E}[\Delta \cdot \Lambda(\Delta)] \mathbb{E}[1 - \exp\{-\Lambda(\Delta)\}] - \mathbb{E}[\Lambda(\Delta)] \mathbb{E}[\Delta(1 - \exp\{-\Lambda(\Delta)\})]}{(1 - \mathbb{E}[\exp\{-\Lambda(\Delta)\}]^2)
$$

(8)

Equation (8) is derived by differentiation of equation (5). Indeed, setting $V(\omega; z) = \mathbb{E}\left[\exp\{-\omega \tau\} z^{N(\tau)}\right]$ we have:

$$
\mathbb{E}[\tau] = -\frac{\partial V}{\partial \omega}(0; 1); \quad \mathbb{E}[N(\tau)] = \frac{\partial V}{\partial z}(0; 1); \quad \mathbb{E}[\tau \cdot N(\tau)] = -\frac{\partial V}{\partial \omega \partial z}(0; 1);
$$

and, hence,

$$
\text{Cov}(\tau, N(\tau)) = -\frac{\partial V}{\partial \omega \partial z}(0; 1) + \frac{\partial V}{\partial \omega}(0; 1) \frac{\partial V}{\partial z}(0; 1).
$$

(9)

Computing the derivatives of the function $V(\omega; z)$, and substituting them into equation (9), yields equation (8).

4 System functionals

In this section we study various functionals of the monitoring system. We begin with path functionals — additive functionals of the system’s sample-path trajectories. We then turn to study the length of the service procedure, under different cases of service regimes.

4.1 Path functionals

Let $Y$ be a functional of the system’s sample path along the random time interval $[0, \tau]$ (i.e., from system initiation till the beginning of the first service procedure). For various path functionals the ‘regenerative representation’

$$
Y \overset{d}{=} X + Y' \cdot I_{\{N(T_1)=0\}}
$$

(10)

holds (in law), where: (i) $X$ is the value of the functional $Y$ along the random time interval $[0, T_1]$; and (ii) $Y'$ is an IID copy of $Y$ which is independent of the random pair $(T_1, N(T_1))$.

We shall henceforth refer to functionals admitting the regenerative representation of equation (10) as additive path functionals.

Taking expectation on both sides of equation (10) yields

$$
\mathbb{E}[Y] = \mathbb{E}[X] + \mathbb{E}[Y] \cdot \mathbb{P}(N(T_1) = 0),
$$

and hence, using equation (3), we obtain that

$$
\mathbb{E}[Y] = \frac{\mathbb{E}[X]}{1 - \mathbb{E}[\exp\{-\Lambda(\Delta)\}]}
$$

(11)
Examples

1. Number of monitoring epochs: \( Y = K \). In this case \( X = 1 \) and hence
\[
\mathbb{E}[X] = 1.
\]

2. Time-to-first-detection: \( Y = \tau \). In this case \( X = T_1 \) and hence
\[
\mathbb{E}[X] = \mathbb{E}[\Delta].
\]

3. Number of particles at service initiation: \( Y = N(\tau) \). In this case \( X = N(T_1) \) and hence
\[
\mathbb{E}[X] = \mathbb{E}[\Lambda(\Delta)].
\]

4. Number of departures: \( Y = D(\tau) \). In this case \( X = D(T_1) \) and hence
\[
\mathbb{E}[X] = \lambda \mathbb{E}[\Delta] - \mathbb{E}[\Lambda(\Delta)].
\]

5. Cumulative particle-load: \( Y = \int_0^\tau N(s)ds \). In this case \( X = \int_0^{T_1} N(s)ds \) and hence
\[
\mathbb{E}[X] = \mathbb{E} \left[ \int_0^\Delta \Lambda(s)ds \right].
\]

6. The ‘particle-free time’: \( \int_0^\tau I\{N(s) = 0\}ds \). In this case \( X = \int_0^{T_1} I\{N(s) = 0\}ds \) and hence
\[
\mathbb{E}[X] = \mathbb{E} \left[ \int_0^\Delta \exp(-\Lambda(s))ds \right].
\]

Remark

It is tempting to use equation (11) in order to compute the joint transform of the random pair \((\tau, N(\tau))\) by setting \( Y = \exp\{-\omega\tau\}z^{N(\tau)} \) (having fixed \( \omega \geq 0 \) and \(|z| \leq 1\)). However, \( Y \) is not an additive functional of the system’s sample-path trajectories, and hence it does not admit the regenerative representation of equation (10). Rather, the random variable \( Y \) satisfies the regenerative representation \( Y \overset{d}{=} X + (Y' - 1) \cdot \mathcal{E}_1 \) where: (i) \( X = \exp\{-\omega T_1\}z^{N(T_1)} \); (ii) \( \mathcal{E}_1 = \exp\{-\omega T_1\}I_{\{N(T_1) = 0\}} \); and, (iii) \( Y' \) is an IID copy of \( Y \) (which is independent of the random pair \((T_1, N(T_1))\)). This representation implies that \( \mathbb{E}[Y] = (\mathbb{E}[X] - \mathbb{E}[\mathcal{E}_1]) / (1 - \mathbb{E}[\mathcal{E}_1]) \). A straightforward computation of \( \mathbb{E}[X] \) and \( \mathbb{E}[\mathcal{E}_1] \) yields, in turn, proposition 2.
4.2 Service procedures

Let $S$ denote the length of a service procedure. Recall that the length of a service procedure, initiated when the monitor detected $n$ particles present in the system, was denoted by $S_n$ ($n = 1, 2, 3 \cdots$). Hence

$$E[S | N(\tau) = n] = E[S_n]. \quad (12)$$

We now analyze three cases of service procedures:

Case (I) – sequential service:

$$S_n = \xi_1 + \cdots + \xi_n, \quad (13)$$

where $\{\xi_n\}_{n=1}^{\infty}$ is an IID sequence of positive-valued random variables with cumulative distribution function $G(\cdot)$, density function $g(\cdot)$, and mean $\mu$. The sequential service model (13) applies to both ‘gated’ and ‘exhaustive’ single-server scenarios. We explain:

In the ‘gated’ scenario the service procedure is conducted sequentially by a single-server, while new-coming particles (arriving after the service has initiated) are rejected. In this scenario $\xi$ represents the time required to process a single particle.

In the ‘exhaustive’ scenario the service procedure is conducted sequentially by a single-server, while new-coming particles (arriving after the service has initiated) are processed as well. The service procedure continues till the server ‘exhausts’ all particles – the ones present at service initiation, as well as the particles arriving while the server is active. In this scenario $\xi$ represents the length of a busy period of an $M/G/1$ queue (see, for example, [7]).

Case (II) – parallel service:

$$S_n = \max\{\xi_1, \cdots, \xi_n\}, \quad (14)$$

where $\{\xi_n\}_{n=1}^{\infty}$ is as above. This service model applies to a scenario where particles present in the system (at service initiation) are processed simultaneously and in parallel, while new-coming particles (arriving after the service has initiated) are rejected. In this scenario $\xi$ represents the time required to process a single particle.

Case (III) – ‘Lévy service’:

$$E[S_n] = \Phi(n) := \int_0^{\infty} (1 - \exp\{-un\})\phi(u)du, \quad (15)$$

where $\phi(\cdot)$ is a non-negative function satisfying the integrability condition

$$\int_0^{\infty} \max\{1, u\}\phi(u)du < \infty.$$
The definition of the function \( \Phi \) is valid, of course, on the entire non-negative half line \( \mathbb{R}_+ \) (rather than merely for the integers \( n = 1, 2, \cdots \)). Readers familiar with the theory of Lévy processes would recognize the function \( \Phi : \mathbb{R}_+ \to \mathbb{R}_+ \) as the Lévy characteristic of a Lévy subordinator (see, for example, Bertoin [2]). The function \( \Phi : \mathbb{R}_+ \to \mathbb{R}_+ \) starts at the origin, is monotone non-decreasing, and concave. A wide class of functions \( \Phi \) is attainable by the ‘Lévy structure’ (15) – to list just a few:

- If \( \phi(u) = (\alpha/\Gamma(1 - \alpha))u^{-1-\alpha} \) (\( 0 < \alpha < 1 \)) then
  \[ \Phi(n) = n^\alpha ; \]

- If \( \phi(u) = \exp\{-au\}/u \) (\( a > 0 \)) then
  \[ \Phi(n) = \ln\{1 + n/a\} ; \]

- If \( \phi(u) = (aa^{-\alpha}/\Gamma(1 - \alpha))\exp\{-au\}u^{-1-\alpha} \) (\( a > 0, 0 < \alpha < 1 \)) then
  \[ \Phi(n) = (1 + n/a)^\alpha - 1 ; \]

- If \( \phi(u) = (a^p/\Gamma(p))\exp\{-au\}u^{p-1} \) (\( a, p > 0 \)) then
  \[ \Phi(n) = 1 - (1 + n/a)^{-p} . \]

In all the three cases presented above the mean service procedure time \( \mathbf{E}[S] \) can be shown to be given by the representation

\[
\mathbf{E}[S] = \mathbf{E}[\Psi(\Lambda(\Delta))] \frac{\mathbf{E}[\exp\{-\Lambda(\Delta)\}]}{1 - \mathbf{E}[\exp\{-\Lambda(\Delta)\}]} ,
\]

where \( \Psi(\cdot) \) is a non-negative valued function. We explain:

In case (I) \( \mathbf{E}[S_n] = \mu n \) and hence \( \mathbf{E}[S] = \mu \mathbf{E}[N(\tau)] \). Using example #3 of the previous subsection thus yields

\[
\mathbf{E}[S] = \frac{\mu \mathbf{E}[\Lambda(\Delta)]}{1 - \mathbf{E}[\exp\{-\Lambda(\Delta)\}]} .
\]

This, in turn, implies that equation (16) holds with \( \Psi(x) = \mu x \).

In cases (II) and (III) the situation complicates, and we have:

**Proposition 3** The function \( \Psi(\cdot) \) admits the ‘Lévy form’

\[
\Psi(x) = \int_0^1 (1 - \exp\{-ux\})\psi(u)du ,
\]

where the ‘Lévy density’ \( \psi(\cdot) \) is given by:

Case (II) – parallel service:
\[
\psi(u) = \frac{1}{g(G^{-1}(1 - u))} ;
\]

Case (III) – Lévy service:
\[
\psi(u) = \frac{\phi(-\ln(1 - u))}{1 - u} .
\]

The proof of proposition 3 is given in the appendix.
5 Performance measures and monitoring optimization

In this last section we introduce performance measures which, as their name implies, quantitatively measure the performance of the monitoring system. Equipped with these performance measures we address the issue of monitoring optimization: how to monitor a system so that to obtain optimal system performance.

5.1 Performance measures

Given a path functional \( Y \), consider the following performance measure \( \Theta(Y) \) associated with it:

\[
\Theta(Y) = \frac{E[Y]}{E[\tau] + E[S]}.
\]  

(21)

The performance measure \( \Theta(Y) \) measures the value of the path functional \( Y \) relative to time. We explain:

Consider the first \( m \) system cycles. Each cycle is composed of a period of length \( \tau \) during which the system is occasionally monitored (but not serviced), and a period of length \( S \) during which the system is serviced. Hence, the cumulative value of the path functional over all the first \( m \) system cycles, relative to the cumulative length of these cycles is given by

\[
\frac{Y_1 + \cdots + Y_m}{(\tau_1 + S_1) + \cdots + (\tau_m + S_m)}.
\]

(22)

Moreover, at the termination of each cycle the system regenerates, and the triplets \((\tau_1, S_1, Y_1), \cdots, (\tau_m, S_m, Y_m)\) are therefore IID. Hence, due to the Law of Large Numbers, the stochastic ratio (22) converges (almost surely), as \( m \to \infty \), to the deterministic limit (21).

Now, if \( Y \) is an additive path functional (i.e., it satisfies the equation (10)), then the mean of \( Y \) is given by equation (11). On the other hand, if the service procedure follows one of the three cases presented in the previous subsection then the mean of \( S \) is given by equation (16). Hence, using equations (11) and (16), we obtain that the performance measure \( \Theta(Y) \) is given by

\[
\Theta(Y) = \frac{E[X]}{E[\Delta] + E[\Psi(\Lambda(\Delta))]}.
\]

(23)

where: (i) \( X \) is the value of the path functional over a single inter-monitoring period \( \Delta \); (ii) the function \( \Psi(\cdot) \) is given by \( \Psi(x) = \mu x \) in case (I); and, (iii) the function \( \Psi(\cdot) \) is given by proposition 3 in cases (II) and (III).

For example, if we take \( Y = \tau \) then \( \Theta(Y) \) is the proportion of time in which the system is in service, and \( 1 - \Theta(Y) \) is, conversely, the proportion of time in which the system is not in service. These proportions are given, respectively,
by

\[ \frac{E[\Psi(\Lambda(\Delta))]}{E[\Delta] + E[\Psi(\Lambda(\Delta))]} \quad \& \quad \frac{E[\Delta]}{E[\Delta] + E[\Psi(\Lambda(\Delta))]} \].

In some systems the length \( S \) of the service procedure is 'load-independent'. That is, the distribution of \( S \) is fixed and does not depend on the number of particles detected in the system. For such systems the counterpart of equation (23) is

\[ \Theta(Y) = \frac{E[X]}{E[\Delta] + E[S](1 - E[\exp\{-\Lambda(\Delta)\}])}. \] (24)

(Again, \( Y \) is assumed an additive path functional and \( X \) is \( Y \)'s value over a single inter-monitoring period. And, equation (24) is obtained from equation (21) using equation (11)).

5.2 Monitoring optimization

With performance measures at hand, we turn to study the issue of optimal monitoring. On one hand monitoring is costly, and hence the system administrator would like to monitor the system as rarely as possible. On the other hand, costs are incurred also when there are particles present in the system and are not being served (by the service procedure). Thus, the system administrator has to balance between these 'opposing' costs by designing an optimal monitoring policy.

To formulate this mathematically, let: (i) \( C_{\text{mon}} \) denote the cost of a single monitoring scan; and (ii) \( Y_{\text{cost}} \) denote the costs incurred during a service cycle. Hence, the overall cost during a service cycle is \( C_{\text{mon}} + Y_{\text{cost}} \). If \( Y_{\text{cost}} \) is an additive path functional (i.e., it satisfies equation (10)), then the performance measure (23) associated with \( Y = C_{\text{mon}}K + Y_{\text{cost}} \) is

\[ \frac{C_{\text{mon}} + E[X_{\text{cost}}]}{E[\Delta] + E[S](1 - E[\exp\{-\Lambda(\Delta)\}])}. \]

where \( X_{\text{cost}} \) is the value of the path functional over a single inter-monitoring period \( \Delta \). Now; taking \( \Delta \) deterministic, the optimal monitoring policy is given by the optimization problem

\[ \inf_{\Delta > 0} \frac{C_{\text{mon}} + E[X_{\text{cost}}]}{\Delta + E[S](1 - \exp\{-\Lambda(\Delta)\})}. \] (25)

Analogously, for systems with 'load-independent' service procedure, the performance measure (24) leads us to the optimization problem

\[ \inf_{\Delta > 0} \frac{C_{\text{mon}} + E[X_{\text{cost}}]}{\Delta + E[S](1 - \exp\{-\Lambda(\Delta)\})}. \] (26)
We give several examples:

5.2.1 Server utilization

Assume that the system has a server which executes the service procedures. The system administrator wishes to utilize the server optimally. Let \( C_{\text{idle}} \) denote the cost, per unit time, of keeping the server idle. Therefore, \( Y_{\text{cost}} = C_{\text{idle}} \tau \) and hence

\[
\mathbb{E}[X_{\text{cost}}] = C_{\text{idle}} \Delta .
\]

5.2.2 Optimal restoration

Assume that the system administrator’s objective is to restore the system back to the ‘particle-free’ state as quick as possible. Let \( C_{\text{pen}} \) denote the penalty cost, per unit time, of service. Therefore, \( Y_{\text{cost}} = C_{\text{pen}} S \) and hence

\[
\mathbb{E}[X_{\text{cost}}] = C_{\text{pen}} \Psi(\Lambda(\Delta)) .
\]

5.2.3 De-contamination

Assume that the particles are hazardous contaminants. Once the monitor detects hazardous particles, the system is quarantined and a de-contamination procedure (=service procedure) initiates. Let \( C_{\text{con}} \) denote the cost of the damage, per unit time, incurred by a single hazardous particle present in the system. Also, let \( C_{\text{decon}} \) denote the cost, per unit time, of the de-contamination procedure. Therefore, \( Y_{\text{cost}} = C_{\text{con}} \int_0^\Delta N(s)ds + C_{\text{decon}} S \) and hence

\[
\mathbb{E}[X_{\text{cost}}] = C_{\text{con}} \int_0^\Delta \Lambda(s)ds + C_{\text{decon}} \Psi(\Lambda(\Delta)) .
\]

5.2.4 Loss minimization

Assume that the particle system represents jobs arriving to a service center. A job that does not begin to receive service during its ‘lifetime’ is lost. The system administrator’s objective is to minimize the number of particles lost. Let \( C_{\text{loss}} \) denote the cost per particle lost, and let \( C_{\text{ser}} \) denote the cost, per unit time, of the service procedure. Therefore, \( Y_{\text{cost}} = C_{\text{loss}} D(\tau) + C_{\text{ser}} S \) and hence

\[
\mathbb{E}[X_{\text{cost}}] = C_{\text{loss}} (\lambda \Delta - \Lambda(\Delta)) + C_{\text{ser}} \Psi(\Lambda(\Delta)) .
\]

6 Appendix

6.1 Proposition 2

Proof. For an inter-monitoring period \( \Delta \) set

\[
U(\omega; z) = \mathbb{E}[\exp\{-\omega \Delta - (1 - z)\Lambda(\Delta)\}] ,
\] (27)
where \( \omega \geq 0 \) and \( |z| \leq 1 \). Also set \( V(\omega; z) = \mathbb{E} \left[ \exp\{-\omega \tau\} z^{N(\tau)} \right] \), and note that

\[
V(\omega; z) = \mathbb{E} \left[ \exp\{-\omega \tau\} z^{N(\tau)} I_{\{N(\tau) = 0\}} \right] + \mathbb{E} \left[ \exp\{-\omega \tau\} z^{N(\tau)} I_{\{N(\tau) > 0\}} \right].
\] (28)

Using equation (4) we have

\[
\mathbb{E} \left[ \exp\{-\omega \tau\} z^{N(\tau)} I_{\{N(\tau) = 0\}} \right] = \mathbb{E} \left[ \exp\{-\omega (T_1 + \tau')\} z^{N(\tau')} I_{\{N(\tau') = 0\}} \right] \] (29)

and

\[
\mathbb{E} \left[ \exp\{-\omega \tau\} z^{N(\tau)} I_{\{N(\tau) > 0\}} \right] = \mathbb{E} \left[ \exp\{-\omega T_1\} z^{N(T_1)} I_{\{N(T_1) > 0\}} \right] - \mathbb{E} \left[ \exp\{-\omega T_1\} I_{\{N(T_1) = 0\}} \right].
\] (30)

Due to proposition 1 the random variable \( N(t) \) is Poisson-distributed with mean \( \Lambda(t) \) (\( t \geq 0 \)). Hence, using equation (28) and the fact that inter-monitoring periods are independent of the inflow process, we have

\[
\mathbb{E} \left[ \exp\{-\omega \Delta\} I_{\{N(\Delta) = 0\}} \right] = \mathbb{E} \left[ \exp\{-\omega \Delta\} \mathbb{P} \left( N(\Delta) = 0 \mid \Delta \right) \right] = \mathbb{E} \left[ \exp\{-\omega \Delta\} \exp\{-\Lambda(\Delta)\} \right] = U(\omega; 0),
\] (31)

and

\[
\mathbb{E} \left[ \exp\{-\omega \Delta\} z^{N(\Delta)} \right] = \mathbb{E} \left[ \exp\{-\omega \Delta\} \mathbb{E} \left[ z^{N(\Delta)} \mid \Delta \right] \right] = \mathbb{E} \left[ \exp\{-\omega \Delta\} \exp\{-(1 - z)\Lambda(\Delta)\} \right] = U(\omega; z).
\] (32)

Now, since \( T_1 \overset{d}{=} \Delta \), combining equations (29) and (31) together yields

\[
\mathbb{E} \left[ \exp\{-\omega \tau\} z^{N(\tau)} I_{\{N(\tau) = 0\}} \right] = U(\omega; 0) \cdot V(\omega; z);
\] (33)
and, from equations (30)-(32) we have
\[ E \left[ \exp\{-\omega \tau\} z^{N(\tau)} I_{(N(T),>0)} \right] = U(\omega; z) - U(\omega; 0) \] (34)

Hence, substituting equations (33)-(34) back into equation (28) gives
\[ V(\omega; z) = U(\omega; 0) \cdot V(\omega; z) + (U(\omega; z) - U(\omega; 0)), \]
which, in turn, yields
\[ V(\omega; z) = \frac{U(\omega; z) - U(\omega; 0)}{1 - U(\omega; 0)}. \]

6.2 Proposition 3

Before we begin the proof, first note that equation (7) (for the z-transform of \( N(\tau) \)) implies that
\[ 1 - E \left[ z^{N(\tau)} \right] = 1 - \frac{E \left[ \exp\{-\lambda(\Delta)\} \right]}{E \left[ \exp\{-\lambda(\Delta)\} \right]}. \] (35)

Proof.

Case (II)
Since \( S_n = \max\{\xi_1, \cdots, \xi_n\} \) we have
\[ E[S_n] = \int_0^\infty P(S_n > s) \, ds = \int_0^\infty (1 - G(s)) \, ds. \]

Hence, using equation (35) yields
\[ E[S] = E[E[S \mid N(\tau)]] \]
\[ = E\left[ \int_0^\infty (1 - G(s))^{N(\tau)} \, ds \right] \]
\[ = \int_0^\infty (1 - E\left[ G(s)^{N(\tau)} \right]) \, ds \]
\[ = \int_0^\infty \frac{1 - E\left[ \exp\{-1 - G(s)\lambda(\Delta)\} \right]}{1 - E\left[ \exp\{-\lambda(\Delta)\} \right]} \, ds \]
\[ = \frac{1}{1 - E\left[ \exp\{-\lambda(\Delta)\} \right]} E\left[ \int_0^\infty (1 - \exp\{-1 - G(s)\lambda(\Delta)\}) \, ds \right] \]
\[ = \frac{1}{1 - E\left[ \exp\{-\lambda(\Delta)\} \right]} E[\Psi(\lambda(\Delta))], \]

where
\[ \Psi(x) := \int_0^\infty (1 - \exp\{-1 - G(s)\}x) \, ds \]
\[ = \int_0^1 (1 - \exp(-ux)) \frac{1}{g(1-u)} \, du \]
(in the last transition we used the change of variables \( u = 1 - G(s) \)).
Case (III)

Using equations (15) and (35) yields
\[
E[S] = E[E[S \mid N(\tau)]]
\]
\[
= E \left[ \int_0^\infty (1 - \exp(-sN(\tau))) \phi(s) ds \right]
\]
\[
= \int_0^\infty (1 - E[\exp(-sN(\tau))] \phi(s) ds
\]
\[
= \int_0^\infty \frac{1 - E[\exp(-sN(\tau)) \Lambda(\Delta)]}{1 - E[\exp(-\Lambda(\Delta))]} \phi(s) ds
\]
\[
= \frac{1}{1 - E[\exp(-\Lambda(\Delta))]} E \left[ \int_0^\infty (1 - \exp(-(1 - \exp(-s))\Lambda(\Delta))) ds \right]
\]
\[
= \frac{1}{1 - E[\exp(-\Lambda(\Delta))]} E[\Psi(\Lambda(\Delta))]
\]

where
\[
\Psi(x) := \int_0^\infty (1 - \exp(-(1 - \exp(-s))x)) \phi(s) ds
\]
\[
= \int_0^1 (1 - \exp(-ux)) \frac{d(\ln(1-u))}{1-u} du
\]
(in the last transition we used the change of variables \( u = 1 - \exp(-s) \)).

References