Algorithmic methods in queueing theory: Assignment 1

January 10, 2017

Remarks

- This assignment should be done in group of two students. If a student could not find a student to work with please inform us so we can find a solution.
- Submit a decent report explaining the followed approach to find the answers to the assignment questions at the end of the document.
- Submit the code of your algorithm used to solve the assignment. You can implement your algorithm in a programming language of you preference. For the teacher, the numerical tools such as Matlab and Mathematica are the most easy to use for the assignment review.

1 Introduction

With the advancement of technology, advanced equipment are becoming more capital intensive. Examples of advanced equipment are large airplanes, lithography systems, and large computing servers. The unplanned downtime of these advanced capital equipment can be extremely expensive. Consequently, these unplanned downtimes should be avoided as much as possible and if they occur, they should be kept as short as possible (by using optimal corrective maintenance policies). The latter implies that malfunctioning parts or components causing the system breakdown are immediately replaced by ready-for-use ones. Such a policy in turn requires high availability of the resources (spare parts, tools and service engineers) that are needed to complete the repair job. However, these resources are also expensive and need high investments. Therefore, an optimal availability of resources in maintenance logistics is necessary to meet the expected operational availability while minimizing the total costs.

2 Assignment description

We consider a service region with a local inventory to store K different types of spare parts. There are different types of repair calls in this service region that arrive according to a Poisson process with a rate λ. Each repair call requires a specific spare part. A repair call is of type-k if it requires one unit of type-k spare part. Let p_k denote the probability that a repair call is of type k. For each repair call, a service engineer is also needed to do the job. A team of service engineers of size E is located in the service region. In this model, we assume that the service time of a repair call of type k (the time between the moment the repair job is assigned to a service engineer and the moment the job will be finished) is exponentially distributed with rate μ_k. The inventory of type-k parts is managed according to the base-stock policy, referred to as $(S_k - 1, S_k)$, which means for every piece of spare parts consumed from stock a replenishment order is issued. In such way, the total number of type-k spare parts in stock and in replenishment is kept equal to S_k. For each type-k spare part the replenishment lead time is exponentially distributed with rate ν_k.

Usually, service times of engineers take much less time than the spare parts replenishment. Therefore, for systems for which a short waiting time is acceptable, it is rational to wait for service engineers if they are not immediately available but use an emergency shipment for spare parts in case of no spare part is available upon arrival of a repair call, called stock-out. When the requested spare part is satisfied by an emergency shipment, there are a number of scenarios for the service engineers that can be applied. In this assignment, we assume that both spare parts and service engineers are satisfied via an emergency channel in case of a spare parts stock-out. In other words, internal service engineers are not responsible for emergency repair calls.

When the spare part is available, the backlogging policy is followed by service engineers. This means that when no service engineer is available upon a repair call request (while the spare part is available), the spare part is reserved and the repair call must wait until a service engineer becomes available. A maximum accepted average waiting time is defined for the total waiting time (not per type of call) W_{max}. The repair calls waiting for engineers will be served by a FCFS policy.

In this assignment, we assume that the emergency replenishment rate is much higher than the regular one, but still finite. This means, when a repair call is satisfied by an emergency channel, there is still a waiting time until the emergency shipment arrives. This waiting time is important for the service policy and is included in W_{max}, the maximum accepted average waiting time of repair calls. For the parts that are requested by emergency shipment, the average waiting time is equal to $\frac{1}{\nu_k^{em}}$.

For each spare part, h_k is the holding cost per piece per time unit. In addition, hiring costs of service engineers and emergency costs (a cost per repair call that is satisfied via an emergency channel) are considered in this model.

3 Input data

We consider a small case consisting of $5(=K)$ different parts with $\lambda = 9.1224$ failures per year. An engineer's service rate is equal to 50 repair calls per year for all types of parts $k = 1, \ldots, 5$. The rate of emergency is 101.3340 calls per year for all types of parts $k = 1, \ldots, 5$. The cost of hiring an engineer per year is equal to 62500\$ and the emergency cost per call is 250000\$.

Table 1: Input data for the case with 5 types of spare parts.

k	1	2	3	4	5
p_k	0.1757	0.0028	0.2435	0.2863	0.2917
ν_k (parts per year)	5.0667	1.9000	5.0667	5.0667	5.0667
h_k (in \$ per part per year)	6924.06	293.29	6262.62	6262.62	6262.62

4 Questions

The above problem can be modeled as a structured infinite state space continuous time Markov chain, in particular as a Quasi-Birth-Death process.

1. Define the Quasi-Birth-Death process of the problem and give the input matrices of the process generator.
2. Derive the stability condition of the system and explain how to find the equilibrium probabilities in this case.
3. For $E = 2$ engineers and $(S_1, \ldots, S_5) = (2, 1, 2, 3, 4)$, compute the emergency probability of an arbitrary repair call, the expected number of repair calls waiting for an available engineer, the expected waiting time of a call served locally or via an emergency channel, and the total expected cost per year in steady state.
4. Starting with $E = 2$ engineers and $(S_1, \ldots, S_5) = (2, 1, 2, 3, 4)$, find which resource (part or engineer) leads to the largest reduction in the expected waiting per invested dollar.
5. Based on the above QBD process, propose a method to compute analytically the coefficient of variation of the inter-arrival times of the calls joining the engineers queue. No calculation is required here only a method description.