European Journal of Operational Research 52 (1991) 103–106
North-Holland

Theory and Methodology

Stable strategies for processor sharing systems

Moshe Haviv
Department of Statistics, Hebrew University of Jerusalem, 91905 Jerusalem, Israel, and Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, BC, Canada V6T 1Y8

Abstract: For a processor sharing model with a Poisson arrival process and general and independent service requirement, under a standard cost structure we look for a join/do not join stable policy for each of the following two cases: (1) when each job knows its service requirement, and (2) when the jobs belong to various classes which differ in their expected service requirement and each job knows its class. For the first case, it is shown that there exists a unique pure stable and symmetric strategy under which jobs join the system if and only if their requirement is smaller than some threshold. A similar phenomenon exists in the second case but randomization may be required. Moreover, the stable policies are computed explicitly. The case of social optimization is considered as well.

Keywords: Stable strategies, balking in queues, processor sharing systems

1. Introduction

When a number of individuals have to share a common resource, one's utility decreases with the number of participants and with the size of their demands. This is the case in recreational areas, in service stations or in computer systems. Moreover, if the system is congested and/or one's demand becomes high, his net gain can become negative and he might be better off by not jointing at all. The model we consider here, the processor sharing model, is an example for such a phenomenon.

Consider the following processor sharing M/G/1 system. Jobs are being produced in accordance with a Poisson process with rate λ per unit time. The service requirement of each of them is random and is independently chosen from a probability function G having a continuous density function g. For simplicity assume that the support of the random variable associated with G is $[0, \infty)$. If n jobs are present in the system, the single processor supplies each of them a service amount of $1/n$ unit time per unit time. Finally each job values its service completion by an amount of R while a cost of C is imposed on it per unit of time in the system.

The following example illustrates a common decision problem. A user needs a job of known length to be processed on a computer. His first option is to run it in the local (expensive) facility which shares its capacity equally between its users. He is then charged an amount of C per unit time that he is logged on. Alternatively, he can forward the job to be processed at a computing center. There he is guaranteed to get the results after a fixed delay to which he attaches a cost R. Based on previous records the user knows (statistically)

Received May 1989

the demand generated for the local facility and will try to minimize his (expected) cost.

This paper consider balking policies from an M/G/1 processor sharing system for jobs which have some information on their service requirements. Section 2 concerns the case of jobs knowing exactly their requirements. Then Section 3 looks at the case of jobs belonging to various classes which differ on their expected service requirement while a job only knows its expected service requirement. We show that the right question one should look at is the question of stable (or Nash-equilibrium) strategies. The case considered in Section 2 leads to a pure stable strategy and the stable strategy in Section 3 may include randomness. The strategies are explicitly given.

2. The case with complete information

Suppose each job knows its service requirement, say x, but otherwise does not know any realization concerning the system. In particular, upon the generation of a job no additional information is revealed to it. However, we assume that the model described above is common knowledge, namely each job knows that the arrival process is Poisson with rate λ and the function G represents its prior with respect to the rest of the jobs. [1] We assume that no one reneges after first joining the system. Suppose a job wishes to maximize its expected net gain. Intuitively speaking, it is clear that the larger its service requirement, the smaller the expected gain from joining the system. Hence, an appealing strategy for each individual is to join the system if and only if x is smaller than some threshold x^*. However, as the best response for a job depends on the behavior of the others, one has to look for stable strategies. We note here that a stable strategy (or a Nash-equilibrium strategy) is a set of strategies, a strategy for each job ('a player') such that no one has a strict positive incentive to unilaterally deviate from what the strategy prescribes him, given that the rest follows their strategies. In Theorem 2.1, it is shown that a stable strategy of the above mentioned threshold shape exists and is unique. Moreover, we show how to compute the corresponding threshold x^* explicitly.

The stable energy imposed by x^* is not necessarily the optimal strategy from the viewpoint of the society. As each job considers only its cost while deciding on joining or not joining the system, it does not concern the externalities that its joining imposes on the rest of the society. However, this is done while looking for an optimal threshold rule from the viewpoint of the entire society. Hence, the unique threshold in the latter case, say x^{**} is smaller than x^*. For more on this relation see Hassin (1985). In Theorem 2.2 we indicate how to compute x^{**} as well. Finally, it is evident that by levying the right toll, namely informing the individuals of a reduced R such that the resulting (new) x^* will coincide with (the old) x^{**}, the individuals' stable strategy coincides with the society optimal behavior. The interested reader is referred to Glasser and Hassin (1983; 1986), to Hassin (1986) or to Whittle (1986, Chapters 9.7–9.9) for some queueing models where stable policies are being involved.

Before stating and proving the two theorems, we quote the following two well known results.

Result 1 (e.g., Ross (1983), p. 174).

Under stationarity, the expected delay in an M/G/1 processor sharing system (when everybody joins) for a job whose service requirement is x is $x/1-\rho$ where ρ, the traffic intensity equals $\lambda\bar{x}$ and where $\bar{x}$ is the expected service requirement under G.

Result 2 (e.g., Ross (1983), p. 172).

Under stationarity, in an M/G/1 processor sharing system the queue size is n with probability $(1-\rho)\rho^n$ for $n \geqslant 0$. In particular, its mean is $\rho/1-\rho$.

Theorem 2.1. *The unique symmetric stable strategy is "join the system if and only if the service requirement x is smaller than or equal to x^*" for some particular x^*.* [2] *Moreover, x^* is the unique solution to the equation (in x)*

$$\frac{Cx}{1-\rho(x)} = R,$$

where $\rho(x) = \lambda \int_{y=0}^{x} y g(y)\, \mathrm{d}y$.

[1] The case where each job knows the queue size but does not know its own requirement is covered by Assaf and Haviv (1990).

[2] Symmetry here means that all jobs follow the same strategy.

Proof. Suppose the entire society adopts some join/do not join strategy. Consider a tagged job who wishes to maximize its expected gain knowing its service requirement and the strategy used by the others. If it joins the system, its expected gain is strictly monotone decreasing with its service length. Hence, its best response should be of the threshold shape. As this is the case for any job and as a stable strategy prescribes a best response for each individual given the rest follow this strategy, we have thus established that a necessary condition for a symmetric strategy to be stable in that it is of a threshold shape.

Now suppose the entire society uses the threshold x^*. Hence, it is easy to see that the resulting system is of the M/G/1-processor sharing type but with $\lambda G(x^*)$ as the Poisson arrival rate and with $g(x)/G(x^*)$ for $0 \leqslant x \leqslant x^*$ and zero elsewhere as the service density function. Then, by Result 1, the expected delay in this system for a job whose length is x, for $0 \leqslant x \leqslant x^*$, is $x/(1-\rho(x^*))$. Suppose $Cx^*/(1-\rho(x^*)) < R$. Then, for $x > x^*$ but sufficiently small, $Cx/(1-\rho(x^*)) < R$ as well. Namely, if the entire society uses strategy x^*, the one who unilaterally deviates to the strategy of balking only for requirements larger than x will receive a gain whose expected value is bounded from below by $R - Cx/(1-\rho(x^*)) > 0$ if its service requirement is x. This is more than it receives from not joining, where gain is zero. Similarly one can contradict the assumption that $Cx^*/(1-\rho(x)) > R$. Hence, $Cx^*/(1-\rho(x^*)) = R$. As $Cx/(1-\rho(x))$ is continuous, unbounded and strictly monotone increasing, x^* is uniquely defined. □

Theorem 2.2. *The socially optimal threshold x^{**} equals*

$$\arg\max_{0 \leqslant x} \left[\lambda G(x) R - \frac{C\rho(x)}{1-\rho(x)} \right].$$

Proof. Suppose the threshold rule x is being adopted by the entire society. Then, the arrival is Poisson and with rate $\lambda G(x)$. Hence the society is being rewarded by $\lambda G(x) R$ per unit time. On the other hand, by Result 2, the expected queue length, under this strategy is $\rho(x)/(1-\rho(x))$. □

Remarks. (1) The social optimal toll, T, is now found by solving for T:

$$R - T = \frac{Cx^{**}}{1-\rho(x^{**})}.$$

Return to the example mentioned in the introduction. A manager representing the society will encourage users to use the computer center by subsidizing each of those using it by an amount of T. In that way he will reduce the load on the local facility. While adopting the corresponding stable strategy the users will then behave in a social optimal way without any further intervention of the manager. Note that in the processor sharing system the expected delay is linear with the length of the job and hence no user has an incentive to disguise himself as two or more users.

(2) The value of x^* can be found by a one-dimensional search procedure such as bisection. In the search of x^{**} one can utilize the observation made above that $x^{**} \leqslant x^*$.

(3) As Results 1 and 2 are valid as well for the M/G/1-system with last come first served with preemption resume queueing discipline (e.g., Kleinrock (1976), p. 171), the same is the case with Theorems 2.1 and 2.2.

3. The case with partial information

Suppose jobs can be partitioned into a countable number of classes. The probability a job belongs to class r is p_r and then its expected service requirement is m_r. Without loss of generality, assume $m_1 < m_2 < \cdots$ and that $p_r > 0$ for all classes. Similar to Section 2 we assume that each job knows its class and that the above is common knowledge.

Before stating our main results concerning this case we need the following notation. For all integers r, $r \geqslant 0$, let $\rho(r)$ be the traffic intensity when all jobs use the strategy to join the system if and only if the index of their class is smaller than or equal to r. Of course, $\rho(r)$ is monotone increasing and it is easy to see that for $r \geqslant 1$,

$$\rho(r) = \lambda \sum_{i=1}^{r} p_i m_i.$$

For completeness let $\rho(0)=0$. Finally, let the integer r^* be defined by

$$r^* \equiv \sup\left\{ r \geqslant 0 \mid \rho(r) < 1 \text{ and } \frac{Cm_r}{1-\rho(r)} \leqslant R \right\}.$$

It is easy to see that if the entire society uses the policy to join the system if and only if the number of stages is smaller than or equal to r^*, then everybody expects a nonnegative gain. However, if r^* is finite $p_{r^*+1} > 0$, then the policy which replaces r^* by r^*+1 as the critical class results is a negative expected gain for those in class r^*+1.

Theorem 3.1. *If $Cm_{r^*+1}/(1-\rho(r^*)) \geqslant R$, then joining the system if and only if the class index is smaller than or equal to r^* is the unique symmetric stable strategy. Otherwise, let θ^* be the solution of the equation (in θ)*

$$\frac{Cm_{r^*+1}}{1-\rho(r^*)-\lambda\theta p_{r^*+1} m_{r^*+1}} = R.$$

Then $0 \leqslant \theta^* < 1$ and the following is the unique symmetric stable strategy: join the system when the class index is smaller than or equal or r^*. If it is r^*+1 join with probability θ^* and if it is larger than r^*+1, do not join.

Proof. First, under any policy adopted by the entire society, the expected net gain for an individual is strictly decreasing with the index of its class (in case of joining). Second, note that a necessary condition for a stable strategy to prescribe randomization for a class is that a job belonging to that class will be indifferent between joining and not joining (given the rest follow the stable strategy). Also, by the former observation, note that there is at most one class for which randomization may be stable. These facts coupled with the definition of r^* immediately indicate that for $r \leqslant r^*$ (resp., $r \geqslant r^*+2$) the stable strategy prescribes to join (resp., not to join).

Finally, for r^*+1, note that if $Cm_{r^*+1}/(1-\rho(r^*)) \geqslant R$, then the one with r^*+1 stages does not expect to have positive gain from joining even under the (favorite) case where the rest from his class and higher indexed classes are not joining. Hence, not joining is its best response. This is not the case if $Cm_{r^*+1}/(1-\rho(r^*)) < R$. However, by the definition of r^*, if all from class r^*+1 join, they all will expect a negative gain and hence one from class r^*+1 will be better off by deviating unilaterally from this policy and not joining. Hence randomization, and with θ^*, will make all those with r^*+1 stages indifferent between joining, not joining or any mixture, like θ^*, between them. Note that $\lambda\theta^* p_{r^*+1} m_{r^*+1}$ is the traffic intensity due to those from class r^*+1. □

Acknowledgement

The author likes to acknowledge helpful discussions with Yair Wand.

References

Assaf, D., and Haviv, M. (1990), "Reneging from time sharing and random queues", *Mathematics of Operations Research* 15, 129–138.

Glazer, A., and Hassin, R. (1983), "?/M/1: On the equilibrium distribution of customer arrivals", *European Journal of Operational Research* 13, 146–150.

Glazer, A., and Hassin, R. (1986), "Stable purchasing in queues", *Operations Research Letters* 41, 285–288.

Hassin, R. (1985), "On the optimality of the last come first serve queues", *Econometrica*, 53, 201–202.

Hassin, R. (1986), "Consumer information in market with random product quality: The case of queues and balking", *Econometrica* 54, 1185–1195.

Kleinrock, L. (1976), *Queueing Systems, Vol. II: Computer Applications*, John Wiley and Sons, New York.

Ross, S.M. (1983), *Stochastic Processes*, John Wiley and Sons, New York.

Whittle, P. (1986), *Systems in Stochastic Equilibrium*, John Wiley and Sons, Chichester.