

Я. М. Агаларов (Москва, ФИЦ ИУ РАН). **Оптимальное ограничение длины очереди в системе $M/M/s$ с неоднородными приборами и общим накопителем.**

УДК 004:942

Резюме: В докладе рассматривается система $M/M/s$ с неоднородными приборами и общим накопителем с возможностью управления верхним порогом длины очереди с целью максимизации среднего предельного дохода. Функция дохода включает плату за успешно обслуженную заявку, штраф за каждую отклоненную заявку, штрафы за единицу времени простоя каждого прибора, штраф за единицу времени ожидания заявки (или превышение допустимого значения времени ожидания заявки, или превышение допустимого значения времени пребывания заявки в системе с однородными приборами), затраты, связанные с техническим обслуживанием мест в накопителе. Доказано свойство выпуклости функции дохода от верхнего порога дли, получены условия существования конечного оптимального верхнего порога длины очереди.

Ключевые слова: система массового обслуживания, оптимизация, накопитель, длина очереди.

Рассматривается СМО с s неоднородными приборами, на которую поступает пуассоновский поток заявок с интенсивностью $\lambda > 0$. Предполагается, что приборы имеют условные номера $1, \dots, N$ и закон распределения времени обслуживания на i -м приборе является показательным с параметром $\mu_i, \mu_1 \geq \dots \geq \mu_s$. Заявки допускаются в систему согласно простой пороговой стратегии следующим образом: поступившая заявка допускается в систему (занимает любое свободное место в накопителе), если в момент ее поступления число заявок в системе меньше порогового значения $k, k \geq s$, в противном случае отклоняется и повторно не возвращается. При распределении заявок на приборы предпочтение отдается наиболее производительным: если в момент поступления в системе есть свободный прибор, то новая заявка занимает свободный прибор с наименьшим номером и начинает сразу обслуживаться, иначе становится в конец очереди на обслуживание. Заявка после окончания обслуживания сразу покидает систему, освободив прибор и место в накопителе. Если в очереди нет заявок и производительность освобожденного прибора строго больше, чем производительность одного из занятых приборов, то заявка с прибора наименьшей среди занятых приборов производительности снимается и пересылается на освободившийся прибор. Параллельное обслуживание заявки на нескольких приборах не допускается.

В качестве показателя эффективности системы используется $D(k)$ — «доход» системы в единицу времени в стационарном режиме работы при стратегии k , где $D(k)$ одна из следующих функций по $k \geq s$:

$$D(k) = C_0 \Lambda_{\text{вых}}(k) - C_1 \Lambda_{\text{отк}}(k) - C_2 W_{\text{ож}}(k) - \sum_{(i=1)}^s C_{3i} Q_{(\text{пр}, i)}(k) - C_4 k, \quad (1)$$

где $\Lambda_{\text{вых}}(k)$ — интенсивность обслуженных заявок, $\Lambda_{\text{отк}}(k)$ — интенсивность отклоненных заявок, $W_{\text{ож}}(k)$ — суммарное среднее время ожидания в очереди обслуженных

в единицу времени заявок, $Q_{(пр,i)}(k)$ — вероятность простоя i -го прибора, $C_0 \geq 0$ — плата, получаемая системой, если поступившая заявка будет обслужена системой, $C_1 \geq 0$ — штраф за отклонение поступившей заявки, $C_2 \geq 0$ — штраф за единицу времени ожидания заявки в очереди, $C_{3i} \geq 0$ — штраф за единицу времени простоя i -го прибора, $C_4 \geq 0$ — стоимость технического обслуживания одного места в накопителе в единицу времени. Ставится задача оптимизации пороговой стратегии:

$$k^* = \arg \max_{k \geq s} D(k). \quad (2)$$

Утверждение. Функция $D(k)$ — выпуклая вверх функция и решение задачи (2) удовлетворяет условиям:

$$k^* = \begin{cases} \infty, & \text{если } C_2 = 0; \\ s, & \text{если } d_{пр}(s+1) - d_{пр}(s) < \lambda [d_{пр}(s) - d_{пр}(s+1)], \end{cases}$$

иначе $0 < k^* < \infty$, где $d_{пр}(k)$ — предельное значение дохода системы без учета штрафов за простой приборов при стратегии k , усредненного по количеству поступивших заявок.

Отметим, что аналогичный результат имеет место и тогда, когда C_2 — штраф за превышение временем пребывания заявки в очереди допустимого предельного значения $t_d > 0$, т. е., если в формуле (1) функцию $W_{ож}(k)$ заменить на $\Lambda_d(k)$ — интенсивность обслуженных заявок, время ожидания которых больше, чем t_d . Утверждение справедливо также в случае однородных приборов и $\Lambda_d(k)$ — интенсивность обслуженных заявок, время пребывания которых в системе больше, чем t_d .

УДК 004:942

Agalarov Ya. M. (Moscow, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences) . **Optimal limitation of the queue length in an M/M/s system with heterogeneous devices and a common storage**

Abstract: The report discusses the M/M/s system with heterogeneous devices and a common storage with the ability to control the upper queue length threshold in order to maximize the average marginal profit. The profit function includes a fee for a successfully serviced request, a fine for rejected request, fine of downtime for each device, a fine for waiting for an request (or for exceeding the allowable request waiting time), costs associated with storage maintains. The property of convexity of the income function is proved; conditions for the existence of a finite optimal upper queue length threshold are obtained.

Keywords: Keywords: queuing system, optimization, storage drive, queue length.