

# Распределение ресурсов в ненаблюдаемых системах с параллельным обслуживанием и ненадежными серверами

М.Г. Коновалов\*, Р.В. Разумчик\*<sup>†</sup>

\* *Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Москва, Россия*

<sup>†</sup> *Московский центр фундаментальной и прикладной математики, Москва, Россия*

Email: mkonovalov@ipiran.ru, rrazumchik@ipiran.ru

Предметом исследования являются стратегии распределения ресурсов (также называемые стратегиями маршрутизации, распределения заданий, диспетчеризации) для систем обслуживания при неполном наблюдении. Рассматривается класс почти ненаблюдаемых систем с параллельной обработкой заданий. Система этого класса состоит из нескольких параллельно и независимо друг от друга работающих серверов, которые выполняют задания, направляемые на них единственным диспетчером. Серверы выходят из строя в случайные моменты времени и после ремонта возобновляют обработку заданий. Предполагается, что задания поступают группами, причем времена между поступлениями, размеры групп и объемы заданий являются независимыми случайными величинами с произвольными законами распределения и конечными средними значениями. Обмен заданиями между серверами невозможен, а применяемая дисциплина обслуживания очереди внутри каждого сервера может быть любой консервативной. Диспетчер, осуществляя выбор сервера для очередного поступающего задания, не имеет никакой информации о текущем состоянии системы. Единственное, что диспетчер может наблюдать (и запоминать) — это свои собственные действия и моменты их совершения. В то же время диспетчеру известна определенная априорная информация. Предполагаются известными: производительности серверов, распределение времени между поступлениями заданий, распределение размера группы и объема заданий, и моменты, в которые изменяются состояния серверов. В указанных условиях возникает проблема построения оптимальной стратегии диспетчеризации с точки зрения минимального стационарного среднего времени пребывания задания в системе или, что то же самое, стационарного среднего времени отклика. Из всех известных из мировой научной периодической печати диспетчеризаций в указанных условиях могут быть реализованы лишь две: случайный выбор [1] и детерминированный выбор [2]. В докладе речь пойдет о недавно разработанном семействе алгоритмов диспетчеризации (по-видимому не поддающихся теоретическому анализу) [3], который свободен от серьезных вычислительных недостатков, присущих существующим диспетчеризациям, и превосходит их при всех допустимых значениях загрузки.

Работа выполнена при поддержке РФФИ (проект № 20-07-00804).

## Литература

1. Bell C. H., Stidham S. Individual versus social optimization in the allocation of customers to alternative servers // *Management Science*, 1983, vol. 29, no. 7, pp. 831-839.
2. Hordijk A., van der Laan D.A. Periodic routing to parallel queues and billiard sequences // *Mathematical Methods of Operations Research*, 2004, vol. 59, no. 2, pp. 173-192.
3. Konovalov M., Razumchik R. A simple dispatching policy for minimizing mean response time in non-observable queues with SRPT policy operating in parallel // *Communications of the ECMS*, 2020, vol. 34, no. 1, pp. 398-402.