

# LR-ЭВРИСТИКА ДЛЯ ЗАДАЧИ БАЛАНСИРОВКИ НАГРУЗКИ НА СЕРВЕРЫ

Н.А. Кочетова

*Институт математики им. С.Л. Соболева СО РАН,  
Новосибирский государственный университет, Новосибирск  
e-mail: nkochet@math.nsc.ru*

Рассматривается задача балансировки нагрузки на серверы, возникающей при организации облачного хостинга веб приложений [1]. Имеется набор серверов, на каждом из которых размещены диски (точнее, их образы). На дисках хранятся интернет сайты с разнородной информацией. Пользователи посещают сайты, что создает определенную нагрузку на серверы. Эта нагрузка меняется со временем и характеризуется несколькими параметрами: CPU, RAM и др. По каждому диску известна активность пользователей в течение планового периода. Эта активность позволяет определить нагрузку на каждый сервер в каждый момент времени по каждому параметру. Если нагрузка по каждому параметру не превосходит заданного порога, то сервер находится в рабочем режиме. В противном случае сервер работает с перегрузкой. Чтобы избежать перегрузки, диски можно перемещать с одного сервера на другой. Такое перемещение требует определенных вычислительных затрат. Будем называть их накладными расходами. Предполагается, что для каждого диска известны накладные расходы по каждому параметру при его изъятии с сервера и подключении к любому другому серверу. Начальное распределение дисков по серверам считается известным. Задача состоит в том, чтобы перераспределить диски по серверам и достичь минимальной суммарной перегрузки на всем плановом периоде при ограничениях на накладные расходы по каждому серверу.

Сформулированная задача может быть представлена в терминах частично-целочисленного линейного программирования, что позволяет использовать для ее решения коммерческое программное обеспечение. Однако из-за большого разрыва целочисленности найти точное решение удастся только на примерах небольшой размерности. В связи с этим предлагается приближенный алгоритм с апостериорной оценкой отклонения от оптимума. Исходная задача заменяется на соответствующую задачу линейного программирования. Ее точное решение позволяет зафиксировать часть дисков на серверах и сократить размерность задачи. Такой подход последовательно применяется до тех пор, пока число переменных не станет приемлемым для точного решения коммерческим пакетом CPLEX. Приводятся результаты вычислительных экспериментов для 20 серверов и 200 дисков.

## ЛИТЕРАТУРА

1. Кочетов Ю.А., Кочетова Н.А. Задача балансировки нагрузки на серверы // Вестник Новосибирского государственного университета. Серия Информационные технологии. 2013. Том 11, вып. 4. С. 71-76.