УДК 338. 32

ОБ ОДНОМ ПОДХОДЕ К РАСПРЕДЕЛЕНИЮ РЕСУРСОВ, МИНИМИЗИРУЮЩЕМ ОБЪЕМ ЗАЯВОК С ЗАДАННЫМ ВРЕМЕНЕМ ОЖИДАНИЯ

Халиков М.А., Максимов Д.А.

ФГБОУ ВПО «Российский экономический университет имени Г.В. Плеханова», Москва, Россия (117997 Российская Федерация, г. Москва, Стремянный пер., 36), e-mail: maksimovdenis@mail.ru

Рассматривается задача распределения ограниченных ресурсов при выполнении запланированных объемов затрат в открытых системах массового обслуживания, например, на участках (блоках) механообработки крупных машиностроительных производств. Показано, что при использовании критерия минимизации объема необработанных к концу планируемого периода заявок с фиксированным временем ожидания исходная задача сводится к задаче линейного программирования большой размерности, для которой актуальными являются исследования устойчивости оптимального решения по интенсивности входного потока заявок. В работе приводятся соответствующие определения и свойства устойчивых решений, исследуется устойчивость задачи оптимального распределения ограниченных ресурсов в зависимости от изменения исходных данных модели и интенсивности входного поступления заявок (количества узлов обслуживания), с целью минимизации времени нахождения заявок в системе.

Ключевые слова: открытые системы массового обслуживания, интенсивность входного потока заявок, распределение ресурсов, время ожидания обработки заявки, задача линейного программирования, устойчивость оптимального решения, свойства оптимальных решений.

ABOUT AN APPROACH TO THE RESOURCE ALLOCATION, MINIMIZING THE VOLUME OF APPLICATIONS WITH A SPECIFIED TIMEOUT

Khalikov M.A., Maksimov D.A.

Plekhanov Russian University of Economics, Moscow, Russia (117977 Russian Federation, Moscow, Stremjannyj pereulok, 36) e-mail: maksimovdenis@mail.ru

The problem of limited resources allocation for implementation of scheduled volume costs in open queuing networks is considered. For example, in the areas of machining process of large engineering industry. It is shown that by using the criterion of minimizing the volume of raw applications with a fixed latency at the end of the planning period, the original problem reduces to linear programming problem of high dimensionality, which relevant to the research of the stability of the optimal solution for the arrival rate of applications. In the research the relevant definitions and attributes of stable solutions are given, the stability problem of optimal resources allocation against the initial data model and the arrival rate of applications is researched in order to minimize the time spent by applications in the system.

Keywords: open queuing networks, arrival rate of applications, resource allocation, waiting time of application, linear programming problem, stability of optimal solution, attributes of optimal solutions.

Одной из актуальных задач для открытых систем массового обслуживания является распределение ограниченных ресурсов при выполнении запланированных объемов работ. Такие задачи встречаются, например, при обработке данных в АСУТП, где поступающая информация (сигналы от датчиков) должна быть обработана в режиме реального времени или при обработке партий деталей по заданной маршрутной технологии на участках взаимозаменяемого оборудования разной производительности.

Общим для этих задач являются ограниченность ресурсов, на которых обрабатываются заявки (сигналы, детали), и цель – минимизация времени нахождения заявок в системе (время ожидания плюс время обслуживания).

В качестве критерия оперативности обработки заявок в открытой системе обслуживания, как правило, рассматривается доля заявок, время ожидания которых не менее заданного τ [1] .

В настоящей работе исследуется задача распределения ресурсов по указанному критерию для случая, когда интенсивность поступления заявок на вход системы подчиняется детерминированному закону или описывается линейной кусочно-постоянной функцией. Для полученного распределения ресурсов вводится понятие устойчивости решения и анализируется устойчивость в зависимости от изменений интенсивности входного потока.

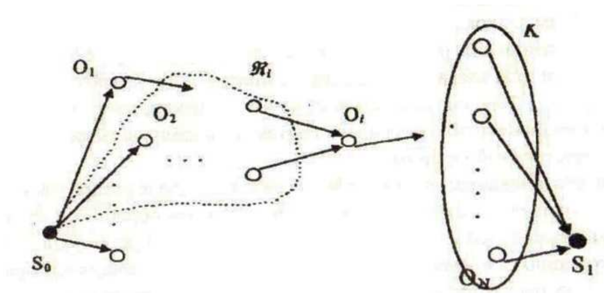

Рис. 1. Граф-структура технологии обработки заявок

Технологию обработки заявок в открытой системе будем представлять ориентированным ациклическим графом G= G(M,N) (рис.1), где:

N – вершины (различные операции обработки заявок);

M – дуги (возможные последовательности обработки заявок на заданном множестве операций) с выделенными вершинами;

S_0 – источник заявок;

S_1 – сток обработанных заявок;

K – множество конечных (завершающих) операций обработки;

$\Re_i$ – множество операций-предшественников для операции O_i.

Фиксируем следующие параметры:

T – установленный для исследуемой системы временной интервал (мин, час, сутки);

τ – (кратное T) — максимальное время пребывания заявки в очереди;

$\Im$ – (кратное T) — продолжительность периода, на котором распределяются ресурсы (директивный период).

Пусть интенсивность поступления заявок со временем ожидания не менее $\theta(\theta > 0)$ на операцию с номером i в интервале $[T(j-1), T(j)]$ задается величиной $u_{ij}^{\theta}(i-1,..., N; j=1,...\Im)$.

Очереди на каждой операции на начало директивного периода задаются величинами v_i^θ (объем очереди на операции i, со временем ожидания заявок не менее θ).

Выполнение обработки операций технологического цикла осуществляется на P единицах оборудования, производительности которых задаются матрицей (α_{ip}) размерности NxP, где α_{ip} – производительность p-го оборудования на i-й операции ($\alpha_{ip} = 0$ в случае, если i-я операция на p-м оборудовании не осуществляется).

Необходимо организовать обработку поступающих заявок таким образом, чтобы на конец периода продолжительностью $\Im$ минимизировать в очереди объем заявок со временем ожидания не менее τ.

Поиск оптимального распределения ресурсов будем осуществлять в предположении, что заявки с операции на операцию передаются мгновенно (без учета «транспортного» лага). Последнее условие накладывает существенные ограничения на исследуемую модель, не искажая, тем не менее, качественной картины.

В такой постановке задача может быть сведена к задаче распределения ресурсов по критерию максимизации общего объема обработанных за директивный период заявок. Для этого достаточно рассмотреть среди всех заявок, находящихся в очередях v_i^θ, и среди заявок, поступающих на обработку u_{ij}^θ. «Возраст» θ которых к окончанию периода [0, $\Im$] будет не менее τ.

Исходя из этого положения примем в качестве входного потока заявок следующий:

$$u_{ij} = \sum_{\theta \ge \tau} u_{ij}^\theta, \quad i = 1,...N; \quad j = 1,...,\Im. \tag{1}$$

Объемы очередей на операциях в начале директивного периода рассчитаем по формуле:

$$v_i = \sum_{\theta > \tau - \alpha} v_i^\theta. \tag{2}$$

Обозначим x_{ij}^p - часть j-го интервала, в течение которого p-е оборудование выполняет i-ю операцию.

Задача минимизации объемов необработанных заявок со временем ожидания к концу директивного периода не менее *m* может быть сформулирована как следующая оптимизационная (относительно x_{ij}^p) задача:

$$\sum_{j=1}^{\Im} \sum_{O \in K} \sum_{p=1}^{P} x_{ij}^p \alpha_{ip} \to \max, \tag{3}$$

при ограничениях

$$v_i + \sum_{i=1}^{j} u_{il} + \sum_{l=1}^{j-1} \sum_{k \in \Re_i} \sum_{p=1}^{P} x_{kl}^{p} \alpha_{kp} - \sum_{l=1}^{j} \sum_{p=1}^{P} x_{il}^{p} \alpha_{ip} \geq 0, \tag{4}$$

i=1,…,N; j=1,…, $\Im$ (условие баланса очередей);

$$\sum_{i=1}^{N} x_{ij}^{p} \leq 1; \quad x_{ij}^{p} \geq 0; \quad i = 1,..., N; \tag{5}$$

j=1,…, $\Im$; p=1,…,P (условие баланса оборудования).

Задача (3)—(5), учитывая линейность функционала и ограничений, относится к задачам линейного программирования (ЛП).

Рассмотрим, как изменяется значение функционала (3) при изменении интенсивности u_{ij} потока входных заявок.

Фиксируем целое ε, большее 0.

<u>Определение 1.</u> Задача (3)—(5) ε – устойчива по интенсивности поступления заявок u_{ij}, если для всех точек ε окрестности ($u_{ij}, u_{ij} \pm \varepsilon$) значение функционала (1) остается неизменным.

<u>Определение 2.</u> Максимальное ε, при котором задача (3)—(5) ε — устойчива по интенсивности поступления заявок, назовем радиусом устойчивости задачи по интенсивности входного потока заявок.

Пусть $\tilde{x}_{ij}^{p}$ – решение задачи (3)-(5) и значение функционала (3) при распределения $\tilde{x}_{ij}^{p}$ равно $\tilde{A}$. Определить диапазон изменения интенсивности входного потока заявок можно, решив следующие задачи ЛП относительно $\tilde{x}_{ij}^{p}$ и ε:

$$\max \varepsilon_1; \tag{6}$$

$$\tilde{A} \geq \sum_{j=1}^{\Im} \sum_{O \in K} \sum_{p=1}^{P} x_{ij}^{p} \alpha_{ip}; \tag{7}$$

$$v_i + \sum_{i=1}^{j} (u_{il} - \varepsilon_1) + \sum_{l=1}^{j-1} \sum_{k \in \Re_i} \sum_{p=1}^{P} x_{kl}^{p} \alpha_{kp} - \sum_{l=1}^{j} \sum_{p=1}^{P} x_{il}^{p} \alpha_{ip} \geq 0, \tag{8}$$

i=1,…,N; j=1,…, $\Im$;

$$\sum_{i=1}^{N} x_{ij}^{p} \leq 1; \quad x_{ij}^{p} \geq 0; \quad i = 1,..., N; \tag{9}$$

j=1,…, $\Im$; p=1,…,P.

$$\min \varepsilon_2; \tag{10}$$

$$\sum_{j=1}^{\Im}\sum_{O\in K}\sum_{p=1}^{P} x_{ij}^{p}\alpha_{ip} \geq \tilde{A}; \tag{11}$$

$$v_i + \sum_{i=1}^{j}(u_{il}+\varepsilon_1) + \sum_{l=1}^{j-1}\sum_{k\in\Re_i}\sum_{p=1}^{P} x_{kl}^{p}\alpha_{kp} - \sum_{l=1}^{j}\sum_{p=1}^{P} x_{il}^{p}\alpha_{ip} \geq 0, \tag{12}$$

i=1,…,N; j=1,…, $\Im$;

$$\sum_{i=1}^{N} x_{ij}^{p} \leq 1; \quad x_{ij}^{p} \geq 0; \quad i=1,...,N, \tag{13}$$

j=1,…, $\Im$; p=1,…,P.

Радиус ρ устойчивости исходной задачи (3)-(5) по интенсивности входного потока заявок определим по формуле:

$$\rho = \min(\varepsilon_1, \varepsilon_2).$$

Рассмотрим случай, когда интенсивности поступления заявок на обработку заданы кусочно-постоянными функциями. Разобьем директивный период на интервалы $[t_j, t_{j+1}), (j=0,...,L-1; t_0=0, t_1=\Im)$, на каждом из которых интенсивность u_{ij} поступления заявок на каждую операцию постоянна.

Задачу оптимального распределения ресурсов будем рассматривать для каждого интервала $[t_j, t_{j+1})$:

$$\sum_{j=0}^{L-1}\sum_{O\in K}\sum_{p=1}^{P} x_{ij}^{p}\alpha_{ip} \to \max; \tag{15}$$

$$v_{ij} + (t_{j+1}-t_j)u_{ij} + \sum_{k\in\Re_i}\sum_{p=1}^{P} x_{kj}^{p}\alpha_{kp} - \sum_{p=1}^{P} x_{ij}^{p}\alpha_{ip} \geq 0, \tag{16}$$

i=1,…,N; j=0,…,L-1;

$$\sum_{p=1}^{P} x_{ij}^{p} \leq (t_{j+1}-t_j); \quad x_{ij}^{p} \geq 0, \quad i=1,...,N, \tag{17}$$

j=0,…,L-1; p=1,…,P.

Примечание. x_{ij}^{p} – часть интервала $[t_j, t_{j+1})$, на котором р-е оборудование используется на i-й операции; v_{ij} – объем очереди на i-й операции в начале интервала $[t_j, t_{j+1})$.

В случае если производительности каждой из Р единиц оборудования на рассматриваемом интервале не превышают интенсивностей поступления заявок, решение задачи (15)—(17) дает оптимальное по производительности распределение оборудования.

Обратно, если производительность каждой единицы оборудования на рассматриваемом интервале $[t_j, t_{j+1})$ превышает интенсивность поступления заявок, то, разрешив кольцевую процедуру переключения оборудования с одной операции на другую [5], можно сохранить распределение соответствующей ЛП-задачи и обработанный объем заявок сделать сколь угодно близким к значению А функционала (13). Докажем этот факт. Выберем $A^* < \tilde{A}$. Для того, чтобы получить распределение оборудования по операциям на интервале $[t_j, t_{j+1})$, которое даст объем обработанных заявок A^*, поступим следующим образом. Разобьем интервал $[t_j, t_{j+1})$ на подинтервалы длиной $[\frac{A-A^*}{A}(t_j - t_{j+1})]$ и на каждом из подинтервалов сохраним распределение, полученное в задаче (15)—(17), сократив время работы каждой единицы оборудования в $\frac{A}{A-A^*}$ раз. В этом случае производительность оборудования на каждом из полученных подинтервалов (за исключением, быть мажет, первого) будет равна $\frac{A}{t_{j+1}-t_j}$, что и доказывает сформулированное утверждение.

Для задачи (15)—(17) можно рассматривать обобщенный критерий оперативности обработки поступающих заявок: насколько велико максимальное время ожидания заявок в системе. В этом случае для получения оптимального распределения может быть применена процедура итерационного решения задачи (15)—(17) при различных интенсивностях входного потока заявок, связанных с выбором τ времени ожидания, по которому производится максимизация обработанного объема заявок.

Список литературы

1. Бродецкий Г.Л., Гусев Д.А. Экономико-математические методы и модели в логистике. Процедуры оптимизации. – М.: Издательский центр «Академия», 2012. – 288 с.

2. Гордон М.П. Логистика товародвижения. – М.: «Центр экономики и маркетинга», 2009. – 195 с.

3. Линдерс М.Р., Фирон Х.Е. Управление снабжением и запасами. Логистика. – М.: Виктория-плюс, 2008. – 768 с.

4. Максимов Д.А., Халиков М.А. Методы оценки и стратегии обеспечения экономической безопасности предприятия. – М.: ЗАО «Гриф и К». 2012. – 220 с.

5. Халиков М.А. Моделирование производственной и инвестиционной стратегии машиностроительного предприятия. – М.: Благовест – В., 2003. – 304 с.

Рецензенты:

Загородников С.Н., д.б.н., профессор, профессор кафедры «Математические методы в экономике» ФГБОУ ВПО «Российский экономический университет имени Г.В. Плеханова», г. Москва;

Титов В.А., д.э.н., профессор кафедры «Информационные технологии» ФГБОУ ВПО «Российский экономический университет имени Г.В. Плеханова», г. Москва.