

А.С. КУИМОВА, аспирантка Волжской государственной академии водного транспорта, Нижний Новгород,

Ю.С. ФЕДОСЕНКО, д.т.н., проф., зав. каф. ИСУиТ Волжской государственной академии водного транспорта, Нижний Новгород

СИНТЕЗ ОГРАНИЧЕННЫХ ПО СТРУКТУРЕ ОПТИМАЛЬНО-КОМПРОМИССНЫХ СТРАТЕГИЙ ОБСЛУЖИВАНИЯ ПОТОКА ОБЪЕКТОВ

Вводится модель однофазного обслуживания детерминированного потока объектов. Формулируется бикритериальная задача синтеза стратегий обслуживания. Показано, что учет ограничений на структуру стратегий обслуживания выводит задачу из класса *NP*-трудных и позволяет построить полиномиальный алгоритм синтеза стратегий, оптимальных по Парето. Приводится пример. Библиогр.: 8 назв.

Ключевые слова: стратегии обслуживания, *NP*-трудность, оптимальность по Парето.

Постановка проблемы и анализ литературы. В работе [1] рассмотрена каноническая модель обслуживания потока объектов стационарным процессором, описывающая практически важную проблему оптимизации использования дискретных ресурсов. Вместе с тем, естественное требование адекватности математического описания реальным ситуациям, возникающим в технических, экономических и организационных системах, нередко предопределяет необходимость различных усложнений канонической модели. Так, в данной работе формулируется значимая для приложений (в частности, в автоматизированных производственных системах и на внутреннем водном транспорте [2]) модель, в которой стратегии обслуживания оцениваются по значениям не одного, а двух независимых критериев. При этом в качестве парадигмы решения сформулированной экстремальной задачи принята концепция Парето [3, 4], предполагающая выявление и предъявление лицу, принимающему решения, всего множества оптимально-компромиссных стратегий обслуживания. Данная оптимизационная задача относится к числу *NP*-трудных [5]. Поэтому актуальной является проблема построения таких модификаций бикритериальной модели обслуживания, которые при сохранении адекватности описания порождают полиномиально разрешимые подклассы задачи синтеза стратегий обслуживания. Одна из таких модификаций формализует ограничение на допустимую величину опережений в обслуживании.

Цель статьи – разработка бикритериальной модели однопроцессорного обслуживания потока объектов, порождающей

полиномиально разрешимую задачу синтеза оптимально-компромиссных стратегий, и построение решающего алгоритма.

Математическая модель и постановка задачи. Задан конечный детерминированный поток объектов $O_n = \{o_1, o_2, \dots, o_n\}$, которые подлежат однофазному обслуживанию на стационарном процессоре P . Для каждого объекта o_i ($i = \overline{1, n}$) определены целочисленные характеристики: t_i – момент поступления в очередь на обслуживание ($0 \leq t_1 \leq t_2 \leq \dots \leq t_n$); τ_i – норма длительности обслуживания; a_i – штраф за единицу времени простоя в ожидании обслуживания; d_i – мягкий директивный срок завершения обслуживания ($d_i \geq t_i + \tau_i$). Если обслуживание объекта o_i начинается в момент времени t_i^* ($t_i^* \geq t_i$), то величина индивидуального штрафа по этому объекту определяется значением линейной функции вида $a_i(t_i^* - t_i)$. Обслуживание объекта o_i ($i = \overline{1, n}$) может быть начато свободным процессором в любой момент времени t ($t \geq t_i$) и осуществляется без прерываний; одновременное обслуживание двух и более объектов и неоправданные простои процессора запрещены; процессор готов к обслуживанию объектов потока O_n в момент времени $t = 0$.

Очевидно, любая стратегия S обслуживания потока O_n представляет собой перестановку $\{i_1, i_2, \dots, i_n\}$ совокупности индексов $N = \{1, 2, \dots, n\}$; при этом объект с индексом i_k в стратегии S обслуживается k -м по очереди ($k = \overline{1, n}$). Введем следующее определение: стратегию S_q назовем q -стратегией, если не существует объекта, который при реализации стратегии S_q допускает пропуск вперед на обслуживание более q других объектов. Множество таких стратегий будем обозначать как $Z(q)$. Пусть $t^*(i_k, S_q)$ и $\bar{t}(i_k, S_q)$ соответственно моменты начала и завершения обслуживания объекта с индексом i_k при реализации стратегии S_q . Считаем, что реализация компактна, т.е. выполняются соотношения

$$t^*(i_k, S) = \max[\bar{t}(i_{k-1}, S_q), t_{i_k}], \quad k = \overline{2, n}; \quad t^*(i_1, S) = t_{i_1};$$

$$\bar{t}(i_k, S_q) = t^*(i_k, S_q) + \tau_{i_k}, \quad k = \overline{1, n}.$$

Качество каждой стратегии будем оценивать по значениям пары критериев $K_1(S_q)$ и $K_2(S_q)$, первый из которых представляет собой суммарный штраф за простои объектов в ожидании обслуживания, а второй – оценивает максимальное по продолжительности нарушение

директивного срока завершения обслуживания среди всех объектов потока O_n . С учетом введенных выше обозначений имеем

$$K_1(S_q) = \sum_{k=1}^n a_{i_k} (t^*(i_k, S_q) - t_{i_k}); \quad K_2(S_q) = \max_{1 \leq k \leq n} (\bar{t}(i_k, S_q) - d_{i_k}, 0).$$

Формализация парадигмы Парето на множестве $Z(q)$ приводит к следующей задаче выделения в плоскости $(K_1(S_q), K_2(S_q))$ полной совокупности эффективных точек и формирования им соответствующих оптимально-компромиссные стратегии обслуживания

$$\{ \min_{S_q \in Z(q)} (K_1(S_q), \min_{S_q \in Z(q)} (K_2(S_q))) \}. \quad (1)$$

Синтез стратегий обслуживания. Алгоритм решения задачи (1) построим на основе идеологии динамического программирования [6, 7]. С этой целью введём операцию \otimes между произвольным вектором $x = (x_1, x_2)$ и множеством Y векторов $y = (y_1, y_2)$ той же размерности: через $Y \otimes x$ обозначим совокупность всех векторов $v = (v_1, v_2)$, где первая компонента представима в виде $v_1 = y_1 + x_1$, а вторая определяется по правилу $v_2 = \max(y_2, x_2)$. Будем обозначать: для произвольного множества двумерных векторов-оценок M через $eff[M]$ максимальное по включению подмножество недоминируемых в M векторов; через $F(t)$ подмножество индексов объектов, которые поступают в систему обслуживания в момент времени t . При обслуживании объектов потока O_n , решения принимаются в те моменты времени, когда процессор свободен и необходимо выбрать следующий объект. Очевидно, что набор (t, j, M_j^t, Q) определяет текущее состояние системы в момент t принятия решения, где j – наименьший из индексов в множестве индексов необслуженных объектов, а M_j^t – множество индексов обслуженных объектов, обошедших на момент времени t объект с индексом j ($|M_j^t| \leq q$), Q – множество индексов объектов, ожидающих обслуживания в этот момент времени. Введем обозначения: $\{M\}_r$ – совокупность индексов объектов из множества M с индексами большими r ; $D(t, \Delta)$ – совокупность индексов объектов, прибывающих в систему на интервале $[t + 1, t + \Delta]$, $\Delta \geq 1$, т.е.

$$D(t, \Delta) = \bigcup_{i=1}^{\Delta} F(t + i).$$

В множество Q считается включенным индекс 0 фиктивного объекта с параметрами $t_0 = t$, $\tau_0 = \min\{\tau \mid D(t, \tau) \neq \emptyset\}$, $a_0 = 0$, $d_0 = \infty$. Выбор такого объекта на обслуживание означает простой процессора с момента времени t до момента поступления в систему очередного объекта. Если в

текущем состоянии фиктивный объект не был выбран на обслуживание, то он исключается из множества Q и в дальнейшем не рассматривается.

Введем оператор R раскрытия состояния системы обслуживания: если (t, j, M_j^t, Q) является некоторым произвольным состоянием, то $R(t, j, M_j^t, Q)$ – множество порождаемых им состояний, т.е. непосредственно следующих за состоянием (t, j, M_j^t, Q) . Последовательно раскрывая все состояния, начиная от начального, получим в итоге совокупность финальных состояний. Заметим, что одно и то же состояние может сформироваться при раскрытии различных состояний. Присвоим начальному состоянию номер 0 и последовательно пронумеруем все состояния в порядке их порождения оператором R . Для одного и того же состояния, если его можно получить несколькими способами, используется только один номер, присвоенный первоначально; каждое рассматриваемое состояние может характеризоваться несколькими оценками.

Пусть $W_q(t, j, M_j^t, Q)$ – задача, получаемая из исходной задачи (1) при условии, что обслуживание начинается в момент времени t ; при этом надо обслужить объекты с индексами из множества Q и все объекты, поступающие в систему позднее момента времени t . Через $E_q(t, j, M_j^t, Q)$ обозначим совокупность эффективных по Парето двумерных оценок, получаемую при решении задачи $W_q(t, j, M_j^t, Q)$. Тогда полная совокупность эффективных оценок задачи (1) будет представлять множество $E_q(0, 1, M_1^0, F(0))$.

Введем обозначения: Q_t – совокупность индексов всех необслуженных к моменту t объектов, т.е. $Q_t = Q \cup \{k : o_k \in O_n \text{ и } t_k > t\}$; $\xi[W]$ – минимальный из положительных индексов объектов множества W . Очевидно, что для любого $\theta \geq 0$, $j = \alpha$ и $Q = \{\alpha\}$, где α – произвольный индекс объекта из потока O_n , имеет место соотношение

$$E_\delta(t_n + \theta, \alpha, M_\alpha^{t_n + \theta}, \{\alpha\}) = (\alpha_\alpha(t_n + \theta - t_\alpha), \max(\bar{t}(\alpha, S_q) - d_\alpha, 0)). \quad (2)$$

Если в состоянии (t, j, M_j^t, Q) в качестве следующего обслуживаемого выбран объект с индексом $\alpha \in Q$ ($\alpha \neq j$) и при этом $|M_j^t| < q$, то очередным моментом принятия решения будет $t + \tau_\alpha$ и выбор индекса следующего объекта для обслуживания будет осуществляться из множества $(Q \setminus \{\alpha\}) \cup D(t, \tau_\alpha)$; при этом $M_j^t \cup \{\alpha\}$ – множество объектов, опередивших объект с индексом j . Если на обслуживание выбран объект с индексом j , то в следующий момент принятия решения $t + \tau_j$, наименьшим индексом объекта будет являться $j^* = \xi[Q \setminus \{j\}]$, множеством его опередивших объектов будет $\{M_j^t\}_{j^*}$, а множество ожидающих

обслуживания объектов определяется как $(Q \setminus \{j\}) \cup D(t, \tau_j)$. Следовательно, при $|M_j^t| < q$ имеет место соотношение

$$\begin{aligned}
 E_q(t, j, M_j^t, Q) &= \text{eff}[\text{eff}[\bigcup_{\alpha \in Q \setminus \{j\}} ((a_\alpha(t - t_\alpha), \max(\bar{t}(\alpha, S_q) - d_\alpha, 0)) \otimes \\
 &\otimes E_q(t + \tau_\alpha, j, M_j^t \cup \{\alpha\}, Q \setminus \{\alpha\} \cup D(t, \tau_\alpha))] \cup ((a_j(t - t_j), \max(\bar{t}(j, S_q) - d_j, 0)) \otimes E_q(t + \tau_j, \xi[Q_t \setminus \{j\}], \{M_j^t\}_{\xi[Q_t \setminus \{j\}]}, Q \setminus \{j\} \cup D(t, \tau_j))]. \quad (3)
 \end{aligned}$$

Если в состоянии (t, j, M_j^t, Q) имеет место равенство $|M_j^t| = q$, то следующим можно обслуживать лишь объект с индексом j . Поэтому

$$\begin{aligned}
 E_q(t, j, M_j^t, Q) &= \text{eff}[(a_j(t - t_j), \max(\bar{t}(j, S) - d_j, 0)) \otimes \\
 &\otimes E_q(t + \tau_j, \xi[Q_t \setminus \{j\}], \{M_j^t\}_{\xi[Q_t \setminus \{j\}]}, Q \setminus \{j\} \cup D(t, \tau_j)]. \quad (4)
 \end{aligned}$$

Выражения (2) – (4) суть решающие рекуррентные соотношения. Непосредственным анализом [8] устанавливается, что вычислительная сложность описанного алгоритма оценивается величиной $O(n^4)$.

Результаты вычислительных исследований и выводы. Целью вычислительных экспериментов являлось определение скоростных и объёмных характеристик разработанного алгоритма. Для значений параметров потока O_n были установлены следующие диапазоны изменения: $\overline{t_{i-1}} \leq t_i \leq \overline{t_{i-1}} + 5$, $1 \leq a_i \leq 11$, $1 \leq \tau_i \leq 15$, $t_i + \tau_i \leq d_i \leq t_i + \tau_i + 4$, $i = \overline{1, n}$. Таким образом была обеспечена достаточно высокая плотность потока и гарантировалось получение заведомо пессимистических оценок характеристик алгоритма. Размерность потока изменялась от 6 до 13 с единичным шагом; аналогично параметр q принимал значения от 1 до $\lfloor n/2 \rfloor$. В соответствующих узлах решетки (n, q) генерировались по равномерному распределению данные для серии из 100 частных задач типа (1).

В результате вычислительных исследований на компьютере с процессором Intel Core 2 Duo, 3,16 ГГц (ОЗУ 2 Гб) установлено, что продолжительность синтеза полной совокупности эффективных оценок и соответствующих им оптимально-компромиссных стратегий обслуживания не превышает 68 с, объем требуемой оперативной памяти не превышает 1,5 Гб, что позволяет рекомендовать разработанный алгоритм для включения в автоматизированные системы диспетчеризации процессов рассмотренного типа.

Список литературы: 1. Коган Д.И. Задача диспетчеризации: анализ вычислительной сложности и полиномиально разрешимые подклассы / Д.И. Коган, Ю.С. Федосенко // Дискретная математика. – 1996. – Т. 8. – Вып. 3. – С. 135-147. 2. Коган Д.И. Проблема синтеза оптимального расписания обслуживания бинарного потока объектов mobile-процессором / Д.И. Коган, Ю.С. Федосенко, А.В. Шеянов / Труды III Международной конференции "Дискретные модели в теории управляющих систем", Москва, 1998. – М.: Изд-во МГУ им. М.В. Ломоносова. 1998. – С. 43-46. 3. Подиновский В.В. Парето-оптимальные решения многокритериальных задач / В.В. Подиновский, В.Д. Ногин. – М.: Физматлит, 2007. – 256 с. 4. Лотов В.А. Многокритериальные задачи принятия решений / В.А. Лотов, И.И. Поспелова. – М.: МАКС Пресс, 2008. – 197 с. 5. Гэри М. Вычислительные машины и труднорешаемые задачи / М. Гэри, Д. Джонсон. – М.: Мир, 1982. – 416 с. 6. Беллман Р. Прикладные задачи динамического программирования / Р. Беллман, С. Дрейфус. – М.: Наука, 1965. – 457 с. 7. Коган Д.И. Динамическое программирование и дискретная многокритериальная оптимизация / Д.И. Коган. – Н. Новгород: Изд-во ННГУ, 2005. – 260 с. 8. Кормен Т. Алгоритмы: построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест, К. Штайн. – М.: Вильямс, 2007. – 1296 с.

УДК 681.31+519.8

Синтез обмежених за структурою оптимально-компрмісних стратегій обслуговування потоку об'єктів / Федосенко Ю.С., Куїмова А.С. // Вісник НТУ "ХПІ". Тематичний випуск: Інформатика і моделювання. – Харків: НТУ "ХПІ". – 2011. – № 17. – С. 70 – 75.

Вводиться модель однофазного обслуговування детермінованого потоку об'єктів. Формулюється бікритеріальна завдання синтезу стратегій обслуговування. Показано, що врахування обмежень на структуру стратегій обслуговування виводить завдання з класу NP-важких та дозволяє побудувати поліноміальний алгоритм синтезу стратегій обслуговування, оптимальних за Парето. Наводиться приклад. Бібліогр.: 8 назв.

Ключові слова: стратегії обслуговування, NP-важкий, оптимальність за Парето.

UDC 681.31+519.8

The synthesis of structure-bounded optimal compromise service policies of object flow / Fedosenko Yu.S., Kuimova A.S. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2011. – №. 17. – P. 70 – 75.

In this paper a model of single-phase service of deterministic objects flow is discussed. A bicriteria problem of service policies synthesis is formulated. It is shown that if restrictions on service policies structure are imposed the problem will become non NP-hard and it will be possible to design a polynomial algorithm for Pareto-optimal service policies synthesis. Example is introduced. Refs.: 8 titles.

Key words: service policies, NP-hardness, Pareto optimality.

Поступила в редакцію 15.02.2011