

Методы повышения эффективности клинической деятельности лечебно-диагностического учреждения на основе оптимизации совокупности медицинских технологических процессов

К. С. Яковлев¹

В работе описывается эвристический алгоритм составления расписаний для множества одновременно выполняющихся медицинских технологических процессов. Приводятся результаты модельных экспериментов.

Введение

Деятельность любого крупного лечебного учреждения характеризуется одновременным параллельным протеканием значительного числа медицинских технологических процессов, использующих общие материальные, клинические, кадровые и иные ресурсы. В качестве примеров ресурсов подобного рода можно рассматривать лабораторное оборудование, медикаменты, специализированное диагностическое оборудование, обслуживающий персонал, коечный фонд и пр.

Анализ ретроспективной прецедентной информации о лечебно-диагностических процессах (ЛДП), хранящейся в медицинских базах данных, свидетельствует о том, что для каждого процесса при определенных условиях существует возможность замены или перестановки лечебно-диагностических мероприятий и ряда иных модификаций. Это означает, что существует возможность оптимизации всей совокупности параллельно выполняющихся ЛДП, приводящей к разрешению коллизий и конфликтных ситуаций, что в свою очередь влечет снижение временных издержек, увеличение эффективности использования оборудования, оборачиваемость коечного фонда и, в конечном итоге, — сокращение сроков пребывания па-

¹ 117312, г. Москва, пр-т 60-летия Октября, 9, Институт Системного Анализа РАН, yakovlev@isa.ru

циентов в стационаре и снижению затрат на каждый из лечебно-диагностических процессов.

1. Формализация понятия лечебно-диагностического процесса

Для формализации понятия лечебно-диагностического процесса введем понятия медицинского технологического процесса (МТП).

Медицинский технологический процесс — это система взаимосвязанных, необходимых и достаточных, научно обоснованных лечебно-диагностических мероприятий, выполнение которых позволяет наиболее рациональным образом провести лечение и обеспечить достижение максимального соответствия научно-прогнозируемых результатов реальным при минимизации затрат [Назаренко и др., 2005].

В самом общем смысле можно рассматривать МТП как некоторую совокупность образующих элементов, а именно — лечебных мероприятий. Абстракцией активной компоненты лечебных мероприятий, приводящей к изменению состояния пациента, являются операторы [Назаренко и др., 2006]. С каждым оператором связывают моменты начала и окончания, периодичность, объем мероприятий и их локализацию.

Применимость операторов может зависеть от выполнения определенных условий, а порядок выполнения операторов — меняться. То есть, уместно говорить о различных маршрутах в МТП. Всего выделяют пять видов маршрутизации МТП: последовательная, параллельная, условная, итеративная и конкурентная маршрутизация [Назаренко и др., 2006]. Графическое представление этих видов маршрутизации приводится на рисунке 1.

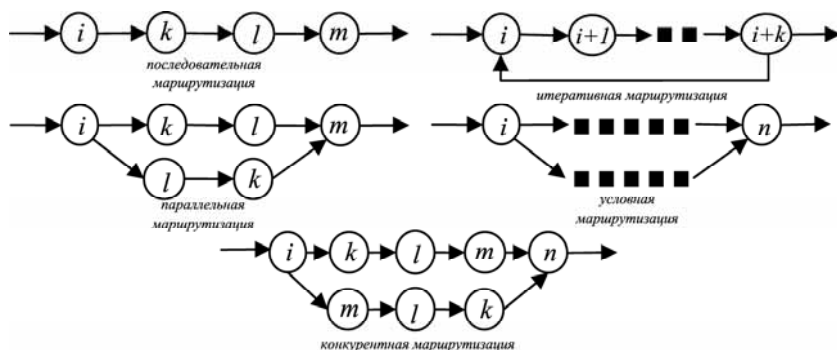


Рис. 1. Графическое представление различных типов маршрутизации МТП

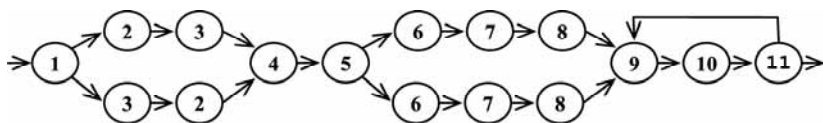


Рис. 2. Представление медицинского технологического процесса в виде ориентированного графа

В общем случае каждый МТП может состоять из совокупности операторов с комбинированной маршрутизацией, т. е. можно говорить об отношении частичного порядка на множестве операторов, которое удобно представлять в виде ориентированного графа, см. рисунок 2.

2. Модель совокупности МТП

Для построения модели совокупности параллельно выполняющихся МТП будем использовать методологию теории расписаний (ТР) — науки, исследующей задачи, в которых необходимо упорядочить или, другими словами, определить последовательность выполнения совокупности работ, использования каких-либо средств, ресурсов и так далее [Коффман, 1984; Танаев и др., 1975]. Предметом ТР являются системы, корректное функционирование которых зависит не только от правильности результатов вычислений, но и от моментов времени к которым данные результаты были получены, что актуально для рассматриваемой предметной области. Будем называть подобные системы — системами реального времени (СРВ), по аналогии с западной литературой [Koch, 1999].

СРВ для совокупности МТП имеет вид:

$$\Lambda = \langle \mathbf{M}, \mathbf{П}, \mathbf{O}, \mathbf{R}, \mathbf{t} \rangle, \quad (*)$$

где:

М — число МТП в системе;

П — совокупность МТП;

О — множество всевозможных операторов МТП (лечебно-диагностических мероприятий);

Р — множество ресурсов;

t — дискретное время.

$$\mathbf{R} = \{R_0, \dots, R_K\}, K < \infty.$$

Дискретное время **t** представляет собой последовательность моментов $t : 0, 1, \dots, \infty$. Момент $t = 0$ — начало отсчёта, $t = \infty$ — конец наблюдения.

Каждый МТП из множества Π представляет собой совокупность лечебных операторов и отношение частичного порядка на этой совокупности:

$$\Pi^{(i)} = \langle \mathbf{O}^{(i)}, \rho^{(i)} \rangle.$$

Здесь:

$$\mathbf{O}^{(i)} = \left\{ \left[O_j^i \right] \right\} = \left\{ O : O \in \mathbf{O}, O \in \Pi^{(i)} \right\}, \quad i = 1, \dots, \mathbf{M}, \quad j = 1, \dots, N^i, \quad O_j^i \text{ — } j\text{-й}$$

оператор i -го МТП, N^i — количество операторов i -го МТП.

$\rho^{(i)}$ — отношение частичного порядка на множестве $\mathbf{O}^{(i)}$, допускающее параллельную, итеративную, условную и конкурентную маршрутизацию.

Каждый оператор из множества \mathbf{O} представляет собой четверку:

$$O = \langle \tau, \bar{r}, w, d \rangle,$$

где:

$\tau = \text{const}$ — время выполнения. Измеряется в дискретных единицах времени \mathbf{t} .

$r = (r_0, r_1, \dots, r_K), r_k = 0, 1$ — вектор ресурсов. Если $r_k = 1$, то для выполнения оператора O требуется ресурс $R_k \in \mathbf{R}$. «0» на k -м месте означает, что для оператора ресурс $R_k \in \mathbf{R}$ не требуется.

$w \in (0, +\infty)$ — удельная стоимость выполнения оператора O . Абсолютная стоимость полагается равной произведению удельной стоимости на продолжительность оператора: $w_{\text{abs}} = w \cdot \tau$.

d — директивный срок. Число единиц времени по истечению, которых оператор должен быть выполнен. Измеряется в дискретных единицах времени \mathbf{t} с начала отсчёта.

Расписанием для системы (*) является упорядоченная последовательность моментов начала выполнения операторов из множества \mathbf{O} :

$$\mathbf{Sh} = \left\{ \left[\begin{array}{c} (i) \\ sh_l^j \end{array} \right] \in \mathbf{t} : l = 0, 1, \dots, L; sh_0 \leq sh_1 \leq \dots \leq sh_L; 1 \leq i \leq \mathbf{M}, 1 \leq j \leq N^i \right\}.$$

Частичное расписание для системы (*) к моменту времени $t \in \mathbf{t}$ — это расписание вида:

$$\mathbf{Sh}(t) = \left\{ sh_l^{(i,j)} \in \mathbf{t} : l = 0, 1, \dots, L(t); sh_0 \leq sh_1 \leq \dots \leq sh_L = t \right\}.$$

Стоимостью расписания \mathbf{Sh} будем называть значение функции $Cost(\mathbf{Sh})$, определяемой следующим образом:

$$Cost(\mathbf{Sh}) = SCost(\mathbf{Sh}) + DCost(\mathbf{Sh}).$$

Здесь:

$$SCost(\mathbf{Sh}) = \sum_{i,j:\exists sh^{(i,j)} \in \mathbf{Sh}} \tau_j^i \cdot w_j^i;$$

$$DCost(\mathbf{Sh}) = \sum_{sh^{(j_{n+1})} \neq sh^{(j_n)} + \tau_{j_n}^i} (sh^{(j_{n+1})} - (sh^{(j_n)} + \tau_{j_n}^i)) * \hat{w};$$

τ_j^i — время выполнения оператора O_j^i расписания \mathbf{Sh} ;

w_j^i — удельная стоимость выполнения оператора O_j^i расписания \mathbf{Sh} ;

\hat{w} — удельная стоимость простоя СРВ.

Задача построения оптимального расписания состоит в таком упорядочивании операторов из множества \mathbf{O} , чтобы:

1. Все операторы были выполнены в срок.
2. Соблюдался порядок выполнения операторов.
3. Конфликты относительно ресурсов были полностью исключены.
4. Стоимость расписания была минимизирована (функция $Cost(\mathbf{Sh})$ достигала минимума).

Расписание, удовлетворяющее требованиям 1), 2), 3) будем называть *допустимым*.

Расписание, удовлетворяющее требованиям 1), 2), 3) и 4) будем называть *оптимальным*.

3. Интерпретация модели

Поставленная выше задача построения оптимального расписания для СРВ (*) имеет простую трактовку в медицинской предметной области. Стоимость оператора СРВ — некоторый интегральный экономический показатель затрат, связанных с выполнением соответствующего лечебно-диагностического мероприятия; время выполнения оператора — время необходимое для проведения лечебно-диагностического мероприятия; директивный срок — максимально допустимый срок, по истечению которого данное мероприятие должны быть завершено (задается экспертом). Стоимость расписания для совокупности МТП — сумма всех затрат связанных с лечением пациентов, которая складывается из двух слагаемых, первое — затраты на проведение лечебных мероприятий, второе — затраты связанные с простоем оборудования и низкой оборачиваемостью койко-мест. Оба этих слагаемых могут рассматриваться независимо друг от друга, что дает возможность более тонкой оптимизации работы лечебно-диагностического учреждения.

4. Эвристический алгоритм составления расписания

Процесс составления расписания для операторов из множества \mathbf{O} будем рассматривать как процесс поиска в пространстве состояний [Audsley, 1991; Parnas *et al.*, 1990]. Это пространство представляет собой дерево. Корень дерева соответствует пустому расписанию, лист — конечному расписанию, промежуточный узел дерева (не корень и не лист) — частично-му расписанию. Задача состоит в том, чтобы найти такой лист дерева, который соответствует оптимальному расписанию.

Для того чтобы не производить полный перебор в пространстве состояний, предлагается использовать эвристический алгоритм составления расписания, который состоит в итерационном расширении частичного расписания до полного с помощью эвристик. Эвристика — это оценивающая функция вида: $H : \mathbf{O} \rightarrow \mathfrak{R}$, которая применяется на каждой итерации алгоритма ко всем операторам, которые еще не добавлены в расписание. Расписание расширяется тем оператором, который минимизирует значение эвристики. Этот оператор считается «лучше» других. Добавляя «лучшие» операторы, в итоге получаем «лучшее» расписание.

Для сокращения пространства поиска на каждой итерации алгоритма будем использовать следующий механизм отсекаания. После добавления

очередного оператора в расписание, получившееся частичное расписание $\mathbf{Sh}(t)$ проверяется на строгую допустимость. Частичное расписание называется *строго допустимым*, если все расписания, полученные из него путем добавления любого из ещё не добавленных операторов из O , являются допустимыми. Если в ходе работы алгоритма выясняется, что полученное на данном шаге частичное расписание $\mathbf{Sh}(t)$ не является строго допустимым, то происходит возврат назад. А именно — из расписания $\mathbf{Sh}(t)$ выкидывается оператор, добавленный последним (осуществляется переход в дереве пространства состояний на одну вершину вверх), и в полученное заново расписание добавляется другой оператор. И так — до тех пор, пока не будет получено строго допустимое расписание. В общем (худшем) случае потребуются столько возвратов, сколько есть вершин на текущем уровне дерева. Если перебраны все возможные продолжения, а строго допустимого расписания не получено, то происходит возврат еще на один уровень выше. И так далее до тех пор, пока или не будет получено требуемое частичное строго допустимое расписание, или поиск не завершится возвратом к корню дерева — пустому расписанию, что будет означать невыполнимость задачи при данных условиях.

В качестве эвристик предлагается использовать следующие функции:

Эвристика	Вид функции H
Минимальное время выполнения ($Min_ \tau$)	$H(O) = \tau$
Максимальное время выполнения ($Min_ \tau$)	$H(O) = -\tau$
Минимальный директивный срок (Min_D)	$H(O) = d$
Минимальное время готовности (Min_S)	$H(O) = EST$
Минимальная «растяжимость» (Min_L)	$H(O) = d - (EST + \tau)$

Эвристика Min_S задается следующим образом. Пусть на i -й итерации алгоритма получено частичное расписание $\mathbf{Sh}(t_{cur})$, $t_{cur} \in \mathbf{t}$ — момент начала выполнения последнего оператора из $\mathbf{Sh}(t_{cur})$. Для данного частичного расписания $\mathbf{Sh}(t_{cur})$ и момента времени $t_{cur} \in \mathbf{t}$ определим вектор готовности ресурсов (EAT -вектор, от Earliest Available Time) следующим образом:

$$\begin{aligned}
 & EAT\left(\mathbf{Sh}\left(t_{cur}\right)\right) = EAT = \\
 & = \left(EAT\left(\mathbf{Sh}\left(t_{cur}\right)_0\right), EAT\left(\mathbf{Sh}\left(t_{cur}\right)_1\right), \dots, EAT\left(\mathbf{Sh}\left(t_{cur}\right)_K\right)\right) = \\
 & = \left(EAT_0, EAT_1, \dots, EAT_K\right)
 \end{aligned}$$

EAT_k — это момент времени $t \in \mathbf{t}$, $t \geq t_{cur}$, по достижению которого ресурс R_k становится свободным для операторов системы.

Для частичного расписания $\mathbf{Sh}\left(t_{cur}\right)$ и момента времени $t_{cur} \in \mathbf{t}$ определим время готовности оператора $O_j^i \in \mathbf{O}$ к выполнению как:

$EST = \max\{EAT_k\}$, k -я компонента вектора ресурсов оператора \bar{r}_j^i равна 1.

```

ЭвристическоеПланирование(ЧастичноеРасписание Sh; var СтрогоДопустимо Feasible)
Begin
     $\tilde{\mathbf{O}} := \text{GetOperations}(\mathbf{Sh})$  1
    If  $\tilde{\mathbf{O}} = \emptyset$  then begin 2
        Result:= $\emptyset$  3
        Exit 4
    End
    If NOT(StronglyFeasible(Sh,  $\tilde{\mathbf{O}}$ )) then begin 5
        Feasible:= false 6
        Result:=  $\emptyset$  7
        Exit 7
    End Else 8
        Feasible:= true 8
    While ( $\tilde{\mathbf{O}} \neq \emptyset$ ) begin 9
        Вычислить  $H(0)$  для всех операторов из  $\tilde{\mathbf{O}}$  10
        Пусть  $O_j^i$  — оператор, минимизирующая  $H$  11
         $\tilde{\mathbf{O}} := \tilde{\mathbf{O}} - O_j^i$  12
        Result:=  $EST_j^i + \text{ЭвристическоеПланирование}(\mathbf{Sh} + EST_j^i, \mathbf{Feasible})$  13
        If Feasible = true then Exit 14
        Result:=  $\emptyset$  15
    End
End

```

Схема 1. Эвристический алгоритм составления расписаний

Описанный алгоритм составления расписания для совокупности МТП имеет вид, показанный на схеме 1.

5. Эксперимент

5.1. Описание

Для анализа пригодности эвристического алгоритма для решения задачи оптимизации лечебных планов было сгенерировано 200 экземпляров систем типа (*), имитирующих работу лечебно-диагностического учреждения, каждая из которых насчитывала 8 операторов, разбитых на МТП случайным образом. Каждому оператору для выполнения требовалось от 1 до 7 общедоступных ресурсов. Вероятность того, что оператору потребуется i -й ресурс ($i = 1, \dots, 7$), равнялась $\frac{1}{2}$.

Генерация данных 200 экземпляров систем типа (*) производилась таким образом, чтобы гарантировать существование оптимального расписания для каждого из них. Моменты окончания выполнения операторов в расписании $(SC_i, i = 1, \dots, 8)$ есть минимально возможные директивные сроки. После того как они получены, реальный директивный срок для i -го оператора устанавливался случайно из интервала $[SC_i, (1 + R) * SC_i]$, где $R \geq 0$ — настраиваемый параметр. Чем меньше R , тем «эффективней» должен работать планировщик¹, чтобы построить допустимое расписание. Если $R = 0$, то построенное расписание (если таковое получено в результате работы планировщика) необходимо совпадает с оптимальным для данного набора операторов. Еще одним настроечным параметром эксперимента являлось число возвратов, необходимых в случае составления алгоритмом ошибочных частичных расписаний — BQ .

В качестве оценки качества работы алгоритма использовалась SR -метрика (от англ. Success Ration), определяемая как процент составленных допустимых расписаний при заданных условиях от общего числа попыток.

5.2. Результаты

А таблице 1 приведены результаты работы алгоритма при $BQ = 0$ (не допускаются возвраты). Это — важный частный случай, т. к. при отсут-

¹ Под планировщиком здесь понимается алгоритм составления расписания (от англ. — scheduler).

Таблица 1

Результаты работы алгоритма без возвратов

Эвристика	R=0.0	R=0.3	R=0.6
<i>Min</i> _τ	0.0%	14.0%	24.5%
<i>Max</i> _τ	0.0%	3.5%	7.0%
<i>Min</i> _D	31.5%	68.5%	91.5%
<i>Min</i> _S	0.0%	13.0%	23.0%
<i>Min</i> _L	13.5%	62.5%	87.5%

Таблица 2

Комбинированные эвристики на основе *Min*_D

Эвристика	Вид функции <i>H</i>
<i>Min</i> _D + <i>Min</i> _τ	$H(O) = d + \alpha \cdot \tau$
<i>Min</i> _D + <i>Max</i> _τ	$H(O) = d - \alpha \cdot \tau$
<i>Min</i> _D + <i>Min</i> _s	$H(O) = d + \alpha \cdot EST$

Таблица 3

Производительность комбинированных эвристик
в задаче составления расписания

Эвристика	R=0.0		R=0.3		R=0.6	
	α	SR	α	SR	α	SR
<i>Min</i> _D + <i>Min</i> _τ	0.0	31.5%	0.0	68.5%	0.0	91.5%
<i>Min</i> _D + <i>Max</i> _τ	0.0	31.5%	0.0	68.5%	0.0	91.5%
<i>Min</i> _D + <i>Min</i> _s	20.0	76.5%	12.0	95.5%	6.0	99.0%

вии возвратов добавление в расписание «не правильно» оцененного эвристикой оператора необходимо влечет невыполнение задачи.

Полученные данные свидетельствуют о том, что к наилучшим результатам приводит использование эвристики *Min*_D. Однако при малых значениях параметра *R* результаты работы алгоритма нельзя считать удовлетворительными. Для исправления этой ситуации на базе эвристики *Min*_D были построены комбинированные эвристики, приведенные в таблице 2, и эксперимент был повторен.

Как видно из таблицы, эвристика *Min*_D является основной, а *Min*_τ, *Max*_τ и *Min*_S — дополнительными с весовым коэффициентом $\alpha, \alpha \geq 0$. Конкретное значение весового коэффициента α подбиралось, исходя из

предположения о том, что эвристика, как линейная функция от α , достигает максимума своей производительности (с точки зрения SR -метрики) в единственной точке. Поэтому, начиная с $\alpha = 0$ и с шагом 0,5 для каждой комбинированной эвристики повторялся эксперимент и измерялся процент успеха, как только процент успеха на очередной итерации эксперимента оказывался хуже, чем на предыдущей — повторения прекращались и значение α фиксировалось. Производительность эвристик именно с этим значением α приведена в таблице 3.

Поставленной цели — значительного увеличения производительность эвристики Min_D при малых значениях R — удалось достичь, путем ее комбинации с Min_S .

Эксперимент был повторен еще раз. Теперь были разрешены возвраты назад (параметр BQ). В качестве эвристики была выбрана комбинированная эвристика $Min_D + Min_s$. Параметру α было присвоено значение 20. Результаты приведены ниже.

Таблица 4

Результаты работы эвристического алгоритма с возвратами

BQ	$R=0.0$	$R=0.3$	$R=0.6$
0	76.5%	95.5%	99.0%
1	77.0%	96.5%	99.0%
2	77.0%	96.5%	99.0%
4	80.5%	96.5%	99.5%
8	83.5%	97.5%	100.0%
16	87.5%	97.5%	100.0%
32	93.5%	99.0%	100.0%
64	98.0%	99.5%	100.0%

Полученные результаты позволяют говорить о том, что при использовании возвратов эвристика $Min_D + Min_s$ позволяет выбирать не только «лучший» оператор на очередной итерации, но также правильно ранжировать все остальные, на случай возврата, и, таким образом, производительность алгоритма в отдельных случаях может достигать 100 процентов.

Проведенные эксперименты показывают высокую эффективность эвристического алгоритма составления расписаний для задач оптимизации множества одновременно выполняющихся лечебно-диагностических процессов.

Литература

1. *Коффман Э. Г.* Теория расписаний и вычислительные машины. М.: Наука, 1984.
2. *Назаренко Г. И., Осипов Г. С.* Основы теории медицинских технологических процессов. Ч. 1. М.: Физматлит, 2005.
3. *Назаренко Г. И., Осипов Г. С.* Основы теории медицинских технологических процессов. Ч. 2. М.: Физматлит, 2006.
4. *Танаев В. С., Шкурба В. В.* Введение в теорию расписаний. М.: Наука, 1975.
5. *Audsley N. C.* Resource Control For Hard Real-Time Systems // University of York, UK, 1991.
6. *Koch B.* The Theory Of Task Scheduling In Real-Time Systems // Universität Hamburg Fachbereich Informatik, 1999.
7. *Parnas D. L., Xu J.* Scheduling Processes With Release Times, Deadlines, Precedence, and Exclusion Relations // IEEE Transactions On Software Engineering. Vol. 16. No. 3. March 1990, 1990.