

Кластеризация состояний объекта при решении задачи выбора оптимальных технологических режимов

А.И. Якимов, Е.М. Борчик, Е.М. Максимов

Аннотация. Предложен метод выбора оптимальных технологических режимов на промышленном предприятии. Для устранения неоднозначности выбора технологических режимов на каждом из этапов технологического процесса поставлена задача многокритериальной оптимизации с использованием принципа оптимальности Беллмана и лексикографическим упорядочением критериев оптимизации по стоимости ресурсов и/или временных затрат для выпуска заданного количества продукции. Предложена методика применения кластерного анализа для оценки состояний технологического процесса, близких в соответствии с заданной мерой близости к оптимальному состоянию.

Ключевые слова: технологический процесс, граф, состояние, кластерный анализ, многокритериальная оптимизация.

Введение

Рассматривается задача многокритериального управления технологическим процессом (ТП) производства готовой продукции. Каждый из этапов технологического процесса допускает выбор одного из допустимых технологических режимов (ТР) обработки продукции. Каждый ТР представляет процесс обработки определенного вида продукции на соответствующем технологическом оборудовании. Себестоимости отдельных ТР складываются из стоимостей используемых ресурсов (например, химикатов, красителей, энергоресурсов и др.), необходимых для обработки единицы готовой продукции на используемом технологическом оборудовании. Пусть исследуемый ТП допускает возможность представления в виде взвешенного ориентированного ациклического графа [1], вершинами которого являются технологические режимы. Нагрузки (стоимости ресурсов, временные затраты) на дуги графа в общем

случае многомерны и имеют стохастический характер.

Для устранения неоднозначности выбора технологических режимов на каждом из этапов ТП сформулирована задача многокритериальной оптимизации с использованием принципа оптимальности Беллмана, где принято лексикографическое упорядочение критериев оптимизации. Например, для минимизации стоимости S ресурсов и/или временных затрат T при выпуске заданного количества продукции имеем:

01) $S \rightarrow \min$, 02) $T \rightarrow \min$ или
01) $T \rightarrow \min$, 02) $S \rightarrow \min$ [2].

Метод многокритериальной оптимизации с использованием принципа оптимальности Беллмана применим для решения практических многокритериальных задач, сводимых к задаче определения кратчайшего/критического пути, или минимальной/максимальной стоимости расхода ресурсов, и/или времени, соответственно, в случае, когда приходится одновре-



менно учитывать несколько критериев, которые должны принимать экстремальные значения.

ТП является управляемой системой, которая может характеризоваться одним из нескольких состояний, изменяемых на каждом этапе ТП в результате управляющего воздействия (управления). Под *состоянием* будем понимать состав технологического оборудования, совокупность используемых ресурсов, определяющих временные и стоимостные затраты на каждом из этапов технологического процесса. Под *управлением* будем понимать выбор состояния на каждом из этапов и последовательности состояний в целом для технологического процесса.

Постановка задачи

Пусть нагрузки (стоимости ресурсов, временные затраты) на дуги графа, в общем случае стохастического характера, уточнены методом Монте-Карло с использованием имитационной модели технологического процесса [3] и/или посредством анализа статистических данных ТП.

При этом для каждого i -го этапа ($i=1, \dots, n$; n - общее количество этапов) определены нормы расхода времени $T_i = (t_{i1}, \dots, t_{im_i})$, $t_{ij} \in R$ ($j=1, \dots, m_i$, m_i - количество допустимых технологических режимов на i -м этапе), затрачиваемого на обработку заданного количества продукции; общие стоимости расхода ресурсов, необходимых для обработки заданного количества продукции $S_i = (s_{i1}, \dots, s_{im_i})$, $s_{ij} \in R$, $i=1, \dots, n$, $j=1, \dots, m_i$, $n, m_i \in N$.

Стоимости $s_{ij} \in S_i$ расхода ресурсов $R(i, j)$ могут рассматриваться как общие стоимости расхода ресурсов $R(i, j)$ либо как стоимости расхода заданных групп ресурсов, например, химических материалов (химикаты, красители), энергоресурсов (пар, газ, электроэнергия).

На практике не всегда удается реализовать оптимальное решение, полученное посредством теоретических расчетов, поэтому необходимо провести поиск рациональных решений, приемлемых для применения в соответствии с условиями регламента данного ТП. Для устранения неоднозначности выбора технологических режимов на каждом из этапов ТП постав-

лена задача многокритериальной оптимизации по стоимости ресурсов и временных затрат для выпуска заданного количества продукции: 01) $S \rightarrow \min$, 02) $T \rightarrow \min$. Требуется определить множество состояний ТП, «близких», в соответствии с заданной мерой близости, к оптимальному состоянию для расширения области применения результатов оптимизации распределения ресурсов и получения максимального экономического эффекта, в зависимости от реальных условий протекания технологического процесса.

Методика решения задачи построения кластеров состояний

Для решения поставленной задачи предлагается процедура анализа выборок данных возможных состояний ТП. При этом общее количество всех возможных последовательностей состояний по графу определяется сочетанием технологических режимов на каждом этапе и составляет [4]:

$$\eta = C_{m_1}^1 \cdot C_{m_2}^1 \cdot \dots \cdot C_{m_n}^1 = m_1 \cdot m_2 \cdot \dots \cdot m_n = \prod_{i=1}^n m_i, \quad (1)$$

где $C_{m_i}^1$ - число сочетаний из m_i технологических режимов, по одному на i -м этапе, $i=1, \dots, n$.

Пусть расход ресурсов на i -м этапе для каждого состояния ТП представлен матрицами вида:

$$X = [x_{ij}], \quad x_{ij} \in R, \quad i=1, \dots, n, \quad j=1, \dots, \max\{m_1, \dots, m_n\}; \quad (2)$$

$$X_T = \psi(\{T_i | i=1, \dots, n\}) = [x_{ij}], \quad x_{ij} = \begin{cases} t_{ij} & | j=1, \dots, m_i, \\ 0 & | j=(m_i+1), \dots, \max\{m_1, \dots, m_n\}, \end{cases} \quad (3)$$

$$X_S = \psi(\{S_i | i=1, \dots, n\}) = [x_{ij}], \quad x_{ij} = \begin{cases} s_{ij} & | j=1, \dots, m_i, \\ 0 & | j=(m_i+1), \dots, \max\{m_1, \dots, m_n\}. \end{cases} \quad (4)$$

Этапы решения задачи

Этап 1. Построение многомерных выборок данных для состояний ТП.

Введем в рассмотрение функции $Z_i : X \rightarrow R$, $i \in \{1, \dots, n\}$ построения возможных состояний ТП на соответствующих i -х этапах:

на первом этапе

$$Z_1(X, j_1) = x_{1j_1}, \quad j_1 \in \{1, \dots, m_1\}; \quad (5)$$

на втором этапе

$$Z_2(X, j_1, j_2) = x_{1j_1} + x_{2j_2}, \quad j_2 \in \{1, \dots, m_2\}, \quad (6)$$

...

на i -м этапе

$$Z_i(X, j_1, j_2, \dots, j_i) = \sum_{k=1}^i x_{kj_k}, \quad (7)$$

$$j_k \in \{1, \dots, m_k\}, \quad k=1, \dots, i, \quad i \in \{1, \dots, n\}$$

Тогда последовательности данных для возможных состояний определяются соотношением вида:

$$y_{li} = (i, Z_i(X, j_1, j_2, \dots, j_i)) \in R^2, \quad (8)$$

$$i=1, \dots, n, \quad j_i=1, \dots, m_i,$$

$$l=1, \dots, \eta, \quad \eta = \prod_{i=1}^n m_i;$$

1) вид одномерных выборок, возможных на n -м этапе состояний ТП:

$$Y = \{y_1, y_2, \dots, y_\eta\}, \quad y_l \in R, \quad l=1, \dots, \eta, \quad \eta = \prod_{i=1}^n m_i, \quad (9)$$

при этом стоимостные затраты (критерий оптимизации 01) $S \rightarrow \min$):

$$y_{Sl} = Z_n(X_S, j_1, j_2, \dots, j_n), \quad j_k=1, \dots, m_k, \quad (10)$$

$$k=1, \dots, n, \quad l=1, \dots, \eta,$$

временные затраты (критерий оптимизации 01) $T \rightarrow \min$):

$$y_{Tl} = Z_n(X_T, j_1, j_2, \dots, j_n), \quad j_k=1, \dots, m_k, \quad (11)$$

$$k=1, \dots, n, \quad l=1, \dots, \eta;$$

2) вид двумерных выборок, возможных на n -м этапе состояний ТП (критерий оптимизации 01) $S \rightarrow \min$, 02) $T \rightarrow \min$):

$$Y = \left\{ \left(y_1^{(1)}, y_1^{(2)} \right), \left(y_2^{(1)}, y_2^{(2)} \right), \dots, \left(y_\eta^{(1)}, y_\eta^{(2)} \right) \right\}, \quad (12)$$

$$\left(y_l^{(1)}, y_l^{(2)} \right) \in R^2,$$

$$l=1, \dots, \eta, \quad \eta = \prod_{i=1}^n m_i,$$

$$\left(y_l^{(1)}, y_l^{(2)} \right) = \quad (13)$$

$$= \left(Z_n(X_T, j_1, j_2, \dots, j_n), Z_n(X_S, j_1, j_2, \dots, j_n) \right),$$

$$j_k=1, \dots, m_k, \quad k=1, \dots, n, \quad l=1, \dots, \eta;$$

3) вид m -мерных выборок всех возможных на n -м этапе состояний ТП определяется с учётом следующего равенства:

$$S = S_{G_{r_1}} + S_{G_{r_2}} + \dots + S_{G_{r_{(m-1)}}}, \quad (14)$$

где $S_{G_{r_k}} \in R$ – стоимость затрат некоторой k -й, $k=1, \dots, (m-1)$ группы ресурсов (например: химикатов, либо энергоресурсов (электроэнергии, пара, газа)), входящей в состав общей стоимости затрат ресурсов $S \in R$ (критерий оптимизации 01) $S_{G_{r_1}} \rightarrow \min$, 02) $S_{G_{r_2}} \rightarrow \min, \dots, m-1$

$$S_{G_{r_{(m-1)}}} \rightarrow \min, m) T \rightarrow \min):$$

$$Y = \left\{ \left(y_1^{(1)}, y_1^{(2)}, \dots, y_1^{(m)} \right), \left(y_2^{(1)}, y_2^{(2)}, \dots, y_2^{(m)} \right), \dots, \left(y_\eta^{(1)}, y_\eta^{(2)}, \dots, y_\eta^{(m)} \right) \right\}, \quad (15)$$

$$\left(y_l^{(1)}, y_l^{(2)}, \dots, y_l^{(m)} \right) \in R^m, \quad l=1, \dots, \eta, \quad \eta = \prod_{i=1}^n m_i;$$

m -мерный вектор с учетом временных затрат

$$\left(y_l^{(1)}, y_l^{(2)}, \dots, y_l^{(m)} \right) =$$

$$= \left(Z_n(X_T, j_1, j_2, \dots, j_n), Z_n(X_{S_{G_{r_1}}}, j_1, j_2, \dots, j_n) \right),$$

$$Z_n(X_{S_{G_{r_2}}}, j_1, j_2, \dots, j_n), \dots, Z_n(X_{S_{G_{r_{(m-1)}}}}, j_1, j_2, \dots, j_n) \right) \in R^m, \quad (16)$$

$$j_k=1, \dots, m_k, \quad k=1, \dots, n, \quad l=1, \dots, \eta.$$

Этап 2. Кластерный анализ многомерных данных.

Требуется проанализировать множество многомерных данных вида (15) ((9), (12), соответственно); разделить (15) на непересекающиеся подмножества (кластеры); определить при этом состояния ТП, наиболее «близкие» $\left(y_\alpha^{(1)}, y_\alpha^{(2)}, \dots, y_\alpha^{(m)} \right) \in R^m$ и наименее «близкие» $\left(y_\omega^{(1)}, y_\omega^{(2)}, \dots, y_\omega^{(m)} \right) \in R^m$, $\alpha, \omega \in \{1, \dots, \eta\}$ к оптимальному (критерий оптимизации, например, 01) $S \rightarrow \min$, 02) $T \rightarrow \min$) и в соответствии с худшим критерием, например, 01) $S \rightarrow \max$; 02) $T \rightarrow \max$.



Для решения данной задачи могут быть использованы методы кластерного анализа K-Means, Tree Clustering, Fuzzy Relation Clustering [5]. Пусть в результате разбиения (15) на кластеры элементам $(y_l^{(1)}, y_l^{(2)}, \dots, y_l^{(m)}) \in R^m$, $l = 1, \dots, \eta$, поставлены в соответствие номера кластеров K_ζ , $\zeta = 1, \dots, |\zeta|$, где $|\zeta|$ – итоговое количество построенных кластеров.

Искомыми являются кластеры K_1, K_2 , в состав которых входят элементы $(y_\alpha^{(1)}, y_\alpha^{(2)}, \dots, y_\alpha^{(m)}) \in K_1 \subset R^m$, $(y_\omega^{(1)}, y_\omega^{(2)}, \dots, y_\omega^{(m)}) \in K_2 \subset R^m$, соответственно.

В случае $|\zeta| > 2$ предлагается объединение элементов, не входящих в состав кластеров K_1, K_2 в некоторый обобщающий кластер K_3 , как прочие элементы (Рис. 1).

Этап 3. Дополнительный анализ структур построенных кластеров.

Этап 3.1. Выявление факторов, оказывающих наиболее существенное влияние на попадание объектов в соответствующие кластеры.

Каждому элементу множества (15) соответствует информация о последовательности состояний ТП – последовательность порядковых номеров вершин $M_B(i, k), i = 1, \dots, n, k = 1, \dots, m_i$ графа (Рис. 2)

$$(M_B(1, j_1), M_B(2, j_2), \dots, M_B(n, j_n)),$$

$$j_i \in \{1, \dots, m_i\}, i = 1, \dots, n, \tag{17}$$

$$M_B(i, k) = \sum_{j=1}^{i-1} m_j + k, i = 1, \dots, n, k \in \{1, \dots, m_i\},$$

$$M_B(i) \stackrel{def}{=} \left\{ \sum_{j=1}^{i-1} m_j + 1, \dots, \sum_{j=1}^i m_j \right\}, i = 1, \dots, n.$$

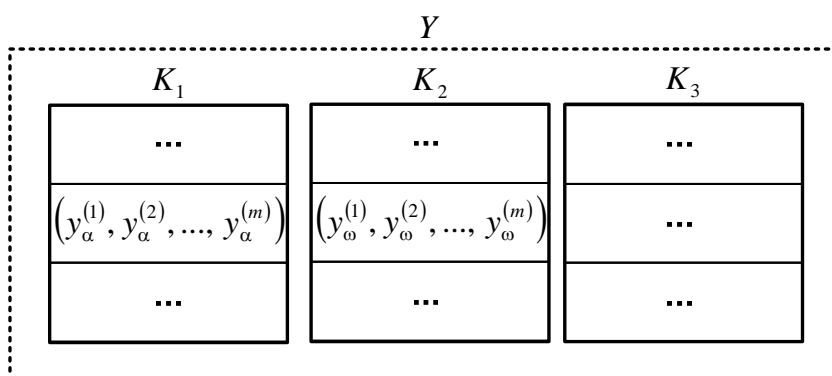


Рис. 1. Разделение исследуемой выборки на кластеры

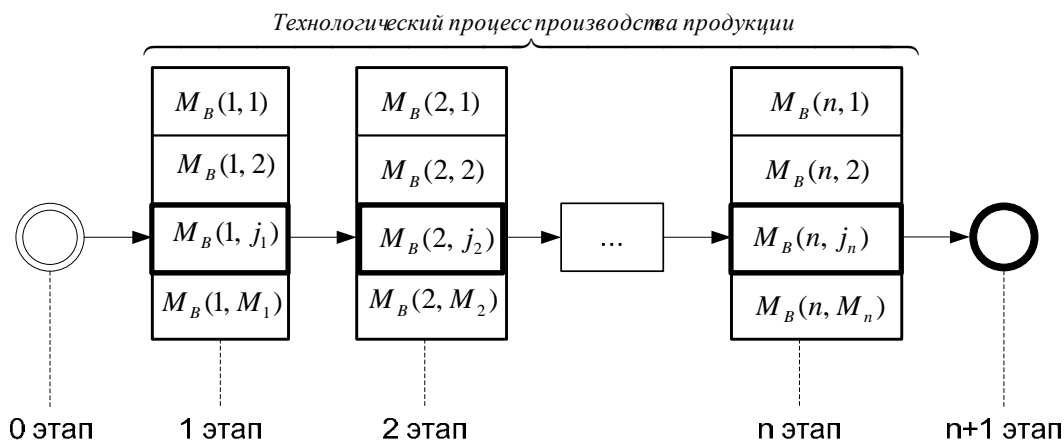


Рис. 2. Последовательность состояний технологического процесса

Предлагается методика качественного анализа кластеров по информации (17), который позволяет подробнее исследовать элементы кластеров, проанализировать состав кластеров, определить их особенности и характеристики. Методика включает следующие шаги:

Шаг 1. Выбор объектов исследования.

В качестве объектов исследования предлагается рассмотрение одиночных независимых состояний (вершин графа ТП с номерами u , $u \in M_B(i)$, $i = 1, \dots, n$). Поскольку граф ТП является связным, допустимо построение любого набора n исследуемых состояний вида (17).

Шаг 2. Построение матрицы вероятностей принадлежности объектов кластерам. Вероятности принадлежности объектов $u \in M_B(i)$, $i = 1, \dots, n$ кластерам рассчитываются как отношение количества случаев попадания объекта в кластер к общему количеству измерений, произведенных над объектом.

Шаг 3. Определение кластеров объектов. Принадлежность объекта кластеру определяется в соответствии со следующим критерием: объект u , $u \in M_B(i)$, $i = 1, \dots, n$, принадлежит кластеру K_{j_0} , $j_0 \in 1, \dots, \zeta'$, $\zeta' \in \{2, 3\}$, тогда и только тогда, когда вероятность принадлежности объекта кластеру в u -й строке матрицы вероятностей принадлежности объектов кластерам максимальна.

При этом интерес представляют номера вершин графа ТП (номера ТР, тип технологического оборудования, соответственно) с наибольшей долей вероятности обеспечивающие попадание итоговых состояний ТП в кластер состояний, наиболее близких к оптимальному $(y_\alpha^{(1)}, y_\alpha^{(2)}, \dots, y_\alpha^{(m)}) \in K_1$.

Этап 3.2. Определение оценочных характеристик для кластеров.

Шаг 1. Выбор оценочных характеристик кластеров.

Для выражения в сжатой форме наиболее существенных особенностей (и возможности последующего качественного анализа) распределений данных в пределах построенных кластеров K_ζ , $\zeta \in 1, \dots, \zeta'$, $\zeta' \in \{2, 3\}$, достаточно указания лишь некоторых их характерных черт

или числовых характеристик, в частности, предлагается определение:

1) математических ожиданий $\Theta_1(K_\zeta)$ выборок значений по каждой k -й компоненте, $k = 1, \dots, m$, векторов состояний ТП $(y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(m)})$, составляющих кластер K_ζ :

$$\Theta_1(K_\zeta) = (\Theta m_\zeta^{(1)}, \Theta m_\zeta^{(2)}, \dots, \Theta m_\zeta^{(m)}) \in R^m, \quad (18)$$

$$\zeta = 1, \dots, \zeta', \quad \zeta' \in \{2, 3\},$$

2) средних квадратических отклонений $\Theta_2(K_\zeta)$ выборок значений по каждой k -й компоненте, $k = 1, \dots, m$, векторов состояний ТП $(y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(m)})$, составляющих кластер K_ζ :

$$\Theta_2(K_\zeta) = (\Theta \sigma_\zeta^{(1)}, \Theta \sigma_\zeta^{(2)}, \dots, \Theta \sigma_\zeta^{(m)}) \in R^m, \quad (19)$$

$$\zeta = 1, \dots, \zeta', \quad \zeta' \in \{2, 3\};$$

3) в соответствии с евклидовой метрикой обобщенных характеристик $\Theta_3(K_\zeta)$ кластеров K_ζ , $\zeta \in 1, \dots, \zeta'$, $\zeta' \in \{2, 3\}$, в зависимости от соответствующих оценок математических ожиданий выборок значений по каждой из k -х компонент, $k = 1, \dots, m$ векторов состояний ТП $(y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(m)})$, составляющих кластер K_ζ :

$$\Theta_3(K_\zeta) = \sqrt{(\Theta m_\zeta^{(1)})^2 + (\Theta m_\zeta^{(2)})^2 + \dots + (\Theta m_\zeta^{(m)})^2} \in R, \quad (20)$$

$$\zeta \in \{1, \dots, 3\}.$$

Шаг 2. Покомпонентное построение и анализ статистических моделей распределения данных об итоговых состояниях ТП внутри кластеров.

Для каждого из построенных на этапе 2 кластеров K_ζ , $\zeta = 1, \dots, \zeta'$, $\zeta' \in \{2, 3\}$, требуется проанализировать многомерные множества вида (15) ((9), (12), соответственно) составляющие данные кластеры. При этом для выборок данных по каждой из k -х компонент, $k = 1, \dots, m$ векторов состояний $(y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(m)})$ ТП, образующих рассматриваемый кластер, необходимо построить статистическую модель распределения выборочных данных (кривую),



наилучшим образом описывающую данную выборку.

Задачу построения статистической модели, в частности, можно решить с использованием семейств универсальных статистических моделей распределений (Пирсона, Джонсона и др. [6, 7]), особенность которых состоит в возможности аппроксимации лишь одномодальных и *U*-образных [7] распределений. При этом для распределений, существенно отличных от нормального, семейства универсальных распределений, по сравнению с множеством «известных» законов распределений [6], обеспечивают большее разнообразие необходимых форм при меньшем общем количестве функций плотностей распределений, составляющих данные семейства [8].

Пусть для каждой из *k*-х компонент $y_i^{(k)}$, $k = 1, \dots, m$ векторов состояний ТП построены статистические модели распределения выборочных данных – функции плотности распределения семейства Пирсона (включающего 7 основных типов и 3 частных случая распределений [9]):

$$f_k(y_i^{(k)}) = \begin{cases} \left(|K_1|/\eta \right) \cdot f_{K_1}^{(k)}(y_i^{(k)}) \mid (y_i^{(1)}, \dots, y_i^{(k)}, \dots, y_i^{(m)}) \in K_1; \\ \dots \\ \left(|K_\zeta|/\eta \right) \cdot f_{K_\zeta}^{(k)}(y_i^{(k)}) \mid (y_i^{(1)}, \dots, y_i^{(k)}, \dots, y_i^{(m)}) \in K_\zeta, \end{cases}$$

$$k = 1, \dots, m, \quad \zeta = 1, \dots, \zeta', \quad \zeta' \in \{2, 3\}, \quad l \in \{1, \dots, \eta\},$$

$$\eta = \prod_{i=1}^n m_i, \tag{21}$$

где $(|K_\zeta|/\eta)$ – нормирующие множители для функций $f_{K_\zeta}^{(k)}(y_i^{(k)})$ плотностей распределения данных каждого из кластеров K_ζ , при этом $|K_\zeta| \geq 30$.

Тогда вероятность [4] достижения некоторого прогнозируемого итогового состояния ТП по соответствующей *k*-й компоненте $y_i^{(k)}$ векторов состояний ТП, $k = 1, \dots, m$ (при $m = 2$: $y^{(1)}$ – время, $y^{(2)}$ – общая стоимость ресурсов ТП) – попадания $y_i^{(k)}$ в заданный промежуток $[\alpha, \beta]$ – равняется площади криволинейной трапеции, ограниченной сверху графиком функции $f_k(y_i^{(k)})$ вида (21), справа и слева – прямыми $x = \alpha$, $x = \beta$ ($\alpha, \beta \in R$):

$$P(\alpha \leq y_i^{(k)} \leq \beta) = \int_\alpha^\beta f_k(y_i^{(k)}) d(y_i^{(k)}), \tag{22}$$

$$l \in \{1, \dots, \eta\}, \quad \eta = \prod_{i=1}^n m_i.$$

Результаты экспериментов. Технологический процесс производства набивной ткани

Рассматривается технологический процесс (ТП) производства готовой продукции (набивной ткани для промышленного применения) [10], представленный в виде взвешенного ориентированного ациклического графа, вершинами которого являются технологические режимы (ТР) (Рис. 3).

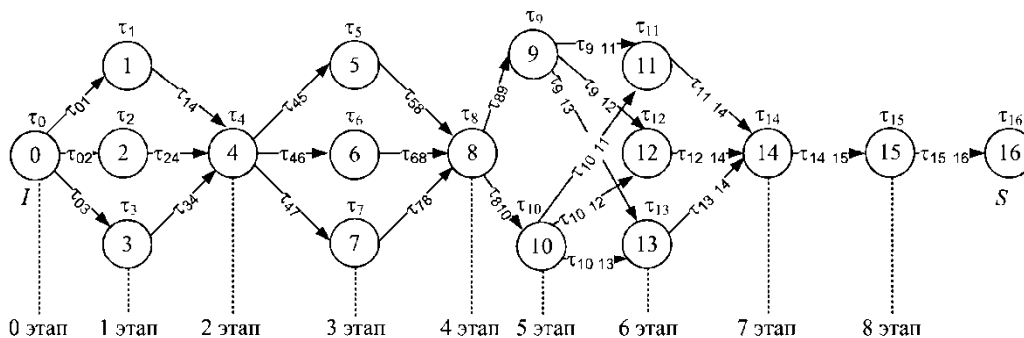


Рис. 3. Ориентированный граф с технологическими режимами

Процесс производства выбранного артикула ткани состоит из $n = 8$ технологических операций (этапов) $Op_i, i = 1, \dots, 8$: Op_1 – отварка; Op_2 – стабилизация; Op_3 – печать на ткани; Op_4 – зреление; Op_5 – промывка; Op_6 – вид отделки «стабилизация»; Op_7 – вид отделки «придание наполненного гифа»; Op_8 – браковка ткани.

При этом каждая из технологических операций допускает выбор одного из следующего количества допустимых технологических режимов – 1 этап: $m_1 = 3$; 2 этап: $m_2 = 1$; 3 этап: $m_3 = 3$; 4 этап: $m_4 = 1$; 5 этап: $m_5 = 2$; 6 этап: $m_6 = 3$; 7 этап: $m_7 = 1$; 8 этап: $m_8 = 1$.

Себестоимости отдельных ТР складываются из стоимостей используемых ресурсов: химикатов, красителей, энергоресурсов (пара, газа, электроэнергии).

Пусть нагрузки на дуги графа ТП (временные затраты; стоимости ресурсов) известны и достоверны. Требуется определить состояния, «близкие» к оптимальному.

Для решения поставленной задачи применяется процедура анализа выборок данных для возможных состояний ТП. При этом общее количество всех возможных последовательностей состояний по графу ТП составляет

$$\eta = \prod_{i=1}^n m_i = 3 \cdot 3 \cdot 2 \cdot 3 = 54.$$

Этапы построения кластеров состояний

Этап 1. Построение многомерных выборок данных для состояний ТП.

Одномерные ($m = 1$) выборки всех возможных на n -м ($n = 8$) этапе состояний ТП строятся по принципу (10). Двумерные ($m = 2$) – по принципу (13).

Пусть общая стоимость ресурсов S состоит из стоимости химикатов S_{Xim} и стоимости энергоресурсов S_{EnR} :

$$S = S_{Xim} + S_{EnR}.$$

Критерий оптимизации:

01) $S_{Xim} \rightarrow \min$, 02) $S_{EnR} \rightarrow \min$, 03) $T \rightarrow \min$.

Худшее состояние определяется критерием: 01) $S_{Xim} \rightarrow \max$, 02) $S_{EnR} \rightarrow \max$, 03) $T \rightarrow \max$.

На Рис. 4 представлены последовательности состояний (случай стоимостей общего расхода ресурсов и составляющих его отдельных групп ресурсов) на этапах ТП.

Трехмерные ($m = 3$) – выборки всех возможных на n -м ($n = 8$) этапе состояний $(y_l^{(1)}, y_l^{(2)}, y_l^{(3)})$, $l \in 1, \dots, \eta$ могут быть построены по принципу (16):

$$\begin{aligned} &(y_l^{(1)}, y_l^{(2)}, y_l^{(3)}) = \\ &= (Z_n(X_T, j_1, j_2, \dots, j_n), Z_n(X_{S_{EnR}}, j_1, j_2, \dots, j_n), \\ &Z_n(X_{S_{Xim}}, j_1, j_2, \dots, j_n)) \in R^3, \\ &j_k = 1, \dots, m_k, k = 1, \dots, n. \end{aligned}$$

В том случае, если стоимость энергоресурсов S_{EnR} складывается из стоимостей электроэнергии S_{ElEn} , пара S_{Par} и газа S_{Gaz} :

$$\begin{aligned} S &= S_{Xim} + S_{EnR} = [S_{EnR} = S_{ElEn} + S_{Par} + S_{Gaz}] = \\ &= S_{Xim} + S_{ElEn} + S_{Par} + S_{Gaz} \end{aligned}$$

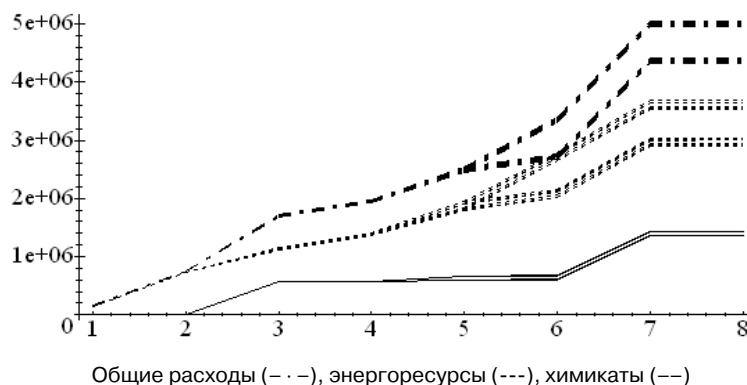


Рис. 4. Последовательность состояний по «стоимостям расхода ресурсов»



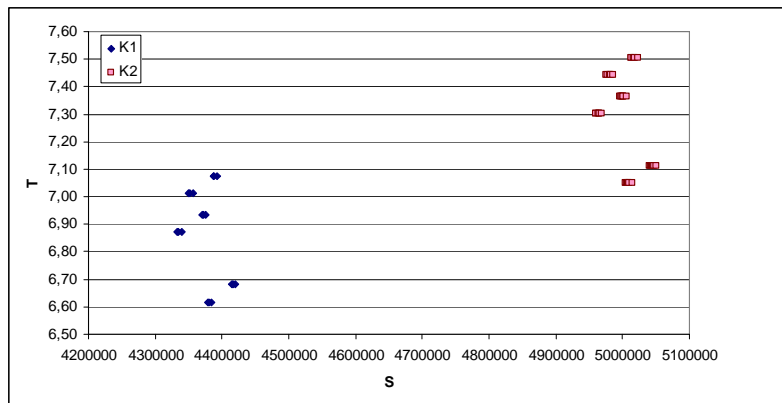


Рис. 5. Выборка всех возможных двумерных состояний на n -м этапе при критерии оптимизации $01)S \rightarrow \min, 02)T \rightarrow \min$

последовательности состояний для соответствующих групп ресурсов на этапах ТП строятся аналогично по принципу (8).

При этом пятимерные выборки состояний $(y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(5)})$, $l=1, \dots, \eta$ построены по принципу (16):

$$\begin{aligned} & (y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(5)}) = \\ & = (Z_n(X_T, j_1, j_2, \dots, j_n), Z_n(X_{SEEn}, j_1, j_2, \dots, j_n), \\ & Z_n(X_{SPar}, j_1, j_2, \dots, j_n), Z_n(X_{SGaz}, j_1, j_2, \dots, j_n), \\ & Z_n(X_{Sxim}, j_1, j_2, \dots, j_n)) \in R^5, \\ & j_k = 1, \dots, m_k, k = 1, \dots, n, n = 8. \end{aligned}$$

Этап 2. Кластерный анализ многомерных данных.

На этапе 1 анализа построены многомерные выборки всех возможных на n -м этапе состояний ТП при обработке 1000 м.п. ткани в соответствии с заданными сочетаниями технологических режимов:

1) двумерная выборка (Рис. 5): $(y_i^{(1)}, y_i^{(2)})$ вида (13), где $y_i^{(1)}$ – итоги по времени, $y_i^{(2)}$ – итоговые стоимости расхода ресурсов, $l=1, \dots, 54$;

2) трёхмерная выборка: $(y_i^{(1)}, y_i^{(2)}, y_i^{(3)})$ вида (16), $l=1, \dots, 54$, где $y_i^{(1)}$ – итоги по времени, $y_i^{(2)}$ – итоговые стоимости расхода

энергоресурсов, $y_i^{(3)}$ – итоговые стоимости расхода химикатов;

3) пятимерная выборка: $(y_i^{(1)}, y_i^{(2)}, y_i^{(3)}, y_i^{(4)}, y_i^{(5)})$ вида (16), $l=1, \dots, 54$, где $y_i^{(1)}$ – итоги по времени, $y_i^{(2)}$ – итоговые стоимости расхода электроэнергии, $y_i^{(3)}$ – итоговые стоимости расхода пара, $y_i^{(4)}$ – итоговые стоимости расхода газа, $y_i^{(5)}$ – итоговые стоимости расхода химикатов ТП.

Требуется проанализировать рассматриваемые многомерные выборки данных, разделить множества данных на непересекающиеся подмножества (кластеры); определить при этом состояния ТП, наиболее «близкие» $(y_\alpha^{(1)}, y_\alpha^{(2)}, \dots, y_\alpha^{(m)}) \in R^m$ и наименее «близкие» $(y_\omega^{(1)}, y_\omega^{(2)}, \dots, y_\omega^{(m)}) \in R^m$, $\alpha, \omega \in \{1, \dots, 54\}$ к оптимальному состоянию.

Результаты применения методов K-Means и Tree Clustering для рассматриваемых множеств эквивалентны. В рассматриваемом примере результаты кластеризации для двумерного, трёхмерного и пятимерного случаев состояний ТП идентичны. В каждом из случаев построено по 2 кластера: K_1, K_2 , где элементы K_1 – состояния ТП, наиболее близкие к оптимальному, элементы K_2 – наиболее удаленные от оптимального. Применения 3-го метода для уточнения результатов кластеризации не требуется.



Этап 3. Дополнительный анализ структур построенных кластеров.

Этап 3.1. Выявление факторов, оказывающих наиболее существенное влияние на попадание объектов в соответствующие кластеры.

Шаг 1. Выбор объектов исследования.

Каждому элементу множества (15) соответствует информация о последовательности состояний ТП – последовательность порядковых номеров вершин графа вида (17). В качестве объектов исследования предлагается рассмотрение одиночных независимых ТР, или вершин графа ТП (Рис. 1) с номерами u , $u = 1, \dots, 15$ (0-я и 16-я вершины графа ТП – исток, сток, соответственно – не учитываются).

Шаг 2. Значения элементов матрицы P вероятностей принадлежности объектов $u \in 1, \dots, 15$ (определённых на шаге 1.) кластерам K_1, K_2 рассчитываются в соответствии с классическим определением вероятностей, как отношение количества случаев попадания объекта в кластер к общему количеству измерений, произведенных над объектом.

$$P^T = P_1^T = P_2^T = \begin{bmatrix} 0,33_{1 \times 10} & 0 & 0 & 1 & 0,33_{1 \times 2} \\ 0,67_{1 \times 10} & 1 & 1 & 0 & 0,67_{1 \times 2} \end{bmatrix}.$$

Шаг 3. Определение принадлежности объектов кластерам в соответствии с предложенным выше критерием.

Рассматриваемые множества многомерных данных разбиваются на 2 кластера. При этом с вероятностью 100% технологический режим, соответствующий вершине № 13 графа ТП принадлежит кластеру K_1 – кластер состояний ТП, наиболее близких к оптимальному.

С вероятностью 100% ТР, соответствующие вершинам № 11, 12 графа ТП, а так же – с вероятностью 67% ТР, соответствующие вершинам № 1-10, 14, 15 графа ТП, принадлежат одному и тому же кластеру K_2 – кластеру состояний, наиболее удаленных от оптимального.

Все сочетания ТР исследуемого ТП, включающие ТР, соответствующий вершине графа № 13, обеспечивают в конечном итоге обработку 1000 м.п. ткани за оптимальное время и с оптимальной стоимостью расхода ресурсов по сравнению с другими сочетаниями ТР ТП.

При этом известно, что ТР, соответствующие вершинам графа № 11, 12 выполняются на

одном и том же технологическом оборудовании; тип оборудования, соответствующего вершине графа № 13 отличен от типа оборудования для вершин графа № 11, 12.

Этап 3.2. Определение оценочных характеристик для кластеров.

Шаг 1. Определение оценочных характеристик кластеров.

Математические ожидания компонент для двумерного случая состояний (T, S) вида (13) для каждого из кластеров:

$$\Theta_1(K_1) = (6,86; 4374824) \in R^2,$$

$$\Theta_1(K_2) = (7,30; 5002508) \in R^2,$$

$$\Theta_3(K_1) = 4374824, \quad \Theta_3(K_2) = 5002508,$$

$$\Theta_3(K_1) < \Theta_3(K_2).$$

Математические ожидания компонент для трёхмерного случая состояний (T, S_{EnR}, S_{Xim}) вида (16) для каждого из кластеров K_1, K_2 :

$$\Theta_1(K_1) = (6,86; 2973701; 1401123) \in R^3,$$

$$\Theta_1(K_2) = (7,30; 3611105; 1391403) \in R^3,$$

$$\Theta_3(K_1) = 3287254, \quad \Theta_3(K_2) = 3869894,$$

$$\Theta_3(K_1) < \Theta_3(K_2).$$

Средняя стоимость энергоресурсов в K_1 (состояния ТП, наиболее близкие к оптимальному) превышает среднюю стоимость энергоресурсов в K_2 (за счет объема потребления); при этом стоимость химикатов и средний расход времени для кластеров K_1, K_2 отличаются не существенно.

Математические ожидания компонент для пятимерного случая состояний $(T, S_{ElEn}, S_{Par}, S_{Gaz}, S_{Xim})$ вида (16) для каждого из кластеров K_1, K_2 :

$$\Theta_1(K_1) = (6,86; 1414644; 1340171; 218886; 1401123) \in R^5,$$

$$\Theta_1(K_2) = (7,30; 1930187; 1610184; 70734; 1391403) \in R^5,$$

$$\Theta_3(K_1) = 2410048, \quad \Theta_3(K_2) = 2873903,$$

$$\Theta_3(K_1) < \Theta_3(K_2).$$

Средняя стоимость газа в K_1 (состояния ТП, наиболее близкие к оптимальному) превышает

среднюю стоимость газа в K_2 (за счет объема потребления), но при этом для K_1 имеет место меньшее потребление более дорогостоящих ресурсов: пара и электроэнергии.

Шаг 2. Покомпонентное построение и анализ статистических моделей распределения внутри кластеров данных об итоговых состояниях ТП.

Количество элементов, образующих кластер K_1 , составляет: $|K_1|=18 \leq 30$, что не образует объема малой выборки элементов, необходимого для построения статистических моделей распределения данных об итоговых состояниях ТП внутри кластеров.

Заключение

Предложенные методы оптимизации режимов ТП на ОАО «Моготекс» (г. Могилев, Беларусь) позволили получить существенный экономический эффект (при производстве 1000 м.п. рассматриваемой группы тканей) за счет уменьшения стоимости расхода ресурсов на 14% и времени на 9%. Результаты кластеризации состояний: в кластере K_1 математическое ожидание экономического эффекта за счет уменьшения стоимости расхода ресурсов – на 13% и времени – на 8% со среднеквадратическим отклонением соответственно 0,53% ($p = 0,95$) и 2,28% ($p = 0,95$); в кластере K_2 математическое ожидание экономического эффекта за счет уменьшения стоимости расхода ресурсов – на 0% и времени – на 3% со средне-

квадратическим отклонением соответственно 0,53% ($p = 0,95$) и 2,25% ($p = 0,95$).

Литература

1. Новиков, Ф.А. Дискретная математика для программистов: учебник / Ф. А. Новиков. – 2-е изд. – СПб.: Питер, 2007. – 368 с.
2. Аверченков, В.И. Многокритериальное управление технологическим процессом с использованием принципа оптимальности Беллмана / В.И. Аверченков, А.И. Якимов, Е.М. Борчик // Известия Волгоградского государственного технического университета. – 2014. – № 3(22). – С. 94–101.
3. Якимов, А.И. Технология имитационного моделирования систем управления промышленных предприятий: монография / А.И. Якимов. – Могилев: Беларус.-Рос. ун-т, 2010. – 304 с.
4. Гмурман, В.Е. Теория вероятностей и математическая статистика: Учеб. пособие для вузов / В.Е.Гмурман. – М.: Высш. шк., 2003. – 478 с.
5. Методы и модели анализа данных: OLAP и Data Mining / А. А Барсегян и [и др.] – СПб: БХВ – Петербург, 2004. – 336 с.
6. Большев, Л.Н. Таблицы математической статистики / Л.Н. Большев, Н. В. Смирнов. – М.: Наука, 1983. – 416 с.
7. Кендалл, М. Теория распределений: пер. с англ. / М. Кендалл, А. Стюарт. – М.: Наука, 1966. – 588 с.
8. Борчик, Е.М. Построение плотностей распределения Пирсона для многомодальных выборок с применением кластеризации / Е.М. Борчик, А.И. Якимов, А.И. Степанов, В.В. Башаримов // Информационные технологии. – 2012. – № 12. – С. 30–35.
9. Кендалл, М. Многомерный статистический анализ и временные ряды / М. Кендалл, А. Стюарт. – М.: Наука, 1976. – 736 с.
10. Отделка хлопчатобумажных тканей: справочник / Н.В. Егоров [и др.]: В 2 ч. – М.: Легпромбытиздат, 1991. – Ч. 1. – 432 с.

Якимов Анатолий Иванович: Доцент ГУ ВПО «Белорусско-Российский университет», г. Могилев. Окончил Могилевский машиностроительный институт в 1979 году. Кандидат технических наук. Количество печатных работ: 120, 2 монографии. Область научных интересов: методы, средства и технологии имитационного моделирования в системах управления промышленным предприятием. E-mail: ykm@tut.by

Борчик Екатерина Михайловна: Ведущий инженер-программист Управления информационных технологий ОАО «Моготекс» г. Могилев. Окончила Могилевский государственный университет им. А.А. Кулешова в 2003 году. Количество печатных работ: 28. Область научных интересов: имитационное моделирование, многокритериальная оптимизация, кластерный анализ. E-mail: katrinb15@gmail.com

Максимов Евгений Михайлович. Старший научный сотрудник. ИПУ РАН. Окончил Московский институт электронного машиностроения в 1975 году. Кандидат технических наук. Количество печатных работ: 78. Область научных интересов: адаптивное управление, идентификация и компьютерное моделирование производственных систем. E-mail: maxfone@yandex.ru

Clustering states of the object in dealing with the problem of choosing the optimal technological modes

A.I. Yakimov, E.M. Borchyk, E.M. Maximov

Abstract. Method selecting optimal technological modes in an industrial plant is proposed. For disambiguate the choice of technological modes at each stage of the technological process the task of multi-criteria optimization using the principle of optimality of Bellman and lexicographical ordering of the optimization criteria for resource costs and / or time costs to release a predetermined amount of products. The method of applying cluster analysis to assess the states of the technological process that are close, in accordance with a predetermined measure of the closeness to the optimal state is proposed.

Keywords: technological process, graph, state, cluster analysis, multi-criteria optimization.

Anatoly I. Yakimov. Associate Professor at SI HPE «Belarusian-Russian University», Mogilev. Ph.D. He graduated the Mogilev Machine Building Institute in 1979. Number of publications: 120 (including 2 monographs). His research interests: methods, tools and simulation technologies in the industrial enterprise management systems.

Ekaterina M. Borchyk, Senior Software Engineer at the Department of Information Technologies, JSC "Mogotex", Mogilev. She graduated the Mogilev State A. Kuleshov University in 2003. Number of publications: 28. Her research interests: simulation, multi-objective optimization, cluster analysis.

Evgeny M. Maximov. Senior Researcher at ICS RAS. Ph.D. He graduated the MIEM in 1975. Number of publications: 78. His research interests: adaptive control, identification and computer modeling of manufacturing systems.

