

УДК 519.7 + 681.513

ПРОЕКТИРОВАНИЕ ТЕХНОЛОГИЧЕСКОГО ОБОРУДОВАНИЯ В СЛУЧАЕ МАЛЫХ ВЫБОРОК НА ОСНОВЕ ОПОРНЫХ ТОЧЕК

¹Болнокин В.Е., ²Ивашов Е.Н., ²Корпачев М.Ю., ²Костомаров П.С.

¹ФГБУН «Институт машиноведения им. А.А. Благонравова РАН», Москва,

e-mail: vitalybolnokin@yandex.ru;

²ФГАОУ ВПО «Московский институт электроники и математики Национального исследовательского университета «Высшая школа экономики», Москва, e-mail: ienmiem@mail.ru

Для «обхода» проблемы малых выборок предлагается рандомизированный подход определения коэффициентов размытости непараметрических решающих правил на основе процедуры их случайного выбора. Реализация «обхода» проблем малых выборок при оценивании плотностей вероятностей обеспечивается не только увеличением объема исходных данных, но и результатами технического моделирования. Предложена модель оптимизации выбора контрольных точек. Минимизируется количество точек, чтобы управление стало наиболее дешевым, а также минимизируются потери от неоптимальности структуры управления.

Ключевые слова: автоматизированное проектирование, задача оптимизации, малые выборки, выбор опорных точек

PROJECTING OF TECHNOLOGICAL EQUIPMENT IN CASE OF SMALL FETCHESBASED ON ANCHOR POINT

¹Bolnokin V.E., ²Ivashov E.N., ²Korpachev M.Y., ²Kostomarov P.S.

¹Federal Budget-Funded Research Institute for Machine Science named after A.A. Blagonravov, Moscow,

e-mail: vitalybolnokin@yandex.ru;

²FGAEU HPE «Moscow institute of electronics and mathematics The National research university «High school of economics», Moscow, e-mail: ienmiem@mail.ru

The randomizing approach of determining the dissipation coefficients of nonparametric deceiving rules is based on the procedure of random selection and is used for «bypass» the problem of small excerpts. Realization of «bypass» of this problems of small fetches (by estimating of the solidity probability) is provided not only by the increasing the volume of the initial data, but also by the results of technical modeling. Model of control points optimization selection is proposed. Number of points should be minimized to make the control cheaper, and to make losses from non optimal management structure also smaller.

Keywords: automatical projection, optimization problems, small fetches, the choice of anchor point

При автоматизированном проектировании элементов, узлов и устройств технологического оборудования разработчик зачастую встречается с проблемой малых выборок, т. е. отсутствия необходимого количества эмпирических данных для принятия рационального конструктивно-технологического решения [1]. Особо остро проблема малых выборок ощущается в те моменты, когда получение эмпирических данных сопровождается существенным увеличением стоимости проектных работ.

Для «обхода» проблемы малых выборок при оценивании плотностей вероятностей $p(x)$ увеличим объем исходных данных $x^i, i = 1, n$, за счет результатов статического моделирования. С этой целью в β -окрестности каждой i -й точки выборки осуществим m имитаций с законом распределения $p_2(x)$. Полученная статическая выборка $x^i + x_2^j, j = 1, m, i = 1, n$, при равновероятных значениях $x^i, i = 1, n$, соответствует смеси плотностей вероятностей:

$$\hat{p}(x) = \frac{1}{n} \sum_{j=1}^n p_2^j(x_2). \quad (1)$$

Нетрудно заметить, что непараметрическая оценка (1) имеет вид

$$\bar{p}(x) = (nmc)^{-1} \sum_{i=1}^n \sum_{j=1}^m \Phi \left(\frac{x - x^i - x_2^j}{c} \right). \quad (2)$$

Существующий парадокс традиционных методов идентификации стохастических моделей состоит в сопоставлении конечной случайной выборки наблюдений переменных изучаемых объектов с конкретным набором параметров модели, оптимальном в некотором смысле. Предлагается принципиально новый рандомизированный подход определения коэффициентов размытости непараметрических решающих правил на основе процедуры их случайного выбора, который рассматривается на примере оптимизации алгоритмов формирования литографических процессов [2].

Формирование случайной последовательности коэффициентов размытости при оценивании плотности вероятности $p(x)$ осуществляется на основе выборки расстояний между исходными наблюдениями $(x^i, i = 1, n)$ и их k – ближайшими соседями.

Рандомизированный метод оптимизации рассмотрим на следующем примере. Пусть $V = (x^i, i = 1, n)$ – выборка из n статистически независимых наблюдателей случайной величины $x \in R^1$ с плотностью вероятности $p(x)$, вид которой неизвестен. Будем считать, что $p(x)$ ограничена и непрерывна со всеми своими производными до второго порядка включительно. В качестве приближения по эмпирическим данным V искомой плотности вероятности $p(x)$ примем статистику типа Розенблатта – Парзена [3]

$$\bar{p}(x) = (nc)^{-1} \sum_{i=1}^n \Phi\left(\frac{x-x^i}{c}\right), \quad (3)$$

где $\Phi(\cdot)$ – ядерные функции, удовлетворяющие условиям положительности, симметричности и нормированности; $c = c(n)$ – последовательность положительных чисел (коэффициентов размытости) таких, что

$$\lim_{n \rightarrow \infty} c(n) = 0, \quad \lim_{n \rightarrow \infty} nc(n) = \infty. \quad (4)$$

Свойства непараметрической оценки плотности вероятности (3), такие как асимптотическая несмещенность, состоятельность, сходимости почти наверное к $p(x)$, подробно рассмотрены в работе [4].

$$\bar{W}_p^{-1}(h) = \int_0^h M(\bar{p}(x) - p(x)) p(c) \sim dc \frac{h^2(t+1)}{2(t+3)} p^{(2)}(x), \quad (8)$$

а его отношение к соответствующему смещению $\bar{W}^{-1}(c)$ для традиционной непараметрической оценки при оптимальных параметрах c^* и h^*

$$\frac{\bar{W}_p^{-1}(h^*)}{\bar{W}^{-1}(c^*)} \sim \frac{t+1}{t+3} \left(\frac{5+t}{t}\right)^{\frac{2}{5}}. \quad (9)$$

Если параметр t плотности вероятности $p(c)$ больше или равен 2, то отношение (9) меньше единицы.

Анализ выражений (5), (8) показывает, что непараметрическая оценка плотности вероятности со случайными значениями коэффициентов размытости (7) обладает

Для сравнения традиционного и рандомизированного метода оптимизации непараметрической оценки плотности вероятности определим отношение $\bar{W}(c^*) / \bar{W}_p(h^*)$ соответствующих им асимптотических выражений среднеквадратических критериев при оптимальных параметрах c и h .

Из условия минимума $W_p(c)$ и $\bar{W}_p(h)$ по c и h нетрудно получить

$$c^* = \left[\frac{\Phi^2}{np^{(2)}(x)^2} \right]^{\frac{1}{5}}, \quad h^* = \left(\frac{5+t}{t} \right)^{\frac{1}{5}} c^*. \quad (5)$$

При оптимальных параметрах c^* , h^* отношение

$$\frac{\bar{W}(c^*)}{\bar{W}_p(h^*)} \sim \frac{(1+5t^{-1})^{\frac{1}{5}}}{1+t^{-1}} \quad (6)$$

меньше единицы при конкретных значениях параметра t закона распределения $p(c)$ коэффициентов размытия ядерных функций.

Однако использование непараметрической оценки со случайными значениями коэффициентов размытия ядерных функций

$$\tilde{p}(x) = n^{-1} \sum_{i=1}^n \frac{1}{c^i} \Phi\left(\frac{x-x^i}{c}\right) \quad (7)$$

позволяет снизить смещение при оценивании плотности вероятностей по сравнению с традиционной статистикой типа (3).

Можно показать, что асимптотическое выражение смещения

свойствами асимптотической несмещенности и состоятельности. Она характеризуется пониженным смещением (9) и несколько большим значением среднеквадратического отклонения (6) по сравнению с непараметрической статистикой (3). Следует ожидать проявления потенциальной эффективности непараметрической оценки плотности вероятности (7) при конечных объемах статистических данных.

Реализация «обхода» проблем малых выборок при оценивании плотностей вероятностей $p(x)$ обеспечивается не только увеличением объема исходных данных $x^i, i = 1, n$, но и результатами технического моделирования, т.к. сложность и многооб-

разие процессов функционирования проектируемых технических систем не всегда позволяют получать для них адекватные математические модели, сформулированные в виде различных аналитических соотношений.

Возникает необходимость изменений структуры системы управления для повышения ее эффективности. Нужно определить контрольные точки – моменты в реализации проекта, в которых будет проводиться анализ эффективности структуры [5].

Можно выделить следующие принципы выбора опорных точек:

- завершение важнейших этапов реализации проекта;
- события передачи существенных результатов в другие подсистемы;

• события, не управляемые в рамках данного комплекса операций.

Введем некоторые обозначения и формализуем наиболее важные ограничения на выбор очередной контрольной точки. Пусть t_g – g -я контрольная точка, $g = 0, N$, где $g_0 = 0$ – начало реализации проекта, $g_N = T$ – срок его завершения. Существует некоторое минимальное D , на величину которого может отстоять контрольная точка от предыдущей. Если эта величина меньше D , то такая точка неэффективна, так как нельзя принять решение в столь короткий срок. Таким образом, имеет место ограничение:

$$\begin{cases} t_{g+1} \geq t_g + \Delta, & \text{если } (t_g + \Delta) < T, \\ t_{g+1} = t_N = T, & \text{если } (t_g + \Delta) \geq T. \end{cases} \quad (10)$$

Если обозначить $t_k(\xi_i)$ как плановые сроки завершения i -й работы, то

$$n^k(\xi_i, \Delta t_g) = n(\xi_i, \Delta t_g) | t_0(\xi_i) \in (t_g, t_{g+1}), \quad (11)$$

где $t_0(\xi_i)$ – плановое начало работы, $n(\xi_i, \Delta t_g)$ – количество работ, выполняемых в период (t_g, t_{g+1}) , а $n^k(\xi_i, \Delta t_g)$ – количество работ, которые завершат свое выполнение в этот период.

$$\delta^k = \frac{n^k(\xi_i, \Delta t_g)}{n(\xi_i, \Delta t_g)}; \delta^0 = \frac{n^0(\xi_i, \Delta t_g)}{n(\xi_i, \Delta t_g)} \quad (12)$$

$$\sum_j \max(\alpha_j(\lambda_j^g - \lambda_j^{g+1}), \beta_j(\lambda_j^{g+1} - \mu_j)) \leq \gamma. \quad (13)$$

Это ограничение определяет границы допустимой неоптимальности системы управления, где γ обозначает суммарные оценочные затраты на проведение процедур анализа и оптимизации организационной структуры.

Кроме того, на неэффективность структуры системы управления указывает также

Показатели (12) можно использовать для приближенного определения степени изменения состава выполняемых работ. Если эти изменения велики, то, возможно, надо изменить и структуру системы. Если же значения этих показателей малы, то структура системы не нуждается в оптимизации [5].

Пусть величины α_j и β_j показывают потери, которые несет система при простое или перегрузке j -го узла. Зная эти величины, можно выразить ограничение на допустимые затраты от неоптимальной структуры:

увеличение числа внешних и внутренних связей в каждом комплексе операций. Необходимо также установить для каждой связи веса λ_{pg} , характеризующие их интенсивность.

Задачу оптимизации выбора контрольных точек можно записать так:

$$N \rightarrow \min \quad (14)$$

$$\sum_{g=1}^{N-1} \left(c(\delta_s^g) + \sum_j \max(\alpha_j(\lambda_j^g - \lambda_j^{g+1}), \beta_j(\lambda_j^{g+1} - \mu_j)) \right) \rightarrow \min \quad (15)$$

$$c(\delta_s^g) + \sum_j \max(\alpha_j(\lambda_j^g - \lambda_j^{g+1}), \beta_j(\lambda_j^{g+1} - \mu_j)) > g = \overline{1, N-1} \quad (16)$$

Таким образом, имеем модель оптимизации выбора контрольных точек (с учетом

ограничений, принятых выше). Минимизируется количество точек, чтобы управление

стало наиболее дешевым, а также минимизируются потери от неоптимальности структуры управления.

Далее рассматривается оптимизация графика работ и распределения ресурсов при управлении интенсивностями отдельных операций.

Рассмотрим методику решения двух типов задач: управление интенсивностью работ с минимизацией затрат и управление объемом работ при ограничениях на имеющиеся ресурсы [5].

На практике чаще всего имеет место дискретная зависимость скорости выполнения операции от затрат, которую можно представить следующим образом:

$$v_i(u_i) = \begin{cases} 0, \text{ при } u_i < \underline{u}_i, \\ v_{i1}, \text{ при } \underline{u}_i \leq u_i < u_{i1}, \\ \dots \\ v_{in}, \text{ при } u_i \geq u_{i,1-n} \end{cases} \quad (17)$$

Будем считать, что скорость выполнения операции не изменяется непрерывно.

Интенсивность v_{ip} является функцией от интенсивности затрат u_{ip} , следовательно величина интервала t_{ip} , за который i -я работа выполнится также является функцией от u_{ip} .

Каждой работе поставим в соответствие величины t_{Hi} и t_{Ki} , которые определяют начало и конец выполнения i -й работы. Если элементарные подработы рассматривать как отдельные работы сетевого графа, то можно представить модель в виде:

$$\begin{aligned} \sum_i t_i(u_i + \Delta u_i) &\rightarrow \min \\ t_i &= \frac{Y_{0i}}{v_i}, \forall_i; \\ \sum_{i \in L_\gamma} t_i &\leq T_0 - \tau(t), \forall_\gamma; \\ t_{Ki} &= t_{Hi} + t_i. \end{aligned} \quad (18)$$

Решение этой задачи является очень сложным, поэтому необходимо ее линеаризовать. Это достаточно легко сделать, если функции зависимости скорости выполнения операции от затрат <<boll55.wmf>> принять в предложенном выше виде. Тогда для каждой выполняемой i -й работы можно установить возможные варианты ее выполнения.

Важной проблемой также является распределение функций между проектной и организационной системами управления, из-за чего часть функций по управлению (например, определение сроков реализации, распределение ресурсов) остаются прерогативой центра, что не всегда эффективно.

Данный тип структуры целесообразно применять в крупных проектах, условия реализации которых не полностью определены. Таким образом, в чистом виде ни одна из существующих структур не является идеальной.

Пусть определен общий объем работ по проекту, на основании которого руководитель проекта разрабатывает график работ. В общем виде такой график может быть представлен в виде системы, состоящей из объектов трех видов.

Будем использовать теоретико-графовое описание системы:

$$\Sigma = Q, U, \epsilon,$$

где Q – множество вершин, U – множество ребер, ϵ – отношение инцидентности, которое каждому ребру из U ставит в соответствие пару вершин из Q [5]:

$$u \in U \Rightarrow (\exists p, q \in Q \times Q)(u \epsilon p, q). \quad (19)$$

Функции описывают поведение элементов системы и представляются набором множеств: $F = R, f$, где $R = (A_i)_{i \in I}$ – семейство некоторых базовых множеств A_i (сигналов, траекторий, ресурсов, т. е. множеств, на которых задаются функции); f – множество всех отображений

$$\prod_{i \in I_1} A_i \rightarrow \prod_{i \in I_2} A_i; I_1, I_2 \subset I, \quad (20)$$

то есть, функций, отражающих определенные задачи реализации проекта. Это связано с затратой ресурсов (денежных средств, сырья и т. д.). В этом случае необходимо учитывать связи, обусловленные наличием ограничений типа

$$\Phi_k \left[f_k \left(\prod_{i \in I} A_i \right) \right] \leq u_k, \quad (21)$$

где u_k – лимит ресурсов, выделенный для реализации k -й функции, Φ_k – потребление ресурсов для реализации функции f_k .

Таким образом, график работ – это структура, вершинам которой поставлены в соответствие функции, а ребрам – базисные множества, на которых эти множества определены. Каждая вершина характеризуется объемом потребляемых ресурсов,

временем выполнения своих функций, а каждое ребро может характеризоваться, например, пропускной способностью [5].

Для построения эффективной системы управления необходимо оптимизировать распределение функций f по узлам системы Q . Введем булевы переменные

$$\xi_{ij} = \begin{cases} 1, & \text{если } i\text{-я функция проекта выполняется в } j\text{-м узле} \\ 0, & \text{в противном случае.} \end{cases} \quad (22)$$

В качестве целевой функции модели распределения функций по узлам можно взять один из следующих функционалов [5]:

минимизация суммарных затрат на выполнение задач

$$\sum_{A_{i,j}} \tilde{n}_{ij} \xi_{ij} \rightarrow \min; \quad (23)$$

минимизация суммарного времени выполнения задач

$$\sum_{A_{i,j}} t_{ij} \xi_{ij} \rightarrow \min; \quad (24)$$

минимизация максимального времени решения задач

$$\max_j \sum_{A_i} t_{ij} \xi_{ij} \rightarrow \min. \quad (25)$$

В зависимости от особенностей системы управления проектами, целочисленная оптимизационная модель распределения отдельных задач проекта по узлам комплектуется из приведенных целевых функций и ограничений.

Заключение

Для повышения эффективности проектирования систем используют модели, сформулированные выше: управления интенсивностью выполнения отдельных работ с минимизацией затрат на интенсификацию, управления объемом выполненных работ при ограничениях на имеющиеся ресурсы. При некоторой модификации данные модели сводятся к типовым задачам, решаемым с использованием стандартных

методов исследования операций применяемых для решения линейных моделей.

Рассматриваемая модель является оптимизационной задачей целочисленного программирования и позволяет объединить структурные элементы проекта в группы, оптимальные с точки зрения минимизации межгрупповых связей. В итоге объединяются два нижних уровня иерархической системы управления, во главе каждого из которых стоит система, управляющая функционированием данного комплекса.

В целом, анализ проблем формирования и реализации последовательности проектов показал, что программа их формирования требует некоего механизма реализации этой программы, что предполагает определение исполнителей и распределение ограниченных ресурсов между ними. Отсюда возникает одна из основных задач – определение оптимальной очередности реализации проектов, включенных в программу, с целью получения максимального эффекта от вложенных средств.

Список литературы

1. Васин В.А., Ивашов Е.Н., Костомаров П.С. и др. Моделирование процессов иммерсионной ультрафиолетовой литографии на этапе аванпроектирования // Нелинейный мир. 2012. Т. 10. № 7. С. 454-459.
2. Лапко А.В. Непараметрические системы обработки неоднородной информации / А.В. Лапко, В.А. Лапко. – Новосибирск: Наука, 2007. – 174 с.
3. Parzen E. On the estimation of a probability density function and mode // Ann. Math. Statist. – 1962. – P. 1065.
4. Епаничников В.А. Непараметрическая оценка многомерной плотности вероятности // Теория вероятности и ее приложения. – 1969. – Т. 14, вып. 1. – С. 156 – 161.
5. Болнокин В.Е. Адаптивное управление на базе нечетких регуляторов и нейросетевой технологии: Монография / В.Е. Болнокин, Хо Д. Лок. – Воронеж: Изд-во «Научная книга», 2012. – 280 с.