



АНАЛИЗ КРИТИЧЕСКИХ ЗНАЧЕНИЙ ПАРАМЕТРОВ И МОДЕЛИРОВАНИЕ УПРАВЛЯЕМОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

И.И. Барладян, А.В. Кузнецов, А.С. Мандель

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Рассмотрена модель оптимального управления выбором числа рабочих устройств обслуживания в многолинейной системе массового обслуживания. На основе качественного анализа выделены существенные характеристические параметры системы. Модель реализована в программной среде MATLAB 7.0. Представлены результаты моделирования. Выделены критические значения параметров системы массового обслуживания и описаны наиболее характерные режимы ее функционирования.

ВВЕДЕНИЕ

В работе [1] предложена модель управляемой системы массового обслуживания, в которой в качестве управляемого и динамически оптимизируемого параметра выступает число рабочих устройств обслуживания в многолинейной системе массового обслуживания. Это число выбирается как функция от изменяющейся во времени интенсивности входящего потока требований. Процесс изменения во времени интенсивности входящего потока является дискретной по времени цепью Маркова с конечным числом состояний. При оптимизации максимизируется суммарная средняя прибыль в заданном интервале планирования. В работе [1] решение задачи оптимального управления сведено к алгоритму дискретного динамического программирования [2, 3].

В настоящей статье проводится качественный анализ формируемых в результате применения предложенного в работе [1] алгоритма дискретного динамического программирования программ оптимального управления. Описываются результаты моделирования алгоритма в среде MATLAB 7.0. Выделены наиболее характерные режимы функционирования рассматриваемой системы массового обслуживания (СМО).

1. МОДЕЛЬ УПРАВЛЯЕМОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

Рассматривается модель функционирования многолинейной СМО на конечном интервале времени $[0, T]$, где $T = N\tau$, N — натуральное число.

В моменты времени $t_n = n\tau$, $n = 1, 2, \dots, N - 1$, с вероятностями $p_{ij}(m_1)$, $i, j = 1, 2, \dots, L$, происходят скачкообразные изменения интенсивности пуассоновского входящего потока от значений λ_i к значениям λ_j , $\lambda_j \in \Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_L\}$. Здесь m_1 — число рабочих устройств обслуживания. Причем в те же моменты времени можно принять решение об изменении числа рабочих устройств обслуживания $m_1 = 1, 2, \dots, M$ путем включения резервных или отключения рабочих устройств обслуживания. Предполагается, что соотношения между предельно возможной суммарной интенсивностью обслуживания $M\mu$, где μ — интенсивность обслуживания на одном рабочем устройстве, и совокупностью возможных значений интенсивностей входящего потока выбраны так, что в каждом из интервалов длительности τ устанавливался стационарный в вероятностном смысле режим.

Критерием выбора числа рабочих устройств обслуживания служит суммарная средняя прибыль СМО в интервале планирования $[0, T]$ в предположении, что расчет прибыли в рассматриваемой СМО обусловлен следующей структурой затрат и доходов:

c_1 — стоимость эксплуатации одного рабочего устройства обслуживания в единицу времени;

c_2 — стоимость содержания одного резервного устройства обслуживания в единицу времени;

A_1 — цена переключения одного устройства из числа резервных в число рабочих;

A_2 — цена переключения одного устройства из числа рабочих в число резервных;

d — стоимость единицы времени пребывания одного требования в очереди;

h — доход, связанный с окончанием обслуживания одного требования.

При указанной структуре параметров затрат в работе [1] выписаны следующие уравнения дискретного динамического программирования, которые позволяют рассчитать оптимальную стратегию управления числом рабочих устройств:

$$\Pi_1^*(\lambda_i, m_1) = \max_{m_{1p}^{(i)} \leq m_{1p} \leq M} C^{(1)}(\lambda_i, m_1, m_{1p}), \quad (1)$$

$$\Pi_s^*(\lambda_i, m_1) = \max_{m_{1p}^{(i)} \leq m_{1p} \leq M} \left\{ C^{(1)}(\lambda_i, m_1, m_{1p}) + \sum_{j=1}^L p_j(m_{1p}) \Pi_{s-1}^*(\lambda_j, m_{1p}) \right\}, \quad (2)$$

где $\Pi_s^*(\lambda_i, m_1)$ — максимальное значение средней прибыли на интервале, который начинается за s шагов до конца периода планирования $[0, T]$, $s + n = N$, $\forall s = \overline{2, N-1}$; m_{1p} — текущее (выбираемое) управляющее решение о числе рабочих устройств; $m_{1p}^{(i)} = \lambda_i / \mu + 1$ — нижнее ограничение на возможное число рабочих устройств, которое, на самом деле, зависит от времени t , поскольку интенсивность λ_i является функцией времени; $\lambda_j, \lambda_j \in \Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_L\}$ (считается, что $\lambda_1 < \lambda_2 < \dots < \lambda_L$) — значения интенсивностей входящего потока в моменты времени за s и за $s-1$ шагов до конца периода планирования соответственно, а $C^{(1)}(\lambda_i, m_1, m_{1p})$ — так называемая одношаговая прибыль, которая равна среднему доходу за один шаг за вычетом средних одношаговых затрат. В работе [1] представлены расчетные соотношения для определения всех компонентов одношаговых доходов и затрат.

В настоящей работе представлены результаты моделирования алгоритма (1), (2) в среде MATLAB 7.0. Прежде чем перейти к описанию результатов моделирования, обсудим наиболее вероятные и естественные варианты соотношений между эконометрическими и статистическими параметрами рассматриваемой СМО.

2. КРИТИЧЕСКИЕ ПАРАМЕТРЫ И ОПИСАНИЕ ПОДХОДА К МОДЕЛИРОВАНИЮ

2.1. Классификация параметров

При моделировании существенны не конкретные значения параметров, а соотношения между ними, а также возможные зависимости между конкретными группами параметров.

Анализ основных уравнений модели позволяет сразу выделить среди параметров модели $M, \mu, \tau, N, c_1, c_2, A_1, A_2, d$ и h те из них, что оказывают наиболее существенное влияние на ее динамику. Наличие зависимостей между параметрами может быть выявлено в ходе эксперимента.

Разделим параметры на две группы:

- $M, \mu, c_1, c_2, A_1, A_2, d$ и h — будем говорить, что они определяют класс системы;
- τ и N — внешние параметры.

Будем считать, что для всех случаев, которые рассматриваются далее, выполнены следующие соотношения:

$$\mu h > c_1, \quad (3)$$

$$h \gg d. \quad (4)$$

Условие (3), по сути своей, означает, что использование данного устройства обслуживания рентабельно. Чтобы пояснить обоснованность предположения (4), рассмотрим подробнее природу возникновения издержек ожидания.

Вследствие длительного пребывания в очереди требования могут покидать систему, так и не дождавись начала их обработки. Клиентам надоедает стоять в очереди, и они уходят (кстати, и в компьютерных сетях при большой нагрузке нередко происходят потери пакетов с данными). Во многих случаях клиенты так и не возвращаются, а находят, вместо этого, другую, конкурентную СМО, способную более быстро удовлетворять их потребности. Это влечет за собой снижение спроса на обслуживание в будущем и, как следствие, потерю прибыли. Издержки на ожидание как раз и учитывают подобную возможность потери прибыли. Напомним, что в соответствии с рассуждениями, проведенными в работе [1], в большинстве случаев значение параметра d на 1—2 порядка меньше значения введенного выше параметра h .

2.2. О «дважды» случайном характере входящего потока

В реальных прикладных задачах входящий поток требований редко бывает чисто пуассоновским. Нередко можно отследить четкие закономерности характера скачкообразных изменений интенсивности входящего потока требований. Перечислим наиболее распространенные из них.

- *Плато* — частота изменений интенсивности входящего потока требований невелика. Поэтому достаточно долгое время рассматриваемая СМО функционирует как чисто пуассоновская, вступая при этом в стационарный (вероятностно) режим.
- *Пик* — входящий поток почти все время стационарен, однако имеют место периоды пиковых скачков нагрузки. Например, системы биллин-



га испытывают пиковые нагрузки утром и вечером, и на фоне этого еще более резкие скачки в праздничные дни. Утренние и вечерние пики также характерны для парков общественного транспорта (днем нагрузка сравнительно невысока).

Исследуемая имитационная модель и семантически, и функционально учитывает отмеченные особенности поведения входящего потока.

2.3. Классификация систем

При моделировании рассматривались два класса систем.

- Класс 1 ($c_1 \sim \mu h$; $A_2 \gg c_1$):
 - высокая стоимость содержания рабочих устройств обслуживания;
 - низкая стоимость перевода резервных устройств обслуживания в рабочий режим;
 - высокая стоимость отключения рабочих устройств.
- Класс 2 ($c_1 \gg \mu h$, $A_2 \approx A_1$):
 - низкая стоимость содержания рабочих устройств обслуживания;
 - издержки на включение и отключение устройств обслуживания невысоки и примерно равны между собой.

2.4. Описание подхода к моделированию

Исследуем зависимость решений от параметров систем при двух типах входящего потока (см. п. 2.2). Для каждого эксперимента представим построенные в результате моделирования в среде MATLAB 7.0 графики зависимостей $m_{1p}(t)$ и $m_{1p}^*(t)$.

Значения N , τ , μ , и c_2 задаются для всей серии экспериментов. При выборе значения μ следует учитывать ограничение $\mu\tau \gg 1$, которое связано с предположением о стационарности режима, устанавливающегося в СМО, а с учетом предположения (3) и того что, как правило, $c_1 \gg c_2$, положим $c_2 = 0$. Для удобства будем считать, что $\tau = 1$.

Для задания «плато» используется матрица переходных вероятностей вида

$$P_{\text{плато}} = \begin{pmatrix} 0,9 & 0,09 & 0,01 & 0 \\ 0,09 & 0,9 & 0,01 & 0 \\ 0 & 0,05 & 0,9 & 0,05 \\ 0 & 0,09 & 0,01 & 0,9 \end{pmatrix}. \quad (5)$$

Наибольший интерес представляют компоненты затрат c_1 , A_1 и A_2 . Именно они и изменяются в ходе проведенных экспериментов. В процессе моделирования множество возможных значений интенсивностей входящего потока задавалось как $\Lambda = \{100; 150; 270; 350\}$.

3. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ И ИХ АНАЛИЗ

На рис. 1–6 зависимость $m_{1p}(t)$ изображена пунктирной линией, а $m_{1p}^*(t)$ — сплошной, по оси абсцисс отложено время, по оси ординат — число рабочих устройств и ограничение на это число.

Рассмотрим зависимость функционала $P_s^*(\lambda_i, m_i)$ от A_1 (обозначим ее $P^*(A_1)$). В данном случае оба класса систем характеризуются схожей динамикой. С ростом компонентов затрат, связанных с переключением устройств обслуживания, решение постепенно перестает «отзываться» на резкие спады интенсивности. Для класса 1 (рис. 7) это начинает наблюдаться при $c_1 : A_2 \approx 1 : 12$, для класса 2 — при $c_1 : A_{1,2} \approx 1 : 6$ (рис. 8). При дальнейшем уве-

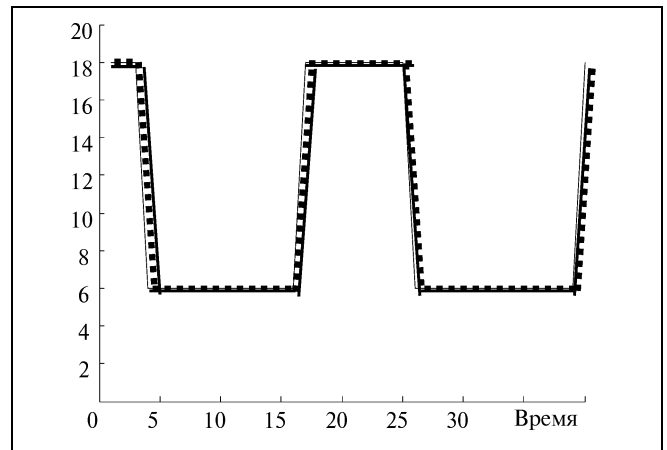


Рис. 1. Классы 1 и 2 ($c_1 : A_2 \approx 1 : 3$)

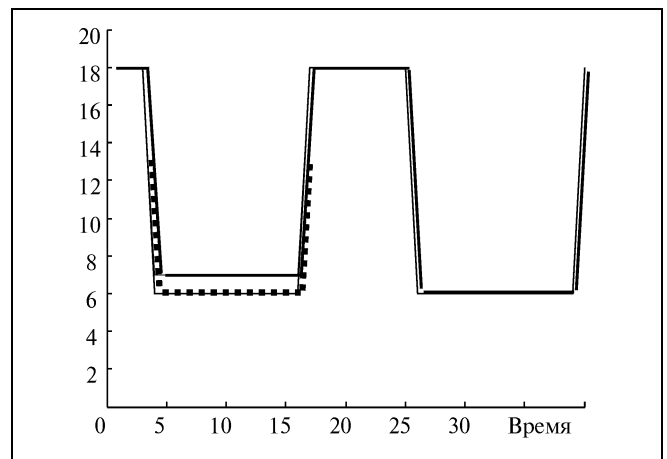
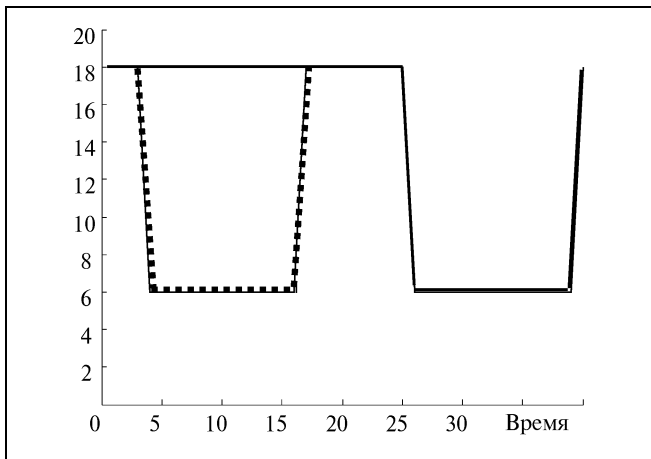
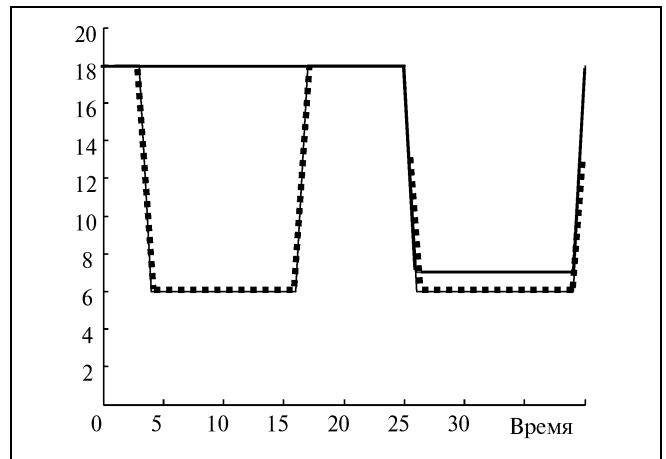
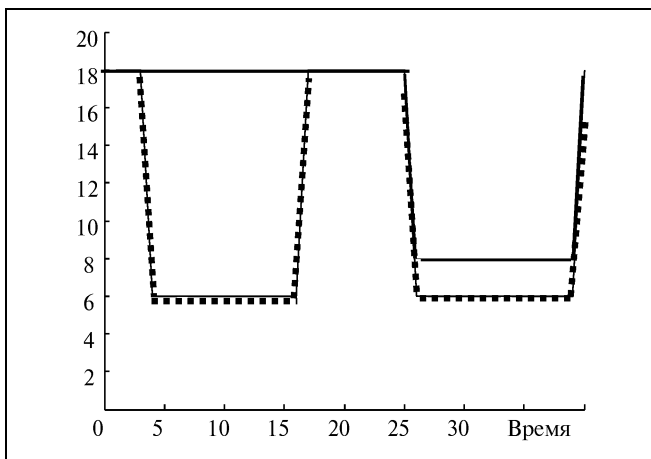
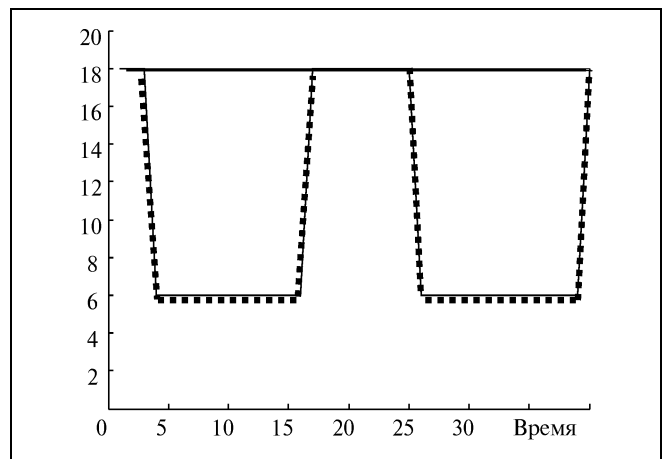
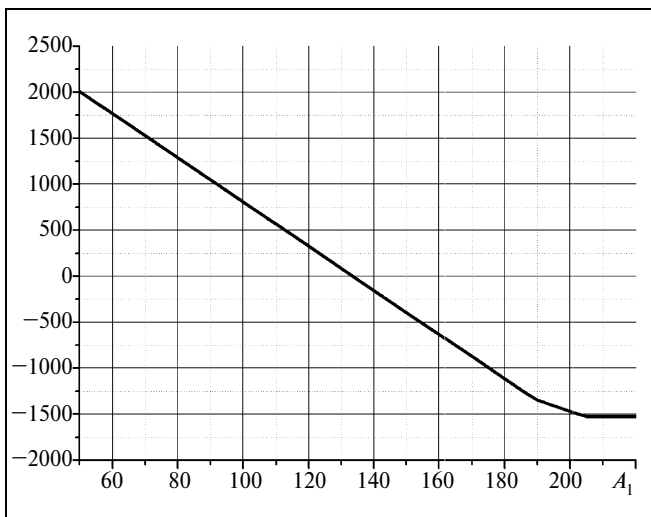
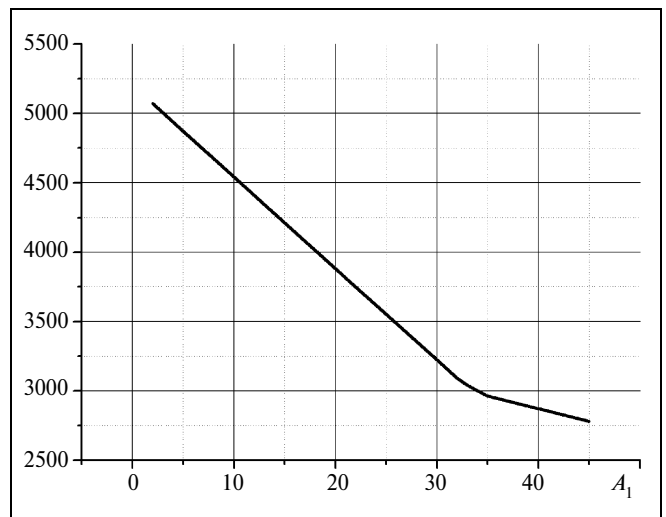


Рис. 2. Класс 2 ($c_1 : A_2 \approx 1 : 6$)


 Рис. 3. Класс 2 ($c_1 : A_2 \approx 1 : 6,3$)

 Рис. 4. Класс 1 ($c_1 : A_2 \approx 1 : 12$)

 Рис. 5. Класс 2 ($c_1 : A_2 \approx 1 : 6,8$)

 Рис. 6. Класс 1 ($c_1 : A_2 \approx 1 : 13$); класс 2 ($c_1 : A_2 \approx 1 : 7$)

 Рис. 7. Зависимость $P^*(A_1)$. Класс 1

 Рис. 8. Зависимость $P^*(A_1)$. Класс 2



личении A_2 относительно c_1 эта тенденция сохраняется (см. рис. 3, 5 и 6). При соотношении $c_1 : A_2 \approx 1 : 13$ для класса 1 и $c_1 : A_{1,2} \approx 1 : 7$ для класса 2 переключение устройств обслуживания становится невыгодным. Оптимальным решением является включение в рабочий режим числа устройств обслуживания, равного максимальной компоненте вектора m_{1p} .

Эластичность функции прибыли по издержкам на включение для систем первого типа (см. рис. 7)

составляет $\left(\frac{\Delta P^*}{\Delta A_1}\right)_1 = 24$, $\left(\frac{\Delta P^*}{\Delta A_1}\right)_2 = 13$, $\left(\frac{\Delta P^*}{\Delta A_1}\right)_3 = 0$.

Для систем второго типа (см. рис. 8) — $\left(\frac{\Delta P^*}{\Delta A_1}\right)_1 = 66$,

$\left(\frac{\Delta P^*}{\Delta A_1}\right)_2 = 22$. Таким образом, эффект от управления

системами класса 2 оказывается весьма заметным.

Значение соотношения $c_1 : A_{1,2}$, при котором достигается тот вид решения, что изображен на рис. 6, конечно же, зависит от картины изменения спроса и может сильно варьировать. Однако само наличие подобного решения дает мощный инструмент для управления системой на базе известных из теории управления запасами [4] двухуровневых стратегий.

ЗАКЛЮЧЕНИЕ

Уравнения динамического программирования, которые определяют оптимальные режимы пере-

ключения устройств обслуживания в многолинейной системе массового обслуживания из рабочих в резервные и обратно, промоделированы в среде MATLAB 7.0.

На основе качественного анализа рассматриваемой системы выделены существенные эконометрические параметры поставленной оптимизационной задачи, а также наиболее характерные режимы ее функционирования.

Сформированы оценки критических соотношений между значениями существенных параметров системы, при которых характер выбираемых оптимальных управлений качественно меняется.

ЛИТЕРАТУРА

1. Кузнецов А.В., Мандель А.С., Токмакова А.Б. Об одной модели управляемой системы массового обслуживания // Проблемы управления. — 2007. — № 5. — С. 39—43.
2. Беллман Р. Динамическое программирование. — М.: ИИЛ, 1960. — 400 с.
3. Арис Р. Дискретное динамическое программирование. Введение в оптимизацию стохастических процессов — М.: Мир, 1969. — 171 с.
4. Хедли Дж., Уайтин Т. Анализ систем управления запасами. — М.: Наука, 1969. — 512 с.

☎ (495) 334-89-69, e-mail: manfoon@ipu.ru

Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым. □

Новые книги

- Абраменко Г.В.** Проектирование сложных наукоемких технических систем. — М., 2006. — 271 с.
- Астапов Ю.М.** Теория управления летательными аппаратами. — М.: МАТИ, 2006. — 276 с.
- Береснев В.Л.** Дискретные задачи размещения и полиномы от булевых переменных. — Новосибирск: Ин-т математики, 2005. — 407 с.
- Бурковский В.Л.** Методы идентификации объектов информационно-управляющих систем. — Воронеж; Невинномысск, 2006. — 206 с.
- Бурцев В.С.** Параллелизм вычислительных процессов и развитие архитектуры супер-ЭВМ. — М.: Торус Пресс, 2006. — 414 с.
- Витяев Е.Е.** Извлечение знаний из данных. — Новосибирск, 2006. — 293 с.
- Губко М.В.** Математические модели оптимизации иерархических структур. — М.: ЛЕНАНД, 2006. — 262 с.
- Гук М.** Интерфейсы устройств хранения: ATA, SCSI и другие. — М.; СПб.: Питер, 2007. — 446 с.
- Емельянова Н.З.** Основы построения автоматизированных информационных систем. — М.: Форум-Инфра-М, 2007. — 415 с.
- Илларионов С.В.** Теория познания и философия науки. — М.: РОССПЭН, 2007. — 535 с.
- Исследования по безопасности.** — М.: Концепт, 2006. — 623 с.
- Итоги науки и техники. Сер. Современная математика и её приложения: Тематические обзоры.** — М.: ВИНТИ, 2006. — Т. 100. — 245 с.; Т. 108. — 264 с.
- Крухмалёв В.В.** Цифровые системы передачи. — М.: Горячая линия — Телеком, 2007. — 351 с.
- Крянев А.В.** Математические методы обработки неопределённых данных. — М.: ФИЗМАТЛИТ, 2006. — 213 с.
- Ларина Н.И.** Государственное регулирование регионального развития. — Новосибирск: ИЭОПП СО РАН, 2005. — 512 с.
- Левин И.И.** Анализ вычислительных процессов и структура на основе CF-сетей. — Махачкала: ДНЦ РАН, 2006. — 253 с.