САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

На правах рукописи

Ивашко Евгений Евгеньевич

ЗАДАЧИ НАИЛУЧШЕГО ВЫБОРА С РАЗЛАДКОЙ

Специальность 01.01.09 — дискретная математика и
математическая кибернетика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Санкт-Петербург – 2009 г.

Работа выполнена в Учреждении РАН Институт прикладных математических исследований Учреждения РАН Карельский научный центр Российской академии наук.

Научный руководитель:	доктор физико-математических наук, профессор Владимир Викторович Мазалов
Официальные оппоненты:	доктор физико-математических наук, профессор Андрей Юрьевич Гарнаев кандидат физико-математических наук Виктория Леонидовна Крепс
Ведущая организация:	Новгородский государственный университет им. Ярослава Мудрого

Защита состоится ".....".................. 200.. г. в .... час. на заседании совета Д.212.232.59 по защите докторских и кандидатских диссертаций при Санкт-Петербургском государственном университете по адресу: 199004, Санкт-Петербург, В.О., Средний пр., д.41/43, ауд. ....

С диссертацией можно ознакомиться в Научной библиотеке Санкт-Петербургского государственного университета по адресу: С.Петербург, Университетская наб., 7/9.

Автореферат разослан "....."..................... 200.. г.

Ученый секретарь
диссертационного совета
доктор физ.-мат. наук, профессор В. Д. Ногин

Общая характеристика работы

Актуальность темы. При построении моделей в области биологии, менеджмента, социологии и других наук часто возникают задачи наилучшего выбора. Эти задачи отражают важные особенности реальных процессов принятия решений в условиях неопределенности. Поэтому актуальным является как рассмотрение новых постановок задач, так и разработка прикладных моделей на их основе. На практике также нередки ситуации, когда в процессе наблюдения вероятностный закон распределения характеристик случайного процесса изменяется, что влечет дополнительные трудности при принятии решений.

В диссертационной работе проводится исследование проблем, принадлежащих классу задач наилучшего выбора с разладкой. Этот класс составляют задачи, в которых необходимо осуществить наилучший выбор в условиях ожидаемого изменения характеристик базового случайного процесса.

Цель диссертационной работы заключается в построении решений задач наилучшего выбора с разладкой методами теории оптимальной остановки и некооперативной теории игр.

В работе исследуются следующие основные задачи:

1. байесовская модель задачи наилучшего выбора с полной информацией с разладкой; наблюдатель использует многопороговую стратегию и байесовскую оценку вероятности разладки для максимизации ожидаемого значения принятой случайной величины.

2. Максимизация ожидаемого выигрыша в задаче наилучшего выбора с полной информацией с разладкой; решение строится в классе многопороговых и однопороговых стратегий.

3. Задача наилучшего выбора с полной информацией с разладкой; наблюдатель, используя однопороговую стратегию, стремится максимизировать вероятность выбора наибольшей величины из случайной последовательности.

4. Теоретико-игровая задача оптимальной остановки для m игроков; рассмотрены три варианта задачи с различными параметрами распределения случайных величин, а также обобщение задачи на случай разладки.

Научная новизна работы заключается в формулировке, исследовании и решении новых постановок задач наилучшего выбора, обобщенных на случай ожидаемого изменения вероятностных характеристик случайного процесса (разладки).

Исследована задача наилучшего выбора с полной информацией с разладкой, в которой наблюдатель с помощью однопороговой стратегии стремится максимизировать вероятность выбора максимального значения из последовательности случайных величин. Получен аналитический вид формулы для вычисления значений оптимальных порогов. Для случая большого количества наблюдений получено явное значение оптимального порога, зависящее от параметров задачи. Показано, что при бесконечно большом числе наблюдений наилучшая стратегия заключается в установке порога бесконечно близкого (снизу) к параметру распределения наблюдений после разладки. На основе представленной задачи построена модель распределения ресурсов вычислительной системы между задачами, для которых объем необходимых вычислений зависит от требуемой точности.

В модели игровой задачи оптимальной остановки исследованы три варианта задачи с различными параметрами распределения случайных величин, а также вариант задачи с разладкой. Найдены оптимальные стратегии принятия наблюдений в игре m игроков.

Рассмотрена многопороговая задача максимизации ожидаемого значения принятой случайной величины в задаче наилучшего выбора с полной информацией с разладкой. Получены рекуррентные формулы для вычисления значений оптимальных порогов при различных значениях параметров.

В задаче максимизации ожидаемого значения принятой случайной величины из последовательности наблюдений исследованы две игровые постановки. Рассмотрены частные случаи абсолютного приоритета одного из игроков и бесконечно большого числа наблюдений; построены функции выигрыша и найдены оптимальные пороговые стратегии.

Исследована байесовская модель задачи наилучшего выбора с разладкой, в которой наблюдатель имеет неполную информацию о наблюдаемых случайных величинах. Рассмотрены модели с дисконтированием выигрыша и платой за наблюдения. Предложена байесовская стратегия порогового вида, в которой порог принятия решения зависит от апостериорной оценки вероятности разладки. Описана модель распределения вычислительных ресурсов, основанная на байесовской постановке задачи наилучшего выбора с разладкой.

Практическую ценность диссертационного исследования составляют построенные математические модели, которые могут послужить основой для разработки эффективных алгоритмов работы информационных систем (в частности, систем распределенных высокопроизводительных вычислений), моделей экономических, биологических и других систем.

Положения, выносимые на защиту. На защиту выносятся следующие положения.

1. Байесовская стратегия в задаче наилучшего выбора с разладкой с неполной информацией о наблюдениях в моделях с дисконтированием выигрыша и платой за наблюдения.

2. Оптимальные стратегии в задаче наилучшего выбора с полной информацией с разладкой в классе однопороговых стратегий для задачи максимизации вероятности выбора максимального значения из последовательности случайных величины и в классе многопороговых стратегий для задачи максимизации ожидаемого принятого значения.

3. Равновесия в теоретико-игровых задачах наилучшего выбора с разладкой, в которых игроки стремятся максимизировать ожидаемое значение принятого наблюдения или вероятность выбора наибольшего значения из последовательности случайных величин.

Связь работ с научными программами, темами. Основные результаты диссертации были получены при проведении исследований в рамках работ по грантам Отделения математических наук РАН (программа "Математические и алгоритмические проблемы информационных систем нового поколения") и Российского фонда фундаментальных исследований.

Апробация работы. Результаты работы докладывались и обсуждались на следующих конференциях:

1. Международная конференция "Optimal Stopping and Stochastic Control", 22–26 августа 2005 г., Петрозаводск.

2. Российско-финская летняя школа "Dynamic Games and Multicriteria Optimization", 2–7 сентября 2006 г., Петрозаводск.

3. V Московская международная конференция по исследованию операций, посвященная 90-летию со дня рождения академика Н. Н. Моисеева, 10–14 апреля 2007 г., Москва.

4. Международная конференция "Теория игр и менеджмент", 28–29 июня 2007 г., Санкт-Петербург.

5. II летняя школа Российского журнала менеджмента, 9–20 июля 2007 г., Санкт-Петербург.

6. Международный семинар "Graduate School Seminar: Citations as impacts of research – who reads our papers?", 5–7 декабря 2007 г., Финляндия–Швеция.

7. VII Международная петрозаводская конференция "Вероятностные методы в дискретной математике", 1–6 июня 2008 г., Петрозаводск.

8. Вторая международная конференция "Теория игр и менеджмент", 26–27 июня 2008 г., Санкт-Петербург.

9. XIII Международный симпозиум "Dynamic Games and Applications", 30 июня–3 июля 2008 г., Вроцлав, Польша.

10. Третья международная конференция "Теория игр и менеджмент", 24–26 июня 2009 г., Санкт-Петербург.

11. Рабочее совещание "Сетевые игры и менеджмент (NGM–2009)", 28–30 июня 2009 г., Петрозаводск.

По материалам диссертации опубликовано 14 работ, из них 7 статей [1, 2, 3, 4, 5, 6, 7] и тезисы 7 докладов [8, 9, 10, 11, 12, 13, 14].

Структура и объем работы. Диссертационная работа состоит из введения, 3 глав, заключения и списка литературы. Общий объем диссертации составляет 107 страниц. Список литературы включает 90 наименований.

Содержание работы

Во **введении** отражена актуальность работы, приведен краткий обзор литературы по теме диссертации, поставлена цель исследования, обоснована новизна работы, сформулированы положения, выносимые на защиту, показана практическая ценность полученных результатов. Кроме того, описана базовая модель задачи наилучшего выбора с разладкой, использующаяся в последующих главах диссертации.

В общем виде задачи наилучшего выбора с разладкой описываются следующим образом (*модель (А)*): рассмотрим производящую систему, генерирующую последовательность из n независимых случайных величин $X_1, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$. В случайный момент θ происходит разладка и закон распределения случайных величин изменяется.

Величины $X_1, X_2, ..., X_{\theta-1}$ описываются абсолютно непрерывной функцией распределения $F_1(x)$, а величины $X_\theta, X_{\theta+1}, ..., X_n$ — абсолютно непрерывной функцией распределения $F_2(x)$. Будем говорить, что до разладки система находится в состоянии S_1, а после — в состоянии S_2.

Момент разладки θ имеет геометрическое распределение с параметром α $(0 \leq \alpha \leq 1)$. То есть изначально система находится в состоянии S_1, а перед генерацией каждой новой случайной величины состояние системы изменяется согласно следующей матрице переходов:

	S_1	S_2
S_1	α	$1-\alpha$
S_2	0	1

Наблюдателю (или нескольким наблюдателям) известны функции распределения случайных величин в состояниях S_1 и S_2, а также вероятность разладки $1-\alpha$, но истинное состояние системы остается неизвестным. После получения каждого из значений случайных величин наблюдателю требуется решить: принять значение случайной величины (и закончить процесс наблюдений) или отклонить. При этом нельзя ни отвергнуть все наблюдения, ни вернуться к наблюдению, которое было отвергнуто ранее. Задача наблюдателя заключается в том, чтобы принять максимальное значение из последовательности случайных величин.

В случае нескольких наблюдателей задачу необходимо исследовать теоретико-игровыми методами. Такую теоретико-игровую модель наилучшего выбора с разладкой будем называть *моделью (B)*.

Особенностью (и дополнительной сложностью) задач наилучшего выбора с разладкой является то, что на неопределенность, связанную со значениями будущих наблюдений, накладывается неопределенность, связанная с текущим состоянием системы (произошла разладка или нет). На основе *моделей (A)* и *(B)* в диссертационном исследовании рассмотрен ряд постановок задач, отличающихся целями наблюдателей, наличием конкурентов при выборе, возможностями варьирования порогов в процессе наблюдения и т.д.

В диссертационном исследовании рассматриваются два класса стратегий наблюдателя: однопороговые и многопороговые стратегии. Многопороговая стратегия — это такая стратегия, при которой наблюдатель на каждом шаге i $(i = 1, ..., n)$ устанавливает порог q_i и принимает случайную величину X_i в случае, если $x_i > q_i$, отвергая ее в противном случае. Однопороговая стратегия — это такая многопороговая стратегия, при которой $q_i = q$, $i = 1, ..., n$.

Первая глава посвящена исследованию байесовской постановки задачи наилучшего выбора с разладкой.

В разделе 1 представлена постановка задачи: пусть наблюдатель действует в условиях *модели (A)*. Информация о наблюдениях неполна — на каждом шаге известно лишь о том, превышает или нет полученная случайная величина установленный наблюдателем порог. Цель наблюдателя — максимизировать среднее выбранное значение из случайной последовательности наблюдений. Решение ищется в классе байесовских пороговых правил следующего вида. Перед шагом k, оценив апостериорную вероятность нахождения системы в состоянии S_1, наблюдатель устанавливает порог $q = q_k(x_1, \ldots, x_{k-1})$ и принимает наблюдение x_k, если оно превышает этот порог. В противном случае наблюдение отвергается. То есть наблюдатель использует многопороговую стратегию с оценкой апостериорной вероятности возникновения разладки.

Описанная задача относится к классу задач наилучшего выбора с разладкой и неполной информацией о наблюдениях.

В разделе 2 рассмотрен вариант задачи, в котором выигрыш наблюдателя на каждом следующем шаге дисконтируется с коэффициентом λ. Согласно условию задачи, наблюдатель не знает истинного состояния системы (S_1 или S_2). Однако эта неопределенность может быть компенсирована с помощью оценки вероятности разладки. По значениям наблюдений вычисляется апостериорная (после получения информации о том, что $x \le q$) вероятность нахождения системы в состоянии S_1:

$$\pi_q = \pi(q) = P\{S_1 | x \le q\} = \frac{P(S_1)P(x \le q | S_1)}{P(x \le q)} = \frac{\alpha \pi F_1(q)}{F_\pi(q)}.$$

Здесь $q = q_i$ – порог, установленный наблюдателем за i шагов до завершения наблюдений, π – априорная вероятность нахождения системы в состоянии S_1 (до получения информации о том, что $x \le q$), а $F_\pi(q) = \pi F_1(q) + \overline{\pi} F_2(q)$, где $\overline{\pi} = 1 - \pi$.

Пусть $v_i(\pi)$ – это выигрыш наблюдателя за i шагов до окончания наблюдений. Тогда уравнение оптимальности с учетом коэффициента дисконтирования λ будет следующим:

$$\begin{cases} v_i(\pi) & = & \max\limits_q \left[\lambda v_{i-1}(\pi_q) F_\pi(q) + \pi E_1(q) + \overline{\pi} E_2(q)\right], \ i \ge 1, \\ v_0(\pi) & = & 0 \ \forall \pi. \end{cases}$$

Здесь $E_k(q) = \int\limits_q^\infty x dF_k(x)$, $k = 1, 2$.

Обозначим выражение в квадратных скобках через $V_i(\pi, q)$. Тогда за i шагов до окончания наблюдений оптимальный порог q_i будет определяться решением оптимизационной задачи $q_i = \arg\max\limits_q V_i(\pi, q)$.

В диссертационной работе мы ограничились построением оптимального правила остановки в классе стационарных стратегий, когда порог q выбора наблюдений зависит только от π и не зависит от номера шага.

Верна следующая теорема о виде функции выигрыша.

Теорема 1.1 *При фиксированном значении q последовательность функций $V_i(\pi, q)$ можно представить в виде*

$$V_i(\pi, q) = \pi G_i(q) + T_i(q),$$

где G_i и T_i удовлетворяют следующим рекуррентным соотношениям:

$$\begin{aligned}
&G_i(q) = \lambda\alpha F_1(q) G_{i-1}(q) + \lambda T_{i-1}(q)(F_1(q) - F_2(q)) + E_1(q) - E_2(q),\ i \geq 1,\\
&T_i(q) = F_2(q)\lambda T_{i-1}(q) + E_2(q),\ i \geq 1,\\
&G_0 = T_0 = 0.
\end{aligned}$$

При $i \to \infty$ компоненты $G_i(q)$ и $T_i(q)$ сходятся соответственно к $G(q)$ и $T(q)$:

$$\begin{aligned}
G(q) &= \tfrac{E_1(q)(1-\lambda F_2(q)) - E_2(q)(1-\lambda F_1(q))}{(1-\lambda\alpha F_1(q))(1-\lambda F_2(q))},\\
T(q) &= \tfrac{E_2(q)}{1-\lambda F_2(q)}.
\end{aligned}$$

Соответственно $V_i(\pi, q) \longrightarrow V(\pi, q)$, где

$$V(\pi, q) = \pi G(q) + T(q).$$

Порог принятия наблюдений $q = q(\pi)$ определяется как

$$q = \arg\max_q V(\pi, q).$$

Верна следующая теорема.

Теорема 1.2 *Для любых непрерывных функций распределения $F_1(x)$ и $F_2(x)$ таких, что $F_1(x)$ стохастически доминирует $F_2(x)$, существует решение байесовской задачи наилучшего выбора с разладкой.*

Используя те же рассуждения, что и в случае дисконтирования выигрыша, в разделе 2 строится байесовская стратегия для случая платы за наблюдения.

Обозначим c — значение, которое наблюдатель платит за каждое новое полученное наблюдение. Цель наблюдателя — максимизировать ожидаемый выигрыш с учетом платы за наблюдения. Выигрыш наблюдателя находим следующим образом:

Повторяя рассуждения, используемые при доказательстве теоремы раздела 2, можно показать, что в классе стационарных стратегий выражение для $V_i(\pi, q)$ имеет аналогичный вид:

$$V_i(\pi, q) = \pi G_i(q) + T_i(q),$$

где $G_i(q)$ и $T_i(q)$ удовлетворяют следующим рекуррентным соотношениям:

$$\begin{aligned}
&G_i(q) = \alpha F_1(q) G_{i-1}(q) + (T_{i-1}(q) - c)(F_1(q) - F_2(q)) + E_1(q) - E_2(q),\ i \geq 1,\\
&T_i(q) = F_2(q)(T_{i-1}(q) - c) + E_2(q),\ i \geq 1,\\
&G_0 = T_0 = 0.
\end{aligned}$$

При этом при $k \to \infty$ компоненты $G_k(q)$ и $T_k(q)$ предельной функции выигрыша сходятся соответственно к $G(q)$ и $T(q)$ вида

$$\begin{aligned}
G(q) &= \tfrac{E_1(q)(1-F_2(q)) - E_2(q)(1-F_1(q)) - c(F_1(q) - F_2(q))}{(1-\alpha F_1(q))(1-F_2(q))},\\
T(q) &= \tfrac{E_2(q) - cF_2(q)}{1-F_2(q)}.
\end{aligned}$$

В разделе 4 на основе байесовской задачи наилучшего выбора представлена модель распределения ресурсов вычислительной системы.

Во **второй главе** рассмотрены две задачи наилучшего выбора с разладкой, критерием оптимальности в которых является максимизация ожидаемого значения принятой случайной величины.

В разделе 1 рассматривается следующая многопороговая задача наилучшего выбора с полной информацией с разладкой. Пусть в условиях *модели (A)* наблюдателю требуется максимизировать ожидаемое значение принятой случайной величины (получаемое им в качестве выигрыша). Известно, что в состоянии S_1 система генерирует случайные величины, равномерно распределенные на отрезке $[0,1]$, в состоянии S_2 — равномерно распределенные на отрезке $[0,b]$.

Решение ищется в классе многопороговых стратегий; значения порогов для каждого шага устанавливаются до начала наблюдений и не могут быть изменены в дальнейшем (т.е. значение порога зависит лишь от номера шага и не зависит от значений полученных наблюдений).

Описанная задача является обобщением на случай разладки проблемы, рассмотренной в работе Мозера (1956).

В п. 1 рассмотрен вариант задачи, в котором при возникновении разладки значения наблюдаемых случайных величин становятся стохастически меньше, чем в исходном состоянии S_1. В п. 2 рассмотрен вариант задачи, в котором в результате разладки новое распределение случайных величин становится стохастически больше первоначального.

Представленные выше выкладки для значений параметров $b > 1$ и $b < 1$ обобщаются в следующей теореме.

Теорема 2.1 *В многопороговой задаче наилучшего выбора с разладкой значение оптимального порога определяется рекуррентными соотношениями*

$$q_n = 0,$$

$$q_{n-k} = \begin{cases} \frac{1-\alpha^{n-k}}{b}\left[I(q_{n-k+1} < b)\frac{b^2+q^2_{n-k+1}}{2} + I(q_{n-k+1} \geq b)bq_{n-k+1}\right] + \alpha^{n-k}\frac{1+q^2_{n-k+1}}{2}, b < 1, \\ I(q_{n-k+1} < 1)\alpha^{n-k}\frac{1+q^2_{n-k+1}}{2} + I(q_{n-k+1} > 1)\alpha^{n-k}q_{n-k+1} + \frac{1-\alpha^{n-k}}{b}\frac{b^2+q^2_{n-k+1}}{2}, b > 1 \end{cases}$$

для $1 \leq k \leq n-1$.

При $\alpha = 1$ или $b = 1$ формулы, определяющие значения оптимальных порогов, сводятся к "последовательности Мозера".

Раздел 2 посвящен игровой задаче наилучшего выбора с разладкой, в которой каждый из двух игроков, используя однопороговую стратегию, стремится максимизировать ожидаемое значение принятой им случайной величины. Игроки соперничают за наблюдения, что выражается вероятностными приоритетами, с которыми каждая новая случайная величина передается тому или другому игроку. Решение ищется в классе однопороговых стратегий. В п. 1 раздела 2 рассмотрена следующая неконкурентная постановка игровой задачи наилучшего выбора с разладкой: пусть в условиях *модели (B)* каждый из двух наблюдателей (игрок I и игрок II) стремится максимизировать ожидаемое значение принятой случайной величины, получаемое им в качестве выигрыша. Известно, что в состоянии S_1 система генерирует случайные величины, равномерно распределенные на отрезке $[0, 1]$, в состоянии S_2 — равномерно распределенные на отрезке $[0, b]$. На каждом шаге i ($i = 1, .., n$) случайная величина X_i передается одному из игроков согласно заданным приоритетам: с вероятностью p наблюдение получает игрок I, с вероятностью $1 - p$ — игрок II. Таким образом, каждый из игроков получает лишь часть случайной последовательности $X_1, X_2..., X_n$. Если один из игроков на некотором шаге k ($1 \leq k < n$) принял наблюдение

X_k, то другому игроку каждое новое наблюдение из оставшейся случайной последовательности $X_{k+1}, ..., X_n$ все также предоставляется с вероятностью, соответствующей его приоритету (это условие задачи и подразумевается под "неконкурентностью"). Если игрок за n шагов не принял наблюдения, то он получает выигрыш, равный нулю.

Решение ищется в классе однопороговых стратегий (пороги q_1 и q_2 для игроков I и II соответственно), значения порогов устанавливаются перед началом наблюдений и не могут быть изменены в дальнейшем.

Предположим, что игрок II имеет больший приоритет выбора наблюдений (то есть $p \leq \frac{1}{2}$); если это не так, то перенумеруем игроков таким образом, чтобы это условие выполнялось. Очевидно, что в таком случае ожидаемый выигрыш (а значит и значение оптимального порога) игрока I также будет меньше ожидаемого выигрыша игрока II.

Оптимальные значения порогов q_1^* и q_2^* будут определяться решением следующей оптимизационной задачи

$$\begin{cases} q_1^* = \arg\max\limits_{q_1} H_n^{S_1}(q_1, q_2^*) \\ q_2^* = \arg\max\limits_{q_2} H_n^{S_1}(q_1^*, q_2). \end{cases}$$

При увеличении количества наблюдений ($n \to \infty$) функция выигрыша игрока I выглядит следующим образом:

$$\begin{gathered} H^{S_1}(q_1, q_2) = \lim_{k\to\infty} H_k^{S_1}(q_1, q_2) = \\ \Big[\Big(\alpha\tfrac{1-q_1^2}{2} + I(q_1 < b)(1-\alpha)\tfrac{b^2-q_1^2}{2(b-q_1)}\Big)\,\tfrac{\alpha(1-p)(1-q_2)}{1-\alpha q_1} + \alpha p\tfrac{1-q_1^2}{2} + \\ \Big(I(q_2 < b)(1-p)(b-q_2)\tfrac{b^2-q_1^2}{2(b-q_1)} + \\ I(q_1 < b)p\tfrac{b^2-q_1^2}{2}\Big)\,\tfrac{(1-\alpha)}{b-(pq_1+(1-p)min(q_2,b))}\Big]\,\tfrac{1}{1-\alpha(pq_1+(1-p)q_2)}. \end{gathered}$$

Тот же предельный переход в формуле функции выигрыша игрока II дает

$$\begin{gathered} H^{S_1}(q_1, q_2) = \lim_{k\to\infty} H_k^{S_1}(q_1, q_2) = \\ \Big[\tfrac{\alpha p(1-q_1)}{1-\alpha q_2}\Big(\alpha\tfrac{1-q_2^2}{2} + (1-\alpha)I(q_2 < b)\tfrac{b^2-q_2^2}{2(b-q_2)}\Big) + \alpha(1-p)\tfrac{1-q_2^2}{2} + \\ \tfrac{I(q_2<b)(1-\alpha)}{b-pq_1-(1-p)q_2}\Big(\tfrac{b^2-q_2^2}{2(b-q_2)}p(b-q_1) + (1-p)\tfrac{b^2-q_2^2}{2}\Big)\Big]\,\tfrac{1}{1-\alpha(pq_1+(1-p)q_2)}. \end{gathered}$$

В диссертационной работе рассмотрен также случай, когда один из игроков имеет абсолютный приоритет принятия наблюдений (то есть $p = 0$). Фактически, это означает, что игрок II действует независимо от своего оппонента. При этом игрок I должен принимать во внимание то, что первое наблюдение с большим значением перейдет игроку II.

В случае, если игрок II имеет абсолютный приоритет, а число наблюдений велико ($n \to \infty$ и $p = 0$) функция выигрыша игрока I принимает следующий вид:

$$H^{S_1}(q_1, q_2) = \lim_{k\to\infty} H_k^{S_1}(q_1, q_2) =$$
$$\left[\left(\alpha \tfrac{1-q_1^2}{2} + I(q_1 < b)(1-\alpha)\tfrac{b^2-q_1^2}{2(b-q_1)}\right)\tfrac{\alpha(1-q_2)}{1-\alpha q_1} + I(q_2 < b)(1-\alpha)\tfrac{b+q_1}{2}\right]\tfrac{1}{1-\alpha q_2}.$$

Оптимальная стратегия игрока II и соответствующий ожидаемый выигрыш определяются следующей теоремой.

Теорема 2.2 *В неконкурентной постановке игровой задачи наилучшего выбора с разладкой при бесконечном числе наблюдений и абсолютном приоритете оптимальная стратегия II игрока определяется как*

$$q_2^* = \begin{cases} \frac{1-\sqrt{1-\alpha^2}}{\alpha} & , \text{ если } \frac{1-\sqrt{1-\alpha^2}}{\alpha} \geq b \\ b-0 & \text{иначе,} \end{cases}$$

при этом ожидаемый выигрыш равен

$$H^{S_1}(q_1, q_2^*) = \max\left(\alpha\frac{1-b^2}{2(1-\alpha b)} + \frac{b(1-\alpha)}{1-\alpha b}, I(b < \frac{1-\sqrt{1-\alpha^2}}{\alpha})\frac{1-\sqrt{1-\alpha^2}}{\alpha}\right).$$

В п. 2 раздела 2 рассмотрена следующая игровая постановка задачи наилучшего выбора с разладкой: пусть каждый из двух наблюдателей (игрок I и игрок II) в условиях *модели (B)* стремится максимизировать ожидаемое значение принятой случайной величины (получаемое им в качестве выигрыша). Известно, что в состоянии S_1 система генерирует случайные величины равномерно распределенные на отрезке $[0, 1]$, в состоянии S_2 — равномерно распределенные на отрезке $[0, b]$. При этом игроки конкурируют за наблюдения — одна и та же случайная величина не может быть принята обоими игроками. Если оба игрока готовы принять наблюдение, то с вероятностью p приоритет выбора предоставляется игроку I, а с вероятностью $(1-p)$ — игроку II. Если один из игроков на каком-то шаге принял наблюдение, то другому игроку предоставляется возможность без конкуренции выбрать наблюдение из оставшейся случайной последовательности.

Решение ищется в классе однопороговых стратегий; значения порогов устанавливаются перед началом наблюдений и не могут быть изменены в дальнейшем.

Описанная постановка игровой задачи наилучшего выбора с разладкой названа конкурентной, так как моделирует поведение двух крупных экономических субъектов (компаний), конкурирующих в одной области (продажа или покупка товаров, оказание услуг и др.). Как правило, экономические субъекты являются неидентичными, то есть имеют ряд отличий (например, уровень цен, узнаваемость бренда, репутация и т.д.), влияющих на занимаемую ими долю рынка — число клиентов, обращающихся в ту или иную фирму. Эта доля рынка в модели выражается вероятностным приоритетом p. В отличие от предыдущей задачи, если уровень запросов клиента не подходит одной компании, то клиент может быть обслужен другой компанией. Также, если одна из фирм не имеет больше возможностей обслуживать новых клиентов (распродан весь товар, занято время оборудования и т.п.), то они обращаются в другую фирму.

Далее будем считать, что игрок II имеет больший приоритет выбора наблюдений ($p \le \frac{1}{2}$), а значит, и значение оптимального порога игрока I также будет меньше значения оптимального порога игрока II.

При большом количестве наблюдений ($n \to \infty$) функция выигрыша игрока I дает

$$
\begin{gathered}
H^{S_1}(q_1, q_2) = \lim_{k\to\infty} H_k^{S_1}(q_1, q_2) = \Big[\alpha \tfrac{q_2^2 - q_1^2}{2} + \alpha p \tfrac{1-q_2^2}{2} + \\
\alpha(1-p)(1-q_2) \Big[\alpha \tfrac{1-q_1^2}{2} + (1-\alpha) I(q_1 < b) \tfrac{b+q_1}{2}\Big] \tfrac{1}{1-\alpha q_1} + \\
(1-\alpha) I(q_1 < b) \tfrac{1}{b} \Big[\Big(I(q_2 < b) \Big(\tfrac{q_2^2 - q_1^2}{2} + p \tfrac{b^2 - q_2^2}{2} + (1-p)(b - q_2) \tfrac{b+q_1}{2}\Big) + \\
I(q_2 > b) \tfrac{b^2 - q_1^2}{2}\Big) \tfrac{1}{b - q_1}\Big]\Big] \tfrac{1}{1-\alpha q_1}.
\end{gathered}
$$

Тот же предельный переход в формуле функции выигрыша игрока II приводит к

$$
\begin{gathered}
H^{S_1}(q_1, q_2) = \lim_{k\to\infty} H_k^{S_1}(q_1, q_2) = \\
\Big[\Big[\alpha \tfrac{1-q_2^2}{2} + (1-\alpha) I(q_2 < b) \tfrac{b+q_2}{2}\Big] \tfrac{1}{1-\alpha q_2} (q_2 - q_1 + p(1-q_2)) + \\
\alpha(1-p) \tfrac{1-q_2^2}{2} + (1-\alpha) I(q_2 < b) \Big(\tfrac{b+q_2}{2}(q_2 - q_1) + \tfrac{b^2 - q_2^2}{2}\Big) \tfrac{1}{b - q_1}\Big] \tfrac{1}{1-\alpha q_1}.
\end{gathered}
$$

В диссертационной работе рассмотрен также случай абсолютного приоритета ($p = 0$). Применительно к задаче конкурентного поведения в экономике это означает, что игрок II доминирует на рынке, и значение его оптимального порога не зависит от поведения игрока I.

В **третьей главе** исследована проблема максимизации вероятности выигрыша в задаче наилучшего выбора с полной информацией с разладкой и в игровой задаче оптимальной остановки.

В разделе 1 рассмотрена задача наилучшего выбора с полной информацией с разладкой, в которой наблюдателю требуется максимизировать вероятность удачного выбора. Постановка задачи следующая: при условиях *модели (A)* наблюдатель стремится максимизировать вероятность выбора наибольшего значения из последовательности случайных величин, используя однопороговую стратегию. Известно, что в состоянии S_1 случайные величины, наблюдаемые игроком, распределены равномерно на отрезке [0,1], а в состоянии S_2 — равномерно на отрезке [0,b] ($b < 1$ — параметр задачи).

Эта задача является упрощенным аналогом классической задачи наилучшего выбора с полной информацией — отличие заключается в том, что в представленной постановке наблюдатель использует однопороговую стратегию принятия наблюдения, а в задаче "о продаже недвижимости" — многопороговую.

Следующие формулы определяют вероятность удачного выбора как функции от параметров n, α и b при значениях порога $q \geq b$ (функция $P_1(n,q,\alpha,b)$) и $q < b$ (функция $P_2(n,q,\alpha,b)$):

$$P_1(n,q,\alpha,b) = \alpha^n \sum_{j=0}^{n-1} \frac{q^j(1-q^{n-j})}{n-j} + \frac{1-\alpha}{1-\alpha q} \sum_{j=1}^{n} \alpha^j \frac{(1-q^j)(1-(\alpha q)^{n-j+1})}{j},$$

$$\begin{aligned} P_2(n,q,\alpha,b) = \alpha^n \sum_{j=0}^{n-1} \tfrac{q^j(1-q^{n-j})}{n-j} + \quad & (1-\alpha)\tfrac{1-(\alpha b)^n}{1-\alpha b} \sum_{j=0}^{n-1} \tfrac{(q/b)^j(1-(q/b)^{n-j})}{n-j} + \\ & (1-\alpha) \sum_{i=0}^{n-2} (\alpha q)^i \sum_{j=1}^{n-i-1} \tfrac{\alpha^j-(\alpha b)^j}{j} \end{aligned}$$

Порог выбора q^*, максимизирующий вероятность принятия наибольшего значения из последовательности случайных величин, определяется как

$$q^* = \arg\max \left[\max_{q \geq b} P_1(n,q,\alpha,b), \max_{q<b} P_2(n,q,\alpha,b) \right],$$

при этом вероятность удачного выбора вычисляется следующим образом:

$$P(n,q^*,\alpha,b) = \max \left(\max_{q \geq b} P_1(n,q,\alpha,b), \max_{q<b} P_2(n,q,\alpha,b) \right).$$

При большом числе наблюдений верна следующая лемма.

Лемма 3.1 *Для любых значений параметров* α *и* b*:* $0 < \alpha, b < 1$ *выполняется следующее неравенство:*

$$P_2(q_2^*, \alpha, b) > P_1(q_1^*, \alpha, b).$$

С учетом леммы, решение задачи при большом числе наблюдений определяется следующей теоремой.

Теорема 3.1 *В задаче максимизации вероятности выбора наибольшего значения из последовательности случайных величин в задаче наилучшего выбора с разладкой при бесконечном числе наблюдений значение оптимального порога определяется как*

$$q^* = b,$$

при этом вероятность удачного выбора

$$P(q^*, \alpha, b) = \frac{1-\alpha}{1-\alpha b}\left[0.517 + ln\frac{(1-b\alpha)}{1-\alpha}\right].$$

В п. 3 представлена модель распределения ресурсов вычислительной системы для задач, допускающих динамическую точность решения. В основу модели положена представленная выше задача наилучшего выбора с полной информацией с разладкой.

В разделе 2 рассмотрена следующая задача: каждый из m игроков независимо от остальных участников получает в качестве наблюдения значение случайной величины X_1. Игрок должен принять решение: остановиться на этом наблюдении или продолжить и получить еще одно наблюдение — значение случайной величины X_2. Побеждает тот игрок, чья сумма очков (равная значению случайной величины X_1, если игрок принял решение остановиться на первом наблюдении, или сумме значений случайных величин X_1 и X_2, если игрок решил продолжить) окажется наиболее близкой, но не превышающей 1. Если очки всех игроков превысили 1, то побеждает тот игрок, чья сумма очков оказалась наименьшей. Игроки не знают ни значений наблюдений, ни решений, принятых другими игроками. Целью каждого из игроков является максимизация вероятности собственного выигрыша. Для принятия решения об остановке или продолжении игроки используют однопороговую стратегию.

При равномерном распределении наблюдений на отрезке $[0,1]$ оптимальное значение порога определяется из уравнения

$$q^{2(m-1)} = \frac{1-q^{2m}}{m(q+1)} + \frac{q^{2m-1}}{2^{m-1}(2m-1)}.$$

В случае, если наблюдения распределены равномерно на отрезке $[0,b]$, где $b>1$, уравнение оптимальности выглядит следующим образом:

$$\left[bq-q+q^2+b(b-1)\right]^{m-1} =$$
$$\frac{1}{b}\left[\frac{b^{2m}-(bq-q+q^2+b(b-1))^m}{m(b+q)} + \frac{(q(b-1+\frac{q}{2})+b(b-1))^m-(\frac{q^2}{2})^m}{(q+b)m} + \frac{q^{2m-1}}{2^{m-1}(2m-1)}\right].$$

При $b=2$ (следовательно, и при $b>2$) значение оптимального порога $q=0$. Это означает, что игрок должен остановиться на первом же наблюдении вне зависимости от полученного значения.

Теперь рассмотрим вариант задачи, в котором случайная величина X_1 имеет равномерное распределение на отрезке $[0,1]$, а X_2 — равномерное распределение на отрезке $[0,b]$, где $b>1$. Тогда значение оптимального порога q определяется из уравнения

$$\left[qb+q^2-q\right]^{m-1} = \frac{1}{b}\left[\frac{b^m-(qb+q^2-q)^m}{m(b+q)} + q^{m-1}\frac{(b-1+\frac{q}{2})^m-(\frac{q}{2})^m}{m} + \frac{q^{2m-1}}{2^{m-1}(2m-1)}\right].$$

При увеличении b, значение оптимального порога стремится к нулю.

Рассмотрим обобщение игры, описанной в *модели B*, на случай разладки. Пусть изначально распределение наблюдений является равномерным на отрезке $[0,1]$. В случайный момент времени (перед первым или перед вторым наблюдением) может произойти разладка, и распределение случайных величин изменится на равномерное на отрезке $[0,b]$, где $b>1$. Момент разладки — это случайная величина, имеющая геометрическое распределение с параметром α. Разладка происходит для всех наблюдателей одновременно.

Решение этой игры можно представить как комбинацию решений трех рассмотренных ранее игр:

$$\alpha^2 q^{2(m-1)} + \alpha(1-\alpha)\left[qb+q^2-q\right]^{m-1} + (1-\alpha)\left[bq-q+q^2+b(b-1)\right]^{m-1} =$$
$$\alpha^2\frac{1-q^{2m}}{m(q+1)} + \alpha(1-\alpha)\frac{1}{bm}\left[\frac{b^m-(qb+q^2-q)^m}{(b+q)} + q^{m-1}\left((b-1+\frac{q}{2})^m-(\frac{q}{2})^m\right)\right] +$$
$$(1-\alpha)\frac{1}{b(q+b)m}\left[b^{2m}-(bq-q+q^2+b(b-1))^m + (q(b-1+\frac{q}{2})+b(b-1))^m-(\frac{q^2}{2})^m\right]$$
$$+\left(\alpha^2-\frac{\alpha^2}{b}+\frac{1}{b}\right)\frac{q^{2m-1}}{2^{m-1}(2m-1)}.$$

Список литературы

[1] Мазалов В. В., Ивашко Е. Е. Байесовская модель в задаче наилучшего выбора с "разладкой" // Вестник СПбГУ. Серия 10. – 2009. – Вып. 4. – С. 142–151, (вклад диссертанта 50%).

[2] Мазалов В. В., Ивашко Е. Е. Задача наилучшего выбора с полной информацией с разладкой // Обозрение прикладной и промышленной математики. – 2007. – Т. 14, вып. 2. – С. 215–224, (вклад диссертанта 50%).

[3] Мазалов В. В., Ивашко Е. Е. Задача выбора оптимального объема запрашиваемых ресурсов в условиях ожидаемой "разладки" сервера // Системы управления и информационные технологии. – 2009. – №1.2(35). – С. 249–252, (вклад диссертанта 50%).

[4] Ивашко Е. Е. Игровая задача наилучшего выбора с разладкой с вероятностными приоритетами // Сборник трудов ИПМИ. – 2008. – Вып. 8. – С. 44–52.

[5] Ивашко Е. Е. Многопороговая задача наилучшего выбора с полной информацией с разладкой // Сборник трудов ИПМИ. – 2007. – Вып. 7. – С. 11–15.

[6] Ивашко Е. Е. Многопороговая модель определения оптимального объема запрашиваемых ресурсов при ожидаемой "разладке" сервера // Информационные технологии моделирования и управления. – 2009. – №2(54). – С. 261–264.

[7] Ivashko E. E. Random priority two-person full-information best-choice game with disorder // Contributions to game theory and management. Collected papers presented on the International Conference Game Theory and Management / Editors Leon A. Petrosjan, Nikolay A. Zenkevich. – SPb.: Graduate Schooll of Management, SPbSU, 2007. – Vol. II. – P. 179–187.

[8] Мазалов В. В., Ивашко Е. Е. Задача наилучшего выбора с неполной информацией с разладкой. Тезисы доклада // Обозрение прикладной и промышленной математики. – 2008. – Т. 15, Вып. 3. – С. 553, (вклад диссертанта 50%).

[9] Ivashko E. E. Random priority zero-sum best choice game with disorder // Networking games and management. Extended abstracts. – Petrozavodsk: IAMR KRC RAS, 2009. – P. 26–27.

[10] Ivashko E., Mazalov V. Best choice game with disorder // Game theory and management. Collected abstracts of papers presented on the Second International Conference Game Theory and Management / Editors Leon A. Petrosjan, Nikolay A. Zenkevich. – SPb.: Graduate School of Management, SPbSU, 2008. –P. 88–89, (вклад диссертанта 50%).

[11] Mazalov V. V., Ivashko E. E. Best-choice problem with disorder // Proceedings of V Moscow International Conference on Operations Research (ORM2007), dedicated to the outstanding Russian scientists Nikita N. Moiseev 90th birthday: Moscow, 2007. – P. 174, (вклад диссертанта 50%).

[12] Mazalov V. V., Ivashko E. E. Best-choice problem with disorder // Proceedings of Dynamic Games and Multicriteria Optimization (DGMO-2006): Petrozavodsk, 2006. – P. 78, (вклад диссертанта 50%).

[13] Mazalov V. V., Ivashko E. E. Best choice problem with disorder and imperfect observations // Scientific papers of the Institute of Mathematics and Computer Science of Wroclaw University of Technology, "13th International Symposium on Dynamic Games and Applications", Wroclaw, Seria: Conferences 5. – N 26. – 2008. – P. 160, (вклад диссертанта 50%).

[14] Mazalov V. V., Ivashko E. E. Random priority zero-sum best choice game with disorder and imperfect observation // Game theory and management. Collected abstracts of papers presented on the Third International Conference Game Theory and Management / Editors Leon A. Petrosjan, Nikolay A. Zenkevich. – SPb.: Graduate School of Management, SPbSU, 2009. –P. 166–167, (вклад диссертанта 50%).