

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

Сысоев Сергей Сергеевич

РАНДОМИЗИРОВАННЫЕ АЛГОРИТМЫ
СТОХАСТИЧЕСКОЙ ОПТИМИЗАЦИИ
И ИХ ПРИМЕНЕНИЕ ДЛЯ ПОВЫШЕНИЯ
ЭФФЕКТИВНОСТИ РАБОТЫ
ВЫЧИСЛИТЕЛЬНЫХ КОМПЛЕКСОВ И
СЕТЕЙ

05.13.11 — Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Санкт-Петербург
2005

Работа выполнена на кафедре системного программирования математико-механического факультета Санкт-Петербургского Государственного Университета.

Научный руководитель: доктор физико-математических наук,
доцент Граничин Олег Николаевич

Официальные оппоненты:

доктор физико-математических наук,
профессор Мелас Вячеслав Борисович
кандидат технических наук
Кудинов Александр Михайлович

Ведущая организация: Санкт-Петербургский институт
информатики и автоматизации РАН

Защита диссертации состоится “ ____ ” _____ 2005 года в ____ часов на заседании диссертационного совета Д 212.232.51 по защите диссертаций на соискание ученой степени доктора наук при Санкт-Петербургском государственном университете по адресу 198504, Санкт-Петербург, Старый Петергоф, Университетский пр., 28, Математико-механический факультет.

С диссертацией можно ознакомиться в Научной библиотеке Санкт-Петербургского государственного университета по адресу: 199034, Санкт-Петербург, Университетская наб. 7/9.

Автореферат разослан “ ____ ” _____ 2005 г.

Ученый секретарь
диссертационного совета
доктор физико-математических наук

Б. К. Мартыненко

Общая характеристика работы

Актуальность темы. Задача поиска минимума (максимума) некоторой функции (функционала)

$$f(x) \rightarrow \min_x$$

известна уже давно. Довольно большой класс задач сводится к ее решению. Часто порядок полученных уравнений и число неизвестных таковы, что аналитический поиск решения становится практически невозможным. На самом деле, практическая ценность аналитического решения не так уж высока — оно все равно будет искажено при использовании (например, за счет ограниченной разрядности вычислительных машин, или неточности измерительных приборов).

В случае непрерывной дифференцируемости функции, задача сводится к поиску корней ее производной (или точек, в которых градиент обращается в ноль). Однако если общий вид исследуемой функции неизвестен, задача становится качественно иной. В тоже время существует тип алгоритмов, позволяющих находить решения довольно широкого класса задач с любой заранее заданной точностью. Применение этих алгоритмов не требует особой изобретательности и не сильно зависит от вида функционала (при условии принадлежности определенному классу применимости). Такого рода универсальность естественным образом влечет простоту реализации этих алгоритмов на вычислительных машинах, а итеративная природа позволяет уточнять полученную оценку с каждой новой итерацией. Речь идет о рекуррентных алгоритмах стохастической оптимизации.

В 1952 году, в работе Кифера-Вольфовица (*Ann. Math. Statist.*, vol. 23), была предложена процедура оптимизации функционала, градиент которого был неизвестен, и наблюдателю были доступны лишь зашумленные измерения y_n . В работе предполагалось, что помехи измерения взаимонезависимы, одинаково распределены и центрированы (в среднем равны нулю). Эти ограничения предоставляют возможность получать оценку функционала f в некоторой точке путем многократного измерения и усреднения результата.

Ранее в 1951 году вышла статья Роббинса и Монро (Ann. Math. Statist., vol. 22), в которой решалась задача поиска корня функции, измеряемой с помехами, на которые накладывались точно такие же ограничения. Роббинс и Монро воспользовались модификацией градиентного метода и отказались от усреднения на каждом шаге, предложив алгоритм

$$x_{n+1} = x_n - \alpha_n y_n.$$

Усреднение в предложенном методе происходило неявным образом, за счет достаточно медленного стремления к нулю параметра алгоритма α_n . Кифер и Вольфовиц также отказались от явного усреднения и применили процедуру Роббинса-Монро к конечноразностной аппроксимации градиента.

Благодаря стремительному развитию вычислительной техники процедуры Роббинса-Монро и Кифера-Вольфовица получили широкое распространение. Большую роль в пропаганде подобных методов сыграл Я.З.Цыпкин. В своих книгах “Адаптация и обучение в автоматических системах” и “Основы теории обучающихся систем” он показал широкую применимость рекуррентных стохастических алгоритмов в задачах оценивания, идентификации, распознавания образов, оптимизации и управления. Позже была оценена эффективность таких алгоритмов и найдены их оптимальные и робастные варианты.

К настоящему времени методика исследования свойств оценок, доставляемых рекуррентными алгоритмами оценивания и оптимизации при зашумленных наблюдениях, приобрела в целом достаточно законченный вид. Основой многих работ по оптимизации сходимости алгоритмов являются работы М.Вазана, М.Б.Невельсона и Р.З.Хасьминского, В.Я.Катковника, Я.З.Цыпкина, Б.Т.Поляка, А.С.Позняка и О.Н.Граничина.

Рассматриваемый в этой диссертации тип алгоритмов, основанных на использовании статистической информации о включаемом в рассмотрение пробном одновременном возмущении, относится к более широкому классу алгоритмов случайного поиска. Значительное использование на практике алгоритмов случайного поиска вызвано потребностью в решении задач оптимизации в условиях, когда свойства ис-

следуемой функции потерь неизвестны, а измерение значений самой функции доступны чаще всего с помехами. В русскоязычной литературе эти алгоритмы исследовались, например, в работах Л.А.Растригина, А.Жилин斯卡са, С.М.Ермакова, В.Б.Меласа и А.А.Жиглявского.

В отличие от предлагавшихся ранее методов рандомизированные алгоритмы стохастической оптимизации обеспечивают состоятельность оценок при существенно менее значительных предположениях о свойствах неизвестной функции потерь и помех в измерении ее значений.

Появление этих алгоритмов было обусловлено требованиями задач реального времени к сокращению количества вычислений функционала потерь на каждой итерации и необходимостью расширить класс допустимых помех. Предполагалась, например, возможная взаимозависимость помех, обусловленная намеренными действиями противника (глушение сигнала, и т. д.). Первые исследования, учитывающие эти условия, начались на рубеже 80–90-х годов прошлого века. Основы этих исследований базируются на работах О.Н. Граничина, Б.Т. Поляка с А.Б.Цыбаковым и А.В.Гольденшлюгером, Дж.Спала.

Цели работы. Расширение границ применимости рандомизированного метода стохастической оптимизации с одним измерением функции потерь на итерации, оценка его работы после конечного числа итераций и применение рандомизированных алгоритмов к некоторым актуальным задачам из области информатики.

Методы исследования. В диссертации применяются методы теории оценивания и оптимизации, функционального анализа и теории вероятностей, а также компьютерное моделирование.

Основные результаты. В работе получены следующие основные результаты:

1. Ослаблены условия сходимости рандомизированного алгоритма стохастической оптимизации с одним измерением функции потерь.
2. Получены оценки точности работы рандомизированного алгоритма стохастической оптимизации с одним измерением функции по-

терь при конечном числе итераций.

3. Разработан метод реализации рандомизированного алгоритма стохастической оптимизации с одним измерением функции потерь на квантовых компьютерах.
4. Предложен способ построения систем с элементами искусственного интеллекта на базе квантовых компьютеров с применением рандомизированных алгоритмов стохастической оптимизации.
5. Разработаны программные имитационные модели, позволяющие анализировать работу рандомизированных алгоритмов стохастической оптимизации на практике.

Научная новизна. Все основные научные результаты диссертации являются новыми.

Практическая и теоретическая ценность. Ослабление условий состоятельности оценок рандомизированного алгоритма стохастической оптимизации расширяет границы применимости алгоритма и дает больше оснований для его использования в тех случаях, когда свойства функции потерь неизвестны. Оценки точности алгоритма при конечном числе итераций позволяют характеризовать его скорость сходимости и, в зависимости от задачи, принимать решение о необходимом количестве итераций. Все описанные в диссертации применения рандомизированных алгоритмов для повышения эффективности работы серверов, балансировщика загрузки, и т. д. реализованы и протестированы на программных моделях и представляют собой самостоятельную практическую ценность.

Апробация работы. По материалам диссертации были сделаны доклады на международных конференциях “Physics and Control — 2003” (Санкт-Петербург), “System Identification and Control Problems — 2004” (Москва), на заседании международной школы-семинара “Адаптивные роботы – 2004” (Санкт-Петербург), на Пятом международном семинаре “5-th St. Petersburg Workshop on Simulation” (Санкт-Петербург,

2005 г.), а также на семинарах кафедры системного программирования математико-механического факультета Санкт-Петербургского государственного университета.

Публикации. Основные результаты диссертации опубликованы в работах [1] — [6].

Структура и объем диссертации. Диссертация состоит из введения, трех глав, заключения и списка литературы, содержащего 79 наименований. Включает 18 рисунков. Общий объем работы — 80 страниц.

Содержание работы

Во **введении** обосновывается актуальность тематики диссертационной работы и кратко излагаются ее основные результаты.

В **первой главе** определяется понятие *функционала среднего риска* и ставится общая задача о его минимизации.

Рассмотрим стандартную формулировку задачи оптимизации некоторого функционала $f(x)$. Параметр x обычно представляет собой вектор сравнительно большой размерности, а вид функционала f часто неизвестен, но доступны его зашумленные измерения y_n в выбранных экспериментатором точках x_n :

$$y_n = f(x_n) + v_n.$$

На практике чаще всего приходится иметь дело с задачами, в которых измеряемое значение функционала зависит еще от некоторого неконтролируемого случайного параметра w :

$$y_n = F(x_n, w_n) + v_n.$$

В таких случаях меняется постановка задачи — функционал минимизируется в среднем относительно случайной составляющей:

$$f(x) = E_w\{F(x, w)\} \rightarrow \min_x,$$

где E_w — условное математическое ожидание. При этом функционал f принято называть *функционалом среднего риска*.

Задачи оптимизации подобного рода возникают, например, когда необходимо выработать или оптимальную стратегию при противодействии противника, или оптимальный диверсификационный пакет при неконтролируемых движениях рынка, или оптимальную стратегию маршрутизации при меняющихся параметрах потока заявок, и т. п. В разделе 1.2 приводятся примеры конкретных задач, которые могут быть сформулированы в терминах оптимизации функционала среднего риска.

Далее, в разделе 1.3, приводится краткий библиографический обзор, касающийся истории развития алгоритмов стохастической оптимизации.

Во **второй главе** уточняется постановка задачи оптимизации функционала среднего риска, описываются основные предположения о классе минимизируемых функционалов, предлагается конкретный вид рандомизированного алгоритма стохастической оптимизации и доказываются результаты относительно состоятельности оценок этого алгоритма и их точности после конечного числа итераций.

Пусть $F(x, w) : \mathbb{R}^q \times \mathbb{R}^p \rightarrow \mathbb{R}^1$ — дифференцируемая по первому аргументу функция, x_1, x_2, \dots — выбираемая экспериментатором последовательность точек измерения (план наблюдения), в которых в каждый момент времени $n = 1, 2, \dots$ доступно наблюдению с аддитивными помехами v_n значение функции $F(\cdot, w_n)$:

$$y_n = F(x_n, w_n) + v_n,$$

где $\{w_n\}$ — неконтролируемая последовательность случайных величин ($w_n \in \mathbb{R}^p$), имеющих одинаковое, вообще говоря, неизвестное распределение $P_w(\cdot)$ с конечным носителем.

Требуется по наблюдениям y_1, y_2, \dots построить последовательность оценок $\{\hat{\theta}_n\}$ неизвестного вектора θ^* , минимизирующего функцию

$$f(x) = E_w\{F(x, w)\} = \int_{\mathbb{R}^p} F(x, w)P_w(dw)$$

типа функционала среднего риска.

Далее мы будем использовать обозначения $\|\cdot\|$ и $\langle \cdot, \cdot \rangle$ для евклидовой нормы и скалярного произведения в \mathbb{R}^q . Пусть $\rho \in (1, 2]$. Введем функцию Ляпунова

$$V(x) = \|x - \theta^*\|^\rho,$$

где θ^* - искомый вектор из пространства параметров, и сформулируем основные предположения.

A.1 Функция $f(x)$ имеет единственный минимум в точке θ^* и

$$(\nabla V(x), \nabla f(x)) \geq \mu V(x), \quad \forall x \in \mathbb{R}^q$$

с некоторой постоянной $\mu > 0$.

A.2 При любом w градиенты функций $F(\cdot, w)$ удовлетворяют условию Гельдера с показателем $\rho - 1$

$$\|\nabla_x F(x, w) - \nabla_x F(y, w)\| \leq M \|x - y\|^{\rho-1}, \quad \forall x, y \in \mathbb{R}^q$$

с некоторой постоянной $M > 0$.

Пусть Δ_n , $n = 1, 2, \dots$ — наблюдаемая последовательность независимых друг от друга случайных векторов в \mathbb{R}^q , называемая в дальнейшем *пробным одновременным возмущением*, компоненты которых независимы между собой и принимают значения ± 1 с вероятностью $\frac{1}{2}$.

Зафиксируем некоторый начальный вектор $\hat{\theta}_0 \in \mathbb{R}^q$ и выберем некоторые последовательности положительных чисел, стремящиеся к нулю: $\{\alpha_n\}$ и $\{\beta_n\}$.

Для построения последовательностей точек измерения $\{x_n\}$ и оценок $\{\hat{\theta}_n\}$ предлагается следующий алгоритм, использующий на каждом шаге (итерации) одно наблюдение:

$$\begin{cases} x_n = \hat{\theta}_{n-1} + \beta_n \Delta_n, & y_n = F(x_n, w_n) + v_n, \\ \hat{\theta}_n = \mathcal{P}_{\Theta_n}(\hat{\theta}_{n-1} - \frac{\alpha_n}{\beta_n} \Delta_n y_n). \end{cases} \quad (1)$$

В этом алгоритме $\mathcal{P}_{\Theta_n}(\cdot)$, $n = 1, 2, \dots$ — операторы проектирования на некоторые выпуклые замкнутые ограниченные подмножества $\Theta_n \subset \mathbb{R}^q$, содержащие, начиная с некоторого $n \geq 0$, точку θ^* . Если заранее

известно замкнутое ограниченное выпуклое множество Θ , содержащее точку θ^* , то можно считать $\Theta_n = \Theta$. В противном случае множества $\{\Theta_n\}$ могут расширяться до бесконечности.

Обозначим $\mathbb{W} = \text{supp}(P_w(\cdot)) \subset \mathbb{R}^p$ — конечный носитель распределения $P_w(\cdot)$; \mathcal{F}_n — σ -алгебру вероятностных событий, порождаемую случайными величинами $\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_n$, формируемыми по предлагаемому алгоритму;

$d_n = \text{diam}(\Theta_n)$ — диаметр множества Θ_n ;

$c_n = 2^\rho \rho q^{\rho/2} \alpha_n^\rho \beta_n^{-\rho}$, $\gamma_n = \alpha_n \mu - 2^{3\rho-2} c_n M^\rho d_n^\rho$,

$\phi_n = 2M\rho d_n^{\rho-1} \alpha_n \beta_n + 2^{\rho-1} c_n \max_{\mathbb{W}} |F(\theta^*, w)|^\rho + 2^{4\rho-2} \rho q^{\frac{3\rho}{2}} \alpha_n^\rho \beta_n^\rho M^\rho$.

Теорема 1. Пусть $\rho \in (1, 2]$ и выполнены условия:

(A.1) для функции $f(x) = E_w\{F(x, w)\}$;

(A.2) для функции $F(\cdot, w) \forall w \in \mathbb{W}$;

функции $F(x, w)$ и $\nabla_x F(x, w)$ равномерно ограничены на \mathbb{W} ;

$\forall n \geq 1$ случайные величины v_1, \dots, v_n и вектора w_1, \dots, w_{n-1} не зависят от w_n, Δ_n , а случайный вектор w_n не зависит от Δ_n ;

$$E\{|v_n|^\rho\} \leq \sigma_n^\rho, \quad n = 1, 2, \dots;$$

$\forall n, 0 \leq \gamma_n \leq 1, \sum_n \gamma_n = \infty, \mu_n \rightarrow 0$, при $n \rightarrow \infty$, где

$$\mu_n = \frac{\phi_n + c_n \sigma_n^\rho}{\gamma_n}, \quad z_n = \left(1 - \frac{\mu_{n+1}}{\mu_n}\right) \frac{1}{\gamma_{n+1}}.$$

Тогда:

1). последовательность оценок $\{\hat{\theta}_n\}$, доставляемых алгоритмом (1), сходится к точке θ^* в следующем смысле:

$$E\{\|\hat{\theta}_n - \theta^*\|^\rho\} \rightarrow 0 \quad \text{при } n \rightarrow \infty.$$

2). Если $\overline{\lim}_{n \rightarrow \infty} z_n \geq z > 1$, то $E\{V(\hat{\theta}_n)\} = \mathcal{O}\left(\prod_{i=0}^{n-1} (1 - \gamma_i)\right)$.

3). Если $z_n \geq z > 1 \forall n$, то $E\{V(\hat{\theta}_n)\} \leq (E\{V(\hat{\theta}_0)\} + \frac{\mu_0}{z-1}) \prod_{i=0}^{n-1} (1 - \gamma_i)$.

4). Если, более того,

$$\sum_n \phi_n + c_n \sigma_n^\rho < \infty$$

тогда $\hat{\theta}_n \rightarrow \theta^*$ при $n \rightarrow \infty$ с вероятностью единица, и

$$P\{\|\hat{\theta}_n - \theta^*\|^\rho \leq \epsilon \forall n \geq n_0\} \geq 1 - \frac{E\{\|\hat{\theta}_{n_0} - \theta^*\|^\rho\} + \sum_{n=n_0}^{\infty} \phi_n + c_n \sigma_n^\rho}{\epsilon}.$$

Доказательство теоремы приведено в последнем разделе второй главы диссертации.

Далее во второй главе рассмотрена проблема выбора вычислительного устройства, наилучшим образом подходящего для выполнения рандомизированного алгоритма стохастической оптимизации с одним измерением функции штрафа на итерации. В работе О.Н.Граничина (АиТ, 2002, №2) был описан способ реализации алгоритма (1) на вычислительных устройствах нового типа — квантовых компьютерах. За прошедшее время терминология и аксиоматика моделей квантовых вычислений стали более четкими. Описанный ранее способ реализации не был похож на типичный “квантовый” алгоритм, и сейчас можно сказать, что выбранный ранее способ представления не был удачным. В разделе 2.6 описан новый способ представления алгоритма на квантовом компьютере, который в большей мере соответствует общей логике квантовых вычислительных алгоритмов.

Заканчивается вторая глава описанием возможной схемы реализации интеллектуальной системы на базе квантовых компьютеров и рандомизированных алгоритмов.

В третьей главе представлены более детальные описания рассмотренных в первой главе примеров задач оптимизации функционалов типа среднего риска, предлагаются методы их решения, основанные на рандомизированных алгоритмах стохастической оптимизации и приводятся результаты имитационного моделирования.

Начинается третья глава с рассмотрения задачи балансировки загрузки. Рассмотрим систему, состоящую из двух серверов, способных параллельно обрабатывать $N^{(1)}$ и $N^{(2)}$ заявок, и балансировщика загрузки — устройства, в реальном времени принимающего решение, к какому из серверов отправить пришедшую в текущий момент времени заявку. Значения $N^{(1)}$ и $N^{(2)}$ могут меняться со временем, и считаются

неизвестными балансировщику. Кроме того, неизвестным предполагается время обслуживания серверами каждой из заявок. При невозможности немедленной обработки запроса на серверах заявка системой отвергается.

Обозначим через $u_t \geq 0$ функцию прихода заявки в момент времени t . Если заявки нет, то $u_t = 0$, если заявка фактически поступила, то u_t - время, которое будет затрачено на ее обработку (об этих величинах нам известна только некоторая индикаторная функция, показывающая есть ли новая заявка на входе балансировщика или нет).

Пусть в каждый момент времени мы можем наблюдать величины z_t , определяемые следующим образом:

- $z_t = 1$, если в момент времени t пришла заявка, балансировщик загрузки направил ее к серверу i , ($i = 1, 2$), и на нем не нашлось свободных каналов, в то время как на другом сервере свободные каналы были,
- $z_t = 0$ во всех остальных случаях.

Таким образом, величины z_t являются для нас измерениями индикаторной функции штрафа. Штраф платится за неверную маршрутизацию, ведущую к замедлению работы в условиях реального времени. Заметим, что штраф не платится в том случае, когда в обоих серверах не оказалось свободных каналов. В такой ситуации отказ в обслуживании не обусловлен недостатками маршрутизации.

Ясно, что функция штрафа в каждый момент зависит от большого количества случайных параметров (величин u_t , $N_t^{(i)}$) и от выбранного алгоритма маршрутизации. В случае когда выбран класс алгоритмов x , можно сформулировать задачу оптимизации:

$$f(x) = E_{\dots} \{z_t(x, \dots)\} \rightarrow \min_x .$$

Здесь математическое ожидание берется по всем случайным параметрам системы, от которых зависит функция штрафа.

В разделе 3.1 рассматривается класс алгоритмов, предлагаемый для решения этой задачи и формулируется метод оптимизации приведенного выше функционала, основанный на рандомизированных алгоритмах

стохастической оптимизации. Там же представлены результаты имитационного моделирования.

Далее, в разделе 3.2, рассматривается задача повышения эффективности работы сервера, который используется для обслуживания очереди заданий, поступающих случайным образом.

Предполагается, что вероятностное распределение времени обслуживания задания сервером зависит от параметра x , который требуется выбрать с целью минимизации среднего времени ожидания клиентами $L(x)$ вместе с некоторой стоимостью использования $q(x)$ параметра x

$$f(x) = q(x) + L(x) \equiv q(x) + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E\{z_i(x)\} \rightarrow \min,$$

где $z_i(x)$ — время, которое задание, законченное i -м по счету, ожидало (“простаивало”) в сервере до момента своего завершения.

В пункте 3.2.2 рассмотрен многомерный вариант задачи.

Для обеих задач рассматриваются варианты применения рандомизированных алгоритмов стохастической оптимизации, для многомерного случая приведены результаты моделирования.

В задаче оценки надежности серверного программного обеспечения (ПО) из пункта 1.2.4 требуется определить оптимальные параметры некоторого эвристического алгоритма, при которых сервер, тестируемый в соответствие с этим алгоритмом, наискорейшим путем попадает в терминальное состояние. Вопросы надежности программного обеспечения имеют колоссальный практический интерес. Решение этих вопросов традиционно разделяется на две большие области — развитие технологии создания ПО и развитие технологии тестирования. В первой области достигнуты весьма впечатляющие результаты — созданы мощные языки и системы программирования, визуализации, отладки, развиты формальные теории — теория языков и грамматик, теория алгоритмов и т. д. Что касается второй области, то теоретические исследования в ней оказались существенно более сложны и малоэффективны — полное тестовое покрытие и формальное доказательство соответствия программы задаче и даже алгоритму в большинстве случаев практически неосуществимы.

В пункте 1.2.4 предлагаются новые критерии качества программного обеспечения и метод их оценки, который также может служить для поиска максимально опасных для сервера траекторий. Для определения оптимальных параметров метода в разделе 3.3 предлагается использовать рандомизированные алгоритмы стохастической оптимизации. Результаты моделирования приведены в пункте 3.3.4.

Последней в третьей главе в разделе 3.4 рассматривается самообучающаяся опознающая система, на вход которой поступают сигналы w принадлежащие одному из l классов (число классов конечно и известно заранее). Сигналы w представляют собой вектора из p компонент, называемых признаками, а последовательность векторов w_1, w_2, \dots является реализацией некоторой последовательности независимых, одинаково распределенных случайных величин.

Пусть заданы функции $Q^k(\mathcal{X}, w)$, $k \in 1, 2, \dots, l$ (l — число классов), определяющих качество классификации и зависящих от матричного параметра $\mathcal{X} = \{x_1, \dots, x_l\}$, $x_i \in \mathbb{R}^q$, который удобно интерпретировать как набор центров классов. Функционал среднего риска в задаче самообучения имеет смысл математического ожидания общих потерь и может быть записан следующим образом:

$$F(\mathcal{X}) = \sum_{k=1}^l \int_{\mathbb{W}^k(\mathcal{X})} Q^k(\mathcal{X}, w) P(dw),$$

где $\{\mathbb{W}^k(\mathcal{X})\}_{k=1}^l$ — разбиение выборочного пространства на l непересекающихся подмножеств

$$\begin{aligned} \mathbb{W}^k(\mathcal{X}) = & \{w : Q^k(\mathcal{X}, w) < Q^j(\mathcal{X}, w), j < k\} \cap \\ & \cap \{w : Q^k(\mathcal{X}, w) \leq Q^j(\mathcal{X}, w), j > k\}, \quad k, j \in 1, 2, \dots, l. \end{aligned}$$

Проблема самообучения распознаванию образов формулируется как поиск некой оценки $\hat{\mathcal{X}}$, минимизирующей предлагаемый функционал среднего риска — иными словами, задача заключается в поиске центров классов, наилучших с точки зрения функций $Q^k(\mathcal{X}, w)$.

В разделе 3.4 предлагается метод решения данной задачи, основанный на рандомизированном алгоритме стохастической оптимизации и приводятся результаты моделирования его работы.

В **заключении** формулируются основные результаты работы.

Работы автора по теме диссертации

- [1] *Granichin O.N., Sysoev S.S.* About Some Characteristics of Computers of New Generation // In Proceedings of the Physics and Control Conference, Saint-Petersburg, 2003, vol. 3, pp. 804-807.
- [2] *Сысоев С.С.* Адаптивное управление распределением загрузки в простейшей вычислительной сети // Труды III международной конференции “Идентификация систем и задачи управления”, М., 2004. С. 1563–1571.
- [3] *Измакова О. А., Сысоев С. С.* Алгоритм стохастической оптимизации с возмущением на входе в задаче самообучения. // Труды Международной школы-семинара “Адаптивные роботы — 2004”. М.-СПб. 2004. С. 49–52.
- [4] *Сысоев С.С.* Рандомизированные алгоритмы стохастической оптимизации, квантовые компьютеры, искусственный интеллект // В сб. “Стохастическая оптимизация в информатике” под ред. О. Н. Граничина. Изд-во С.-Петербур. ун-та. 2005. С. 206–221.
- [5] *Граничин О. Н., Сысоев С.С., Чуйко Д. С.* Проблемы тестирования сервера как задачи о моделировании редких событий // В сб. “Стохастическая оптимизация в информатике” под ред. О. Н. Граничина. Изд-во С.-Петербур. ун-та. 2005. С. 48–72.
- [6] *Chuyko D., Granichin O.N., Sysoev S. S.* Simulation of rare events and probability estimation // In: Proceedings of the 5-th St. Petersburg Workshop on Simulation. St. Petersburg, 2005, pp. 215–220.

Подписано в печать 15.09.2005 г. Формат бумаги 60X90 1/16. Бумага офсетная.
Печать ризографическая. Объем 1 усл. п. л. Тираж 100 экз. Заказ N 737.
Отпечатано в ПК "МИКРОМАТИКС" с оригинал-макета заказчика.
199178, Санкт-Петербург, Большой пр. В.О., д. 55.