

Севастопольский государственный технический университет

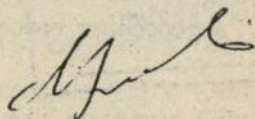
На правах рукописи

ЦУКАНОВ Александр Викторович

СТАТИСТИЧЕСКОЕ И ИНФОРМАЦИОННО-ЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ  
В ИДЕНТИФИКАЦИИ СЛОЖНЫХ ОБЪЕКТОВ УПРАВЛЕНИЯ

05.13.03 - системы и процессы управления

Автореферат диссертации на соискание ученой степени  
доктора технических наук



Севастополь 1995

718 33.000  
Диссертацией является

ЛНБ України ім.В.Стефаніка

Работа выполнена  
техническом универ



00761278 (V)

Официальные оппоненты:

1. Академик НАН Украины Скурихин Владимир Ильич
2. Доктор физико-математических наук,  
профессор Тимченко Игорь Евгеньевич
3. Доктор технических наук,  
профессор Пряшников Федор Дмитриевич

Ведущая организация: Институт проблем машиностроения НАН Украины (г. Харьков)

Защита состоится 19 октября 1995 г. в 14 час на заседании специализированного совета Д 11.03.01 в Севастопольском государственном техническом университете по адресу: 335053, Севастополь, Стрелецкая бухта, студгородок.

С диссертацией можно ознакомиться в библиотеке университета

ЛНБ ім. В. Стефаніка  
АН України

Автореферат розослан "15" сентября 1995 г.

Ученый секретарь специализированного  
совета

Шерешевский А. Н.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Идентификация объектов управления (ОУ) является одним из наиболее важных и интенсивно разрабатываемых разделов теории управления. Формализация и решение многих задач в этом направлении стали возможными благодаря интенсивному развитию теории математического моделирования по эмпирическим данным и планирования эксперимента (Валник В.Н., Ивахненко А.Г., Федоров В.В., Кокс Д., Шибата Р., Акаике Г. Аткинсон А., Аллен Д.М. и др.). Математические методы построения моделей в условиях ограниченного объема экспериментальных данных и неопределенной структуры модели все больше приобретают комплексный характер и учитывают множество факторов для построения вычислительных и алгоритмических процедур, в том числе результаты вычислительных экспериментов (ВЭ) над имитационными объектами и априорную информацию об ОУ (Кипнис В., Трунов В.Г., Эфрон Б., Романов В.Л. и др.).

Большой объем информации, необходимой для оптимизации процесса моделирования в задачах автоматизации структурной идентификации регрессионных моделей сложных ОУ, характеризующихся большим числом составляющих элементов, слабоструктурированностью, смешанностью качественных и количественных характеристик, неполнотой информации, ошибками в данных и наличием множества моделей и языков для своего описания, делает актуальной проблему использования достижений теории и практики информационных систем и систем с искусственным интеллектом для повышения эффективности методов и систем идентификации (Гэйл В.А., Кузин Е.С., Поспелов Д.А., Скурихин В.И., Семенов Н.А., Сильвестров А.Н. и др.).

По мере развития теории и практики идентификации обнаружались новые принципиальные и важные нерешенные вопросы.

Это, во-первых, проблема корректного выбора эффективного для реальной практики построения систем управления метода идентификации ОУ и, во-вторых, проблема связанная с комплексным подходом к методам идентификации и способам хранения и поиска информации о внутренней структуре ОУ, о статистических характеристиках приложенных к нему возмущений, о свойствах методов идентификации и др. Такого рода информация обычно хранится в базе данных (БД), при этом качество процесса идентификации во многом зависит от эффективности и надежности БД, что в свою очередь определяется информационно-логической моделью (ИЛМ) данных. Современная теория построения ИЛМ статистических данных не может в полной мере удовлетворить требованиям практики построения высокоэффективных систем управления с БД (Особенно для сложных технологических объектов в машиностроении, биопромышленности и энергетике).

В связи с вышеизложенным разработка методов комплексного статистического и информационно-логического моделирования в идентификации сложных ОУ является актуальной научной проблемой, имеющей важное народнохозяйственное значение.

Объектом исследования в диссертации являются автоматизированные системы идентификации сложных ОУ с БД и знаний.

Цель диссертации состоит в разработке концепции и методов комплексного статистического и информационно-логического моделирования в идентификации сложных ОУ, осуществляющих выбор, испытание и применение математических методов структурной идентификации (МСИ) регрессионных моделей и планирования эксперимента, реализация которых обеспечивает эффективное решение задач идентификации сложных ОУ.

Для достижения цели исследования в диссертации решены следующие задачи:

- исследована роль и обосновано использование ВЭ для задач анализа и синтеза МСИ и планирования эксперимента в системах рассматриваемого класса;

- разработаны и систематизированы критерии оптимальности процесса моделирования, исследовано влияние имеющейся информации об ОУ и данных ВЭ на выбор эффективных МСИ регрессионных моделей и планов эксперимента с целью построения информационной технологии идентификации сложных ОУ;

- разработаны и исследованы методы и алгоритмы использования ВЭ для задач анализа и синтеза МСИ и планирования эксперимента;

- разработаны концепция и теоретические основы информационно-логического моделирования данных о свойствах МСИ и планов эксперимента;

- разработано информационное и программное обеспечение, позволяющее реализовать предложенные в диссертации методы в виде расчетно-логических и экспертных систем.

Методы исследования. В работе используются теории математической статистики, оптимального эксперимента, идентификации систем, принятия решений, реляционных БД, искусственного интеллекта, а также методы имитационного моделирования и информационных технологий.

Научная новизна. К наиболее существенным научным результатам работы относятся следующие:

1. Предложена концепция, позволяющая на основе ВЭ и имеющейся априорной информации об уровне шума и об ограничениях на управляемые переменные и функцию отклика разрабатывать методы испытания и выбора МСИ и планов эксперимента, которая включает в себя:

- новые критерии оптимальности выбора МСИ и плана экс-

перимента, дающие возможность провести сравнительный анализ ряда МСИ, основанных на обобщенных информационных критериях селекции регрессионных моделей и критериях скользящего котроля и определить области эффективного использования соответствующих МСИ;

- новые методы и алгоритмы испытания с помощью ВЭ МСИ и планирования эксперимента, основанные на определении, построении и анализе тестовых функциональных зависимостей и аппроксимационных моделей целевых функций;

- новые критерии оптимальности планов эксперимента для случайных функционалов качества процедуры построения регрессионных моделей, к которым относятся задачи планирования эксперимента при структурной неопределенности модели ОУ и при наличии случайных регрессоров в модели;

2. Разработаны основы теории реляционных БД для систем идентификации сложных ОУ, в рамках которой получены и доказаны следующие результаты:

- расширена реляционная алгебра за счет введения специальных арифметических операций над отношениями, доказаны свойства этих новых операций;

- за счет введения расширенной реляционной алгебры показано, что большинство операций векторного и матричного исчисления реализуется как формулы реляционной алгебры и могут быть реализованы в виде соответствующих запросов к БД, что позволило создать ИЛМ данных, описывающих МСИ, результаты ВЭ и реальный ОУ, и предложить новую информационную технологию идентификации ОУ ;

- определен и исследован новый класс С-зависимостей между атрибутами ИЛМ, который позволил описывать и анализировать сложные классификационные экспериментальные данные;

- исследовано влияние нормализации реляционных ИЛМ на надежность БД в случае ошибок манипулирования данными при оперативной идентификации ОУ и определены условия, при которых нормализация отношений БД ведет к повышению ее надежности.

Практическая ценность и реализация работы. На основе теоретических результатов работы было создано научно-методическое, математическое и программное обеспечение гибридных экспертных систем экспериментального моделирования (ЭСЭМ). Указанное обеспечение создавалось под научным руководством и при непосредственном участии автора по заданию 04.09 Комплексной целевой научно-технической программы Минвуза Украины "Автоматизированные системы научных исследований и испытаний сложных систем", в рамках проекта "Ветровое волнение" Общегосударственной программы исследования и использования Мирового океана в интересах науки и народного хозяйства, а также при выполнении ряда хоздоговорных и госбюджетных работ.

Ряд результатов работы внедрен в учебном процессе Севастопольского государственного технического университета.

На защиту выносятся следующие научные результаты, полученные лично автором: 1. Концепция комплексного статистического и информационно-логического моделирования в идентификации сложных ОУ; 2. Метод опорных функций для испытания и тестирования МСИ и планов эксперимента; 3. Свойства планов эксперимента для идентификации сложных ОУ при случайном функционале качества, зависящим от неизвестных параметров; 4. Новые операции реляционной алгебры для систем идентификации с БД и их свойства; 5. Теорема об оптимальности нормализации реляционных отношений при наличии ошибок в БД; 6. Разработанные на основе полученных результатов методики и алго-

ритмы создания интегрированных расчетно-логических систем идентификации.

Апробация работы. Основные результаты работы доложены и обсуждены на Ежегодной конференции статистиков Великобритании, на заседании Королевского общества, на научных семинарах в университетах Лондона, Глазго, Единбурга, Кингстона (Великобритания, 1983 г. и 1991 г., Канада, 1992 г.), на I Всемирном конгрессе общества математической статистики и теории вероятностей им. Бернули (Ташкент, 1986 г.), на международной конференции "Статистические методы в экспериментальных исследованиях" (Варна, Болгария, 1987), на научных семинарах в Королевском университете (Кингстон, Канада, 1992 г.), на 6, 8, 9 и 10 Всесоюзных конференциях по планированию и автоматизации эксперимента в научных исследованиях (1980, 1989 и 1992 гг. - Москва, 1986 г. - Ленинград), на 1 и 2 Всесоюзных конференциях "Перспективы и опыт внедрения статистических методов в АСУ ТП" (1981, 1984 гг - Смоленск), на Всесоюзной конференции "Технические средства изучения и освоения Мирового океана" (1981 г. - Севастополь), на 1 и 2 Всесоюзных конференциях "Перспективные методы планирования и анализа экспериментов при исследовании случайных полей и процессов" (1982 г. - Нальчик, 1985 г. - Севастополь), на Всесоюзном научно-техническом семинаре "Программное обеспечение новых информационных технологий" (октябрь 1991, Тверь), на Республиканской конференции "Повышение качества программного обеспечения ЭВМ" (1986 г. - Севастополь), на Республиканском семинаре "Прикладные интеллектуальные системы" (1992 г. - Москва).

Публикации. По теме диссертации опубликовано более 40 печатных работ, в том числе 12 работ в зарубежных изданиях,

две брошюры, ряд статей в центральных изданиях по перечню ВАК, тезисы и доклады на различных международных и республиканских конференциях.

Структура и объем диссертации. Диссертация состоит из введения, семи глав, заключения, списка литературы, содержащего 314 наименований и приложений. Основной текст диссертации занимает 256 страниц. Работа содержит 53 рисунка и 19 таблиц.

#### ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы и излагается перечень вопросов, исследованию которых посвящена диссертационная работа, формулируется цель исследования, а также защищаемые автором основные положения.

В первой главе сформулирована задача выбора и оптимальной настройки МСИ регрессионных моделей сложных ОУ на основе тестовых испытаний МСИ на ЭВМ. Показано, что решение задач идентификации широкого класса ОУ требует проведения многократных имитационных экспериментов, использования большого объема оперативно изменяющихся экспериментальных данных и разнообразной априорной информации о свойствах МСИ, что приводит к необходимости использования технологии БД. Предложена общая концепция информационно-логического моделирования статистических данных в идентификации сложных ОУ, на основе которой указаны возможности и определены направления исследования и разработки МСИ с использованием БД.

В работе рассматривается ОУ с точки зрения его внешних характеристик безотносительно к особенностям его внутренней структуры и к внутренним процессам, который описывается следующими характеристиками: выходной переменной  $y$ ; вектором управляемых независимых входных переменных  $x^T = [x_1, x_2, \dots, x_k]$ ,

где  $k$  - число входных переменных; вектором неуправляемых, но измеряемых входных переменных (возмущений)  $\mathbf{z}^T = [z_1, z_2, \dots, z_l]$ , где  $l$  - число неуправляемых входных переменных; и множеством качественных признаков (категорийных переменных)  $A_j$ .

Рассматриваются квазирегулярные ОУ, для которых в пределах интервала регулярности и постоянных значений категориальных переменных существуют функциональные зависимости вида  $y = \eta(\mathbf{x}, \mathbf{z}) + \varepsilon$ , где  $\eta(\mathbf{x}, \mathbf{z})$  - функция, вид которой до эксперимента неизвестен,  $\varepsilon$  - случайная ошибка, вызванная влиянием неучтенных факторов и ошибкой измерений. Относительно случайной ошибки предполагается, что имеется априорная информация о законе распределения ошибки и о некоторых параметрах этого распределения, например, ошибка  $\varepsilon$  может иметь нулевое среднее, неизвестную дисперсию  $\sigma^2$  и нормальное распределение. Также предполагается, что имеется априорная информация о пределах изменения функциональной зависимости  $\eta(\mathbf{x}, \mathbf{z})$ :

$a \leq \eta(\mathbf{x}, \mathbf{z}) \leq b$ , о множестве возможных значений вектора входных переменных  $\mathbf{x} \in W$  и о законе распределения вектора случайных возмущений  $\mathbf{z}$ .

Идентификация математической модели ОУ проводится на основании ограниченного объема экспериментальных данных (случай малой выборки). При этом  $\mathbf{Y}^T = [y_1, y_2, \dots, y_N]$  - вектор значений функции отклика в  $N$  опытах,  $\mathbf{X}^T = [x_1, x_2, \dots, x_N]$  - матрица плана эксперимента размером  $N$  строк и  $k$  столбцов и  $\mathbf{Z}^T = [z_1, z_2, \dots, z_N]$  - матрица измеренных значений вектора случайных переменных  $\mathbf{z}$  в  $N$  опытах. В диссертации предполагается, что экспериментальные данные, значения категориальных переменных и априорная информация  $I_2$  хранятся в реляционной БД, структура которой задается информационно-логической моделью (ИЛМ)  $R_{\text{ОИ}} = \{R_i\{A_i\}\}$ ,  $i=1, 2, \dots, n_r$ , состоящей из множества

отношений  $R_1$ , каждое из которых описывается множеством атрибутов  $\{A_i\}$ , задающими схему отношения со своими множествами доменов, функциональных и статистических связей между атрибутами. В дальнейшем  $r(R)$  обозначает конкретное отношение,  $t$  - кортеж отношения и  $t(A)$  значение атрибута  $A$  в кортеже  $t$ . В ИЛМ данных об ОУ также входят множество словарей данных  $W_{\text{ОИ}}$  и множество ограничений на целостность данных  $L_{\text{ОИ}}$ . В процессе обмена данными с БД возникают случайные ошибки манипулирования данными, структура и вероятностные свойства которых задается моделью распределения вектора ошибок  $\theta \sim r(R_{\text{ОИ}}, Y, X, Z)$ . Организация БД для идентификации ОУ отличается от традиционной организации БД для экономических и социальных приложений за счет большого числа статистических метаданных, сгруппированных, повторяющихся и динамически изменяющихся значений атрибутов, и множества классификационных зависимостей.

В работе анализируется специальный случай, когда на классе моделей ОУ задана структура вложенных друг в друга моделей:

$$\eta_j(x, \alpha_j) \in S_j, \quad (j = 1, 2, \dots, q), \quad (1)$$

где  $\alpha_j^T = [\alpha_1, \alpha_2, \dots, \alpha_{k_j}]$  - вектор параметров, а  $k_j$  - число параметров для модели  $j$ , при этом  $S_1 \subset S_2 \subset \dots \subset S_q$ .  $S_j$  - множество возможных моделей для класса  $j$  ( $j=1, 2, \dots, q$ ),  $q$  - количество классов. Для линейных по параметрам моделей, структура задается в зависимости от числа членов модели. В этом случае каждый класс  $S_j$  задается моделью:

$$\eta_j(x, \alpha_j) = f_j^T(x) \alpha_j \quad (j = 1, 2, \dots, q). \quad (2)$$

где  $f_j^T(x) = [f_1(x), f_2(x), \dots, f_{k_j}(x)]$  - вектор известных функций.

Оценки вектора  $\alpha_j$  для каждой модели  $j$  определяются с

помощью функционала оценивания  $\hat{\alpha}_j = \Phi_j(X, Z, Y)$ , где  $\Phi_j$  - функционал от выборки  $Y, Z$ , плана эксперимента  $X$  и от класса моделей  $S_j$ . В качестве основного функционала для оценивания параметров модели используется функционал метода наименьших квадратов (МНК). Оценки вектора  $\alpha_j$  также зависят от ошибок записи и коррекции данных в БД  $\Phi$ .

МСИ осуществляет выбор класса модели  $j$  (селекция моделей) с помощью критерия селекции. Так для линейных по параметрам классов моделей большинство квадратических критериев селекции модели ОУ по экспериментальным данным можно представить в форме:

$$Cr(j) = \beta_j Y^T G_j Y + \gamma_j \quad (j= 1, 2, \dots, q) \quad (3)$$

где  $Cr$  - рассматриваемый критерий,  $G_j$  - матрица канонической формы критерия для модели класса  $S_j$ ,  $\beta_j$  и  $\gamma_j$  - константы. Прогнозирующие свойства моделей ОУ, построенных с помощью конкретного МСИ зависят от вида матриц  $G_j$  для каждого класса моделей  $j$ , которые являются функциями плана эксперимента  $X$ , экспериментальных данных  $Z$ , констант  $\beta_j$  и  $\gamma_j$ , априорной информации об ОУ  $I_a$ , объема выборки  $N$ , класса моделей, вероятностных свойств ошибок и, что наиболее существенно, от истинной модели ОУ и ее параметров.

Выбор МСИ является сложной задачей, решение которой в настоящее время носит в основном эмпирический характер и малоформализовано. Первое направление исследований в этой области заключается в построении аналитических зависимостей вероятностных характеристик критериев (3) при выполнении ряда заранее заданных условий. Как правило, в этом случае удается получить аналитическое решение только для асимптотических свойств сравнительно простых МСИ, для случая больших выборок и только частных задач моделирования при малом числе

наблюдений. Второе направление заключается в сравнении критериев селекции  $C_T$  и планов  $X$  с помощью ВЭ на ЭЕМ, под которым понимается испытание МСИ на тестовых моделях и сгенерированных методом Монте-Карло экспериментальных данных. Результаты исследований в области планирования ВЭ для тестирования МСИ в настоящее время носят разрозненный характер и нуждаются в систематизации и дополнительном изучении, что является одной из задач настоящей диссертации. При выборе критерия селекции  $C_T$  и плана  $X$  необходимо использовать максимум имеющейся информации о свойствах МСИ. Значительный объем такой информации приводит к необходимости организации БД о МСИ, задаваемой тройкой множеств  $(R_M, W_M, I_M)$ . Так как информация в такой БД оперативно корректируется и дополняется и подвержена ошибкам  $\delta$ , то структура ИЛМ оказывает значительное влияние на эффективность и надежность процесса идентификации, что приводит к необходимости исследования особенностей ИЛМ данных в идентификации ОУ.

В диссертации проведен анализ общих свойств и методов тестирования МСИ и на его основе предложена концепция комплексного анализа и построения статистических функциональных зависимостей между переменными, описывающими ОУ, и ИЛМ данных, что позволяет на единой основе разработать методы и программные средства эффективных в прикладном плане и надежных расчетно-логических систем идентификации сложных ОУ.

Во второй главе предлагаются обобщенные вероятностные меры эффективности МСИ, исследуются методы выбора и испытания критериев и процедур селекции регрессионных моделей, использующие БД с априорной информацией об ОУ. Особое внимание уделяется предложенному в диссертации методу опорных функций (МОФ).

Для линейных по параметрам регрессионных моделей предполагается, что старший класс моделей  $S_q$  содержит истинную модель, то есть имеет место выражение

$$Y = P_q \alpha_q + \epsilon, \quad (4)$$

где  $\epsilon$  - вектор ошибок и часть параметров вектора  $\alpha_q$  может иметь нулевые значения.

В качестве цели идентификации ОУ рассматривается получение модели с наилучшей прогнозирующей способностью, которую для любой модели  $j$  с фиксированным значением вектора параметров  $\alpha_q$  можно оценить следующим функционалом

$$I(j, \alpha_j) = \int_{W_p} E((y - \eta_j(x, \alpha_j))^2) dx / \int_{W_p} dx, \quad (j=1, 2, \dots, q). \quad (5)$$

Здесь  $E()$  - оператор математического ожидания.

Если параметры модели  $j$  оцениваются с помощью МНК по формуле  $\hat{\alpha}_j = (P_j^T P_j)^{-1} P_j^T Y$ , где  $P_j^T = [f_j(x_1), f_j(x_2), \dots, f_j(x_N)]$  - невырожденная матрица значений вектора  $f_j$  в  $N$  экспериментальных точках, тогда  $I(j, \hat{\alpha}_j)$  является случайной величиной, и качество структуры модели ОУ из класса  $S_j$  можно оценить средним значением этой величины, которое для постоянной дисперсии  $\sigma^2$  имеет вид

$$EI(j) = \sigma^2 + \sigma^2 \int_{W_p} f_j^T(x) (P_j^T P_j)^{-1} f_j(x) dx / \int_{W_p} dx + \\ + \alpha_{q-j}^T [-C, I_{q-j}] \int_{W_p} f_q(x) r_q^T(x) dx [-C, I_{q-j}]^T \alpha_{q-j} / \int_{W_p} dx, \quad (6)$$

где  $C = P_{q-j}^T P_j (P_j^T P_j)^{-1}$  и  $P_q = [P_j P_{q-j}]$ . Матрица  $P_{q-j}$  это часть матрицы  $P_q$  без первых  $j$  столбцов и  $\alpha_{q-j}$  - вектор параметров модели  $q$  без параметров модели  $j$ .

Выбор класса модели с помощью МСИ является случайным событием, что приводит к необходимости исследования параметров распределения индекса  $j$  - номера выбранной модели.

В работе вводится следующая случайная индикаторная функция:

$$v_j(Y, Cr) = \begin{cases} 1, & \text{если выбрана модель из класса } S_j; \\ 0 & \text{во всех остальных случаях.} \end{cases} \quad (7)$$

На основании функции распределения случайной величины  $v_j(Y, Cr)$  определяются следующие меры эффективности МСИ:

1.  $E\{v_t(Y, Cr)\}$  - вероятность выбора правильной или истинной структуры модели. Здесь  $t$  - номер класса, содержащий истинную модель или модель, наиболее близкую к истинной.
2.  $\sum_{j=1}^q E\{v_j(Y, Cr)I(j, \hat{\alpha}_j)\} - \sigma^2$  - математическое ожидание средне-квадратической ошибки прогноза поверхности отклика.
3.  $\sum_{j=1}^q E\{v_j(Y, Cr)\} \cdot EI(j) - \sigma^2$  - математическое ожидание средне-квадратической ошибки прогноза поверхности отклика при условии выбора соответствующей модели.
4.  $P\left(\sum_{j=1}^q v_j(Y, Cr) \cdot I(j, \hat{\alpha}_j) > I_0\right)$  - доверительный уровень превышения среднеквадратической ошибки прогноза заданного значения  $I_0$ . Здесь  $P(\cdot)$  - вероятность случайного события.

Предложенные меры эффективности МСИ зависят от значений вектора неизвестных параметров  $\alpha_q$ . При каждом фиксированном значении вектора  $\alpha_q$  можно найти структуру модели  $OY$ , для которой рассматриваемая функция потерь  $L(Cr)$  принимает оптимальное значение

$$L_{\min}(X, \alpha_q, \sigma^2) = \min_j L(j, X, \alpha_q, \sigma^2). \quad (8)$$

Эта величина не зависит от вектора наблюдений  $Y$  и для любого  $\alpha_q$  может быть вычислена заранее.

Для каждого конкретного МСИ в диссертации предлагается исследовать поведение функции упущенных возможностей (ФВВ)

$$\partial L(Cr, \alpha_q, X, \sigma^2) = L(Cr, \alpha_q) - L_{\min}(X, \alpha_q, \sigma^2). \quad (9)$$

При этом в качестве обобщенного критерия оптимальности МСИ

выбирается функционал  $\min_{Cr \alpha_q} \max_{\alpha_q} \Delta L(Cr, \alpha_q)$ , который зависит от множества моделей  $S_q$ , плана эксперимента  $X$ , вектора неизвестных параметров  $\alpha_q$  и дисперсии  $\sigma^2$ .

Предлагается и исследуется метод опорных тестовых функций (МОФ) для определения областей предпочтительности критериев селекции регрессионных моделей. Основная идея метода заключается в том, чтобы сравнивать эффективность различных процедур построения моделей не во всей области возможных значений параметров  $\alpha_q$ , а в специальных точках этой области, которые соответствуют определенным тестовым моделям, названных в диссертации опорными функциями.

В МОФ задается расстояние между двумя моделями  $\eta_i(x)$  и  $\eta_j(x)$  в виде функционала:

$$d_{ij} = \left[ \int_W (\eta_i(x) - \eta_j(x))^2 dx \right]^{\frac{1}{2}}. \quad (I0)$$

Среди всех моделей класса  $S_j$  в качестве опорной функции выбирается та модель, для которой минимальное расстояние до моделей из класса  $S_{j-1}$  будет максимальным, то есть удовлетворяющую следующему выражению:

$$\max_{\eta_j} \min_{\eta_{j-1}} d_{j, j-1}, \quad j=2, \dots, q. \quad (II)$$

В частности, для вложенных друг в друга множеств полиномов условию (II) удовлетворяют ортогональные полиномы Чебышева.

Предполагается, что при имеющейся априорной информации условия эксперимента можно задать величиной обратной отношению сигнал-шум  $NSR = \sigma / \Delta\eta$ , где  $\Delta\eta = \max(\eta(x)) - \min(\eta(x))$ . Преобразованием выходной переменной  $y$  в новую переменную  $y' = (2(y - \min(\eta(x))) / \Delta\eta - 1) / NSR$  задача анализа сводится к случаю единичной дисперсии и новым ограничениям на функцию отклика  $-1 / NSR \leq \eta'(x) \leq 1 / NSR$ , при этом  $\eta'(x) = E(y')$ , зависимости

$L(Cr, \alpha_q, X)$  сохраняют свои характерные особенности и меняются только масштабы у переменных  $L$  и  $\alpha_q$ . Знание области  $W$  определения входных переменных  $x$  и пределов изменения функции отклика  $\eta$  позволяет найти область изменения параметров  $\Omega$ , ( $\alpha_q \in \Omega$ ) и спланировать ВЭ для испытания МСИ. При этом опорные функции соответствуют точкам плана ВЭ в области  $\Omega$ .

В диссертации показано, что при заданном числе наблюдений  $N$ , известном плане эксперимента  $X$  и известном NSR всегда можно с помощью ВЭ определить максимально возможную сложность структуры моделей  $q$  и оптимальный МСИ. При этом зависимость  $\partial L(Cr, \alpha_q, X, \sigma^2)$  достаточно гладкая для полиномиальной аппроксимации и имеет характерные особые точки. К ним относятся граничные точки, точки максимума функций  $L(Cr, \alpha_q, X, \sigma^2)$  и точки равенства пар функций  $L(Cr, \alpha_q, X, \sigma^2)$ . Предлагается решить проблему повышения эффективности технических систем оперативной идентификации и корректного использования имеющейся априорной информации путем разработки и исследования соответствующих систем БД для хранения результатов ВЭ и полученных аппроксимационных зависимостей для  $L(Cr, \alpha_q)$ .

С помощью предложенного МОФ исследованы следующие два класса критериев селекции регрессионных моделей:

1. Обобщенные информационные критерии  $C_p$  Малюса

$$C_j = \text{RSS}_j + \gamma \cdot n_j \cdot \hat{\sigma}^2, \quad (12)$$

где  $\text{RSS}_j$  - остаточная сумма квадратов для модели из класса  $S_j$ ,  $\hat{\sigma}^2$  - оценка дисперсии  $\sigma^2$ , полученная по модели наивысшего ранга  $q$  и  $\gamma$  - регулируемый параметр критерия;

2. Различные модификации критерия скользящего контроля

$$T_j = \frac{1}{N} \sum_{u=1}^N (y_u - \eta_j(x_u, \hat{\alpha}_{j(-u)}))^2, \quad (13)$$

где  $\hat{\alpha}_{j(-u)}$  - оценки вектора параметров, полученные по выборке,

из которой удалено наблюдение с номером  $u$ .

Результаты исследования показали, что существуют области, где один критерий является более предпочтительным чем другой. Так, если дисперсия ошибки меньше 0.6, то критерий  $C_p$  при  $\gamma=2$  лучше критерия  $T_j$  для всех опорных моделей степени выше первой. Для дисперсии ошибки больше 0.6 критерий  $C_p$  лучше критерия  $T_j$  для опорных моделей степени больше второй.

В третьей главе на основе разработанных критериев качества МСИ, ВЭ и МОФ рассматриваются задачи планирования эксперимента для случаев, когда функционал качества плана является случайной величиной и зависит от параметров модели, что имеет место при структурной неопределенности модели ОУ и при наличии в модели случайных регрессоров  $z$ .

Получить аналитическую зависимость функции распределения случайных величин  $v_j$ ,  $j=1,2,\dots,q$  от плана эксперимента  $X$  чрезвычайно сложно. В связи с этим в работе предлагается аппроксимировать распределение случайных величин  $v_j$  с помощью параметров распределения значения критерия  $C_r$ . Так для селекции вложенных моделей с помощью критерия скользящего контроля  $T_j$  предлагается аппроксимировать распределение величин  $v_j$  при фиксированном значении вектора параметров  $\alpha_q$  средними величинами  $T_j$  и в качестве эмпирического критерия качества плана эксперимента использовать величину

$$C(\alpha_{j-1}) = (E(T_1|j) - E(T_1|l)) / (E(T_j|j) - E(T_1|l)) \quad (14)$$

где  $j$  - номер класса моделей, относительно которого определяется величина  $OL$ ,  $l$  - номер класса моделей с минимальным средним значением критерия  $T_j$  для оставшихся классов. Предлагается считать план  $X_*$  "лучше" плана  $X$ , если  $C(X_*) > C(X)$ . Показано, что критерий эффективности плана эксперимента (14) при больших значениях параметра нецентрально-

сти достаточно хорошо аппроксимирует обобщенные критерии качества МСИ, что подтверждается примерами ВЭ для сравнения различных планов.

Ставится и решается задача планирования эксперимента для критерия (9). Предложенная методика рассматривается в основном на примере критерия  $C_p$ . При этом развивается предыдущая методология, основанная на МОФ. Показывается, что зависимость  $EI(X, \alpha_q, \sigma^2)$  представляет собой функцию, определенную на множестве рассматриваемых моделей  $S_q$ , планов эксперимента, векторов неизвестных параметров  $\alpha_q$  и значений дисперсии ошибки  $\sigma^2$ .

Пусть  $X^*$  - план эксперимента, который минимизирует функцию потерь  $\min_j L(j, X)$ ,  $j=1, 2, \dots, q$  для фиксированного вектора параметров  $\alpha_q$  и для фиксированной дисперсии  $\sigma^2$ , тогда можно вычислить функцию

$$L_{\min}(\alpha_q, \sigma^2) = \min_X \min_j L(j, X, \alpha_q, \sigma^2). \quad (15)$$

Функция  $L_{\min}(\alpha_q, \sigma^2)$  представляет собой минимально возможное значение ошибки прогноза для известных параметров и известной дисперсии и значение этой функции не зависит от экспериментальных данных  $Y$ . Для планов эксперимента определяется функция упущенных возможностей

$$\delta L(\alpha_q, \sigma^2) = L(\alpha_q, \sigma^2) - L_{\min}(\alpha_q, \sigma^2), \quad (16)$$

где  $\delta L$  есть функция плана  $X$ , вектора неизвестных параметров  $\alpha_q$ , дисперсии ошибок  $\sigma^2$  и критерия селекции. В работе предлагается для поиска оптимального плана использовать критерии вида

$$\min_X \max_{\alpha_q} \delta L(\alpha_q) \quad \text{или} \quad \min_X \int_{\Omega} \delta L(\alpha_q) d\alpha_q / \int_{\Omega} d\alpha_q.$$

Рассматривается решение ряда задач сравнительного анализа с помощью ВЭ планов для моделей с одной входной незави-

симой переменной и области прогноза, заданной набором дискретных точек, которые подтверждают возможность поиска точных оптимальных планов для фиксированного числа опытов.

При распространении предложенной методики на задачи планирования эксперимента при наличии сопутствующих случайных переменных  $x$  исследуется случай, когда на стадии планирования эксперимента значения неуправляемых переменных неизвестны, а известен только вид закона их распределения и некоторые его параметры, такие как вектор средних и ковариационная матрица. Рассматривается смешанная регрессионная модель вида  $y = \eta(x, z) + \varepsilon$ , где  $x \sim N(x^T L, \Sigma)$ . Предполагается, что вектор неуправляемых, но измеряемых переменных  $x$ , может в свою очередь линейно зависеть от вектора управляемых переменных  $z$ . Для МНК оценок параметров аддитивной модели

$$Y = F \cdot \alpha_1 + Z \cdot \alpha_2 + \varepsilon \quad (17)$$

можно записать условную ковариационную матрицу

$$V \left\{ \begin{matrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{matrix} \right\} \Big| Z = \begin{bmatrix} F^T F & F^T Z \\ Z^T F & Z^T Z \end{bmatrix}^{-1} \sigma^2 = \begin{bmatrix} V(\hat{\alpha}_1) & \dots \\ \dots & V(\hat{\alpha}_2) \end{bmatrix}, \quad (18)$$

где  $Z$  - матрица измерений значений вектора случайных переменных  $x$  в  $N$  опытах.

Для рассматриваемой модели функционал качества плана, также как и рассмотренные в предыдущем параграфе функционалы качества планов для МСИ, является случайными величинами, принимающими при реализации эксперимента конкретные значения. Предлагается планировать эксперимент так, чтобы среднее значение функционала качества было минимальным. В работе определяются и анализируются D- и A-оптимальные в среднем планы эксперимента, которые минимизируют соответственно функционалы  $E(\det V(\hat{\alpha}))$  и  $E(\text{sp} V(\hat{\alpha}))$  и показывается, что для модели (17)

$$E\{ |V(\hat{\alpha}_2)| \} = |\Sigma^{-1}| / \prod_{i=1}^1 (N-k-1-1) \quad \text{и} \quad (19)$$

$$E\{ |V| \} = |F^T F|^{-1} |\Sigma|^{-1} / \prod_{i=1}^1 (n-k-1-1). \quad (20)$$

Из формул (19) и (20) следует, что, во-первых, средняя обобщенная дисперсия оценки вектора  $\alpha_2$  не зависит от матрицы плана, а зависит только от ковариационной матрицы  $\Sigma$  и от параметров  $N, k, l$ , во-вторых, D-оптимальные планы для модели без учета переменных  $x$  совпадают с D-оптимальными в среднем планами для оценивания всех коэффициентов модели (17).

Выражение для математического ожидания обобщенной дисперсии оценки вектора  $\alpha_1$  имеет вид

$$E\{ |V(\hat{\alpha}_1)| \} = |F^T F|^{-1} |\Sigma|^{-1} \left( \sum_{r=0}^1 C_r \sum_{p \in Q(r, l)} |G_p| \right) / \prod_{i=1}^1 (N-k-1-1), \quad (21)$$

где  $C_r = \prod_{i=1}^1 (N-1-1) / \prod_{i=1}^r (n-1-1)$ ,  $\prod_{i=1}^1 (N-1-1) = I$ ,  $Q(r, l)$  при

$l < r \leq 1$  обозначает совокупность всех строго возрастающих последовательностей, составленных из чисел  $1, 2, \dots, l$  по  $r$  чисел в каждой, и матрица  $G_p$  представляет собой матрицу  $\Sigma$ , в которой строки последовательности  $Q(r, l)$  заменены на соответствующие строки матрицы  $L^T F^T F L$ . Из формулы (21) следует что функционал качества плана зависит от значений матриц  $L$  и  $\Sigma$ .

Так как рассмотренные функционалы качества критериев селекции и планов эксперимента зависят от неизвестных параметров и, как правило, аналитического решения не имеют, то для поиска оптимальных статистических стратегий и планов эксперимента используется ВЭ на ЭВМ. При этом предполагается, что для функционала качества процедуры идентификации можно оценить обобщенную модель

$$L(\alpha_q) = \Phi(\alpha_q, \beta) + \omega(\alpha_q), \quad (22)$$

где  $\Phi(\alpha_q, \beta)$  - модель зависимости функционала качества от параметров метода селекции или плана эксперимента,  $\beta$  - вектор оцениваемых параметров функционала по результатам ВЭ.

$\omega(\alpha_q)$  - случайный процесс, отражающий гладкость реальной функции отклика. Если процесс  $\omega(\alpha_q)$  является белым шумом, то к построению расчетной модели (22) можно подходить как к построению обычной регрессионной зависимости. В работе предлагается для построения целевой функции использовать подход, в котором случайный процесс  $\omega(\alpha_q)$  предполагается имеющим нулевое математическое ожидание и взаимно ковариационную функцию  $\sigma_L^2 R(\alpha_q^1, \alpha_q^2)$  между двумя случайными процессами  $\omega(\alpha_q^1)$  и  $\omega(\alpha_q^2)$ , где  $\alpha_q^1$  и  $\alpha_q^2$  - два векторных параметра (две точки параметрического пространства),  $\sigma_L^2$  - дисперсия случайного процесса. Ковариационная функция отражает степень гладкости поверхности отклика и выбирая соответствующим образом эту функцию можно с достаточной степенью точности аппроксимировать поверхность отклика.

В четвертой главе рассматриваются вопросы информационно-логического моделирования данных для МСИ и ВЭ. Вводятся новые операции расширенной реляционной алгебры, которые позволяют решить проблемы, связанные с особенностями использования данных в МСИ.

Первоначально определяется новая операция сложения двух отношений, которая выполняется над отношениями, содержащими как минимум один общий атрибут с доменом из действительных чисел.

Пусть имеется два отношения  $\gamma$  и  $\nu$  со схемами  $R = \{A_1, A_2, \dots, A_k, B_1, B_2, \dots, B_l, C_1, C_2, \dots, C_m\}$  и  $S = \{A_1, A_2, \dots, A_k, B_1, B_2, \dots, B_l, D_1, D_2, \dots, D_n\}$ , в которых атрибуты  $A_1, A_2, \dots, A_k$  имеют действительные значения,  $m \geq l$  и  $Q = R \cup S$ . Суммой отношений  $\gamma$  и  $\nu$  по действительным атрибутам  $A_1, A_2, \dots, A_k$ , входящим в каждое отношение, и записываемой как  $q = \gamma + (A_1, A_2, \dots, A_k)\nu$  является отношение, содержащее все кортежи  $\tau_q$  над  $Q$  такие, что суще-

ствуют кортежи  $t_r \in \Gamma$  и  $t_s \in \Sigma$ , для которых выполнены условия, что для любого  $C_i$   $t_q(C_i) = t_r(C_i)$  ( $i=1,2,\dots,m$ ), для любого  $D_i$   $t_q(D_i) = t_s(D_i)$  ( $i=1,2,\dots,n$ ), для любого  $B_i$   $t_q(B_i) = t_r(B_i) = t_s(B_i)$  ( $i=1,2,\dots,l$ ) и для любого  $A_i$   $t_q(A_i) = t_r(A_i) + t_s(A_i)$  ( $i=1,2,\dots,k$ ).

Аналогичным образом определяются операции разности  $r-s$ , умножения  $r \cdot s$  и деления  $r/s$  для двух отношений.

Показано, что все введенные операции отвечают условиям коммутативности, ассоциативности при одинаковом множестве действительных атрибутов, участвующих в операции, и дистрибутивности для операции сложения относительно операций умножения и деления.

Основными информационными объектами в МСИ являются вектора и матрицы. Предлагается вектора представлять в виде отношения со схемой  $(I, X)$ , где атрибут  $I$  задает номер координаты вектора, а атрибут  $X$  задает значение координаты вектора, матрицы представлять в виде отношения со схемой  $(I, J, X)$ , где атрибут  $I$  задает номер строки, атрибут  $J$  задает номер столбца и атрибут  $X$  задает значение элемента матрицы.

В главе показано, что большинство операций векторного и матричного исчисления реализуются как операции введенной расширенной реляционной алгебры. Приведены формулы реляционной алгебры для сумм и произведений векторов и матриц, для вычисления следа, определителя матрицы и для других матричных операций. Так произведение двух совместимых для операции умножения матриц  $A$  размером  $[M \times N]$  и  $B$  размером  $[N \times K]$ , представленных в виде отношений  $r$  и  $s$ , может быть определено через реляционные операции в виде

$$rs = \sum_{i=1}^N (X) \pi(\sigma(\delta(r; J_1 \rightarrow J) \cdot (X) \delta(s; I_1 \rightarrow I); I_1 = 1 \& J_1 = 1); (I, J, X)). \quad (23)$$

Здесь  $\sigma(\gamma; P(A_1, A_2, \dots, A_n))$  - операция выбора по условию, задаваемому предикатом  $P$ ;  $\pi(\gamma; A)$  - операция проекции ( $A$  - подмножество схемы отношения  $R$ );  $\delta(\gamma; A \rightarrow B)$  - операция переименования атрибута  $B$  в атрибут  $A$ ;  $\&$  - логическая операция "И".

Показано, что через введенные арифметические операции над отношениями легко выражаются вертикальные функции над отношениями, такие как сумма всех значений атрибута, среднее значение, дисперсия, количество элементов и др., что позволяет для повышения скорости вычислений хранить соответствующие суммы и суммы произведений значений случайных переменных как метаданные.

Для описания и анализа аддитивных, мультипликативных и иерархических связей между признаками и переменными вводятся взаимно-однозначные функциональные зависимости, которые означают одновременно имеющие место две функциональные зависимости  $X \rightarrow Y$  и  $Y \rightarrow X$  и обозначаются в виде  $X \leftrightarrow Y$ .

Показано, что имеют место следующие правила вывода для взаимно-однозначных зависимостей ( $\leftrightarrow$  - означает "влечет за собой"): рефлексивность ( $X \leftrightarrow X$ ), аддитивность ( $X \leftrightarrow Y$  и  $X \leftrightarrow Z \leftrightarrow X \leftrightarrow YZ$ ), транзитивность ( $X \leftrightarrow Y$  и  $Y \leftrightarrow Z \leftrightarrow X \leftrightarrow Z$ ) и псевдотранзитивность ( $X \leftrightarrow Y$  и  $YZ \leftrightarrow W \leftrightarrow XZ \leftrightarrow W$ ).

Введен новый класс зависимостей между атрибутами, которые в диссертации названы классификационными зависимостями или  $S$ -зависимостями.

Пусть  $\gamma$  - отношение со схемой  $R$ ,  $X$  и  $Y$  - различные подмножества  $R$  и в множестве  $R$  существуют такие три различных атрибута  $J_1, J_2$  и  $J_3$ , тогда множество атрибутов  $X$  связано с множеством атрибутов  $Y$  аддитивно ( $X \leftrightarrow Y$ ), если имеют место взаимно-однозначные функциональные зависимости  $Y \leftrightarrow J_1$  и  $X \leftrightarrow J_2$  и при этом не существуют другие зависимости  $YXZ \leftrightarrow J_3$ , где

$Z \in R \setminus XY$ ; иерархически  $(X \cdot Y)$ , если имеют место взаимно-однозначные функциональные зависимости  $Y \leftrightarrow J_1$  и  $XY \leftrightarrow J_2$  и при этом не существует зависимости  $X \leftrightarrow J_3$ ; аддитивно-мультипликативно  $(X \cdot Y)$ , если имеют место взаимно-однозначные функциональные зависимости  $Y \leftrightarrow J_1$ ,  $X \leftrightarrow J_2$  и  $XY \leftrightarrow J_3$ ; мультипликативно  $(X \cdot Y)$ , если имеет место взаимно-однозначная функциональная зависимость  $XY \leftrightarrow J_1$  и при этом не существует зависимости  $X \leftrightarrow J_2$  и  $Y \leftrightarrow J_3$ .

Показано, что имеют место следующие правила формирования корректных наборов формул для описания введенных зависимостей между атрибутами (знак  $\sim$  - означает эквивалентную запись): коммутативность  $(X+Y \sim Y+X, X \cdot Y \sim Y \cdot X$  и  $X \cdot Y \sim Y \cdot X)$ , ассоциативность аддитивной зависимости  $(X+Y, Y+Z \sim X+Z, Y+Z \sim X+Z, Y+X)$ , ассоциативность мультипликативной зависимости  $((XY) \cdot Z \sim (XZ) \cdot Y \sim (YZ) \cdot X)$  и ассоциативность аддитивно-мультипликативной зависимости

$$((XY) \cdot Z, X \cdot Z, Y \cdot Z \sim (XZ) \cdot Y, X \cdot Y, Z \cdot Y \sim (YZ) \cdot X, X \cdot Y, Z \cdot X). \quad (24)$$

Существуют запрещенные комбинации С-зависимостей, например, такие как  $X+Y, X \cdot Y; (XY)+Z, X+Z$  и другие. Показано, что с помощью С-зависимостей можно описать большинство классификационных связей между признаками. Например, С-зависимости  $(XY) \cdot Z, Y \cdot Z, X \cdot Z$  задают два признака с перекрестной классификацией  $X$  и  $Y$ , которые сгруппированы внутри уровня признака  $Z$ .

Определяется операция иерархического соединения двух отношений  $r$  со схемой  $R = \{A_1, A_2, \dots, A_k, B, C_1, C_2, \dots, C_l\}$  и с со схемой  $S = \{C_1, C_2, \dots, C_l, D_1, D_2, \dots, D_m\}$ , где домен атрибута  $B$  содержит имя отношения  $s$ .

Иерархическим соединением отношений  $r$  и  $s$  по атрибуту  $B$  и записываемым как  $q = r \rightarrow (B)s$  является отношение со схемой

$Q=RS$ , содержащее все кортежи  $t_q$  над  $Q$  такие, что существуют кортежи  $t_r \in r$  и  $t_s \in s$ , для которых выполнены условия, что  $t_q(B) = s$ , для любого  $A_i$   $t_q(A_i) = t_r(A_i)$  ( $i=1, 2, \dots, k$ ), для любого  $D_i$   $t_q(D_i) = t_s(D_i)$  ( $i=1, 2, \dots, m$ ) и для любого  $C_i$   $t_q(C_i) = t_r(C_i) = t_s(C_i)$  ( $i=1, 2, \dots, l$ ).

На основании введенных операций в главе анализируются системы данных (СД), позволяющие хранить информацию о теоретических свойствах и результатах испытания МСИ с помощью ВЭ.

Первоначально рассматривается структурированная система теоретических данных о МСИ:  $SD_M^i = (R_M^i, W_M^i, I_M^i, (D_{ao}, D_{am}, D_{ae}, D_{ac}, D_{xm}^i, D_{\Pi}), U_M^i, M_C)$ . Здесь  $(D_{ao}, D_{am}, D_{ae}, D_{ac}, D_{xm}^i, D_{\Pi})$  - множество, которое включает в себя следующие СД: об абстрактном ОУ  $D_{ao} = (R_{ao}, W_{ao}, L_{ao}, U_{ao})$ , об абстрактной математической модели  $D_{am} = (R_{am}, W_{am}, L_{am}, U_{am})$ , об абстрактном эксперименте  $D_{ae} = (R_{ae}, W_{ae}, L_{ae}, U_{ae})$ , об абстрактной цели моделирования  $D_{ac} = (R_{ac}, W_{ac}, L_{ac}, U_{ac})$ , о теоретических характеристиках МСИ  $D_{xm}^i = (R_{xm}^i, W_{xm}^i, L_{xm}^i, U_{xm}^i)$  и о программных процедурах  $D_{\Pi} = (R_{\Pi}, W_{\Pi}, L_{\Pi}, U_{\Pi})$ ;  $U_M^i$  - актуальное состояние БД, то есть подмножество всех возможных состояний БД.

В работе предлагается для создания универсальной классификации математических моделей использовать комбинированную ИЛМ данных, описывающую математические модели. Основу такой ИЛМ составляет реляционная модель, которая дополняется текстовой моделью дескрипторного типа. При этом в систему словарей-справочников  $W_{am}$  предлагается ввести дополнительно тезаурус ключевых слов (дескрипторов) и гипертекст с поясняющей текстовой информацией.

На следующем этапе исследования рассматривается структурированная СД о ВЭ

$$SD_{ВЭ} = (R_{ВЭ}, W_{ВЭ}, L_{ВЭ}, (D_{ИО}, D_{ТМ}, D_{ИЭ}, D_{ФЦ}, D_{ХМ}^2), U_{ВЭ}, M_{Т}).$$

Здесь  $\{D_{ИО}, D_{ТМ}, D_{ИЭ}, D_{ФЦ}, D_{ХМ}^2\}$  - множество, которое включает в себя следующие СД: об имитационном ОУ  $D_{ИО} = (R_{ИО}, W_{ИО}, L_{ИО}, U_{ИО})$ , о тестовой математической модели  $D_{ТМ} = (R_{ТМ}, W_{ТМ}, L_{ТМ}, U_{ТМ})$ , об имитационном эксперименте  $D_{ИЭ} = (R_{ИЭ}, W_{ИЭ}, L_{ИЭ}, U_{ИЭ})$ , о целевых функциях моделирования  $D_{ФЦ} = (R_{ФЦ}, W_{ФЦ}, L_{ФЦ}, U_{ФЦ})$  и о характеристиках МСИ, полученных из ВЭ  $D_{ХМ}^2 = (R_{ХМ}^2, W_{ХМ}^2, L_{ХМ}^2, U_{ХМ}^2)$ ;  $U_{ВЭ}$  - актуальное состояние БД о ВЭ.

Интеграция двух рассмотренных СД осуществляется путем создания структурированной СД о МСИ

$$SD_M = (R_M, W_M, L_M, (SD_M^1, D_{ХМ}^2), U_M).$$

На основе разработанных СД предлагается технология формирования знаний о результатах испытания МСИ.

В пятой главе рассматриваются вопросы оптимального построения реляционных моделей данных для процессов идентификации ОУ при наличии ошибок загрузки и модификации данных.

Пусть при записи или модификации данных происходят ошибки с вероятностью совершения ошибки при модификации одной записи из  $n$  записей равной  $\rho$  и пусть  $L(e)$  функция потерь, которая зависит от числа ошибок  $e$  в БД и не зависит от конкретной записи, в которой произошла ошибка. Так как число  $e$  и функция  $L(e)$  являются случайными величинами, то в качестве критерия эффективности схемы БД используется среднее значение функции потерь  $E(L(e))$ .

Рассматриваются БД, хранящие  $k$  различных значений переменной  $x$  в  $n$  опытах ( $k \leq n$ ), задаваемые атрибутом  $X$ , что особенно характерно для спланированного эксперимента с фиксированными уровнями факторов. В этом случае возможное количество ошибок  $e$  меняется в пределах от 0 до  $n$ . В диссертации до-

казана следующая теорема.

**Теорема.** Пусть рассматриваются две следующие схемы базы данных: схема  $R_1$ , состоящая из одного отношения  $r(X)$ , и схема  $R_2$ , состоящая из двух отношений, являющихся результатом выполнения операции нормализации отношения. Если ошибки в данных распределены по биномиальному закону распределения,

$$f(e | n, \rho) = \begin{cases} C_n^e \rho^e (1-\rho)^{n-e}, & e=0, 1, \dots, n \\ 0 & \text{в остальных случаях.} \end{cases} \quad (C_n^e = \frac{n!}{e!(n-e)!})$$

и функция потерь выпукла вверх по числу ошибок  $e$ ,

$$L(e) \geq L(e_1) + \frac{(e - e_1)}{(e_2 - e_1)} (L(e_2) - L(e_1)) \quad (e_1 \leq e \leq e_2),$$

тогда имеет место следующее неравенство  $E(L(R_1)) \geq E(L(R_2))$ .

На основании теоремы показано, что при выпуклости функции потерь по количеству ошибок в БД приведение реляционной схемы данных к нормальным формам путем декомпозиции отношений является эффективной операцией.

В работе показано, что для задачи вычисления среднего и дисперсии по экспериментальным данным с аддитивными ошибками в БД и функционалами качества

$$L(\mu, e) = E\left(\sum_{i=1}^n (y_i + \delta_i) / n - \mu\right)^2 = (\sigma_1^2 + e\sigma_2^2) / n^2 \quad (25)$$

$$\text{и } L(\hat{\sigma}_1^2, e) = E\left(\sum_{i=1}^n (y_i + \delta_i)^2 - \left(\sum_{i=1}^n (y_i + \delta_i)\right)^2 / n\right) / (n-1) - \sigma_1^2. \quad (26)$$

неравенство в теореме обращается в равенство и обе реляционные схемы данных эквивалентны.

Пусть для моделей вида (4) и МНК качество модели оценивается функционалом

$$L = E\{(\hat{\alpha} - \alpha)^T (\hat{\alpha} - \alpha)\}. \quad (27)$$

В случае возникновения ошибок в записи уровня управляемой переменной вместо матрицы  $F$  в функционале оценивания будет использоваться случайная матрица  $F_e$ , в которой записа-

но е ошибочных значений уровня управляемой переменной. В этом случае функция потерь принимает вид

$$L(e) = \alpha^T E(A) \alpha + \sigma^2 E(\text{sp}(F_e^T F_e)^{-1}) \quad (28)$$

где  $D_e = F_e - F$  и  $A = D_e^T F_e (F_e^T F_e)^{-2} F_e^T D_e$ .

Получить аналитическое выражение для функционала (28) в общем виде не удастся. В связи с этим для аппроксимации функционала (28) используется ВЭ и моделирование с помощью метода Монте-Карло.

В главе исследуется структура СД о планах эксперимента  $D_{пе} = (R_{пе}, W_{пе}, L_{пе}, U_{пе})$ , основанная на использовании каталогов планов. СД содержит как информацию о традиционных критериях оптимальности планов эксперимента, таких как  $D$ ,  $A$ ,  $E$ -оптимальность, так и ограничения на величину функции отклика  $\eta(x)$ , класс моделей для селекции  $S_q$ , данные о методе селекции, множество значений критерия качества для МОФ, полученные по результатам ВЭ с тестовыми моделями, и аппроксимационные зависимости для критериев оптимальности процедур селекции.

В работе предлагается интегрировать СД о планах эксперимента со структурированными СД, описывающих МСИ и результаты ВЭ. Интеграция на уровне данных осуществляется с помощью обобщенной реляционной схемы, задающей связи между СД  $D_{пе}$  и  $D_{ме}$ .

В работе показано, что с помощью введенной расширенной реляционной алгебры реализуются следующие специальные операции над данными:

1. Вычисление минимального значения функционала качества  $L_{\min}(\alpha_q, S_q, Cr)$  при заданном классе моделей  $S_q$ , критерии селекции и векторе неизвестных параметров  $\alpha_q$  по подмножеству планов эксперимента  $X(S_q) \subseteq X$ , допустимых для класса моделей

$S_q$ ;

2. Построение аппроксимационного функционала качества для конкретного плана эксперимента  $\Phi(\alpha_q, \beta, S_q, X_i, Cr)$  на основе ВЭ над тестовыми моделями;

3. Вычисление непараметрического функционала качества для множества планов  $X \omega(\alpha_q, X, S_q, Cr)$  и оценок максимального правдоподобия параметров ковариационной функции  $R(\alpha_1, \alpha_2, X, S_q)$  на основании данных вновь проведенных ВЭ;

4. Вычисление прогнозируемого значения функции упущенных возможностей  $\delta L(\alpha_q, X_i, Cr)$  для каждого плана;

5. Выбор при заданных ограничениях на функцию отклика  $\eta(x)$  и дисперсию ошибки  $\sigma^2$  плана эксперимента по критерию  $\min_X \max_{\alpha_q} \delta L(\alpha_q, X_i, Cr)$ ;

6. Выбор плана эксперимента при заданных ограничениях на функцию отклика  $\eta(x)$  и дисперсию ошибки  $\sigma^2$  по критерию  $\min_X \int \delta L(\alpha_q, X_i, Cr) d\alpha_q$ ;

7. Вычисление средних значений функционалов качества для моделей со случайными переменными (Г7) при заданных коэффициентах линейной связи между переменными  $L$  и матрицей ковариаций между случайными переменными  $\Sigma$ ;

В главе приведены примеры ИЛМ различного типа планов эксперимента, которые подтверждают возможность реализации предложенной структуры СД с помощью существующих программных средств.

В шестой главе рассматриваются вопросы информационной технологии идентификации ОУ на основе создания единой базы знаний о МСИ и планах эксперимента. Определены и исследованы основные структуры данных и процессы преобразования информации. Сформулированы принципы, позволяющие

создавать расчетно-логические системы идентификации сложных ОУ.

Первоначально исследуется метасистема ВЭ, состоящая из структурированной СД о методах моделирования  $SD_M$  и структурированной СД о ВЭ  $SD_{ВЭ}$ . Метасистему ВЭ можно формально определить в виде:  $MSD_{МВЭ} = (R_{МВЭ}, W_{МВЭ}, L_{МВЭ}, (SD_{ВЭ}, SD_M), K_{МВЭ})$ . Здесь  $K_{МВЭ}$  - система продукций, задающая правила перехода от одной подсистемы к другой и осуществляющая интеграцию на уровне процедур двух структурированных СД.

Для организации интерфейса со специалистами, загружающими и использующими базу знаний, предлагается организовать СД о сценариях диалогов  $D_d = (R_d, W_d, L_d, U_d)$ , которая является неотъемлемой частью информационной технологии моделирования.

Формально технологическую операцию тестирования методов можно записать в виде следующей процедуры:

$$(SD_{ВЭ}, SD_M, D_d) \xrightarrow{M_T, K_{ВЭ}} U_{ВЭ}. \quad (29)$$

Здесь состояние СД о ВЭ  $U_{ВЭ}$  формируется на основании всей семантической информации, содержащейся в метасистеме, с помощью системы продукций  $K_{ВЭ}$  и методов тестирования  $M_T$ .

В главе рассматривается обобщенная структура информационной технологии идентификации ОУ. С этой целью анализируется структура информационной среды МСИ, связанная с использованием метода для обработки реальных экспериментальных данных. Эта информационная среда задается структурированной СД об ОУ  $SD_{ОИ} (R_{ОИ}, W_{ОИ}, L_{ОИ}, (D_О, D_{км}, D_в, D_ц, D_{км}), U_{ОИ})$ . Здесь в структуру системы включены следующие подсистемы данных, которые являются нижним уровнем информационной технологии: система семантического описания конкретного ОУ  $D_О = (R_О, W_О, L_О, U_О)$ ; СД о результатах конкретных экспериментов,

проведенных над объектом  $D_{\theta} = (R_{\theta}, W_{\theta}, L_{\theta}, U_{\theta})$ , которая содержит собственно экспериментальные данные; СД о целях моделирования  $D_{\zeta} = (R_{\zeta}, W_{\zeta}, L_{\zeta}, U_{\zeta})$ , которая включает в себя информацию о конкретной цели использования будущей модели, о функциях предпочтения, риска, потерь и другую информацию, необходимую для принятия решения о выборе наилучшей модели; СД о конкретных математических моделях, построенных по реальным экспериментальным данным для ОУ,  $D_{\text{ЮМ}} = (R_{\text{ЮМ}}, W_{\text{ЮМ}}, L_{\text{ЮМ}}, U_{\text{ЮМ}})$ ; СД о характеристиках модели, связанных с МСИ и планом эксперимента,  $D_{\text{ХМ}} = (R_{\text{ХМ}}, W_{\text{ХМ}}, L_{\text{ХМ}}, U_{\text{ХМ}})$ , которую имеет смысл выделить в отдельную СД, так как она позволяет осуществлять сравнительный анализ моделей, полученных различными МСИ.  $R_{\text{ОИ}}, W_{\text{ОИ}}$  и  $L_{\text{ОИ}}$  задают семантические конструкции для создания структурированной СД и объединяют все СД об ОУ в одну.

Метасистема данных расчетно-логической системы моделирования или база знаний задается в виде совокупности множеств  $MSD_{\text{ОМ}} = (R_{\text{ОМ}}, W_{\text{ОМ}}, L_{\text{ОМ}}, \{SD_{\text{ОИ}}, SD_{\text{М}}\}, K_{\text{ОМ}})$ . Метасистема состоит из двух структурированных систем  $SD_{\text{ОИ}}$  и  $SD_{\text{М}}$ , дополнительных интегрирующих семантических схем  $R_{\text{ОМ}}, W_{\text{ОМ}}, L_{\text{ОМ}}$  и системы продукций, осуществляющих правила передачи информации между системами  $K_{\text{ОМ}}$ . Включение в систему идентификации ОУ автоматизированного ВЭ обуславливает создание интегрированной метасистемы, основанной на концепции базы знаний о методах селекции,  $MSD_{\text{И}} = (R_{\text{И}}, W_{\text{И}}, L_{\text{И}}, \{MSD_{\text{ОМ}}, MSD_{\text{МВЭ}}\}, K_{\text{И}})$ .

Технологию функционирования такой автоматизированной системы идентификации можно описать следующим образом:

Шаг I. Описывается ОУ. То есть происходит следующее преобразование информации в системе:

$$(U_{\text{д}}, U_{\text{вО}}, (R_{\text{О}}, W_{\text{О}}, L_{\text{О}})) \xrightarrow{Q_1(K_{\text{ОМ}})} Q_2(U_{\text{О}}) \quad (30)$$

Здесь  $Q_1()$  и  $Q_2()$  - обозначают операторы формирования подмножества. На этом шаге используется СД о сценариях диалогов  $U_d$ , текущее состояние БД об абстрактных ОУ  $U_{ao}$  и семантическое описание данных об ОУ. Управляющая информация задается в подмножестве продукций  $Q_1(K)$ , описывающих правила описания объектов.

Шаг 2. В систему загружаются существующие экспериментальные данные.

$$(U_o, U_d, U_{av}, (R_v, W_v, L_v)) \xrightarrow{Q_2(K_{om})} (Q_4(U_v), Q_5(U_{om})) \quad (31)$$

Управление процессом осуществляется подмножеством продукций расчетно-логической системы моделирования  $Q_2(K_{om})$ . В результате изменяются состояния СД об экспериментальном материале и ОУ.

Шаг 3. Определение цели моделирования. На этом шаге формируются целевые функции, определяется назначение математической модели.

$$(U_v, U_{oi}, U_d, U_{ac}, (R_c, W_c, L_c)) \xrightarrow{Q_6(K_{om})} Q_7(U_c) \quad (32)$$

Здесь используется вся информация об ОУ, полученная на предыдущих шагах и имеющаяся информация в системе об абстрактных целях моделирования  $U_{ac}$ .

Шаг 4. После того как описан ОУ, определена цель моделирования и имеются экспериментальные данные определяется класс моделей и критерии оптимальности модели:

$$(U_c, U_v, U_{oi}, U_d, U_{am}, (R_{am}, W_{am}, L_{am})) \xrightarrow{Q_8(K_{om})} Q_8(U_{am}). \quad (33)$$

В этом случае происходит формирование подмножества данных об абстрактных математических моделях, которое отвечает поставленной цели и ОУ.

Шаг 5. Зная класс моделей и критерии оптимальности можно решить вопрос о выборе МСИ и методов оценивания парамет-

ров моделей.

$$(U_{ам}, U_{ц}, U_{в}, U_{ои}, U_{д}, U_{м}, (R_{м}, W_{м}, I_{м})) \xrightarrow{Q_{10}(K_{ом})} Q_{11}(U_{м}) \quad (34)$$

Шаг 6. Построение модели.

$$(SD_{ои}, D_{п}) \xrightarrow{Q_{12}(K_{ом}), M_{о}} (Q_{13}(U_{км}), Q_{14}(U_{ои})). \quad (35)$$

Этот шаг может иметь итерационный характер. При каждой итерации происходит поиск и выполнение соответствующих алгоритмов и формируются параметры, которые заносятся в БД о конкретных математических моделях  $U_{км}$ , содержащей оценки параметров модели и значения соответствующих статистических критериев. Одновременно меняется состояние СД об ОУ  $U_{ои}$ .

Шаг 7. Если необходимо получить дополнительные экспериментальные данные, то выполняется этап планирования эксперимента.

$$(SD_{м}, SD_{ои}, SD_{вв}, D_{пв}) \xrightarrow{Q_{15}(K_{и})} (Q_{16}(U_{пв}), Q_{17}(U_{ои})). \quad (36)$$

После этого шага проводится эксперимент с ОУ и возможен возврат к предыдущим шагам технологического цикла.

Шаг 8. На этом шаге происходит проверка выполнения поставленной цели, принятие решения о дальнейшем моделировании, после чего весь технологический цикл может быть повторен вновь.

Возможность практической реализации предложенной информационной технологии проиллюстрирована в главе примерами основных отношений реляционной схемы и систем продукции.

В седьмой главе диссертации описываются созданные на основе теоретических результатов работы программные комплексы: экспертная система экспериментального моделирования (ЭСЭМ), интеллектуальное рабочее место по планированию эксперимента (ИРМПЭ) и экспертная система по выбору методов моделирования (ЭСММ). ЭСЭМ представляет из себя интегрирован-

ную систему хранения и обработки экспериментальных данных на ПЭВМ. ЭСЭМ имеет интерфейс с пакетами прикладных программ, осуществляющих селекцию регрессионных моделей и ВЭ по испытанию МСИ. ИРМПЭ представляет собой программно-технический комплекс на базе ПЭВМ, содержащий реляционную БД, в которую загружен каталог планов, экспертную систему производственного типа с правилами по выбору оптимального плана и систему интегрированного интерфейса с пользователем. ЭСММ представляет собой расчетно-логическую систему по выбору методов синтеза и оптимизации моделей сложных динамических объектов, в число которых включены также методы регрессионной идентификации. В системе использованы предложенные в работе структуры СД и правила выбора МСИ.

В главе приводятся результаты идентификации процесса отказов автоматизированного склада на Мелитопольском моторном заводе и процесса биологической очистки морской воды от нефтяного загрязнения, осуществленные с помощью созданных программных комплексов.

В заключении сформулированы выносимые на защиту результаты диссертационной работы.

В приложениях представлены список основных обозначений и акты о внедрении результатов диссертации.

#### ОБЩИЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

В диссертации предлагается для решения актуальной научной проблемы автоматизации процессов идентификации сложных ОУ использовать "машинемкие" методы статистического моделирования, основанные на соединении статистической теории с методом Монте-Карло, и информационные технологии, в которых широко используются БД и базы знаний.

Исследование роли ВЭ для анализа МСИ и планирования экс-

перимента показало, что применение БЭ позволяет снизить ошибку выбора МСМ и плана эксперимента, повысить прогнозирующую точность синтезируемых моделей и ускорить включение новых методов и алгоритмов в автоматизированные системы идентификации сложных ОУ.

Применение информационно-логического моделирования для организации данных ведет к снижению избыточности данных за счет использования механизма нормализации отношений и повышения надежности данных путем анализа специальных взаимосвязей между данными в реляционной БД, что в конечном итоге повышает качество процесса идентификации.

В работе получены следующие результаты, обладающие научной новизной:

1. Разработаны и исследованы новые критерии оптимальности процедуры идентификации, использующие усредненное по области возможных значений параметров моделей или максимальное значение функции упущенных возможностей.

2. Разработан и исследован метод опорных тестовых функций для определения областей предпочтительности критериев селекции регрессионных моделей и планов эксперимента, позволяющий формализовать выбор тестовых функций для испытания методов моделирования и планов. С помощью этого метода исследованы обобщенные информационные критерии селекции моделей и класс критериев скользящего контроля при разном объеме априорной информации об ОУ.

3. Разработан метод построения аппроксимационных моделей для оценивания и прогноза целевых функций, основанный на применении регрессионного и корреляционного анализа к данным БЭ. Показано, что эффективность метода может быть значительно увеличена за счет использования БД с результатами испыта-

ний.

4. Сформулированы задачи построения оптимальных в среднем планов и исследованы свойства  $\bar{U}$  и  $A$  оптимальных в среднем планов, что позволило разработать методику и алгоритмы поиска планов эксперимента для моделей с сопутствующими случайными переменными, основанные на широком использовании данных БД.

5. Предложена и исследована концепция комплексного статистического и информационно-логического моделирования в системах автоматизированной идентификации  $\bar{U}Y$ , с помощью которой удалось показать, что эффективность процесса идентификации  $\bar{U}Y$  во многом зависит от организации и технологии использования БД.

6. В рамках концепции разработаны основы теории реляционных БД для систем автоматизированной идентификации. Основой теории являются предложенные в диссертации арифметические операции над отношениями, которые расширяют реляционную алгебру, и новые типы классификационных зависимостей между атрибутами, что дает возможность информационно-логического моделирования сложных структур экспериментальных и модельных данных.

7. Показано, что с помощью введенной расширенной реляционной алгебры большинство операций матричного исчисления, используемых в методах идентификации могут быть реализованы в виде запросов к БД, что позволяет анализировать весь процесс получения данных и их использования в рамках единой теории.

8. Сформулированы задачи анализа ИЛМ данных для БД, в которых возникают ошибки при записи и коррекции информации. Предложено надежность и эффективность БД оценивать по ста-

статистическим характеристикам моделей  $\bar{U}$ , получаемым по информации, хранящейся в БД. Исследованы процессы нормализации отношений с ошибками в БД. Получены условия, при которых нормализация отношений повышает качество процесса идентификации.

9. Введены дополнительные операции над данными, связанные с построением и корректировкой аппроксимационных зависимостей для функционалов качества методов моделирования и планов эксперимента. Показано, что эти операции могут быть реализованы в виде выражений расширенной реляционной алгебры.

10. На основе разработанных теоретических положений предложена информационная технология построения регрессионных моделей и планирования эксперимента, базирующаяся на концепции базы знаний о МСИ на этапе его создания, испытания с помощью ВЭ и эффективного применения к обработке реального эксперимента. Разработаны и исследованы семантические конструкции в виде реляционных моделей данных, описывающие методы моделирования, результаты ВЭ и реальный  $\bar{U}$ .

11. Создано на основе теоретических результатов работы научное методическое, математическое и программное обеспечение конкретных экспертных и расчетно-логических систем идентификации  $\bar{U}$ , которое используется в ряде организаций, показало свою работоспособность, позволило решить ряд разнообразных технологических и технических задач идентификации и обеспечило соответствующий экономический и технический эффект.

Подходы и методы, предложенные в диссертации могут использоваться не только для класса систем, описываемых линейными по параметрам регрессионными моделями, но и для других

систем построения моделей по экспериментальным данным.

Совокупность разработанных в диссертации теоретических положений и методов может быть квалифицирована как новое крупное достижение в развитии перспективного направления исследования автоматизированных систем идентификации сложных объектов управления.

ОСНОВНЫЕ ПОЛОЖЕНИЯ ДИССЕРТАЦИИ ОТРАЖЕНЫ В ПУБЛИКАЦИЯХ:

1. Цуканов А.В. Оптимальное моделирование и планирование эксперимента.- Киев : Знание, 1986. - 20 с.
2. Цуканов А.В. Экспертные системы экспериментального моделирования в энергетике.- Киев: Знание, 1989.-19 с.
3. Подвинцев В.В., Цуканов А.В. Об одном методе активно-пассивного поиска экстремума// Изв. АН СССР.- Сер. Техническая кибернетика.- 1976.- №1.- С.55-59.
4. Новиков В.Э., Цуканов А.В., Копытов Ю.П. Снижение размерности при построении математической модели биосистемы// Биология моря.- Киев: Наукова думка.-1977.- Вып.40.- С.51- 54.
5. Цуканов А.В. О критериях оптимальности планов эксперимента для одной регрессионной модели со случайными переменными// Теория вероятностей и ее применения.- 1981.- №1.- С.178 - 182.
6. Бондарев В.Н., Виленчик Д.В., Гольденберг С.В., Цуканов А.В. Информационная технология построения математических моделей объектов// Электронное моделирование.- 1992.- Т.14.- №1.- С. 90- 94.
7. Цуканов А.В., Виленчик Д.В., Бондарев В.Н. Организация обмена знаниями между различными системами с искусственным интеллектом// Программные продукты и системы.- 1992.- № 3.- С. 35-40.
8. Hertzberg A.M., Tsukanov A.V. The Monte-Carlo Comparison

of two Criteria for the Selection of Models// J. Statist. Comput. Simul.- 1985.- Vol. 22.- P.113-126.

9. Herzberg A.M., Tsukanov A.V. The design of Experiments for Model Selection with the Jackknife Criterion// Utilitus Mathematica.-1985.-Vol. 28.-P. 243-253.

10. Herzberg A.M., Tsukanov A.V. A note on Modifications of the jackknife Criterion for Model Selection// Utilitus Mathematica.- 1986.- Vol.29.- P. 209-216.

11. Herzberg A.M., Tsukanov A.V. The problem of the joint choice of the experimental design and the criterion for the selection of regression models// Proceedings of the conference "Statistical methods in experiment research and quality control". Varna (Bulgaria).-1986.- Vol.3.- P.23-25.

12. Herzberg A.M., Tsukanov A.V. The design of experiments for model selection// Proceedings of the 1 -st World Congress of the Bernoulli Society.- 1987.- Vol.2.- P. 175-178.

13. Tsukanov A.V. ESEM - Expert system of experimental modeling// Computer in the USSR.- 1991.- N1.- P.10-11.

14. Tsukanov A.V., Zhuravets A.I. The one method of the experimental design construction for a class of models in the presence of uncontrolled variables// Advances in modeling & analysis.- 1993.- Ser.B.- V.27.- N4.- P.57-63.

15. Tsukanov A.V., Tsepkova N.A. The Monte-Carlo comparison of the two strategies of the model identification// Advances in modeling & analysis.-1995.-Ser.B.-V.32.-N3.- P.55-63.

16. Tsukanov A.V., Potanina M.V. The choice of the estimation method for the selection of the optimal regression model of a complex object// Advances in modeling & analysis.- 1995.- Ser.B.- V.32.- N3.- P.55-63.

17. Tsukanov A.V., Tsepkova N.A. The comparison of the two

approximation schemes efficiency for the model selection procedure // Advances in modeling & analysis.- 1995.- Ser.A.- V.31.- N1.- P.29-37.

18. Tsukanov A.V. Integrated statistical and information models of systems with databases // Advances in modeling & analysis.- 1995.- Ser.B.- V.34.- N2.- P.27-34.

19. Herzberg A.M., Tsukanov A.V. The design of experiments for Model Selection: Minimization of the expected Mean-Squared Error // Utilitatus Mathematica.-1995.-vol.47.- P.85-96.

20. Цуканов А.В. Планирование D-оптимального в среднем эксперимента для одного вида регрессионных моделей со случайными переменными / Севастопольский приборостроит. ин-т.- Севастополь, 1980.- 20 с.- Деп. в ВИНИТИ, N4265-80Деп.

21. Цуканов А.В. База экспериментальных данных в среде СУБД ПАЛЬМА / Севастопольский приборостроит. ин-т.- Севастополь, 1988.- 28 с.- Деп. в УкрНИИТИ, N2001-Ук88.

22. Цуканов А.В., Потанина М.В. База программных модулей в среде СУБД ПАЛЬМА / Севастопольский приборостроит. ин-т.- Севастополь, 1989.- 26 с.- Деп. в УкрНИИТИ, N497-Ук89.

23. Цуканов А.В., Бурилова Н.Р., Журавец А.И. База данных сценариев диалогов / Севастопольский приборостроит. ин-т.- Севастополь, 1989.- 50 с.- Деп. в УкрНИИТИ, N2659-Ук89.

24. Технология обмена знаниями между различными системами с искусственным интеллектом / Бондарев В.Н., Виленчик Д.В., Журавец А.И., Цуканов А.В., Севастопольский приборостр. ин-т.- Севастополь, 1993.- 50 с.- Деп. в УкрИНТЭИ, N587-Ук93.

25. Цуканов А.В., Журавец А.И. О выборе плана эксперимента для построения прогнозирующей линейной по параметрам регрессионной модели при наличии сопутствующих случайных переменных / Севастопольский приборостроит. ин-т.- Севастополь,

- 1994.- 17с.- Деп. в ГНТБ Украины, NI654-Ук94.
26. Цуканов А.В. Интегрированные статистические и информационные модели систем с базами данных/ Севастопольский приборостроит. ин-т.- Севастополь, 1994.-15 с.- Деп. в ГНТБ Украины, NI814-Ук94.
27. Цуканов А.В., Потанина М.В. Сравнительный анализ эффективности процедуры селекции переменных и ридж-регрессии/ Севастопольский приборостроит. ин-т.- Севастополь, 1994.- 17 с.- Деп. в ГНТБ Украины, NI652-Ук94.
28. Цуканов А.В., Цепкова Н.А. Двухступенчатая процедура селекции моделей/ Севастопольский гос. техн. ун-т.- Севастополь, 1995.-10 с.-Деп. в ГНТБ Украины, N84-Ук95.
29. Цуканов А.В., Цепкова Н.А. Две схемы аппроксимации для сравнения эффективности процедуры селекции моделей / Севастопольский гос. техн. ун-т.- Севастополь, 1995.-12с.-Деп. в ГНТБ Украины, N83-Ук95.
30. Цуканов А.В. К построению теории статистических баз данных/ Севастопольский гос. техн. ун-т.- Севастополь, 1995.- 29с.- Деп. в ГНТБ Украины, N 1534-Ук95.
31. Цуканов А.В. Экспериментальная оптимизация одного класса технологических процессов с контролируруемыми случайными возмущениями// Тез. докл. Всесоюз. конф.: Перспективы и опыт внедрения статистических методов в АСУ ТП (май 1981, Смоленск).- Смоленск, 1981.- С.144-146.
32. Новиков В.Э., Цуканов А.В. Планирование эксперимента при сравнительном анализе многомерных наблюдений// Тез. докл. Всесоюз. конф.: Перспективные методы планирования и анализа экспериментов при исследовании случайных полей и процессов (ноябрь 1982, Нальчик).- М., 1982.-Ч.1.- С. 210-211.
33. Дуганов Г.А., Патрикеев Л.Я., Цуканов А.В. Моделирование

- систем отключения импульсного питания с программным управлением и классификацией утечек шахтной контактной сети// Тез. докл. 3-й Всесоюз. конф.: Безопасность и надежность горно-рудных предприятий (1982, Днепропетровск).- 1982.- С. 95-98.
34. Цуканов А.В. О критериях выбора оптимальной модели технологического процесса// Тез. докл. 2 Всесоюз. конф.: Перспективы и опыт внедрения статистических методов в АСУ ТП (май 1984, Смоленск).- Смоленск, 1984.- С. 71-72.
35. Новиков В.Э., Цуканов А.В. Планирование экспериментов при обнаружении изменчивости характеристик сложного объекта // Тез. докл. 2 Всесоюз. конф.: Перспективные методы планирования и анализа экспериментов при исследовании случайных полей и процессов.- М., 1985.- Ч.2.- С. 48-49.
36. Цуканов А.В. Построение концептуальной модели базы данных на этапе предварительного описания объекта для автоматизированной консультационной системы по прикладной статистике// Тез. докл. Респуб. конф.: Повышение качества программного обеспечения ЭВМ ( сентябрь 1986, Севастополь).- Севастополь, 1986.- С. 68-69.
37. Новиков В.Э., Цуканов А.В. Планирование экспериментов для выявления разладки// Тез. докл. 8 Всесоюз. конф.: Планирование и автоматизация эксперимента в научных исследованиях. Секц. I; Математическая теория планирования эксперимента ( октябрь 1986, Ленинград ).- М., 1986.- С.9.
38. Цуканов А.В. Особенности подготовки специалистов в области автоматизированных банков данных// Тез. докл. научно-методического семинара по компьютеризации учебного процесса (1989, Севастополь).- Севастополь; СВВМИУ, 1989.- С.44-46.
39. Потанина М.В., Цуканов А.В. База данных о программных модулях в АСНИ// Тез. докл. зональной н.т. конф.: Обработка

информации в АСНИ (1989, Пенза ).- Пенза, 1989.- С.31-32.

40. Цуканов А.В., Потанина М.В. Реляционная база данных по планированию многомерных факторных экспериментов// Тез. докл. 9 Всесоюз. конф.: Планирование и автоматизация эксперимента в научных исследованиях (сентябрь 1989, Москва).- М.;МЭИ, 1989.- Ч.2.-С.64.

41. Цуканов А.В. Цепкова Н.А. Гипертекст на основе реляционной СУБД// Всесоюзный н.т. семинар: Программное обеспечение новых информационных технологий ( октябрь 1991, Тверь ).- Тверь, 1991, С. 23-24.

42. Журавец А.И., Цуканов А.В. Планирование активно-пассивного эксперимента при неопределенности структуры модели// Тез. докл. X научн. конф.: Планирование и автоматизация эксперимента в научных исследованиях (октябрь 1992, Москва).- М.: АНТАЛ, 1992.- С. 17-18.

43. Новиков В.Э., Цуканов А.В. Рабочее место по планированию эксперимента в структуре интеллектуальной системы научных исследований (ИСНИ)// Матер. Респ. семинара: Прикладные интеллектуальные системы ( апрель 1992, Москва).- М.: ЦРДЗ, 1992.- С. 66-67.

44. Новиков В.Э., Цуканов А.В. Интеллектуальное рабочее место по планированию эксперимента и обработке данных// Тез. докл. X научн. конф.: Планирование и автоматизация эксперимента в научных исследованиях (октябрь 1992, Москва).- М.: АНТАЛ, 1992.- С. 59.

45. Цуканов А.В., Херзберг А.М. Планирование эксперимента при многокритериальном методе селекции регрессионных моделей// Тез. докл. X научн. конф.: Планирование и автоматизация эксперимента в научных исследованиях (октябрь 1992, Москва).- М.:АНТАЛ, 1992.- С.32.

46. Цуканов А.В., Потанина М.В. Совместный выбор метода оценивания и метода селекции при построении оптимальной регрессионной модели сложного объекта // Тез. докл. международного семинара: Экспериментально-статистическое моделирование в компьютерном материаловедении (май 1993, Одесса). - Киев: Знание, 1993. - С.27.

47. Цуканов А.В., Журавец А.И. Об одном методе построения плана эксперимента для класса моделей в присутствии неконтролируемых переменных // Тез. докл. международного семинара: Экспериментально-статистическое моделирование в компьютерном материаловедении (май 1993, Одесса). - Киев: Знание, 1993. - С.28.

Цуканов А.В. Статистичні та інформаційно-логічні моделювання в ідентифікації складних об'єктів керування. Дисертація на здобуття наукового ступеня доктора технічних наук по спеціальності 05.13.03 - системи та процеси керування. Севастопольський державний технічний університет. Севастополь, 1995.

Захищається рукопис на базі 47 робіт, що містять результати досліджень проблеми вибору метода структурної ідентифікації та плану експерименту при наявності апріорної інформації про дисперсії помилки спостережень та галузі можливих значень змінних. Вибір ґрунтується на методі чисельного експерименту по тестуванню критерію селекції моделей та плану експерименту. Доказується, що ефективність системи ідентифікації може бути поліпшена за рахунок використання технології баз даних. Розроблено теорію реляційних баз даних для систем ідентифікації з можливими похибками в даних.

Tsukanov A.V. Statistical and information-logical modeling in the identification of complex control objects. Doctor of technical science thesis, speciality 05.13.03 - control systems and processes. Sevastopol State Technical University, Sevastopol, 1995.

The manuscript based on the 47 articles is defended. It contains the results of the investigations of the problem of the choice of the structural identification method and the design of experiment when there is the prior information about the variance of the error of observations and regions of possible values of variables. The choice is based on the computer simulation method for testing the selection models criteria and the design. It was proved that the data base technology improve the efficiency of the identification system. The theory of relation data base for identification systems with possible errors in data is presented.

Ключеві слова: структурна ідентифікація, план експерименту, чисельний експеримент, реляційні бази даних, інформаційно-логічні моделювання, програмні системи.

Сдано в набор 12.09.95. Подписано в печать 11.09.95.  
Формат бумаги 60x84 1/16. Бум.тип. №2. Офсет. печать.  
Усл.печ.л. 2.29. Усл.изд.л. 2. Тираж 100. Заказ №137.

---

КМУ СГТУ. Севастополь, Гоголя, 14.

11/09/95

AB 33.090