

Л.С. Ломакина, Д.В. Ломакин, С.А. Зеленцов

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
ДИАГНОСТИРОВАНИЯ МНОГОМЕРНЫХ
ОБЪЕКТОВ**

Монография

**Воронеж
Издательство «Научная книга»
2018**

УДК 004.67
ББК 32.81
Л 74

Рецензенты:

- Соустова И.А.** д-р физ.-мат. наук, профессор (Институт прикладной физики РАН);
- Чиркова М.М.** д-р физ.-мат. наук, профессор (Волжский государственный университет водного транспорта, Федеральное агентство морского и речного транспорта)

Л 74 Ломакина, Л.С. Информационные технологии диагностирования многомерных объектов: Монография / Л.С. Ломакина, Д.В. Ломакин, С.А. Зеленцов. – Воронеж: Издательство «Научная книга», 2018. – 180 с.

ISBN 978-5-98222-953-3

В книге рассматриваются основные подходы диагностирования многомерных объектов и базовые вопросы функционирования систем диагностирования и анализа сложных объектов с учетом их структурных особенностей. Большое внимание уделено описанию универсальных методов анализа многомерных объектов, которые могут применяться для решения прикладных задач в различных областях. Подробно описан обширный класс алгоритмов классификации, в основе которых лежит принцип индуктивного вывода логических закономерностей. Приводятся примеры применения нейросетевых технологий и диагностических экспертных систем.

Книга рассчитана на научных и инженерно-технических работников в области диагностирования сложных систем, а также полезна для студентов, магистрантов и аспирантов направления «Информатика и вычислительная техника».

Рис. 62. Табл. 23. Библиогр.: 103 назв.

УДК 004.67
ББК 32.81
Л 74

ISBN 978-5-98222-953-3

**© Ломакина Л.С., Ломакин Д.В.,
Зеленцов С.А., 2018**

СОДЕРЖАНИЕ

Введение.....	5
1. Методологические аспекты диагностирования многомерных объектов.....	7
1.1. Описание и формализация проблемы	7
1.1.1. Различие форм представления проблем и его влияние на компьютеризацию	7
1.1.2. Формы представления информации на различных этапах понимания	9
1.1.3. Различия в постановке проблемы из-за различия в подходах к объекту	13
1.1.4. Классификация типов проблем и методы их решения	16
1.1.5. Типы проблем и возможности компьютеризации.....	20
1.2. Одномерное статистическое моделирование	21
1.2.1. Одномерный статистический анализ.....	21
1.2.2. Модель пропорциональных интенсивностей Кокса	36
1.3. Многомерное статистическое моделирование	38
1.3.1. Виды многомерного анализа	40
1.3.2. Включение независимых переменных в модель	41
1.3.3. Взаимодействие между переменными	43
1.3.4. Анализ качества модели.....	43
1.3.5. Чувствительность, специфичность и точность.....	44
1.4. Принцип нечеткой логики	47
1.4.1. Косвенные наблюдения и корреляция.....	47
1.4.2. Скрытые структуры данных	49
1.4.3. Способы определения критериев различия	50
1.4.4. Способы описания объектов.....	50
2. Теоретические аспекты диагностирования многомерных объектов	52
2.1. Основные аспекты систем представления и обработки знаний	52
2.1.1. Формализация систем решения проблем	52
2.1.2. Основы построения систем обработки знаний.....	85
2.1.3. Методы создания систем обработки знаний.....	96
2.1.4. Особенности представления знаний в экспертных системах	106
2.2. Основные аспекты алгоритмов диагностирования на основе решающих деревьев.....	129
2.2.1. Основные понятия и определения	129
2.2.2. Этапы построения деревьев решений.....	131
2.2.3. Random forest.....	134
2.3. Основные аспекты алгоритмов диагностирования на основе нейросетевых технологий	136
2.3.1. Искусственный нейрон. Однослойный перцептрон.....	136

2.3.2.	Многослойный персептрон	138
2.3.3.	Нейронная сеть Ворда.....	142
2.3.4.	Нейронная сеть Кохонена.....	142
2.3.5.	Методы обучения искусственных нейронных сетей	145
3.	Практические аспекты диагностирования многомерных объектов....	149
3.1.	Анализ исходных данных. Модель внешней среды	149
3.2.	Пример построения диагностической экспертной системы.....	153
3.3.	Примеры диагностирования многомерных объектов на основе решающих деревьев.....	159
3.4.	Примеры диагностирования многомерных объектов на основе нейросетевых технологий	164
4.	Заключение	173
5.	Список литературы	174

ВВЕДЕНИЕ

В природе практически все объекты (физические, биологические, технические, социальные) описываются множеством физических величин или параметров (переменных), которые в совокупности образуют модель объекта как структурированный состав, т.е. систему.

Анализ свойств модели на основе системного подхода и экспериментальная проверка результатов анализа являются основной целью современных научных исследований. Объект (образец) обнаруживает себя во вне как совокупность наблюдаемых переменных, значения которых интерпретируются как свойства объекта, его состояния.

Таким образом, состояние объекта описывается значениями измеряемых переменных $\xi = (\xi_1, \xi_2, \dots, \xi_n)$, последовательность которых можно рассматривать как вектор в n -мерном евклидовом пространстве, пространстве переменных. В литературе переменные часто называют компонентами, поскольку в совокупности они образуют систему — модель объекта. Обнаружить структурные закономерности (свойства) системы т.е. различного рода зависимости между компонентами, по единственному состоянию системы не представляется возможным, поскольку зависимость отражает характер совместного изменения значений переменных. Для решения поставленной задачи необходимо располагать по крайней мере, множеством состояний, каждое из которых изображается точкой в n -мерном пространстве переменных.

Целью настоящей работы является рассмотрение с единых позиций общих вопросов, связанных с подходами к моделированию и диагностированию многомерных объектов при решении разнообразных прикладных задач. Также важной особенностью работы является обсуждение возможностей и перспектив современных информационных технологий при анализе многомерных объектов.

Работа написана с позиций современных информационных технологий и содержит изложение различных аспектов методологии, технологии и реализации решения основных задач диагностирования многомерных объектов.

Первая глава в доступной для понимания форме описывает логику и методологию современной аналитической статистики, применяемой в мировой практике. Основные механизмы, которые требуются для описания проблем и получения решения, различаются в зависимости от явлений и объектов реального мира, в котором возникают эти проблемы, а также от точки зрения и уровня понимания тех людей, которые берутся за решение этих проблем. Поскольку аспекты оказывают влияние на методы понимания проблем и использования информационных технологий, то их необходимо тщательно проанализировать.

Вторая глава дает общее представление о методах многомерного анализа, который разработан для оценки одновременного влияния более чем

одного фактора на результат. В отличие от одномерной статистики, которая дает оценку того, как каждая (одна) переменная связана с интересующим нас результатом, многомерная статистика дает информацию о степени влияния на исход каждой из (многих) переменных, а также об эффекте взаимодействия этих переменных между собой. В отечественной литературе многомерный анализ часто называют многофакторным анализом.

В третьей главе описаны результаты исследований авторов по интеллектуальным методам анализа многомерных данных и методам представления знаний в интеллектуальных системах. Обсуждены правила логического вывода в системах, основанных на знаниях, а также вопросы, связанные с формализацией проблем, способы представления информации на различных этапах решения проблемы и модели решения. Рассмотрены основные структуры систем обработки знаний.

Среди интеллектуальных подходов при диагностировании многомерных объектов по способу обучения можно выделить два основных направления: системы с дедуктивным обучением и системы с индуктивным обучением.

Дедуктивное обучение предполагает формализацию знаний экспертов и их перенос в компьютер в виде базы знаний. К дедуктивным системам относят экспертные системы (ЭС) различного назначения. В общем случае они состоят из базы знаний (БЗ), механизма вывода и интерфейса, обеспечивающего связь пользователя и эксперта с системой.

Системы индуктивного обучения основываются на выявлении закономерностей в эмпирических данных. К индуктивным подходам относят методологию нейронных сетей (НС) и генетических алгоритмов (ГА).

В главе «Практические аспекты диагностирования многомерных объектов» рассматриваются примеры практического применения представленных в предыдущих главах моделей и алгоритмов. Авторами было разработано специализированное программное обеспечение, нацеленное на диагностирование многомерных объектов и позволяющее автоматизировать данный процесс. В наглядной форме приведены проведенные с помощью данного программного обеспечения вычислительные эксперименты, и проведен анализ их результатов.

Монография будет полезна широкому кругу специалистов, занимающихся решением научно-технических проблем, специалистам в области инженерии знаний, а также всем, кто интересуется новейшими областями применения информационных технологий.

96. Swets JA. Measuring the accuracy of diagnostic systems. *Science* (1988) 240 1285-93.
97. Tarjan RE, Yannakakis MY. Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs. *SIAM Journal on Computing* (1984) 13 iii 566-79.
98. Teach RL, Shortliffe EH. An analysis of physician attitudes regarding computer-based clinical consultation systems. *Computers and Biomedical Research* (1981) 14 vi 542-58.
99. Titterton DM. A comparative study of kernel-based density estimates for categorical data. *Technometrics* (1980) 22 ii 259-68.
100. Tutz G. An alternative choice of smoothing for kernel-based density estimates in discrete discriminant analysis. *Biometrika* (1986) 73 ii 405-11.
101. Velleman P. F., Wilkinson, L. Nominal, ordinal, interval, and ratio typologies are misleading. // *The American Statistician*, 1993. Vol. 47, p. 65-72.
102. Weiss SM, Kulikowski CA, Amarel S, Safir A. A model-based method for computer-aided medical decision-making. *Artificial Intelligence* (1978) 11 145-72.
103. Zadeh LA. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics* (1973) 3 i 28-44.

Научное издание

Ломакина Любовь Сергеевна
Ломакин Дмитрий Викторович
Зеленцов Сергей Александрович

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ДИАГНОСТИРОВАНИЯ
МНОГОМЕРНЫХ ОБЪЕКТОВ**

Монография

Издание публикуется в авторской редакции

Дизайн обложки С.А. Кравец

Подписано в печать 26.01.2018. Формат 60x84 1/16
Усл. печ. л. 12,9. Заказ 000. Тираж 1000 экз.

ООО Издательство «Научная книга»
394077, Россия, г. Воронеж, ул. 60-й Армии, 25-120
<http://www.sbook.ru/>

Отпечатано с готового оригинал-макета
в ООО «Цифровая полиграфия»
394036, Россия, г. Воронеж, ул. Ф. Энгельса, 52
Тел. (473) 261-03-61