

Ерома Алексей Анатольевич

**СРЕДСТВА ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА СТРУКТУРЫ
СВЯЗЕЙ ПО ДАННЫМ МОНИТОРИНГА СЛОЖНЫХ
ТЕХНОЛОГИЧЕСКИХ ОБЪЕКТОВ**

Специальность 05.13.01 - "Системный анализ, управление и
обработка информации (промышленность)"

Автореферат

на соискание учёной степени

кандидата технических наук

Москва

2004

Диссертация выполнена на кафедре "Автоматизированные системы обработки информации и управления" (ИТ-7) Московской Государственной Академии Приборостроения и Информатики (МГАПИ).

Научный руководитель: кандидат технических наук,
профессор Р.А. Ашинянц

Официальные оппоненты: доктор технических наук, зав. каф.,
Анохин Алексей Никитич,
кандидат технических наук,
Леонов Александр Александрович.

Ведущая организация: Всероссийский Научно-Исследовательский Институт по
эксплуатации Атомных Электростанций (ВНИИАЭС)

Защита диссертации состоится «21» декабря 2004 года в 12 часов на заседании диссертационного Совета № Д212.119.02 Московской Государственной Академии Приборостроения и Информатики по адресу: 107076, Москва, ул. Стромынка, 20.

С диссертацией можно ознакомиться в библиотеке МГАПИ.

Автореферат разослан «17» декабря 2004

Учёный секретарь
диссертационного Совета

М.В. Ульянов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы.

Существует класс сложных технологических объектов, для которых в настоящее время отсутствуют надёжные методы анализа данных оперативного мониторинга. Этот класс объектов характеризуется:

1. Невозможностью организации активного эксперимента с целью создания математической модели;
2. Большим числом доступных для измерения параметров (сотни);
3. «Дрейфом» во времени отдельных параметров;
4. Отсутствием обоснований для необходимости измерений значительного числа параметров;
5. Отсутствием точных данных о взаимосвязях между параметрами и их влиянии на качество наблюдаемых процессов.

Объектами данного класса являются и атомные электростанции (АЭС), для которых оперативный мониторинг радиационной и технологической обстановки является чрезвычайно важной задачей при эксплуатации АЭС и неотъемлемым условием обеспечения безопасной и устойчивой работы.

Задаче анализа потока данных оперативного мониторинга сложных технологических систем, до настоящего времени не уделено должного внимания. В то время как во многих других сферах деятельности, таких как, телекоммуникации и т.п., массивы информации, накопленные в базах данных, активно используются для выявления скрытых внутренних закономерностей с последующим практическим применением извлечённых знаний. В качестве обзорной работы в направлении интеллектуального анализа данных можно назвать Holsheimer M, Siebes (1994). Примером успешного применения методов интеллектуального анализа к проблеме мониторинга в телекоммуникационных сетях служит работа Lane T., Bradley C.E. (1997).

В связи с этим актуальной является задача разработки методов и алгоритмов для системного анализа вектора данных оперативного мониторинга сложных технологических объектов с применением современных методов интеллектуального анализа данных, в целях усовершенствования характеристик существующих систем мониторинга. Эффективный анализ данных мониторинга возможен при создании средства интеллектуального анализа данных с использованием методов индукции статистических моделей, позволяющих выявить внутреннюю структуру вектора параметров оперативного мониторинга. Создание подобных систем «интеллектуального» мониторинга и средств автоматизированного анализа данных



мониторинга позволит повысить эффективность системы мониторинга и будет способствовать повышению уровня безопасности при эксплуатации АЭС.

Разработка методов и алгоритмов для извлечения знаний из массива данных оперативного мониторинга сложных технологических объектов с использованием современных методов искусственного интеллекта и интеллектуального анализа данных, с последующим приложением разработанных алгоритмов и программ применительно к объектам атомной энергетики делают тему диссертации актуальной как в прикладном отношении, так и в научно-техническом плане.

Цели и задачи исследования.

Целями настоящего исследования являются:

1. Создание методов интеллектуального анализа данных сложных технологических объектов;
2. Создание средств для системного анализа и обработки данных мониторинга;
3. Использование программных средств для обучения специалистов.

Для достижения этих целей нами поставлены следующие задачи:

1. Провести обзор методов интеллектуального анализа данных;
2. Разработать модифицированный алгоритм индукции, использующий знания экспертов предметной области в качестве эвристики при построении модели;
3. Разработать модифицированный алгоритм индукции нелинейных моделей, оценить сложность алгоритма;
4. Разработать методику мониторинга на основе индукции статистических моделей по данным оперативного мониторинга;
5. Разработать программный комплекс для оценки эффективности и точности алгоритмов методики на вычислительных экспериментах и данных наблюдения.

Методы исследований, примененные в диссертационной работе, опираются на теоретические и методологические основы теории статистики, теории машинного обучения и теории баз данных и методы интеллектуального анализа данных.

Научная новизна исследования.

1. Модифицированный алгоритм индукции, использующий знания экспертов предметной области;
2. Модифицированный алгоритм индукции нелинейных моделей;
3. Методика интеллектуального оперативного мониторинга состояния сложных технологических объектов на базе индукции модели объекта по данным наблюдения;

Практическая значимость результатов исследования и их внедрение в практику.

На базе разработанных методик и алгоритмов интеллектуального оперативного мониторинга создан программный комплекс «Пионер» для тренинга специалистов по поддержке принятия решений в случае радиационных аварий на АЭС. Разработанная система представляет собой эффективный инструмент для проведения тренинга служб радиационной разведки. Программный комплекс внедрён в практику работы Технического кризисного центра (ТКЦ) ИБРАЭ РАН, о чём имеется соответствующий акт.

Предложенная в настоящей работе методика интеллектуального мониторинга позволяет повысить эффективность систем мониторинга АЭС и способствует решению важных прикладных задач в области обеспечения экологической безопасности и безопасности персонала АЭС.

Апробация работы.

Основные результаты диссертационной работы докладывались на конференциях:

1. V Международной научно-практической конференции "Фундаментальные и прикладные проблемы приборостроения, информатики, экономики и права" — Москва, 2002;
2. Всероссийской научно-технической конференции "Новые информационные технологии" — Москва, 2003;
3. IV конференции стипендиатов ИБРАЭ РАН, Москва, 2003;
4. Конференции Школы молодых ученых ИБРАЭ РАН, Москва, 2004.

Публикации.

По теме диссертации опубликовано 6 научных работ, в том числе, 1 препринт (№ИБРАЭ-2004-05). Исследования по диссертации отражены в научно-исследовательских отчётах Института проблем безопасного развития атомной энергетики РАН (ИБРАЭ РАН).

Объём и структура диссертации.

Диссертационная работа изложена на 167 страницах машинописного текста, состоит из введения, 4 глав, отражающих результаты исследований автора, заключения, библиографического указателя, включающего 93 источника, иллюстрирована 35 рисунками и 19 таблицами.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы, поставлены цели и задачи исследований.

В первой главе приведён обзор и сравнение существующих методов в области интеллектуального анализа данных применительно к задаче анализа вектора данных мониторинга. Отмечены трудности, возникающие, при использовании существующих методов интеллектуального анализа в связи со спецификой данных оперативного

мониторинга состояния сложных технологических объектов. Вид знаний, которые мы ожидаем получить в результате анализа, определяет как возможные формы представления знаний, так и методы интеллектуального анализа, которые могут быть использованы для извлечения знаний соответствующего вида из данных. С прикладной точки зрения, большой интерес представляют знания о закономерностях внутренней структуры объекта мониторинга, извлекаемые методами индукции. На основе анализа работ, посвященных системам представления знаний и методам извлечения знаний из данных, сделан вывод, что методы индукции статистических моделей могут быть успешно применены как для задачи структуризации вектора параметров оперативного мониторинга, так для задачи анализа данных оперативного мониторинга с целью выявления структурных отклонений. В конце первой главы сформулирована постановка задачи исследования.

Вторая глава посвящена разработке оригинальной методики интеллектуального мониторинга состояния сложных технологических объектов на основании выделения статистических моделей. От анализа значений отдельных показателей предлагается перейти к анализу изменений в системе взаимосвязей объекта мониторинга. Отличия между типичной и предлагаемой системами мониторинга показано на рис. 1.

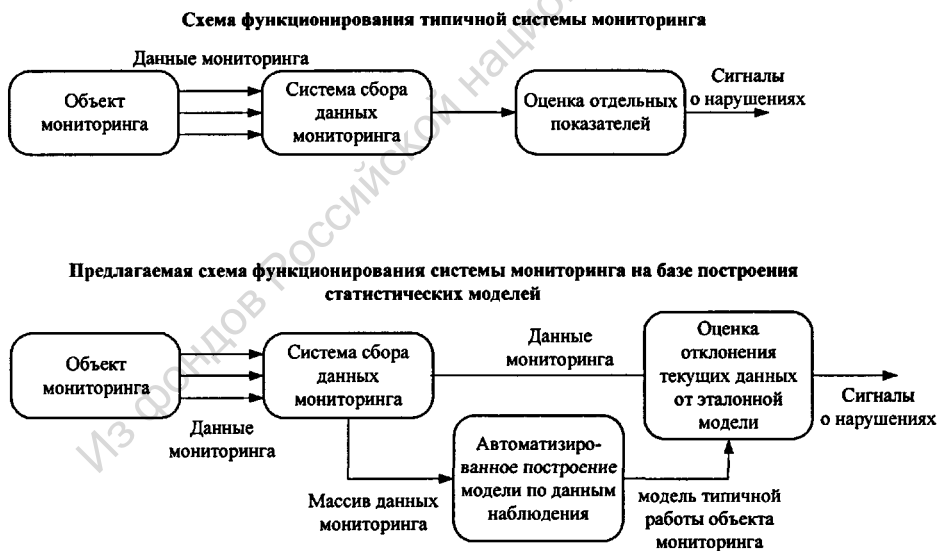


рис. 1. Отличия типичной и предлагаемой систем мониторинга.

Предложена обобщённая схема процесса интеллектуального мониторинга отклонений на основе индукции статистических моделей, а также показана роль методов индукции в процессе мониторинга - см. рис. 2.

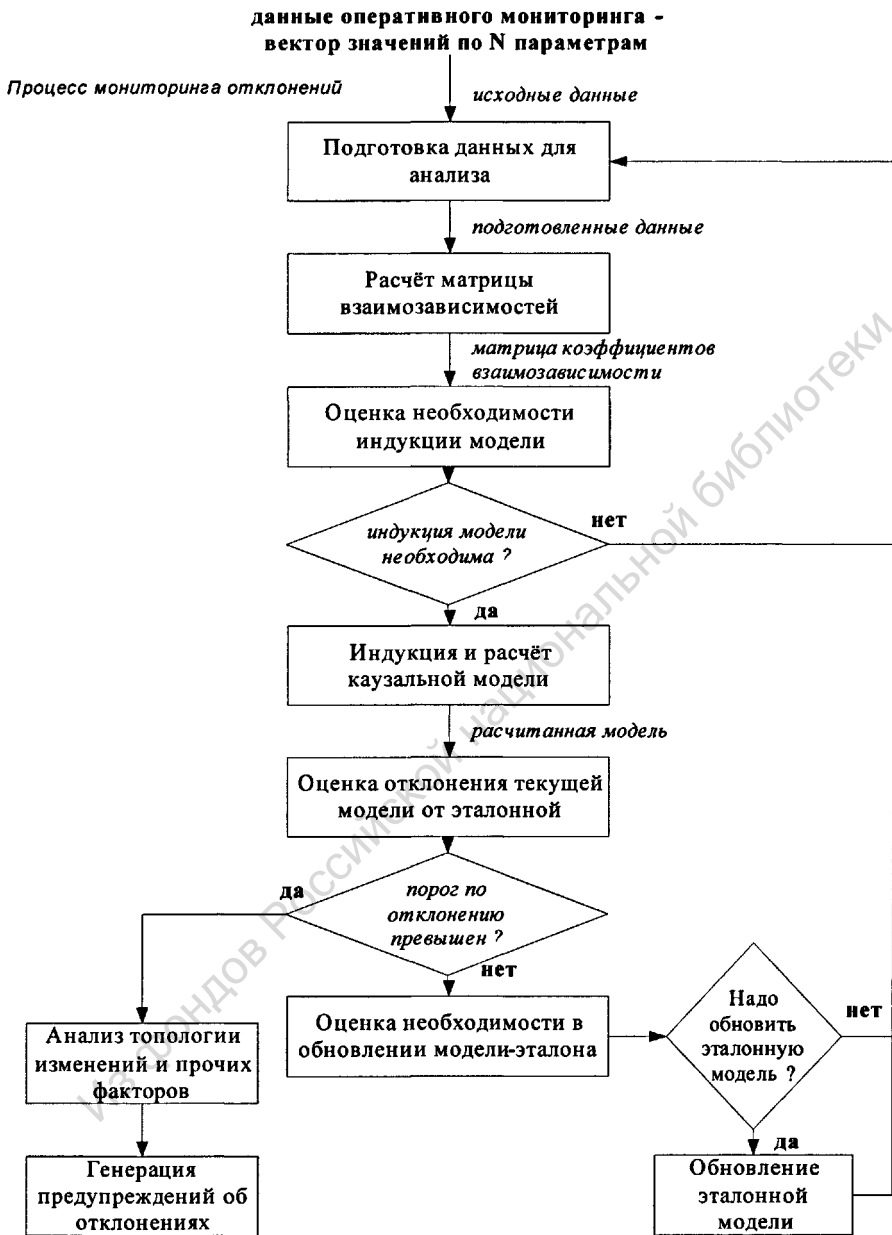


рис. 2 Схема процесса интеллектуального мониторинга отклонений на основе индукции статистических моделей

В настоящей методике в качестве инструмента для построения статистических шаблонов, вместо традиционных методов интеллектуального анализа данных, используются методы индукции статистических моделей. Методы индукции позволяют извлечь из данных наблюдения не только знания об определённых закономерностях функционирования объекта мониторинга, но и знания о структуре объекта. Использование в целях мониторинга извлечения знаний подобного вида, с учётом специфики исходных данных мониторинга, принципиально отличают методику интеллектуального мониторинга, предлагаемую в настоящей работе, от аналогичных методик в других областях. Место «шаблона» занимает статистически выведенная структурная модель объекта мониторинга.

Интеллектуальный мониторинг определим как итеративный процесс извлечения знаний о внутренней структуре объекта по данным оперативного мониторинга, поступающим в реальном времени, с последующей оценкой отклонений текущей структуры данных относительно структуры соответствующей нормальному функционированию объекта. Под внутренней структурой данных понимается совокупность статистических зависимостей, определённых на исходном множестве параметров.

Методика интеллектуального мониторинга имеет статистическую основу, что накладывает определённые ограничения на данные и предполагает наличие этапа *подготовки данных для анализа*.

Индукция каузальной модели представляет собой процесс эвристического поиска модели, в достаточной степени, согласующейся с исходными статистическими данными.

Индукция модели объекта проводится на основании статистической информации о взаимосвязях между параметрами, поэтому на следующем этапе осуществляется *расчёт матрицы взаимозависимостей*. В общем случае выбор критерия оценки зависимости двух параметров определяется в зависимости от специфики данных, в настоящей работе за основу взят коэффициент корреляции Пирсона, в качестве меры зависимости двух параметров.

Нет необходимости в выполнении индукции модели для каждого нового вектора значений параметров, поскольку, при большой длине окна мониторинга отдельные вектора значений не будут оказывать существенного влияния на корреляционную матрицу, и, как следствие, модели, построенные и на $(I + 1)$ -ОМ шаге, не будут отличаться.

Предпочтительным является проведение повторной индукции тогда, когда проверка ранее выведенной модели по текущим данным, показывает значительные расхождения. Это означает, что когда наблюдаются существенные отличия текущей корреляционной матрицы от корреляционной матрицы, прогнозируемой моделью. Эта проверка выполняется на этапе *оценки необходимости повторной индукции модели*.

На этапе *индукции и расчёта структурной модели объекта* по данным мониторинга осуществляется построение модели объекта. В общем случае, модель, полученная при индукции, будет отличаться от фактической модели объекта. Отличия между индуцированной и аналитической моделью определяются статистической природой процесса индукции модели. Так, например, малозначащие функциональные связи и редко проявляющие себя связи с большой вероятностью вообще не будут представлены в модели. С другой стороны, одна и та же совокупность исходных данных может одинаково хорошо описываться не одной единственной моделью, а целым множеством различных моделей. Не располагая какими-либо темпоральными или экспертными данными о направленности взаимосвязей параметров, невозможно сделать выбор в пользу одной модели из этого множества. Структурная вычислительная модель здесь представляет собой ацикличную аддитивную модель и определяется следующим образом:

$$X_i = \sum_{\substack{y_j \neq i \\ X_j \rightarrow X_i}} X_j * a_{y_j} + \xi_i \quad (1)$$

где a_{y_j} — коэффициенты регрессии или прямого влияния параметра j на параметр i , ξ_i — необъяснённая составляющая параметра (ошибка). Для алгоритмов индукции принимается: исходная модель является *марковской моделью*, т.е. её граф не содержит направленных циклов, а ошибки некоррелированы между собой $M[\xi_i, \xi_j] = 0$.

Индукцию модели объекта по данным мониторинга можно проиллюстрировать на следующем примере. Имеется набор данных мониторинга по пяти параметрам. Корреляционная матрица на множестве этих параметров представлена в табл. 1. Индукция модели позволяет нам перейти от полностью связанного графа изображённого на рис. 3 а) к частично-ориентированному графу рис. 3 б). При наличии дополнительной информации об ориентации остальных связей можно получить полностью ориентированный граф и, применяя аппарат множественной регрессии, перейти к рассчитанной модели объекта мониторинга рис. 3 с).

* A	B	C	D	E
A 1.00				
B 0.00	1.00			
C 0.71	0.00	1.00		
D 0.50	0.50	0.35	1.00	
E 0.00	0.71	0.00	0.71	1.00

табл. 1. Корреляционная матрица по параметрам мониторинга для примера индукции модели объекта.

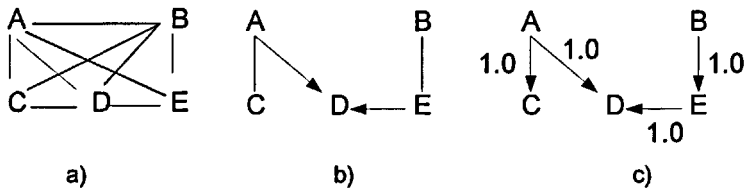


рис. 3. Индукция модели объекта, матрица по которому представлена в табл. 1. а) – исходный полностью связанный граф, б) – частично-ориентированный граф, полученный из исходного графа в результате индукции, в) – структурная модель объекта, построенная с привлечением дополнительных знаний для уточнения направленности отдельных связей

На следующем этапе производится оценка отклонения текущей модели от эталонной модели. Под эталонной моделью понимается некоторая, сгенерированная ранее, модель объекта, относительно которой оцениваются все текущие структурные изменения. Можно выделить следующие подходы оценки отклонений от эталонной модели:

- *Равномерная оценка.* Различные связи в модели полагаются равнозначными. При этом в качестве оценки может выступать сумма квадратов отклонений весов связей в текущей модели от эталонной модели или соответствующие максимальные значения;
- *Взвешенная оценка.* Каждой связи в модели соответствует некоторый вес, определяющий степень важности изменений этой связи по сравнению с остальными. При этом можно выделить два подхода к выработке соответствующих весов:
 - *Экспертная взвешенная оценка.* В рамках этого подхода эксперты предметной области количественно определяют степень важности каждой выявленной или потенциально возможной взаимосвязи.
 - *Взвешенная оценка по сети.* Каузальная модель представляет собой сеть связанных параметров, поэтому оценку важности связи можно определить по степени её влияния на систему в целом.

В главе проведено исследование вопросов определения и предварительной подготовки вектора исходных данных мониторинга. Обоснована необходимость применения методов анализа внутренней структуры данных для эффективного выявления структурных отклонений в функционировании объекта мониторинга, приведён пример использования методов индукции для выявления структурных отклонений. По результатам численных экспериментов получены оценки взаимосвязи времени реакции и точности системы мониторинга.

До настоящего момента предполагалось, что связи объекта мониторинга носят относительно стабильный характер. В действительности связи могут изменяться не только под воздействием внештатных причин, но и в штатном порядке - в соответствии с

различными технологическими и организационными процессами. В данном разделе проведена адаптация методики интеллектуального мониторинга для объектов с динамически изменяющимися связями. Процесс мониторинга с использованием нескольких эталонных моделей и их динамическим обновлением представлен на рис. 4.



рис. 4. Методика интеллектуального мониторинга с использованием множества эталонных моделей и динамическим обновлением эталонов.

Наиболее приемлемым решением в данном случае нам представляется идентификация возможных состояний объекта в процессе инициализации (обучения) системы мониторинга с последующим мониторингом отклонений относительно выявленных допустимых состояний.

Под *состоянием объекта* понимается модель взаимосвязей объекта с рассчитанными коэффициентами. Для идентификации возможных состояний объекта мониторинга из множества генерируемых моделей в процессе инициализации системы мониторинга модели предпочтительным является использование методов кластеризации для выделения кластеров моделей и их наиболее характерных представителей. Мерой расстояния между соответствующими моделями объекта в конкретный момент времени, является выбранный ранее критерий оценки отклонений текущей модели от эталонной модели. После идентификации возможных состояний объекта, в процессе мониторинга отклонений. Наиболее характерные представители классов моделей, выделенные на этапе кластеризации моделей, играют роль эталонных моделей.

Проведённый на вычислительных экспериментах, анализ факторов, влияющих на точность и скорость реакции на структурные изменения системы мониторинга, показал:

1. При наличии достаточного объема данных, обеспечивается надёжная работа (т.е. выдерживаются приемлемые частоты ложных срабатываний и пропусков структурных отклонений) системы мониторинга.
2. Для конкретного объекта мониторинга и объёма выборки («окна» индукции), можно выбрать оптимальное значение порогового коэфф., которое обеспечивает минимальные частоты ложных срабатываний и пропусков событий.
3. С увеличением объёма выборки (длины окна индукции), скорость реакции системы мониторинга на отклонения снижается.

Пример зависимости частоты ложных срабатываний и частоты игнорирования событий от порогового коэффициента для модели из 5 переменных и объёма выборки (длины окна индукции) $W = 500$ приведён на рис. 5. Зависимость времени обнаружения отклонения от объёма выборки приведена на рис. 6.

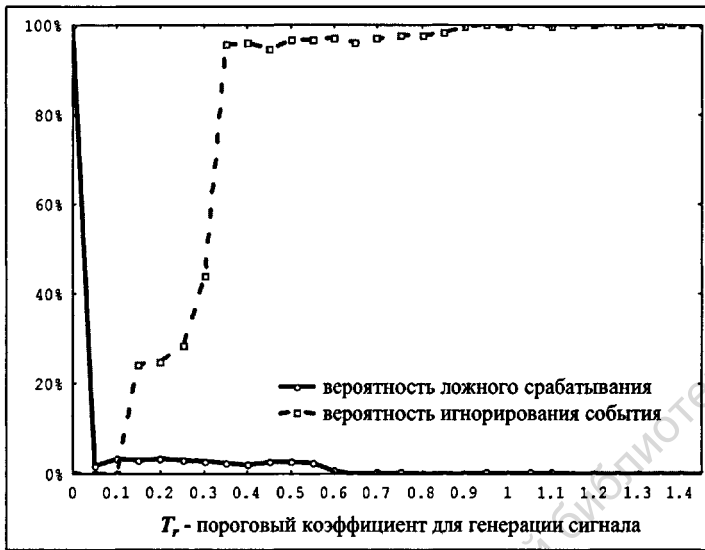


рис. 5. Зависимость частоты ложных срабатываний и частоты пропусков изменений в зависимости от порога значимости T_r , $WND = 500$, для модели из 5 параметров.

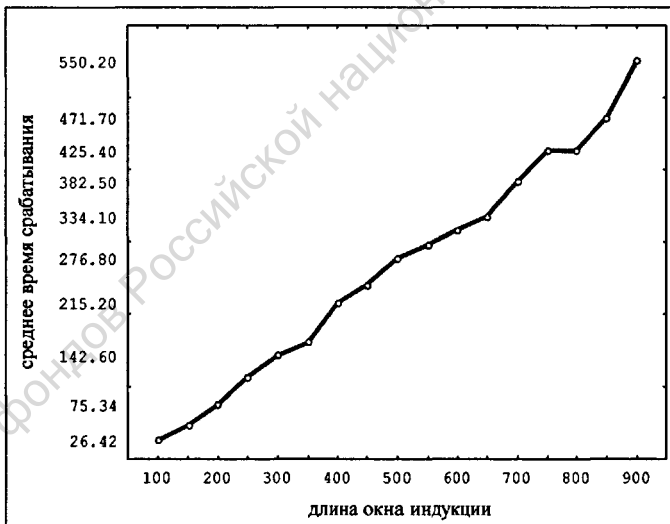


рис. 6. Зависимость времени обнаружения изменения структуры объекта от объёма выборки (длины окна индукции) W при пороге $tr=0.1$ для модели из 5 параметров.

Это ставит перед нами дилемму: при малой длине окна индукции появляются проблемы с ложными срабатываниями и игнорированием событий, но при этом мы можем относительно

быстро реагировать на изменения. При большей длине окна мы выигрываем в точности мониторинга, но теряем в возможности быстрого прогнозирования отклонений.

Конкретные значения параметра длины окна индукции и порога значимости определяются спецификой объекта наблюдения и должны выявляться на этапе предварительной настройки системы мониторинга.

Третья глава посвящена вопросам построения модели объекта по эмпирическим данным оперативного мониторинга. Рассмотрен вопрос о привлечении знаний экспертов предметной области для решения задач идентификации исходной модели в процессе индукции. Включение знаний экспертов предметной области в процесс индукции моделей позволяет решить проблемы статистической эквивалентности и идентифицируемости модели. Для этого предложен следующий формализм представления знаний экспертов:

Предикат	Матем. формулировка в терминах связей графа G	Смысловая нагрузка
$d(x, y)$	$\{x \rightarrow y, x \leftarrow y, x - y\} \cap G \neq \emptyset$	Существует ли прямая связь $x - y$.
$n(x, y)$	$\exists Z \{ \{x \rightarrow Z, Z \rightarrow y\} \in G, Z \in S_{xy}$	Существует направленная опосредованная или прямая связь направленная $x \rightarrow \dots \rightarrow y$.
$p(x, y)$	$\{x \rightarrow y\} \in G$ или $d(x, y) \cap n(x, y)$	Существует ли прямая связь между параметрами x, y , направленная $x \rightarrow y$.
$b(x)$	$\forall x \neq y, b(x) \rightarrow \neg p(y, x)$	Параметр x не зависит от других и является экзогенным (внеш.) по отношению к системе.
$e(x)$	$\forall x \neq y, e(x) \rightarrow \neg p(x, y)$	Параметр x является выходным параметром и не влияет напрямую на др. параметры системы.

Знания эксперта предметной области, выраженные с использованием формализма, представляют собой базовые знания о структуре объекта. Принимается *допущение* о том, что любые изменения в структуре объекта согласуются со знаниями эксперта. Т.е. знания экспертов полагаются абсолютными и верными при любых изменениях в структуре объекта.

Разработанный модифицированный алгоритм индукции с использованием дополнительных знаний выглядит следующим образом:

1. Для каждой пары переменных (a, b) , производится поиск множества наборов переменных S_{ab} таких, что $\forall C \in S_{ab}$ имеет место $I(A, B | C)$. Если $S_{ab} = \emptyset$ и

выполняются $\neg I(A, B | \emptyset)$ и $\neg d(A, B) \notin F$, то поместить неориентированную дугу \overline{ab} между этими переменными¹.

2. Для каждой пары несвязанных напрямую переменных (a, b) с общей соседней вершиной c , если $c \notin C, \forall C \in S_{ab}$ и ориентация стрелок в сторону c не противоречит базе фактов (т.е. для базы фактов выполняется $\{\neg p(a, c); \neg p(b, c); \neg n(a, c); \neg n(b, c); n(c, a); n(c, b); \neg e(a); \neg e(b)\} \cap F = \emptyset$), то направить стрелки в сторону c , т.е. $a \rightarrow c \leftarrow b$.

3. Для каждой ненаправленной связи \overline{ab} в графе, если выполняется $\{\neg p(b, a); \neg n(b, a); n(a, b); b(a); e(b)\} \cap F \neq \emptyset$, то направить связь $a \rightarrow b$.

4. Далее рекурсивно производится ориентация дуг графа в соответствии со следующими двумя правилами:

а если для ненаправленной связи \overline{ab} в графе существует направленный путь из a в b , то направить $a \rightarrow b$;

б. если для пары несвязанных напрямую переменных (a, b) существует вершина c , такая что $a \rightarrow c, c \rightarrow b$, то если $\{\neg n(c, b); n(b, c); b(b); e(c)\} \cap F = \emptyset$, направить связь $c \rightarrow b$, иначе удалить связь $c \rightarrow b$.

В четвёртой главе проведена адаптация существующих методов индукции нелинейных моделей.

Линейная модель в отдельных случаях достаточно грубо описывает систему взаимосвязей объекта и приближения, даваемого методикой мониторинга с использованием метода индукции линейных аддитивных моделей, в её исходном виде, будет явно не достаточно для эффективной работы системы мониторинга. Это требует развития существующих алгоритмов индукции для получения адекватных структурных моделей объекта по данным мониторинга, и для эффективной работы системы мониторинга, в целом, в условиях нелинейности взаимосвязей и наличия временных задержек.

Под моделью с нелинейными взаимосвязями будем понимать ацикличную модель с некоррелированными аддитивными ошибками, описываемую уравнениями вида:

¹ При этом знания эксперта считаются более приоритетными, чем статистический вывод о независимости, из-за возможного проявления ошибок индукции, например, в моделях с избыточными связями

$$X_i = \sum_{\substack{\forall j, i \neq j \\ X_j \rightarrow X_i}} \varphi_y(X_j) * v_j + \xi_i \quad (2)$$

где v_j — коэффициенты регрессии или прямого влияния параметра j на параметр i , $\varphi_y(X_j)$ — функция передачи влияния j на i , ξ_i — случайная сост. параметра (ошибка).

$$\varphi_y(X_j) = \sum_{k=1}^K p_k(X_j) * \alpha_{yk} \quad (3)$$

где α_{yk} — коэффициент при k -ом члене разложения передаточной функции φ_y при передаче возмущения от параметра X_j к параметру X_i , $p_k(X_j)$ — k -ый член разложения передаточной функции φ_y норм. по условию $M[p_k(X_j)] = 0, D[p_k(X_j)] = 1$, а ξ_y — ошибка аппроксимации параметра X_i параметром X_j .

Для разработки модифицированного алгоритма индукции, за основу взят не исходный IC-алгоритм индукции (Pearl J., Verma TS., “A Statistical Semantics for Causation”, Preprint Cognitive Systems Laboratory, Computer Science Department, University of California, 1991), основанный на принципе условной независимости, а предложенный ранее в настоящей работе модифицированный алгоритм индукции с использованием знаний экспертов предметной области. Это позволяет применять в процессе индукции дополнительные знания в виде различных эвристик, основанных на обнаружении асимметрий в исходном статистическом распределении. Для получения модифицированного алгоритма индукции необходимо предложить:

- критерий условной независимости двух параметров;
- эвристику для ускорения процесса индукции модели (если это возможно);

Для нелинейного сл. определим предикат условной независимости $I^M(X, Y | S_{XY})$:

$$I^M(X, Y | S_{XY}) = \begin{cases} TRUE, \rho_{XY|S}^M < k_{DM} \\ FALSE, \rho_{XY|S}^M \geq k_{DM} \end{cases} \quad (4)$$

где $\rho_{XY|S}^M$ — коэффициент условной корреляции между переменными X, Y при исключении влияния переменных S , в условиях нелинейности взаимосвязи.

$$\rho_{XY|Z}^M = R[X_{|z}, \gamma_{yz}(Y_z)] \quad (5)$$

где $X_{|z}, Y_{|z}$ — переменные X, Y при исключении влияния переменной Z , а γ_{yz} — функция влияния переменной Y на переменную X , при исключении влияния Z .

$$X_{|z} = (X - \varphi_{xz}(Z)) \quad (6)$$

$$Y_{|z} = (Y - \varphi_{yz}(Z)) \quad (7)$$

Используется эвристика: если наблюдается $\forall x, y \left| R_{yx}^M \right| \gg \left| R_{xy}^M \right|, \left(R_{yx}^M \right)^2 \geq \xi_R \rightarrow n(x, y)$ и влияние технологического фактора при этом исключено, то взаимосвязь между переменными x, y направлена $x \rightarrow y$. Условие $\left(R_{yx}^M \right)^2 \geq \xi_R$ отсекает случайные вариации коэффициенты корреляции, возникающие вследствие ограниченности выборки.

Разработанный модифицированный метод индукции нелинейных моделей, который позволяет точнее описывать нелинейные структурные взаимосвязи объекта мониторинга, и приводит к более корректным моделям, что и расширяет сферу применения предложенной методики мониторинга.

В настоящем исследовании решена задача распространения методики мониторинга на случай нелинейности взаимосвязей, случай наличия временных задержек и для смешанного темпорально-нелинейного случая. Разработан модифицированный метод индукции моделей с временными задержками, что позволяет учитывать специфику объекта мониторинга и получать корректные модели с временными задержками. Приведены оценки временной сложности модифицированных алгоритмов индукции. По результатам численного моделирования сделан вывод относительно практической применимости модифицированных алгоритмов индукции. Предложен метод индукции для обобщенного темпорально-нелинейного случая, и приведены оценки его временной сложности и практической применимости.

ЗАКЛЮЧЕНИЕ

Основные результаты исследования проведенного в настоящей диссертации состоят в следующем:

Разработанный модифицированный алгоритм индукции использует знания экспертов предметной области. Знания выражаются в форме предикатов первого порядка. Это позволяет повысить скорость индукции модели по исходным данным и уменьшить неопределенность получаемой модели.

Разработанный модифицированный алгоритм индукции нелинейных моделей делает возможным индукцию моделей при наличии нелинейной взаимосвязи между параметрами. Предложенная для этого алгоритма эвристика, основанная на наличии асимметрии вследствие нелинейности, способствует ускорению процесса индукции, уменьшению неопределенности и уточнению получаемой модели. Проведенная оценка временной сложности нового алгоритма индукции показала возможность индукции небольших моделей (десятки параметров) на современных компьютерах в реальном времени;

Предложенная новая комплексная методика интеллектуального оперативного мониторинга состояния сложных технологических объектов на базе индукции

статистической модели объекта позволяет, в отличие от существующих методик, выявлять структурные отклонения в технологических цепях по данным наблюдения;

Проведённый на вычислительных экспериментах, анализ факторов, влияющих на точность и время реакции системы мониторинга, построенной на базе статистической индукции внутренней модели объекта, показал возможность надёжной работы (т.е. выдерживаются приемлемые частоты ложных срабатываний и пропусков структурных отклонений) подобной системы мониторинга, при условии достаточности данных наблюдения;

Проведённый на вычислительных экспериментах, анализ точности и скорости алгоритмов индукции моделей в зависимости от количества рассматриваемых параметров модели и других характеристик исходных моделей, показал, что для групп параметров небольшой размерности (десятки параметров) индукция модели и, соответственно, процесс интеллектуального мониторинга могут проводиться в режиме реального времени на современных компьютерах;

На базе разработанных методик и алгоритмов интеллектуального оперативного мониторинга создан программный комплекс «Пионер» для тренинга специалистов по поддержке принятия решений в случае радиационных аварий на АЭС. Разработанная система представляет собой эффективный инструмент для проведения тренинга служб радиационной разведки. Программный комплекс практику работы Технического кризисного центра (ТКЦ) ИБРАЭ РАН, о чём имеется соответствующий акт.

СПИСОК ОСНОВНЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Ашинянц Р.А., Ерома А.А., Формальный концептуальный анализ в задачах извлечения знаний // Сборник трудов V Международной научно-практической конференции "Фундаментальные и прикладные проблемы приборостроения, информатики, экономики и права", Кн. "Информатика", М.: МГАПИ, 2002.С.6-10.
2. Ерома А.А. Индукция причинно-следственных связей в интеллектуальном анализе данных // Сборник трудов VI Всероссийской научно-технической конференции "Новые информационные технологии", М: МГАПИ, 2003.Т.2.С. 112-115.
3. Ерома А.А. "Автоматизированный анализ данных оперативного мониторинга" // Сборник трудов VI Всероссийской научно-технической конференции "Новые информационные технологии", М.: МГАПИ, 2003.Т.2.С. 115-117.
4. Ерома А.А., Никишин Д.А., Разработка автоматизированной информационной системы «Экологический мониторинг» // Сборник трудов IV конференции стипендиатов ИБРАЭ РАН (24-25 апреля 2003), М.: ИБРАЭ, 2003.С. 17-19.
5. Ашинянц Р.А., Ерома А.А. Методика интеллектуального оперативного мониторинга состояния сложных технологических объектов // Известия РАН. Энергетика. 2004.№5.С.117-122.
6. Ерома А.А. Развитие методов интеллектуального анализа данных мониторинга сложных технологических объектов. Препринт №ИБРАЭ-2004-05. М.:ИБРАЭ,2004.35с.

№24181

Из фондов Российской национальной библиотеки