

УДК 004.02; 004.04

ИССЛЕДОВАНИЕ МЕТОДОВ МНОГОМЕРНОГО АНАЛИЗА ДАННЫХ ДЛЯ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМ ПРОЦЕССОМ КОКСОВАНИЯ

Власов А.А., Мартыненко Т.В.

Донецкий национальный технический университет
кафедра автоматизированных систем управления, г. Донецк.

E-mail: cvetliy_e@mail.ru

Аннотация

Власов А.А., Мартыненко Т.В. Исследование методов многомерного анализа данных для системы поддержки принятия решений управления технологическим процессом коксования. Разработана математическая модель, описывающая зависимость температуры коксования от параметров технологического процесса коксования. Исследованы и протестированы реализации актуальных методов многомерного анализа данных. Была проведена сравнительная характеристика реализаций выбранных методов анализа многомерных данных. Был выбран наиболее подходящий метод для решения поставленной задачи снижения размерности признакового пространства.

Общая постановка проблемы.

Процесс коксования угля является одним из наиболее древних технологических процессов. Технология коксования, эксплуатируемое технологическое оборудование, контрольно-измерительные приборы и автоматика введены в действие, в основном, в начале второй половины прошлого столетия и не претерпели к настоящему времени существенных изменений. В тоже время технологический процесс содержит значительные энергетические резервы, использование которых для Украины чрезвычайно важно, т.к. уголь остается фактически единственно доступным энергетическим ресурсом. Одна из возможностей снижения энергоемкости процесса коксования – это внедрение средств контроля и управления на основе микропроцессорной техники и современных методов управления.

Тенденции развития современных систем управления химико-технологическими процессами показывают, что они должны быть адаптивными, интеллектуальными. Адаптивность указанных систем обеспечивается возможностью настройки их на: различные типы сырья, вид выпускаемой продукции, аппаратурно-технологическое оформление. Опыт управления сложными, инерционными, потенциально опасными объектами, к классу которых относится процесс коксования, показывает, что при управлении и подготовке персонала необходимо учитывать не только требования к протеканию процесса в эксплуатационном режиме, но и передовой опыт высококвалифицированных операторов и знания химиков-технологов, ученых, являющихся экспертами в данной предметной области. Таким образом, разработка для целей управления процессом коксования системы поддержки принятия решений, включающей интеллектуальные компоненты, позволяющие получить семантическое решение неформализованной задачи, цифровую математическую модель, а также подсистему прогнозирования температуры коксования, является актуальной.

Разработка математической модели.

Формально, математическую модель, описывающую зависимость температуры коксования от параметров коксования, можно выразить через формулу [1]

$$T = f(X_m, U_m, A_m) \quad (1)$$

Вектор входных параметров

$$X_m = \{X_1, X_2\} \quad (2)$$

где X_1 и X_2 – характеристики шихты и отопительного газа соответственно.

$$X_1 = \{T_{ш}^H, G_{ш}, C_{ш}^{cm}, \alpha_{ш}^{cm}, \gamma_{ш}\} \quad (3)$$

$$X_2 = \{C_{оз}, \alpha_{оз}, T_{оз}^H, T_{оз}^K\} \quad (4)$$

Вектор параметров и коэффициентов модели,

$$A_m = \{A_1, A_2\} \quad (5)$$

где A_1 и A_2 – характеристики печи и шихты соответственно.

$$A_1 = \{K, F_{ст}, L, \lambda_{ст}, \delta_{ст}\} \quad (6)$$

$$A_2 = \{\alpha_{ш}\} \quad (7)$$

Вектором управляющего воздействия данной модели U_m является расход отопительного газа $G_{оз}$.

$$U_m = \{G_{оз}\} \quad (8)$$

где λ - теплопроводность;

K - коэффициент теплопередачи греющей стенки;

L - ширина камеры;

$F_{ст}$ - площадь греющей стенки;

$\alpha_{оз}$ - коэффициент температуропроводности отопительного газа;

$\lambda_{ст}$ - теплопроводность стенки;

$\delta_{ст}$ - толщина стенки;

$\alpha_{ш}$ - коэффициент температуропроводности шихты;

$C_{ш}$ - теплоемкость отопительного газа;

$G_{ш}$ - расход шихты;

$T_{ш}^{ст}$ - температура шихты у стенки;

$T_{ш}^H$ - начальная температура шихты;

$T_{оз}^H$ - начальная температура отопительного газа;

$T_{оз}^K$ - конечная температура отопительного газа;

$\alpha_{ш}^{cm}$ - коэффициент температуропроводности шихты у стенки;

$C_{ш}^{cm}$ - теплоемкость шихты у стенки.

Задача снижения размерности признакового пространства.

Чтобы упростить математическую модель и задачу прогнозирования температуры коксования, необходимо снизить размерность признакового пространства. Такое снижение размерности возможно, поскольку в большинстве случаев признаки взаимосвязаны и, следовательно, данные избыточны с точки зрения информации, и эта избыточность полностью определяется корреляционной матрицей исходных переменных [2].

Снижение размерности сводится к преобразованию p -мерного исходного пространства данных X в другое пространство Y , в котором можно выбрать подмножество латентных (ненаблюдаемых) переменных меньшей размерности $r < p$ без существенной потери информации.

Исследование методов анализа многомерных данных.

В ходе исследования наиболее перспективными методами решения поставленной задачи оказались метод главных компонент, факторный анализ, группа методов экстремальной группировки параметров и класс методов многомерного шкалирования.

Метод главных компонент (МГК) является наиболее распространенным методом анализа многомерных данных и одним из методов факторного анализа.

МГК применяется для таких задач, как визуализация данных, компрессия изображений и видео, подавление шума на изображениях, индексация видео. Также метод широко используется в таких отраслях точных наук, как биоинформатика, хемометрика, геодезии, гидрологии, а также в ряде отраслей естественных наук – эконометрике, политологии, психодиагностике, социологии и др. [3].

Факторный анализ. Факторный анализ, как и анализ главных компонент, предназначен для преобразования p -мерной случайной векторной переменной в факторы, число которых существенно меньше p . При этом под фактором понимается гипотетическая переменная, в той или иной мере связанная с наблюдаемыми переменными [4].

Факторный анализ является более точным, и при этом более сложным в реализации методом, чем МГК. Впервые он возник в психометрике и в настоящее время широко используется не только в психологии, но и в нейрофизиологии, социологии, политологии, в экономике, статистике и других науках [5].

Методы экстремальной группировки параметров (МЭГП). Эти методы предназначены для структурного анализа набора параметров, описывающих исследуемое множество объектов [6]. Главная идея этих методов – выделение групп «похожих» параметров на основе выбранной меры близости между параметрами.

Реализация МЭГП представлена в [7]. Ее недостатками являются неоптимальное начальное разбиение параметров на группы, неспособность определения оптимального числа групп, а начальное вычисление некоторых параметров метода затруднительно. Авторы реализации пытались упростить вычисление этих параметров, но это снижает эффективность группировки.

Многомерное шкалирование. Основная идея этого класса методов состоит в представлении каждого объекта точкой геометрического пространства (обычно размерности 1, 2 или 3), координатами которой служат значения скрытых (латентных) факторов, в совокупности достаточно адекватно описывающих объект. При этом отношения между объектами заменяются отношениями между точками – их представителями [8].

В отличие от факторного анализа, методы многомерного шкалирования могут быть применимы, пока сохраняет смысл порядок следования рангов сходств. Также они могут быть основаны на прямом оценивании сходств между стимулами субъектов, тогда как факторный анализ требует, чтобы субъекты были оценены через их стимулы по некоторому списку атрибутов [9].

На основании проведенного исследования методов, а также реализаций их алгоритмов и результатов работы с реальными данными [10][11], была построена таблица сравнительной характеристики методов анализа многомерных данных (табл. 1).

Таблица 1. Сравнительная характеристика методов анализа многомерных данных

Методы анализа	Критерии оценки			
	Сложность реализации	Точность анализа	Распространенность	Спектр охвата задач
МГК	Средняя	Высокая	Очень высокая	Очень широкий
Факторный анализ	Высокая	Очень высокая	Высокая	Широкий
МЭГП	Высокая	Выше среднего	Высокая	Широкий

Таблица 1. Сравнительная характеристика методов анализа многомерных данных. Продолжение

Методы анализа	Критерии оценки			
	Сложность реализации	Точность анализа	Распространенность	Спектр охвата задач
Многомерное шкалирование	Средняя	Выше среднего	Высокая	Очень широкий

Заключение. В заключении хотелось бы отметить, что процесс коксования отличается своей сложностью и точностью при построении математической модели. Любые ошибки и большие погрешности могут привести к тому, что по окончании процесса будет получен кокс низкого качества, что поведет за собой убытки при продаже. Поэтому необходимо выбрать наиболее точный метод для решения задачи снижения размерности признакового пространства, чтобы минимизировать потерю информации, даже если придется пожертвовать ресурсоемкостью процесса и затратами на разработку реализации алгоритма.

В ходе исследования было решено использовать факторный анализ как наиболее точный метод анализа многомерных данных для дальнейшего решения основной задачи о разработке системы поддержки принятия решений управлением технологическим процессом коксования. К преимуществам данного метода можно отнести его широкую распространенность и способность охватывать широкий спектр задач. Наличие большого числа литературы с подробным описанием факторного анализа, его способов применения и примерами алгоритмов позволяет минимизировать затраты времени и сил на реализацию данного метода, что является несомненным плюсом.

Список литературы

1. Чистякова Т.Б., Бойкова О.Б., Бабина Е.В., Нестеров Д.М. Имитационная модель для управления и изучения температурного режима коксования. Санкт-Петербургский государственный технологический институт (Технический университет).
2. Большаков А.А., Каримов Р.Н. Методы сжатия информации. Саратов: Саратовский политехнический институт, 1991. 88с.
3. Метод главных компонент. [Электронный ресурс] / Интернет-ресурс. – Режим доступа: http://ru.wikipedia.org/wiki/Метод_главных_компонент
4. Факторный анализ. [Электронный ресурс] / Интернет-ресурс. – Режим доступа: http://ru.wikipedia.org/wiki/Факторный_анализ
5. Большаков А.А., Каримов Р.Н. Методы обработки многомерных данных и временных рядов: учебное пособие для вузов. М.: Горячая линия – Телеком, 2007. 522с.
6. Браверман Э.М., Мучник И.Б. Структурные методы обработки эмпирических данных. М.: Наука, 1983.
7. Лушельский В.Я. Группировка параметров на основе квадратичной матрицы связи. — Автоматика и телемеханика, 1970, № I, с. 133-143
8. Орлов А.И. Прикладная статистика. М.: Издательство «Экзамен», 2004.
9. Многомерное шкалирование. [Электронный ресурс] / Интернет-ресурс. – Режим доступа: <http://www.statsoft.ru/home/textbook/modules/stmulsc.html>
10. В. В. Мокеев, А. В. Мокеев. Многофакторный и прогнозирование продаж на основе метода главных компонент. Челябинск: Известия Челябинского научного центра, вып. 4 (42), 2008.
11. А. В. Мокеев. О точности и быстродействии метода синтеза главных компонент. Челябинск: Бизнес-информатика №3(13), 2010.