

УДК 004.65:621.921.34

А.А. Лебедева, Е.М. Чистяков, К.З. Гордашник, кандидаты технических наук

Институт сверхтвердых материалов им. В.Н. Бакуля НАН Украины, г. Киев

К ВОПРОСУ О ПРОГНОЗИРОВАНИИ СВОЙСТВ СВЕРХТВЕРДЫХ МАТЕРИАЛОВ НА ОСНОВЕ СИСТЕМНОГО АНАЛИЗА ИНФОРМАЦИИ

Рассмотрен подход к прогнозированию свойств сверхтвердых материалов на основе системного анализа информации. Данные о свойствах веществ рассматриваются как информационный ресурс с определенными требованиями к выбору информационных технологий. Показано, что прогнозирование свойств сверхтвердых материалов требует использования интеллектуального анализа данных и нейронно-сетевых технологий.

Ключевые слова: *системный анализ, система управления знаниями, предметная область, интеллектуальный анализ, сверхтвердый материал, физико-механические свойства.*

Разнообразие областей применения сверхтвердых материалов (СТМ) во многом определяется уникальным сочетанием физико-механических (особенно теплофизических) и химических свойств этих материалов. Благодаря столь ценным свойствам СТМ все шире применяют в различных отраслях промышленности и современной техники как конструкционные материалы.

Физико-механические свойства СТМ могут быть целенаправленно изменены технологией производства. Например, возможно получение синтетических алмазных порошков повышенной прочности, хрупких, повышенной вязкости и так далее. Причем, важно оптимальное сочетание определенных свойств, особенно в изменяющихся рабочих условиях.

На основе достоверной и достаточно полной информации о свойствах СТМ и режимах их спекания можно построить комплексную систему связи свойств и параметров технологических процессов производства.

В настоящее время в Институте сверхтвердых материалов им. В.Н. Бакуля НАН Украины накоплен обширный фактический материал (базы данных) об СТМ [1]. Однако само по себе наличие данных не решает проблему выбора оптимальных параметров технологии создания материалов этого класса. Кроме того, зачастую информация представляется в виде среднего значения определенных свойств без их статистической оценки, что не позволяет обоснованно оценить достоверность анализируемых показателей. Поступающая новая информация, как правило, позволяет решить конкретную задачу.

Таким образом, актуальной является задача прогнозирования свойств СТМ и выбора параметров процессов их производства и применения на основе базы знаний, включающей как материаловедческие базы данных, так и операции их обслуживания.

Процедура прогнозирования предусматривает использование системы хранения, статистической обработки и оценки свойств СТМ.

В процессе развития информационных технологий, а также систем сбора и хранения данных возникла проблема анализа больших массивов данных, которые практически невозможно обработать вручную. Следовательно, необходимо представить исходную информацию в более компактном виде [2], что позволит использовать ее для построения информационной модели материала.

Для решения поставленной задачи массив информации об СТМ разделим на категории:

- исходные данные – необработанные массивы данных, получаемые в результате измерения физико-механических свойств СТМ в конкретных условиях получения;
- информация – обработанные данные, которые представляют информационную ценность (например, статистически обработанные исходные данные);
- знания — содержат ноу-хау, отображают скрытые взаимосвязи объектов и не являются общедоступными (например, функциональные зависимости физико-механических свойств СТМ от конкретных параметров спекания).

Системный анализ информации об СТМ предполагает детальное изучение промышленного синтеза, областей применения, возможностей и методов управления физико-механическими свойствами СТМ, а также анализ информационных источников. Такой анализ целесообразно проводить в рамках системного анализа информации предметной области (ПрО) «СТМ», которая описывается с помощью онтологии. Онтология включает словарь терминов (тезаурус) ПрО и логические выражения, описывающие взаимосвязи понятий [3; 4]. Тезаурус содержит точные и однозначные определения терминов ПрО. Таким образом, онтология ПрО представляет собой тезаурус ее понятий, обеспечивающий возможность толкования терминов ПрО посредством интерпретации таких типов парадигматических отношений как, «часть—целое», «класс—подкласс» и некоторых видов ассоциативных связей. Каждое понятие имеет имя и может иметь атрибуты, каждый атрибут может иметь значение с учетом специфики ПрО. Информационные элементы онтологии и тезауруса соответствуют реальным объектам ПрО «СТМ» [1].

Результаты анализа ПрО «СТМ» показали, что описывающая ее информация характеризуется низким уровнем формализации и следующими особенностями:

- отсутствием корректной классификации для определения однотипных объектов;
- многозначностью терминов или использованием различных терминов для определения одних и тех же объектов;
- использованием различных комплексов параметров для описания однотипных объектов;
- использованием признаков, которые воспринимаются неоднозначно;
- различием методов испытаний для определения одних и тех же свойств и параметров;
- одновременным использованием количественной и качественной информации и др.

Все это приводит к неопределенности формирования правил в базе знаний, необходимости отслеживания и сохранения нескольких значений одного показателя (например, твердости). Существуют также особенности информацию об СТМ, относящиеся к логической структуре, требований к процессу обработки, форме представления и т.п.

Для ПрО «СТМ» в настоящий момент не существует априорных моделей или корреляций, и опытные данные являются единственно значимым источником информации, т. е. накоплен огромный фактический материал о структуре, свойствах и эксплуатационных характеристиках веществ в отсутствие работоспособных моделей. Именно в этом случае оправдано использование статистических методов, методов нечеткой логики и интеллектуального анализа, чтобы из «сырых» данных вычленить закономерности типа «структура – свойство», «параметры спекания – свойство», «добавка – свойство».

Информационной основой интеллектуального анализа данных является база знаний (БЗ) – система материаловедческих и корпоративных баз данных (БД) и объединяющих их программных процедур сбора, хранения, передачи и обработки информации. В систему БД

ИСМ входят 15 корпоративных и 12 материаловедческих БД, в том числе БД о комплексных характеристиках СТМ. Необходимым элементом БЗ является программный блок смыслового (семантического) анализа полнотекстовых данных. Метод интеллектуального анализа данных в рамках большого их комплекса для поиска моделей состава, структуры и свойств СТМ включает статистические и корреляционные методы. Эти модели можно применить для прогнозирования неизвестных характеристик как существующих, так и новых материалов.

Важным этапом интеллектуального анализа является построение информационной модели объекта исследования. Информационная модель СТМ представляет собой совокупность структурированных данных, достаточно полно описывающих наиболее существенные характеристики СТМ. Класс СТМ широкий и допускает систематизацию по различным признакам. Классификацию СТМ по группам осуществляют путем оценки и анализа свойств, проверки статистических гипотез, анализа значимости влияния на свойства материалов различных технологических факторов, что позволяет построить информационные модели материалов различных классов.

При формировании структуры информационной модели СТМ разделили на группы. В каждой группе находятся объекты, которые можно описать одним комплексом данных. Наиболее оправдано, на наш взгляд, разделить СТМ на следующие группы:

- порошки синтетических алмазов;
- поликристаллы на основе синтетических алмазов;
- порошки кубического нитрида бора;
- поликристаллы на основе кубического нитрида бора;
- композиты на основе природных и синтетических алмазов и кубического нитрида бора.

Наиболее полно исследовали группу «Поликристаллы на основе синтетических алмазов» (ПКА). Для примера опишем модель этой группы. Как информационные ресурсы объекты ПКА можно описать такими полями:

- исходный материал (марка порошка, зернистость, состав шихты и др.);
- добавки и наполнители;
- аппарат высокого давления;
- параметры спекания (давление, температура, время выдержки)
- характеристика изделия (форма, типоразмер, структура и др.);
- методы исследований;
- физико-механические свойства;
- применение.

Выбор входных переменных обусловлен онтологией и тезаурусом ПрО «ПКА» как части всей области «СТМ».

После описания всех полей необходимо решить сложную задачу установления зависимостей между входными параметрами спекания, физико-механическими характеристиками и областями применения.

В настоящее время, особенно с появлением нанотехнологий, возросла потребность в разработке теории и методов управления структурой материала.

Как известно, прогнозировать свойства материалов невозможно без детального изучения зависимости их физико-механических свойств от микроструктуры. Таким образом, возникла потребность в разработке и исследовании моделей сред с микроструктурой. Структура среды и, в частности, размер зерна - один из важнейших показателей качества материалов, непосредственно влияющих на их физико-механические характеристики. В настоящее время

инструментов теоретического прогнозирования физико-механических свойств в зависимости от структуры не существует [5]. В то же время с использованием интеллектуального анализа, в частности, нейронно-сетевых технологий, на основе информационного моделирования решение задач прогнозирования с учетом структурных подходов вполне реально.

Достаточность данных анализируют с помощью описания состава исходных данных для формирования классов рассматриваемых СТМ, алгоритмов анализа и наполнения данных, а также с учетом используемых алгоритмов расчета параметров. На основе этих знаний анализируют достаточность информации в БД и подбирают необходимые параметры для адекватного описания модели.

На сегодняшний день использование при подготовке справочных данных о свойствах СТМ методов интеллектуального анализа данных и технологий управления знаниями является скорее исключением, чем правилом. Как отмечалось, соотношение данных и знаний по их объему и значимости сильно изменяется в зависимости от ПрО.

Таким образом, на основании сказанного можно утверждать, что для плохо формализуемых задач с недостаточным количеством исходных данных, их противоречивостью и искажениями аналитические зависимости установить довольно сложно, а иногда невозможно. В таком случае выявить и использовать «скрытые» в исходной информации знания можно только с помощью анализа и обработки больших массивов накопленной как числовой, так и текстовой информации. Решают такие задачи с применением специальных методов и средств, в частности, нечеткой логики и нейронных сетей, позволяющих осуществлять интеллектуальную обработку данных с поиском закономерностей.

Отмеченные особенности информации в ПрО «СТМ» показывают, что работа с этим информационным ресурсом требует таких средств интеллектуального анализа данных, как инженерия знаний, нечеткая логика, значительно выходящих по своим возможностям за рамки стандартных процедур вычислительной математики и статистики. На этом пути можно надеяться на эффективное совмещение традиционной обработки с плохо формализуемыми экспертными сведениями о качестве и надежности данных, а в конечном итоге на вычленение из «сырых» опытных данных знаний со всесторонним учетом их достоверности.

В статті розглянуто підхід до прогнозування властивостей надтвердих матеріалів на основі системного аналізу інформації. Дані про властивості речовин розглядають як інформаційний ресурс з певними вимогами до вибору інформаційних технологій. Показано, що розв'язання задач прогнозування властивостей надтвердих матеріалів вимагає використання інтелектуального аналізу даних та нейронно-мережових технологій.

Ключові слова: системний аналіз, система управління знаннями, предметна область, інтелектуальний аналіз, надтвердий матеріал, фізико-механічні властивості.

The paper discusses a systemic-data-analysis-based approach to the forecasting of properties of superhard materials. Data concerning properties of materials are presented as an information resource with specific requirements set for information technology selection. It is demonstrated that to forecast properties of superhard materials intellectual data analysis and neural network methods need be applied.

Key words: Systems analysis, knowledge management system, subject area, intellectual analysis, superhard material, physical and mechanical properties.

Литература

1. Интегрированная модель предметной области "Сверхтвердые материалы"/ В. Н. Кулаковский, А. А. Лебедева, И. В. Скворцов и др.// Сверхтвердые матер. – 2009. – № 5. – С. 90–91.
2. Краковецкий А. Извлечение данных из веб-ресурсов. – Режим доступа: <http://msug.vn.ua/Posts/Details/4213>.
3. Разработка высокоэффективных средств создания и обработки онтологических баз знаний / В. А. Филатов, С. С. Щербак, А. А. Хайрова. – Режим доступа: <http://shcherbak.net>.
4. Лебедева А. А., Фидаров Т. З., Скворцов И. В. Подход к построению предметной онтологии для экспертной системы «Сверхтвердые материалы» // Породоразрушающий и металлообрабатывающий инструмент – техника и технология его изготовления и применения: Сб. науч. тр. – К.: ИСМ им. В.Н. Бакуля НАН Украины, 2007. – Вып. 10. – С. 89–91.
5. Свойства веществ и материалов: возможности и перспективы информационных технологий в подготовке и распространении справочных данных / В. Ю. Зицерман, Г. А. Кобзев, Л. Р. Фокин. – Режим доступа: <http://www.thermophysics.ru>.

Поступила 21.06.12

УДК 001:004.6

В. Н. Колодницкий, канд. физ.-мат. наук

Институт сверхтвердых материалов им. В. Н. Бакуля НАН Украины, г. Киев

АНАЛИЗ ЦИТИРОВАНИЯ ЖУРНАЛА «СВЕРХТВЕРДЫЕ МАТЕРИАЛЫ» НА ОСНОВЕ ДАННЫХ НАНОМЕТРИЧЕСКИХ БАЗ В ЦЕЛЯХ ПОВЫШЕНИЯ ЕГО ИМПАКТ-ФАКТОРА

Приведены наукометрические критерии индексов цитирования журнала «Сверхтвердые материалы», а также анализ цитирования опубликованных в нем научных статей через базу Google Scholar в целях прогнозирования и повышения его импакт-фактора.

Ключевые слова: научный журнал, наукометрическая база данных, цитирование, импакт-фактор, журнал «Сверхтвердые материалы».

Показателем цитирования научных журналов, который определяет их информационную значимость, является импакт-фактор (Impact Factor). Это один из формальных критериев, по которому сравнивают уровень научных исследований в смежных областях знаний. При присуждении грантов, научных степеней и званий эксперты непременно обращают внимание на наличие у соискателя публикаций в журналах, охватываемых библиометрическим справочником Journal Citation Reports (JCR), который выпускает компания Thomson Reuters (Институт научной информации – The Institute for Scientific Information/ISI), г. Филадельфия, США.