

**АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ
БАЗЫ ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНОЙ
ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНОЙ
СИСТЕМЫ ДЛЯ ОПРЕДЕЛЕНИЯ
ТЕПЛОФИЗИЧЕСКИХ СВОЙСТВ
МАТЕРИАЛОВ**

А.Ю. Любимов, З.М. Селиванова

ГОУ ВПО «Тамбовский государственный технический университет», г. Тамбов

Рецензент д-р техн. наук, профессор А.А. Арзамасцев

Ключевые слова и фразы: база знаний; интеллектуальная система; информационно-измерительная система; теплофизические свойства.

Аннотация: Предложены информационная среда создания базы знаний, математическая модель реализации запроса информации пользователем, решение задачи классификации исследуемых материалов, позволяющие разработать алгоритм информационного поиска для извлечения знаний и принятия решений в процессе функционирования интеллектуальной информационно-измерительной системы контроля.

Функционирование интеллектуальной информационно-измерительной системы (ИИИС) неразрушающего контроля (НК) теплофизических свойств (ТФС) материалов основано на базе знаний. При разработке базы знаний необходимо обеспечить: оперативность поиска требующейся информации пользователю, точность, полноту поиска, непротиворечивость информации.

Разработано алгоритмическое обеспечение базы знаний, позволяющее повысить оперативность извлечения информации из базы знаний, осуществить точность и полноту поиска информации и принятия решений при контроле ТФС.

Задачи исследования:

а) разработать информационную среду создания базы знаний для ИИИС НК ТФС материалов;

б) решить задачу классификации исследуемых материалов (ИМ), информация о которых хранится и извлекается из базы знаний;

Любимов А.Ю. – магистрант кафедры «Конструирование радиоэлектронных и микропроцессорных систем»; Селиванова З.М. – доктор технических наук, профессор кафедры «Конструирование радиоэлектронных и микропроцессорных систем», e-mail: Selivanova@mail.jesby.tstu.ru, ТамбГТУ, г. Тамбов.

в) разработать математическую модель реализации запроса информации пользователем;

г) сформировать продукционные правила, на основе которых реализуется механизм вывода информации;

д) разработать алгоритм извлечения требуемой информации из базы знаний в процессе функционирования ИИИС НК ТФС материалов.

Новизна заключается в создании алгоритма извлечения знаний, реализованного в ИИИС НК ТФС материалов, на основе применения интеллектуальной процедуры выбора класса ИМ (распознавание образов) по доминирующему признаку и разработки математической модели запроса информации пользователем.

При решении задач исследования использованы методы теории принятия решений, информационных технологий и искусственного интеллекта.

Использование современных информационных технологий позволяет создать информационную среду разработки базы знаний ИИИС НК ТФС материалов как совокупность знаний и данных, позволяющих принимать решение при информационном поиске для контроля ТФС материалов [1]. Под информационной средой понимается структурированная информация, необходимая для обеспечения оперативности извлечения знаний, точности и полноты получаемой информации пользователем. Для повышения эффективности функционирования ИИИС НК ТФС материалов информационную среду создания базы знаний необходимо постоянно пополнять, совершенствовать и развивать с использованием приобретенных знаний. Структурная схема компонентов информационной среды создания базы знаний приведена на рис. 1. Пополнение базы знаний на этапе проектирования осуществляется экспертом, а на этапе эксплуатации в ИИИС – пользователем.

Постановка и решение задачи классификации исследуемых материалов [2]. Задача классификации ИМ заключается в том, чтобы всю совокупность ИМ $Q = \{Q_i, i = 1, \dots, N\}$, представленную в виде матрицы Q , разбить на однородные классы.

Обозначим множество ИМ, свойства которых определяются ИИИС, через

$$J = \{I_i, i = \overline{1, N}\}, \quad (1)$$

здесь I_i – идентификатор (имя) i -го материала; N – число различных ИМ.

Каждый ИМ $I_i \in J$ характеризуется набором m свойств (параметров)

$$J_i = (q_{i1}, q_{i2}, \dots, q_{im}), \quad i = \overline{1, N}, \quad (2)$$

где q_{ij} – идентификатор j -го свойства для i -го материала.

Диапазон возможных числовых значений параметра (характеристики) для q_{ij} обозначим следующим образом

$$d_{ij} = [d_{ij}^H, d_{ij}^B], \quad i = \overline{1, N}, \quad j = \overline{1, m}, \quad (3)$$

где d_{ij}^H, d_{ij}^B – нижняя и верхняя границы интервала соответственно.



Рис. 1. Структурная схема информационной среды

Пара (q_{ij}, d_{ij}) содержит информацию о виде ИМ, наименовании свойства и интервале возможных значений этого свойства.

На основе данных (1) – (3) формируется $N \times (m+l)$ – расширенная матрица информации об ИМ и их свойствах

$$Q_P = \begin{pmatrix} (qd)_{11} & \cdots & (qd)_{1m} & (qd)_{1(m+1)} & \cdots & (qd)_{1(m+l)} \\ (qd)_{21} & \cdots & (qd)_{2m} & (qd)_{2(m+1)} & \cdots & (qd)_{2(m+l)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (qd)_{N1} & \cdots & (qd)_{Nm} & (qd)_{N(m+1)} & \cdots & (qd)_{N(m+l)} \end{pmatrix} = \begin{pmatrix} Q_1^P \\ Q_2^P \\ \vdots \\ Q_N^P \end{pmatrix}, \quad (4)$$

где l – число факторов, учитываемых при идентификации измерительной ситуации; Q_i^P – i -я строка матрицы Q_P .

В самом общем виде задача классификации ИМ формулируется следующим образом. Множество ИМ (1), характеризуемое матрицей (4), требуется разбить на k классов K_1, K_2, \dots, K_k таких, что выполняются условия: ИМ, принадлежащие одному классу, близкие между собой по всей совокупности свойств, при этом в качестве меры близости рассматривается некоторое расстояние $d(Q_\nu, Q_\mu)$ между Q_ν и Q_μ в m -мерном пространстве;

$$\bigcup_{i=1}^k K_i = J; \quad \forall i, j = \overline{1, N}, \quad i \neq j: K_i \cap K_j = \emptyset; \quad k < N.$$

Алгоритмическое обеспечение. Решение задач классификации ИМ по данным информационной матрицы (4) рассматривается в некотором признаковом пространстве π , в котором строки матриц $Q_i, i = \overline{1, N}$ представлены точками [2]. В общем случае, пространство π имеет размерность $\pi \in \{m, m+l, p\}$, координатные оси обозначим q_1, q_2, \dots, q_π . Геометрическая близость двух Q_i, Q_j или нескольких точек в этом пространстве соответствует близости «физических» свойств ИМ, их однородности. В этом случае задача классификации заключается в разбиении анализируемой совокупности точек на классы таким образом, чтобы объекты, принадлежащие одному классу, находились на небольших расстояниях друг от друга. Для оценки близости ИМ использовано взвешенное евклидово расстояние

$$d(Q_i, Q_j) = \sqrt{\sum_{k=1}^{\pi} v_k (\bar{d}_{ik} - \bar{d}_{jk})^2}, \quad (5)$$

где v_k – весовой коэффициент k -го показателя для i -го и j -го материалов соответственно. Если показатель задан интервальным значением, то в качестве \bar{d}_{ik} берется его середина.

Алгоритм решения задачи классификации ИМ с использованием меры близости (5) включает следующие этапы.

1. Методом экспертных оценок определяются число классов k и ИМ – представители этих классов, обозначим их $\bar{Q}_i, i = \overline{1, k}$. Представители классов берутся из информационной матрицы (4) или задаются в виде некоторых виртуальных материалов, характеризующих соответствующий класс ИМ. В последнем случае для \bar{Q}_i допускается задание не всех показателей (признаков), а лишь основных, наиболее влияющих на определение измерительной ситуации.

2. Определяется один или два показателя ИМ из числа основных, по значениям которых классы можно проранжировать и каждому классу присвоить соответствующее имя.

3. Для каждого класса определяются пороговые значения расстояний $d_{\text{пор}}^s, s = \overline{1, k}$.

4. Производится расчет расстояний $d(Q_i, \bar{Q}_s)$, $i = \overline{1, N_1}$, $s = \overline{1, k}$ по формуле (5) и принимается решение о закреплении ИМ матрицы Q за определенными классами, где N_1 – число ИМ, оставшихся в матрице Q после выделения представителей классов (обычно $N_1 = N - k$). Данный этап включает следующие процедуры:

4.1. Из матрицы Q берется строка Q_i с показателями очередного ИМ и рассчитываются расстояния

$$d(Q_i, \bar{Q}_j) = \sqrt{\sum_{j=1}^{\pi} v_j (\bar{d}_{ij} - \bar{d}_{sj})^2}, \quad s = \overline{1, k}. \quad (6)$$

4.2. Определяется минимальное расстояние

$$d_{\min}(Q_i, \bar{Q}_m) = \min_s \{d(Q_i, \bar{Q}_s), s = \overline{1, k}\}. \quad (7)$$

4.3. Расстояние $d_{\min}(Q_i, \bar{Q}_m)$ сравнивается с пороговым и принимается решения об отнесении i -го ИМ к соответствующему классу, то есть

$$d_{\min}(Q_i, \bar{Q}_m) \left\{ \begin{array}{l} \leq d_{\text{пор}}^m, Q_i \in K_m, \\ \leq d_{\text{пор}}^m, Q_i \text{ включается в } K_m \text{ условно} \end{array} \right\}. \quad (8)$$

5. После выполнения процедур (6) – (8) для всех ИМ производится анализ результатов классификации. Если все ИМ закреплены за определенными классами и нет ИМ, условно включенных в какой-либо класс, то на этом решение задачи заканчивается. Если какой-либо класс содержит условно включенные в него материалы, то эти материалы выделяются в отдельную группу K_{k+1} .

Во многих случаях на практике на результаты классификации решающее влияние оказывает один показатель (признак) в модели ИМ. Назовем этот признак доминирующим. При наличии доминирующего признака q_d целесообразно использовать эвристический алгоритм классификации:

а) определяется диапазон возможных значений признака q_d для всех ИМ, то есть находятся границы интервала $[q_d^H, q_d^B]$;

б) полученный диапазон $[q_d^H, q_d^B]$ разбивается на k подинтервалов, и для каждого подинтервала выбираются ИМ – представители соответствующих классов;

в) для каждого ИМ определяются расстояния до представителей классов и по значению минимального расстояния материал относится к соответствующему классу.

Математическая модель запроса информации пользователем может быть представлена зависимостью

$$P_3 = f(X_M, Re_i, Y_{\Pi}, J),$$

где $X_M = \{X_{МИ}, X_{МП}, X_{МС}, X_{ММет}\}$ – множество описаний классов ИМ ($X_{МИ}$ – класс изоляционных материалов; $X_{МП}$ – класс полимерных материалов; $X_{МС}$ – класс строительных материалов; $X_{ММет}$ – класс металлов); $Re_{di} = \{Re_{d1}, \dots, Re_{d4}\}$ – множество продукционных правил извлечения информации о классах материалов соответствующего диапазона теплопроводности d ; $Y_{\Pi} = \{Y_i, i = 1, \dots, n\}$ – множество результатов поиска искомой информации; $J = \{I_i, i = \overline{1, N}\}$ – множество материалов (I_i – идентификатор (имя) i -го материала, N – число ИМ).

Вектор запроса W искомой информации из базы знаний о конкретном материале составляется по правилу [3]

$$W = \left(\sum_{i=1}^m qd_{im} \right).$$

Вектор запроса класса исследуемых материалов W_k представлен в виде

$$W_k = \left(\sum_{i=1}^N (qd)_{i1}, \sum_{i=1}^N (qd)_{i2}, \dots, \sum_{i=1}^N (qd)_{im} \right).$$

Алгоритм извлечения знаний необходим при разработке базы знаний ИИИС НК ТФС твердых материалов (строительных, полимерных, композитных, теплоизоляционных и др.) и изделий из них [1]. База знаний включает совокупность баз данных, которые, в свою очередь, представляют собой множества декларативных знаний о ТФС материалов, определенным образом формализованных и представленных в памяти ИИИС. При построении базы знаний учитываются характеристики ИМ и их зависимости от воздействующих факторов в рассматриваемой предметной области.

Диалог пользователя и эксперта с базой знаний осуществляется через пользовательский интерфейс, одним из назначений которого является реализация алгоритма извлечения знаний из базы знаний. Эта система осуществляет обработку и редактирование данных, содержащихся в базе знаний [4], а также их извлечение и представление пользователю в понятной, заданной форме.

Таким образом, база знаний включает в себя набор правил, перебором которых управляет интерпретатор правил. Интерпретатор правил выполняет две функции: во-первых, просмотр существующих фактов из рабочей памяти (базы данных) и правил из базы знаний и добавление в рабочую память новых фактов и, во-вторых, определение порядка просмотра и применения правил. Этот алгоритм извлечения знаний управляет процессом взаимодействия с пользователем, сохраняя для него информацию о полученных данных и запрашивает, если это необходимо, информацию,

когда для выполнения очередного правила в рабочей памяти оказывается недостаточно данных. Механизм вывода реализуется программой и включает в себя два компонента: один – реализует вывод, другой – управляет этим процессом. Действие компонента вывода основано на применении продукционных правил. Например, при определении класса материала по заданной величине теплопроводности продукционные правила Re_i можно представить в следующем виде:

- 1) ЕСЛИ $\lambda \in d_1, d_1 = 0,02 \dots 0,2$, ТО $M \in X_{MI}$;
- 2) ЕСЛИ $\lambda \in d_2, d_2 = 0,21 \dots 0,3$, ТО $M \in X_{MP}$;
- 3) ЕСЛИ $\lambda \in d_3, d_3 = 0,31 \dots 0,8$, ТО $M \in X_{MC}$;
- 4) ЕСЛИ $\lambda \in d_4, d_4 = 0,81 \dots 500$, ТО $M \in X_{MMeT}$.

где λ – теплопроводность, Вт/(м·К); d_i – диапазон теплопроводности для i -го класса материалов ($i = 1, \dots, 4$); M – материал.

Продукционные правила реализуются, если условия и данные запроса пользователя соответствуют левой части правил. Управляющий компонент механизма вывода определяет порядок применения правил и выполняет четыре функции:

а) сопоставление – образец правила сопоставляется с имеющимися данными запроса пользователя;

б) выбор – если в конкретной ситуации одновременно может быть применено несколько правил, то из них выбирается одно, наиболее соответствующее заданному критерию, позволяющему пользователю получить запрашиваемую информацию при эксплуатации ИИИС НК ТФС материалов;

в) срабатывание – если образец правила при сопоставлении совпал с какими-либо фактами из рабочей памяти, то правило выполняется;

г) действие – оперативная память подвергается изменению путем добавления в нее заключения сработавшего правила.

Интерпретатор правил работает циклически. В каждом цикле он просматривает все правила, чтобы выявить те, которые удовлетворяют фактам в рабочей памяти. После выбора образца правило срабатывает, его заключение заносится в рабочую память, и, затем, цикл повторяется заново. В одном цикле может сработать только одно правило. Если сразу несколько правил удовлетворяют фактам, то интерпретатор производит выбор в соответствии с критерием единственного правила, которое срабатывает в данном цикле. Таким образом, алгоритм работы интерпретатора можно представить в виде блок-схемы (рис. 2).

Представленный алгоритм позволяет осуществлять извлечение интересующих пользователя знаний в ИИИС НК ТФС материалов в соответствии с заданным критерием выбора продукционных правил и вывод результата в заданной форме.

Для создания базы знаний используется программный продукт компании Borland – C++ Builder. База знаний, написанная на C++ компилируется через Assembler с целью последующей записи отдельных модулей в память микроконтроллера ИИИС.

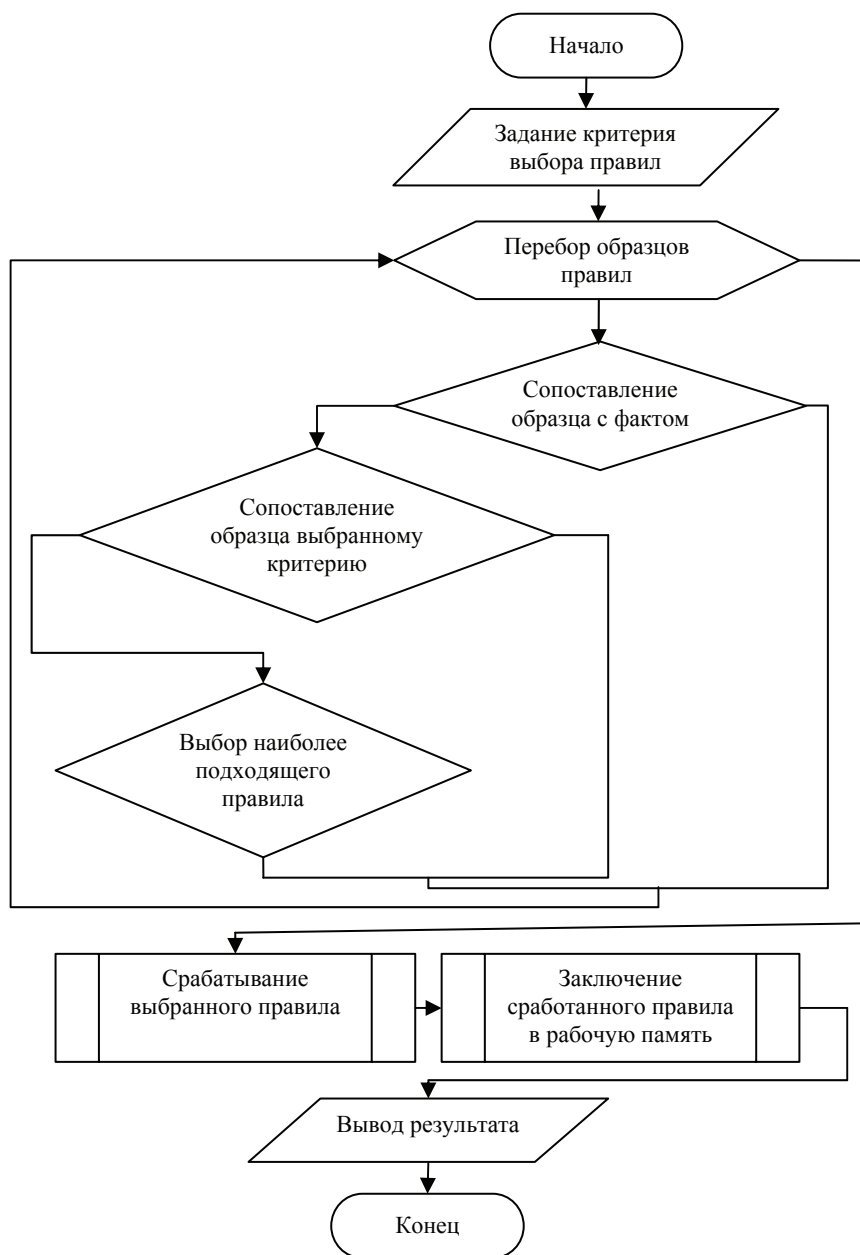


Рис. 2. Алгоритм работы интерпретатора

Выводы. На основе разработанных информационной среды создания базы знаний, математической модели запроса информации пользователем, решения задачи классификации ИМ, продукционных правил для выполнения механизма вывода информации и программы реализации базы знаний создано алгоритмическое обеспечение базы знаний, позволяющее осуществлять оперативный информационный поиск для извлечения знаний и принятия решений при функционировании ИИИС НК ТФС материалов.

Список литературы

1. Селиванова, З.М. Интеллектуальная информационно-измерительная система для определения теплофизических свойств твердых материалов / З.М. Селиванова // Приборы и техника эксперимента. – 2006. – № 4. – С. 153–154.
2. Айвазян, С.А. Прикладная статистика. Основы эконометрики : учеб. для вузов : в 2 т. Т. 1 : Теория вероятностей и прикладная статистика / С.А. Айвазян, В.С. Мхитарян. – М. : Юнити-Дана, 2001. – 656 с.
3. Информационные системы : учеб. пособие / О.Л. Голицина [и др.]. – 2-е изд., перераб. и доп. – М. : Форум ; ИНФРА-М, 2007. – 496 с.
4. Любимов, А.Ю. База знаний теплофизических характеристик материалов для информационно-измерительной системы / А.Ю. Любимов, Д.И. Колбасин // Информац. системы и процессы : сб. науч. тр. – Тамбов ; М. ; СПб. ; Баку ; Вена : Нобелистика, 2008. – Вып. 7. – С. 49–52.

Algorithmic Maintenance of the Knowledge Base of Intellectual Information-Measuring System for Determining Thermo-Physical Properties of Materials

A.Yu. Lubimov, Z.M. Selivanova

Tambov State Technical University, Tambov

Key words and phrases: intellectual system; information-measuring system; knowledge base; thermo-physical properties.

Abstract: The paper introduces the information environment for creating the knowledge base and mathematical model for information inquiry by the user as well as the solution to the problem of classification of the investigated materials, allowing to develop the algorithm of information search for extraction of knowledge and decision-making in the course of functioning of the intellectual information-measuring monitoring system.

© А.Ю. Любимов, З.М. Селиванова, 2010