

УДК 519.95

ФОРМАЛИЗАЦИЯ ПРОБЛЕМЫ РАЗРАБОТКИ ЭКСПЕРТНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ С РАЗВИВАЮЩИМСЯ ИНТЕЛЛЕКТУАЛЬНЫМ ЯДРОМ НА БАЗЕ ИНС-МОДЕЛЕЙ

© А.А. Арзамасцев, Н.А. Зенкова, А.В. Неудахин

Ключевые слова: автоматизированная экспертная система, развивающееся интеллектуальное ядро, нейросетевые модели, математическая формализация задач накопления информации.

Статья посвящена разработке автоматизированной технологии построения экспертных систем, сочетающих в себе функциональные преимущества развивающегося интеллектуального ядра и сетевых технологий взаимодействия. Приведены результаты математической формализации задач накопления информации и построения на ее основе развивающегося интеллектуального ядра.

Необходимость в экспертных системах (ЭС), как в интеллектуальных помощниках, способных решать трудноформализуемые задачи, существует во многих сферах человеческой деятельности: промышленности, научных исследованиях, социальной и экономической сферах.

Ранее нами была разработана автоматизированная технология построения ЭС с интеллектуальным ядром на основе нейросетевых моделей [1–2]. Концептуальная модель таких ЭС приводится на рис. 1. Основными отличительными компонентами данной модели является веб-интерфейс, база данных, используемая для долговременного хранения технической информации, а также интеллектуальное ядро, включающее развивающуюся систему представления знаний экспертов, основанную на нейросетевых моделях.

Для разработки автоматизированной технологии построения ЭС, соответствующих разработанной концептуальной модели, была проведена математическая формализация задач накопления информации и построения на ее основе развивающегося интеллектуального ядра.

Основные принципы формализации. Пусть имеется распределенная система получения исходной информации, включающая N терминалов, каждый из которых предназначен для ввода n независимых параметров, имеющих m уровней каждый. Работа терминалов организована таким образом, что каждый из них работает лишь часть времени так, что интервалы их простоя являются случайными величинами, заданными плотностями распределений $p_1(\tau)$, $p_2(\tau)$, ..., $p_N(\tau)$. Время ввода информации не зависит от номера терминала, но линейно зависит от n , т. е. $\tau_{\text{input}} = k_1 n$, где k_1 – некоторый коэффициент пропорциональности. Будем считать, что очередная порция информации добавляется к базе данных в виде одной записи, когда на i -м терминале ($i = 1, \dots, N$) каждому независимому параметру x_j ($j = 1, \dots, n$) присвоен соответствующий уровень l_k ($k = 1, \dots, m$).

Будем считать, что передача информации с любого терминала в базу данных осуществляется мгновенно,

т. е. это время на несколько порядков меньше времени, затрачиваемого пользователем на ввод информации.

Тогда среднее время одного цикла работы i -го терминала (ввод + ожидание следующего ввода информации), соответствующее добавлению в базу данных одной записи, может быть вычислено как:

$$\bar{\tau}_i = k_1 n + \int_0^{t_i^{\max}} \tau p_i(\tau) d\tau.$$

Поэтому за период времени T (например $T = 24$ ч) на i -м терминале могут быть осуществлены r_i циклов ввода, соответствующие добавлению в базу данных r_i записей:

$$r_i = \frac{T}{\bar{\tau}_i} = \frac{T}{k_1 n + \int_0^{t_i^{\max}} \tau p_i(\tau) d\tau}.$$

Соответственно, общее число записей информации в БД, полученное со всех терминалов за период T , составит:

$$R = \sum_{i=1}^N r_i = T \sum_{i=1}^N \frac{1}{k_1 n + \int_0^{t_i^{\max}} \tau p_i(\tau) d\tau}.$$

Выполним оценку периода первичного накопления данных в БД – t_1 , предшествующего первому обучению интеллектуального ядра (ИНС-модели).

В соответствии с теоремами А.Н. Колмогорова [1–2] о представимости функций нескольких переменных с помощью суперпозиций и сумм функций одного переменного, можно утверждать, что каждая непрерывная функция n переменных, заданная на единичном кубе n -мерного пространства, представима в виде:

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} h_q \left[\sum_{p=1}^n \varphi_q^p(x_p) \right],$$

где функции $h_q(u)$ непрерывны, а функции $\varphi_q^p(x_p)$, кроме того, еще и стандартны, т. е. не зависят от выбора функции f .

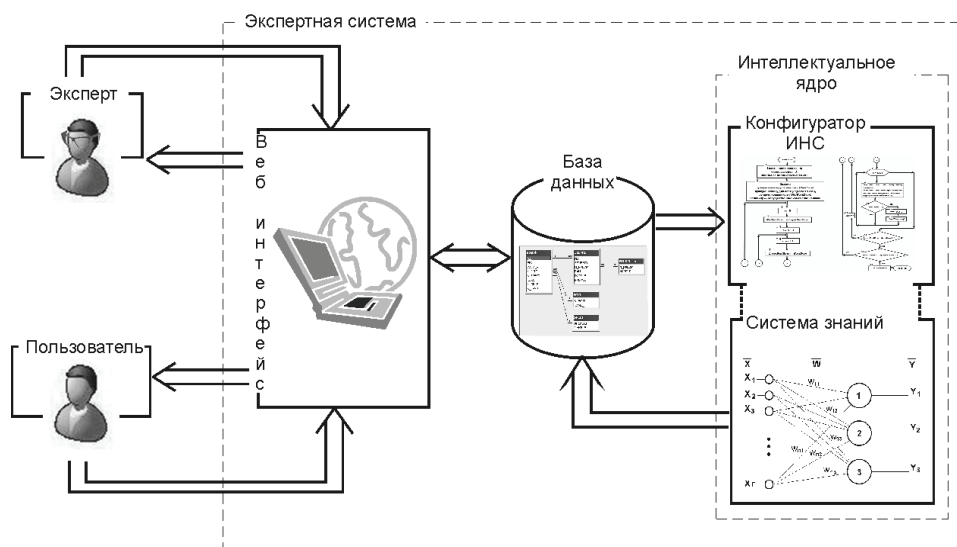
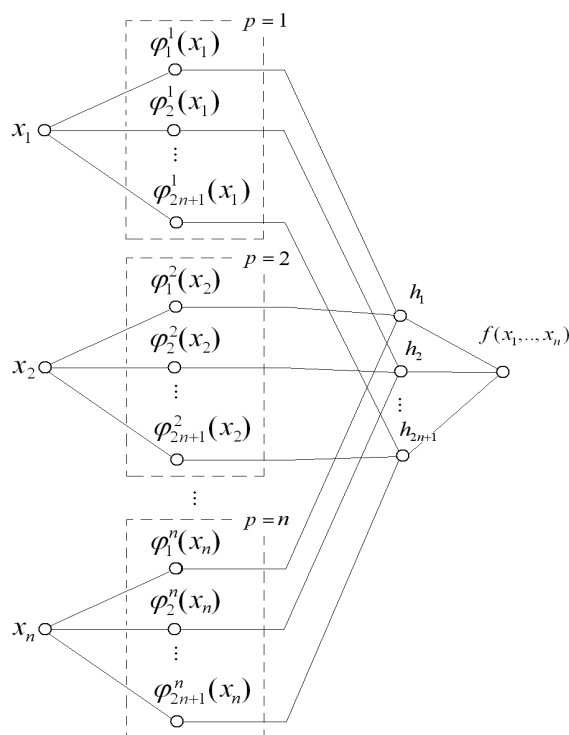
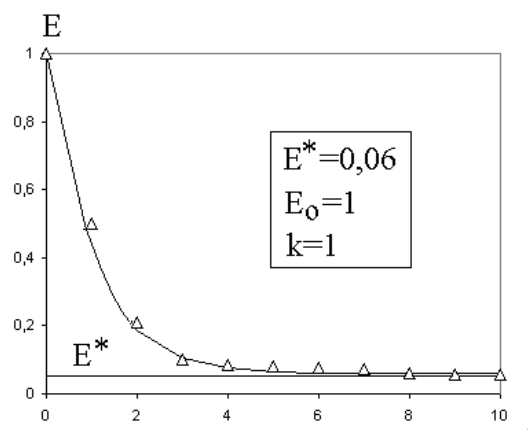


Рис. 1. Концептуальная модель веб-ориентированных нейросетевых ЭС

Рис. 2. Представление функции n переменных в виде ИНС-модели

Фактически это означает, что «минимальная» структура ИНС-модели, позволяющая аппроксимировать функцию n переменных, должна иметь следующий вид: входной слой, состоящий из n нейронов, первый скрытый слой, включающий $(2n+1) \cdot n$ функциональных нейронов, второй скрытый слой, состоящий из $(2n+1)$ функциональных нейронов, и выходной, суммирующий нейроны (рис. 2). Указанная нейронная сеть не является полносвязанной и имеет $(2n+1)n + (2n+1)n + 2n + 1 = (2n+1)^2$ связей (степеней свободы при обучении ИНС-модели).

Рис. 3. Корреляция эмпирических данных по среднеквадратичной погрешности обучения сети – E (треугольники) от времени обучения с результатами расчета по уравнению (6). На графике показаны приведенные безразмерные значения

Необходимо заметить, что теоремы А.Н. Колмогорова не несут никакой информации о типе нелинейности функций φ и h . Так, в интеллектуальном ядре экспертной системы, базируемой на ИНС-модели, мы используем активационную функцию

$$f(s) = \frac{1}{1 + e^{-\alpha s}} \text{ и полносвязанную сеть – многослойный персептрон.}$$

Для этого случая число степеней свободы ИНС-модели будет $(2n+1) \cdot (3n^2 + n + 1)$. Учитывая, что число записей в БД должно быть не меньше, чем число степеней свободы, получим оценку для первичного периода накопления данных:

$$t_1 \geq K \cdot \left(\sum_{i=1}^N \frac{1}{k_1 n + \int_0^{t_i^{\max}} \tau p_i(\tau) d\tau} \right)^{-1} \quad (1)$$

где $K = (2n+1)^2$ для ИНС-модели, построенной в соответствии с теоремой А.Н. Колмогорова, и

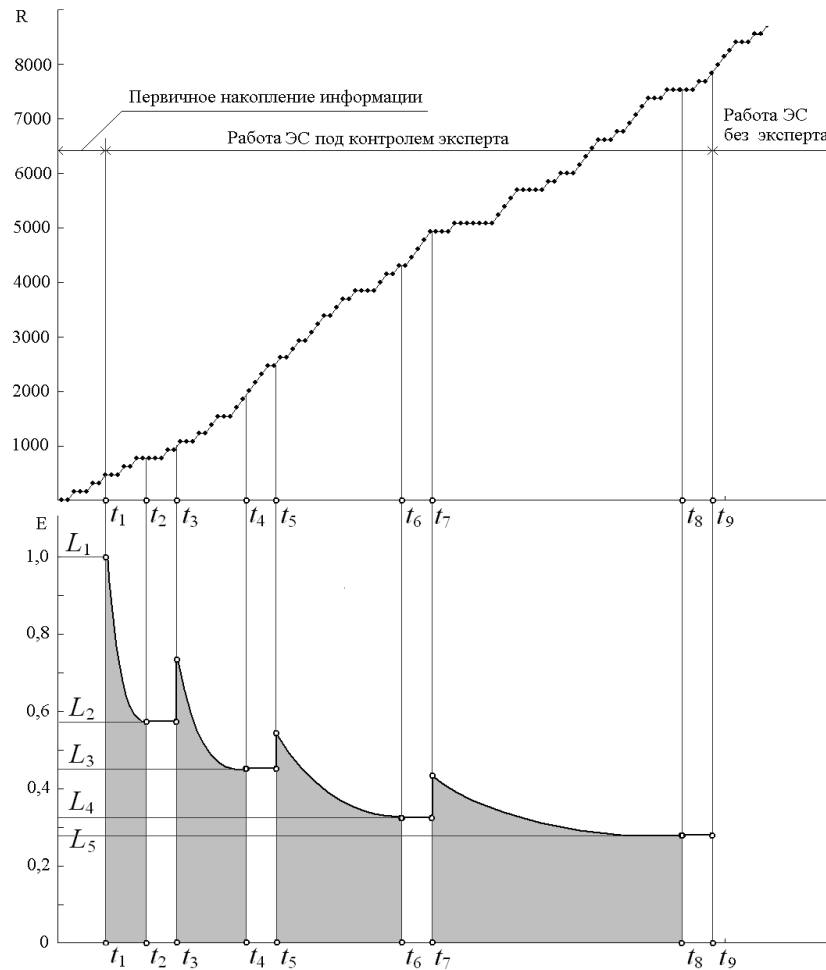


Рис. 4. График накопления первичной информации (верхняя часть рисунка); R – число записей в базе данных, полученных с терминалов. Процесс обучения ИНС-модели, составляющей интеллектуальное ядро ЭС. E – приведенная среднеквадратическая ошибка ИНС-модели. По оси абсцисс – безразмерное время

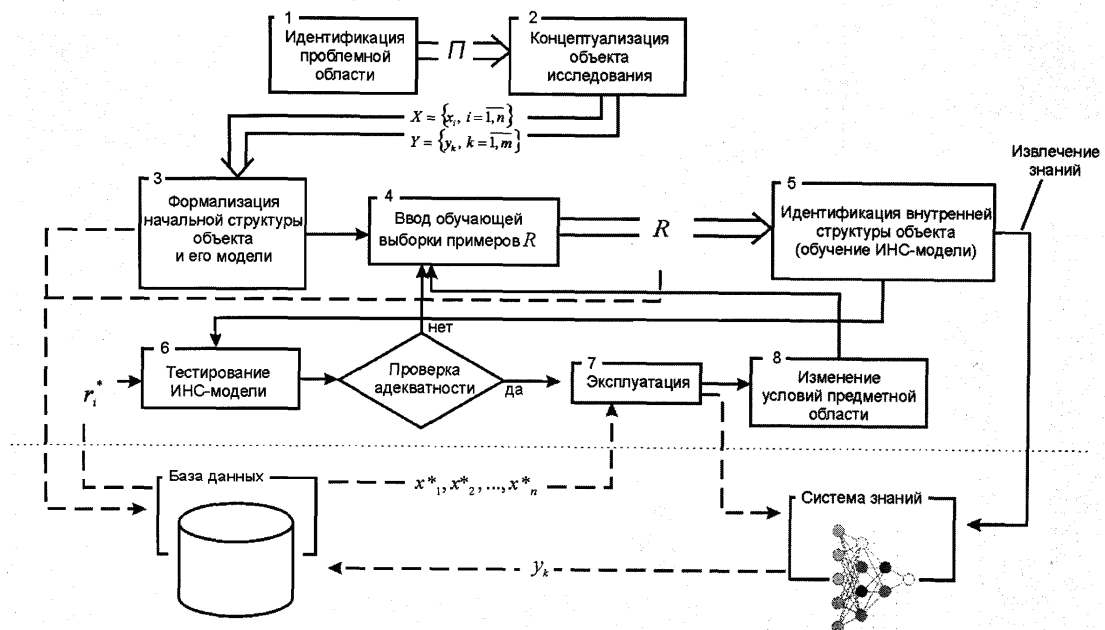


Рис. 5. Функциональная модель жизненного цикла построения ЭС, обладающих интеллектуальным ядром на основе ИНС-моделей и веб-интерфейсом

$K = (2n+1)(3n^2 + n + 1)$ для полносвязанной ИНС-модели.

Процесс обучения интеллектуального ядра ЭС заключается в минимизации функционала вида:

$$E\{\mathbf{w}[S(n)]\} = \|\Delta\| = \|\mathbf{Y}^{tabl} - \mathbf{Y}^{net}(\mathbf{w})\| = \sqrt{\sum_{i=1}^R \sum_{j=1}^n (Y_{ij}^{tabl} - Y_{ij}^{net}(\mathbf{w}))^2} \rightarrow \min_{\mathbf{w}} \quad (2)$$

Вектор $\mathbf{w}$ определяется структурой сети S , которая, в свою очередь, связана с числом ее входов n так, что имеется однозначное соответствие n и $w(S)$. Поэтому вектор $\mathbf{w}^*$, зависящий от структуры сети и минимизирующий функционал (2), соответствует оптимальной ИНС-модели интеллектуального ядра и может быть определен как:

$$\mathbf{w}^* = \arg \min_{\mathbf{w} \in \Omega} F(\mathbf{w}) \quad (3)$$

Минимизацию функционала (2) осуществляем двумя методами: градиентным, так что последующие значения вектора $\mathbf{w}$ вычисляются по формуле $w^{t+1} = w^t - h(t)grad(E(w^t))$, и методом сканирования. Для указанных методов оценки времени обучения могут быть получены следующим образом: $\Theta_{grad} = k_2 n R$ и $\Theta_{skan} = k_3 R \alpha^N$. Здесь k_2 и k_3 – коэффициенты, зависящие от технических характеристик используемого оборудования, α – число разбиений диапазона изменения переменной.

Процесс обучения ИНС-модели, составляющей интеллектуальное ядро ЭС, зависит от количества и качества данных в обучающей выборке и может быть описан следующим дифференциальным уравнением:

$$\frac{dE(t)}{dt} = k_4 [E^* - E(t)],$$

с начальным условием:

$$E(0) = E_0 \quad (4)$$

где $E(t)$, E_0 и E^* – значения приведенной среднеквадратичной погрешности, ее начального значения и уровня, на котором она будет зафиксирована в конце цикла обучения ИНС-модели; k_4 – параметр удельной скорости обучения, зависящий от n , объема обучающей выборки и используемого метода минимизации невязки (2); t – безразмерное время.

Разделяя переменные в уравнении (4) и интегрируя, получим:

$$\int \frac{dE(t)}{k(E^* - E(t))} = \int dt$$

или

$$-\frac{1}{k} \ln[k(E^* - E(t))] = t + c. \quad (5)$$

Определив постоянную интегрирования из начальных условий (4), получим решение в виде:

$$E(t) = E^* (1 - e^{-kt}) + E_0 e^{-kt}. \quad (6)$$

Предложенная математическая формализация на практике позволяет решить следующие задачи: оценить первичный период накопления информации, необходимой для построения ИНС-модели исследуемого объекта; определить начальную структуру ИНС-модели на основе входных характеристик объекта исследования; сделать оценку количества циклов обучения интеллектуального ядра в присутствии эксперта.

На основе полученных результатов была разработана функциональная модель (рис. 5) жизненного цикла процесса построения ЭС с интеллектуальным ядром на основе нейросетевых моделей и поддерживающих механизмы удаленного доступа.

ЛИТЕРАТУРА

1. Арзамасцев А.А., Зенкова Н.А., Неудачин А.В. Автоматизированная технология построения экспертных систем с интеллектуальным ядром на основе ИНС-моделей // Открытое образование. 2008. № 3 (68). С. 35-39.
2. Арзамасцев А.А., Зенкова Н.А., Неудачин А.В. Технология построения медицинской экспертной системы на основе аппарата искусственных нейронных сетей // Информационные технологии. 2009. № 8. С. 60-63.
3. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных // Докл. АН СССР. 1956. Т. 108. № 2. С. 179-182.
4. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного // Докл. АН СССР. 1957. Т. 114. № 5. С. 953-956. (Цит. по книге: Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996).

Поступила в редакцию 11 июня 2009 г.

Arzamastsev A.A., Zenkova N.A., Neudachin A.V. Formalisation of a problem of working out of expert informational system with developing intellectual kernel on the basis of ANN-models.

Article is devoted to working out of the automated technology of construction of the expert systems combining functional advantages of the developing intellectual kernel and network technologies of interaction. Results of mathematical formalization of problems of accumulation of the information and construction on its basis of a developing intellectual kernel are resulted.

Key words: automated expert system; developing intellectual kernel; neuro-net models; mathematical formalization of problems of accumulation of the information.