

A requirement for stable algorithm creation based on a canonical method of integration of rigid-body dynamics equations is represented.

Keywords: canonical method, dynamic equation, stability, solid body

Получено 19.05.10

УДК 519.853.4

В. А. Тененёв, доктор физико-математических наук, профессор;
А. В. Тененёва, студентка
Ижевский государственный технический университет
ОБУЧЕНИЕ НЕЧЕТКИХ НЕЙРОННЫХ СЕТЕЙ
ГЕНЕТИЧЕСКИМ АЛГОРИТМОМ

Для обработки данных рассматривается применение нечеткой нейронной сети, основанной на правилах, построенных методом деревьев решений. Для оптимизации параметров сети применяется гибридный генетический алгоритм с вещественным кодированием. Установлено преимущество оптимизированных нечетких сетей над обычными многослойными нейронными сетями при решении задач аппроксимации.

Ключевые слова: обработка данных, нечеткие нейронные сети, генетический алгоритм, аппроксимация

Для извлечения закономерностей из набора эмпирических данных широко используются нейронные сети различных типов и адаптивные системы нечеткого вывода. Объединение принципов нечеткого логического вывода и нейросетевой структуры приводит к нечетким нейронным сетям [1]. Рассмотрим один из вариантов нечеткой сети – нечеткую нейронную сеть Такаги – Сугено – Канга (TSK), содержащую модель вывода с пятью слоями.

Нечеткая сеть основана на нескольких правилах M . Под правилом понимается логическая конструкция, представленная в виде *if A then B* ($A \rightarrow B$). Для вектора $\mathbf{x} = [x_1, \dots, x_n]$, где $A_i, i = \overline{1, M}$ – это лингвистические переменные с соответствующими им функциями принадлежности $\mu_{A_i}(x)$, построенные в пространстве входных значений. Выходной результат правила представлен в виде $y_i(\mathbf{x}) = p_{i0} + \sum_{j=1}^n p_{ij}x_j$. Функция принадлежности представляется обобщенной

функцией Гаусса $\mu_{A_i}(x_j) = \frac{1}{1 + \left(\frac{x_j - c_i}{\sigma_i}\right)^{2b_i}}$. Результат *if A then B* задается алгебраическим произведением

$$\mu_{A_i}(x) = \prod_{j=1}^n \frac{1}{1 + \left(\frac{x_j - c_i}{\sigma_i}\right)^{2b_i}}.$$

Агрегированный выходной сигнал такой сети рассчитывается по формуле

$$y(\mathbf{x}) = \frac{1}{\sum_{k=1}^N w_k} \sum_{i=1}^M w_i \left(p_{i0} + \sum_{j=1}^n p_{ij} x_j \right), \quad (1)$$

где $\mathbf{P} = [p_{ij}]$, $i = \overline{1, M}$; $j = \overline{0, n}$ – веса, подбираемые в процессе обучения; M – число правил вывода, задаваемых пользователем. Коэффициенты принадлежности $\mathbf{W} = [w_i]$ определяются значениями соответствующих функций $w_i = \mu_{A_i}(\mathbf{x})$.

Неизвестные в сети параметры $c_{ij}, \sigma_{ij}, b_{ij}$, $\mathbf{P} = [p_{ij}]$, $i = \overline{1, M}$; $j = \overline{0, N}$ делятся на две части $c_{ij}, \sigma_{ij}, b_{ij}$ и $\mathbf{P} = [p_{ij}]$.

Для обучения используются обучающие выборки (\mathbf{x}^l, d^l) , $l = \overline{1, L}$. При фиксированных коэффициентах $w_i^l = \prod_{j=1}^n \frac{1}{1 + \left(\frac{x_j^l - c_{ij}}{\sigma_{ij}} \right)^{2b_{ij}}}$ запишем $y(\mathbf{x}^l) = d^l$, $q = \overline{1, L}$ в виде

системы линейных алгебраических уравнений:

$$\begin{bmatrix} w_1^1 & w_1^1 x_1^1 & \cdots & w_1^1 x_n^1 & \cdots & w_M^1 & w_M^1 x_1^1 & \cdots & w_M^1 x_n^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_1^L & w_1^L x_1^L & \cdots & w_1^L x_n^L & \cdots & w_M^L & w_M^L x_1^L & \cdots & w_M^L x_n^L \end{bmatrix} \begin{bmatrix} p_{10} \\ \vdots \\ p_{1n} \\ \vdots \\ p_{M0} \\ \vdots \\ p_{Mn} \end{bmatrix} = \begin{bmatrix} d^1 \\ \vdots \\ d^L \end{bmatrix}$$

или $\mathbf{GP} = \mathbf{d}$. Если количество уравнений H больше, чем число неизвестных $M(n+1)$, то значения $\mathbf{P} = [p_{ij}]$ получаются в результате применения операции псевдоинверсии

$$\mathbf{P} = \mathbf{G}^+ \mathbf{d},$$

где $\mathbf{G}^+ = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T$.

Выходные значения сети находятся из (1) или

$$\mathbf{y} = \mathbf{GP}, \quad \mathbf{y} = [y^1, \dots, y^H]^T.$$

Для корректировки параметров $c_{ij}, \sigma_{ij}, b_{ij}$, $i = \overline{1, M}$, $j = \overline{1, n}$ можно применять оптимизационный генетический алгоритм минимизации целевой функции

$$E = \frac{1}{2} \sum_{h=1}^H (y(\mathbf{x}^h) - d^h).$$

Нечеткая сеть TSK, как показало тестирование, лучше аппроксимирует сложные функции, чем нейронные сети типа многослойного персептрона.

В работе [2] предложено получать набор нечетких правил

$$R_r : \text{if } \bigcap x_i \in A_{ir} \text{ then } y \text{ is } B_r, r = \overline{1, K_R} \quad (2)$$

на основе построения дерева решений. Деревья решений – это способ представления правил в иерархической, последовательной структуре, где каждому объекту соответствует единственный узел, дающий решение.

Условие $x_i \in A_{ir}$ соответствует условию разделения множества объектов $x_i \begin{matrix} > \\ < \end{matrix} (w_{ij}), j = \overline{1, m-1}; i = \overline{1, n}$ и означает попадание величины x_i в нечеткий интервал w_{ij}^\pm с функциями принадлежности:

$$\mu^+(x_i) = \begin{cases} \frac{1}{1 + \exp(-\beta(x_i - w_{iq}))} + \frac{1}{1 + \exp(-\beta(1 - w_{iq}))} \frac{x_i - w_{iq}}{1 - w_{iq}} & | x_i \geq w_{iq}; \\ \frac{1}{1 + \exp(-\beta(x_i - w_{iq}))} \frac{x_i}{w_{iq}} & | x_i < w_{iq}, \end{cases} \quad (3)$$

$$\mu^-(x_i) = \begin{cases} \frac{1}{1 + \exp(\beta(x_i - w_{iq}))} + \left(1 - \frac{1}{1 + \exp(-\beta w_{iq})}\right) \frac{w_{iq} - x_i}{w_{iq}} & | x_i \geq w_{iq}; \\ \frac{1}{1 + \exp(\beta(x_i - w_{iq}))} \frac{1 - x_i}{1 - w_{iq}} & | x_i < w_{iq}. \end{cases}$$

Функция принадлежности $\mu^+(x_i)$ соответствует условию $x_i \geq (w_{iq})$, а $\mu^-(x_i)$ условию $x_i < (w_{iq})$.

Величина β характеризует размытость интервала. При $\beta \rightarrow \infty$ нечеткий интервал переходит в обычный.

При заданном векторе \mathbf{x} определяются степени истинности каждого правила (2): $\alpha_r, r = \overline{1, K_R}$. Степени истинности соответствуют значениям функций принадлежности левых частей (предпосылок):

$$\alpha_r = \min_k (\mu_k^\pm), k = \overline{1, g_r},$$

где g_r – количество условий в данном правиле R_r .

В результате агрегированный выходной сигнал определяется по формуле

$$y(\mathbf{x}) = \frac{1}{\sum_{r=1}^K \alpha_r} \sum_{r=1}^{K_R} \alpha_r \left(p_{r0} + \sum_{j=1}^n p_{rj} x_j \right). \quad (4)$$

Дерево решений строится по следующему алгоритму. Пусть целевая переменная соответствует некоторым классам $k = \overline{1, K}$, на которые разбито множество данных. Требуется отыскать некоторое классифицирующее правило, позволяющее разбить множество данных на эти классы [2]. В процессе поиска классифицирующего правила проводится перебор всех независимых переменных и отыскивается наиболее представительное правило на данном этапе. В бинарных деревьях решений применяются предикаты вида $x \leq w$, $x > w$. Данные разбиваются на две группы в соответствии со значением этого предиката. После этого процесс повторяется для каждой из этих групп до тех пор, пока получающиеся подгруппы содержат в себе представителей классов и включают в себя достаточно большое количество точек для того, чтобы статистически значимо быть разбитыми на меньшие подгруппы.

Обозначим $P_{iq}^k, k = \overline{1, K}; i = \overline{1, m-1}; q = \overline{1, n}; l = 1, 2$ – вероятность принадлежности классу k по атрибуту i и q -му пороговому значению $x_i \begin{matrix} > \\ < \end{matrix} (w_{ij}), j = \overline{1, m-1}; i = \overline{1, n}$, а P_l^k – вероятность попадания в класс k . В качестве меры математического ожидания информации, необходимого для определения класса объекта из множества X , рассматривается энтропия Шеннона

$$H_l = - \sum_{k=1}^K P_l^k \log_2 P_l^k. \quad (5)$$

Величина энтропии характеризует степень нечеткости системы данных. Критерием выбора является выражение, соответствующее максимальному упорядочиванию данных по классам

$$H_{iq} = - \sum_l \sum_{k=1}^K P_{liq}^k \log_2 P_{liq}^k \rightarrow \min.$$

Минимальное значение энтропии H_{iq} отвечает максимуму вероятности появления одного из классов. Выбранный номер переменной x_i и пороговое значение w_{iq} , минимизирующие H_{iq} ,

$$(i, w_{iq}) = \text{ArgMin } H_{iq}$$

используются для проверки значения переменной x_i по условию $x_i \begin{matrix} > \\ < \end{matrix} (w_{iq})$ и дальнейшее движение по дереву производится в зависимости от полученного результата.

Пороговые величины для входной переменной определяются выражением

$$w_{ij} = x_i^{\min} + (x_i^{\max} - x_i^{\min}) \frac{j}{M_i}, i = \overline{1, n}; j = \overline{1, M_i - 1},$$

где x_i^{\max}, x_i^{\min} – максимальные и минимальные значения переменной; $M_i = m$ – количество элементов разбиваемого подмножества.

Правило остановки разбиения узла основано на том, что разбиение должно быть нетривиальным, т. е. получившиеся в результате узлы должны содержать не менее заданного количества примеров N_l^{\min} .

Степень нечеткости системы данных может определяться либо по вероятностному способу (5), либо по возможностному [2]. Вместо вероятности определим возможность F_l^k принадлежности классу k F_{liq}^k , $k = \overline{1, K}$; $i = \overline{1, n}$; $q = \overline{1, m}$; $l = 1, 2$ – по атрибуту i и q -му пороговому значению $x_i \underset{<}{>} (w_{ij})$. Мера возможности определяется

$$\text{выражением } F = \frac{f_k}{\max_k f_k}.$$

Возможностная мера нечеткости системы данных определяется по формулам [3]:

$$U(F) = \frac{1}{L_F} \int_0^{L_f} \log_2 |c(F, L)| dL \quad \text{или} \quad U(F) = \frac{1}{L_F} \sum_{k=1}^{K-1} (L_{k+1} - L_k) \log_2 |c(F, L_{k+1})|.$$

Функция $c(F, L) = \{F_k \geq L\}$ называется функцией уровня L . Возможностный подход для определения критерия разбиения бывает предпочтительнее в случае ограниченного количества точек в обучающей выборке.

На структуру дерева и, следовательно, на точность прогноза оказывают влияние различные факторы: параметр функций принадлежности β в формулах (3); минимальное количество элементов в разделяющемся множестве N^{\min} ; количество классов выходной переменной K ; количество интервалов разбиения входной переменной m . Под точностью прогнозирования будем понимать среднеквадратичное отклонение рассчитываемых и фактических значений на обучающей выборке

$$\Phi(\beta, N^{\min}, K, m) = \sum_{l=1}^L (y(\mathbf{x}^l) - d^l)^2.$$

Параметры $\mathbf{X} = (\beta, N^{\min}, K, m)^T$ уточняются в результате решения оптимизационной задачи

$$\Phi(\mathbf{X}) \rightarrow \min. \quad (6)$$

Получение значения целевой функции (6) предполагает построение дерева решений, вычисление параметров нечеткой сети и расчет выходного значения (4). Для решения задачи оптимизации применяется гибридный генетический алгоритм с несколькими операторами скрещивания [3].

BLX- α оператор.

Для скрещивания выбираются две особи:

$$\mathbf{x}^{(1)} = (x_1^{(1)}, \dots, x_i^{(1)}, \dots, x_n^{(1)}), \quad \mathbf{x}^{(2)} = (x_1^{(2)}, \dots, x_i^{(2)}, \dots, x_n^{(2)}).$$

Значение нового гена определяется как линейная комбинация $x_i = a_{BLX} x_i^{(1)} + b_{BLX} x_i^{(2)}$. Коэффициенты a_{BLX}, b_{BLX} определяются следующими соотношениями:

$$a_{BLX} = (1 + \alpha - u(1 + 2\alpha)), \quad b_{BLX} = (u(1 + 2\alpha) - \alpha),$$

где число $\alpha \in [0, 1]$; $u \in (0, 1)$ – случайное число.

Оператор *Bin1*, имитирующий скрещивание при бинарном кодировании.

$$x_i = a_{Bin1}x_i^{(1)} + b_{Bin1}x_i^{(2)},$$

$$a_{Bin1} = \frac{(1+z)}{2}, \quad b_{Bin1} = \frac{(1-z)}{2},$$

$$z = \begin{cases} (2u)^{\frac{1}{1+\gamma}} & | u \leq 0,5; \\ (2(1-u))^{-\frac{1}{1+\gamma}} & | u > 0,5. \end{cases} \quad \text{Коэффициент } \gamma > 1.$$

Оператор скрещивания *Bin2*.

Так же, как в бинарном кодировании, преобразуем вещественное число r_i на отрезке $[A_i, B_i]$ в целое число $g_i = (2^L - 1) \frac{(r_i - A_i)}{B_i - A_i}$. В случае бинарного скрещивания

при двоичном представлении числа g_i случайным образом определяется положение раздела хромосомы k для скрещивания. После этого производится обмен выделенными частями хромосом. Если имеются два числа $g^{(1)} = \sum_{j=0}^L l_j^{(1)} 2^j$, $g^{(2)} = \sum_{j=0}^L l_j^{(2)} 2^j$,

то результат скрещивания $g = \sum_{j=0}^k l_j^{(1)} 2^j + \sum_{j=k+1}^L l_j^{(2)} 2^j$.

Результатом имитации данной операции в вещественном варианте является выражение

$$x_i = a_{Bin2}x_i^{(1)} + b_{Bin2}x_i^{(2)}.$$

Коэффициенты a_{Bin2} , b_{Bin2} определяются следующим образом:

$$b_{Bin2} = \frac{2^{\xi}}{2^L} = 2^{\xi-L} = 2^{-1} (2^{-u})^{L-1}, \quad a_{Bin2} = 1 - b_{Bin2},$$

где $\xi \in [0, L)$ – случайное число, соответствующее позиции скрещивания;

$u = 1 - \frac{\xi}{L-1}$ – случайное число и $u \in [0, 1]$.

В качестве оператора мутации применялась случайная мутация, при которой g_i , подлежащий изменению, принимает случайное значение из диапазона своего изменения $[A_i, B_i]$.

Кроме оператора мутации применяется оператор инверсии, который для вещественного кода имеет вид:

$$\bar{\mathbf{g}} = \begin{cases} (\bar{g}_i = g_{i-\xi}), i = \overline{1, N-\xi}; \\ (\bar{g}_i = g_{i-n+\xi}), i = \overline{N-\xi+1, N}, \end{cases} \quad \xi \in (1, N).$$

Перед началом процесса при $t = 0$ формируется популяция. Проводится эволюция популяции на итерации $t = t + 1$. Отбор особей для скрещивания осуществляется турнирным методом: случайным образом выбираются две особи и особь с лучшим качеством (с минимальным значением $F(\mathbf{x})$ или с максимальным $\bar{F}(\mathbf{x})$) привлекается для скрещивания при заданной вероятности скрещивания. Для решения задачи (6) применяется гибридный алгоритм, основанный на параллельной работе генетических операторов и какого-либо дополнительного градиентного или прямого метода [3]. В популяции, созданной генетическим алгоритмом, выбирается лучшая особь – лидер. Этот лидер обучается отдельно по дополнительному методу. Если его качественный показатель при этом лучше, чем у всех остальных особей в популяции, то он вводится в популяцию и участвует в воспроизводстве потомков. Если же в популяции появляется особь, полученная в результате эволюции, с лучшим показателем, то лидером становится она. В качестве дополнительного метода целесообразно также применить генетический алгоритм, но в более узкой области поиска. Границы узкой области поиска определяются следующим образом. Если $\mathbf{x} = (x'_i)$, $i = \overline{1, N}$ – лучшая в популяции особь, то границы области дополнительного поиска равны:

$$A'_i = x'_i - \Delta_i, B'_i = x'_i + \Delta_i; \Delta_i = \delta(B_i - A_i).$$

Проверка алгоритма оптимизации параметров нечеткой сети на основе деревьев решений проведена для двух наборов данных, полученных с использованием функций:

$$f_1(\mathbf{x}) = \sum_{i=1}^4 x_i^2 \sin(2x_i), \quad x_i \in [-3, 3], \quad L = 500 \text{ и}$$

$$f_2(x_1, x_2) = (1 - x_1^2 - x_2^2) \exp\left(-\frac{x_1^2 + x_2^2}{2}\right), \quad x_i \in [-3, 3], \quad L = 1089.$$

Объем обучающей выборки составил 80 %.

Результаты обучения нечеткой нейронной сети сравниваются с многослойной сигмоидной нейронной сетью.

Для первой функции получены следующие параметры нечеткой сети: $m = 5$; $K = 9$; $N^{\min} = 17$; $\beta = 5,8$. Количество правил равно 37. Среднеквадратичная ошибка составила $\varepsilon = 5,68\%$. Ошибка, полученная на обычной нейронной сети, равна 7,9 %.

Вторая функция является сложной для аппроксимации многослойной нейронной сетью сигмоидного типа. На рис. 1 показано сравнение фактических данных d с расчетными y после 2000 глобальных итераций.

Удовлетворительное согласование наступает после 10 000 итераций (рис. 2).

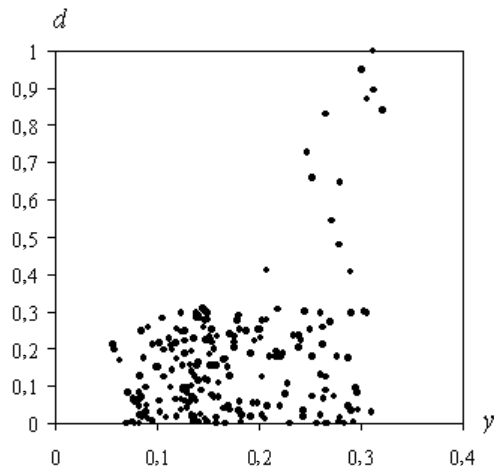


Рис. 1. Сравнение фактических данных с расчетными (2 000 итераций)

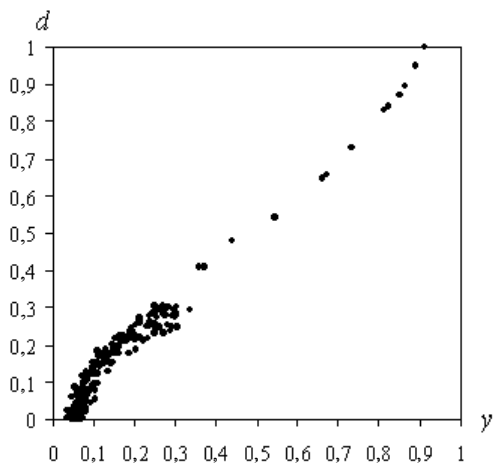


Рис. 2. Сравнение фактических данных с расчетными (10 000 итераций)

Дальнейший процесс обучения не повышает возможности аппроксимации обычной нейронной сети.

Для неоптимизированной нечеткой сети сравнение расчета с фактическими данными показано на рис. 3.

Оптимизированная нечеткая сеть для второй функции имеет следующие параметры: $m = 3$; $K = 9$; $N^{\min} = 12$; $\beta = 23/7$. Количество правил равно 76. Средне-квадратичная ошибка составила $\varepsilon = 0,57\%$. Результаты сравнения приведены на рис. 4.

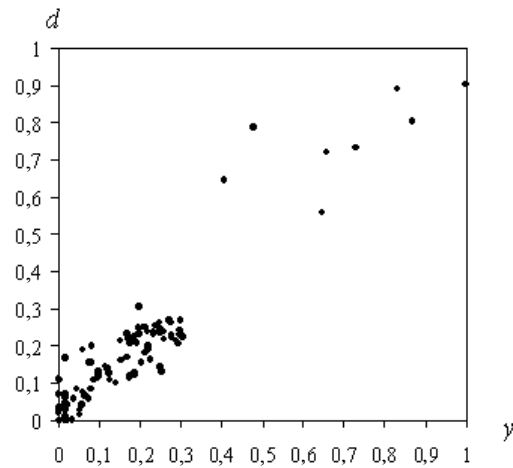


Рис. 3. Сравнение фактических данных с расчетными (неоптимизированная нечеткая нейронная сеть)

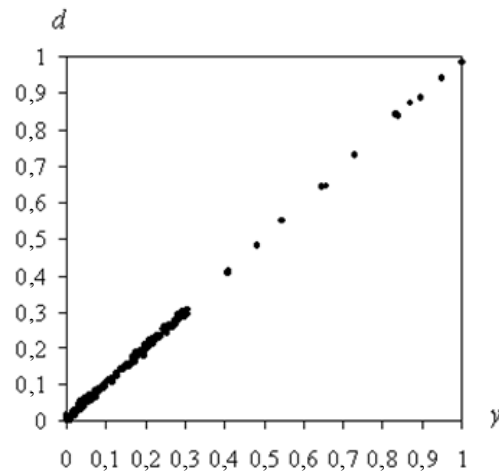


Рис. 4. Сравнение фактических данных с расчетными (оптимизированная нечеткая нейронная сеть)

Оптимизированная нечеткая сеть на основе дерева решений обеспечивает очень хорошую аппроксимацию тестовой функции. Коэффициент корреляции составляет 0,9995. Обычная нейронная сеть дает минимальную ошибку 3,5 %. Все сравнительные результаты получены на данных из тестовой выборки, не участвовавших в обучении сетей.

Вывод

Предложен алгоритм обучения нечеткой сети на основе деревьев решения, обеспечивающий высокие аппроксимирующие свойства.

Список литературы

1. Осовский С. Нейронные сети для обработки информации / пер. с пол. И. Д. Рудинского. – М. : Финансы и статистика, 2002. – 344 с.
2. Тенев В. А., Ворончак В. И. Решение задач классификации и аппроксимации с применением нечетких деревьев решений // Интеллектуал. системы в пр-ве. – 2005. – № 2. – С. 46–69.
3. Тенев В. А. Применение генетических алгоритмов с вещественным кроссовером для минимизации функций большой размерности // Интеллектуал. системы в пр-ве. – 2006. – № 1. – С. 93–107.

* * *

V. A. Tenenev, Doctor of Physical and Mathematical Sciences, Professor, Izhevsk State Technical University

A. V. Teneneva, Student, Izhevsk State Technical University

Teaching Fuzzy Neural Networks the Genetic Algorithm

The application of fuzzy neural networks to data processing based on decision tree rules is considered. The hybrid genetic algorithm with real coding is used for optimization of the network parameters. The advantage of the optimized fuzzy networks over usual multi-layer neural networks for an approximation problem solution is established.

Keywords: data processing, fuzzy neural networks, genetic algorithm, approximation

Получено 05.04.10

УДК 699.844

А. П. Тюрин, кандидат технических наук, доцент
Ижевский государственный технический университет

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ЗВУКОПОГЛОЩЕНИЯ ВАКУУМИРОВАННЫХ СОТОВЫХ КОНСТРУКЦИЙ*

Рассмотрен конечно-элементный подход оценки коэффициентов звукопоглощения вакуумированных и невакуумированных сотовых конструкций. Решение задачи осуществлялось в нахождении комплексных значений звукового давления в микрофонных позициях импедансной трубы с последующим вычислением коэффициентов звукопоглощения. Сравнение результатов численного и физического экспериментов показали удовлетворительное совпадение.

Ключевые слова: конечные элементы, звуковое давление, передаточная функция, коэффициент звукопоглощения

В работе рассматривается создаваемый аэродинамическим путем звук, который является продуктом воздушного потока и отделен от звука, производимого вибрациями твердых тел. Звуком будем называть малые колебания сжимаемого газа.

Уравнения распространения звука получаются из основных уравнений динамики сжимаемого газа, которые включают в себя:

– уравнение неразрывности

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v_i}{\partial x_i} = 0; \quad (1)$$

© Тюрин А. П., 2010

* В написании работы участвовал аспирант К. Ю. Замотин (кафедра «Математическое и программное обеспечение высокопроизводительных вычислений», Санкт-Петербургский государственный политехнический университет).