

УДК 004.272.26: 004.93

Олійник А. О.¹, Скрупський С. Ю.², Субботін С. О.³, Благодарьов А. Ю.⁴, Гофман Є. О.⁵

¹Канд. техн. наук, доцент кафедри програмних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

²Канд. техн. наук, доцент кафедри комп'ютерних систем та мереж, Запорізький національний технічний університет, Запоріжжя, Україна

³Д-р техн. наук, завідувач кафедри програмних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

⁴Аспірант кафедри програмних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

⁵Канд. техн. наук, с.н.с. кафедри програмних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

ПЛАНУВАННЯ РЕСУРСІВ ПАРАЛЕЛЬНОЇ ОБЧИСЛЮВАЛЬНОЇ СИСТЕМИ ПРИ СИНТЕЗІ НЕЙРО-НЕЧІТКИХ МОДЕЛЕЙ ДЛЯ ОБРОБКИ ВЕЛИКИХ ДАНИХ

Вирішено задачу планування ресурсів паралельних комп'ютерних систем при синтезі нейро-нечітких мереж. Об'єкт дослідження – процес синтезу нейро-нечітких моделей. Предметом дослідження є методи планування ресурсів паралельних комп'ютерних систем. Мета роботи полягає в побудові моделі планування ресурсів паралельних комп'ютерних систем, що здійснюють вирішення прикладних завдань на основі паралельного методу синтезу нейро-нечітких мереж. Запропоновано модель планування ресурсів паралельних комп'ютерних систем при синтезі нейро-нечітких мереж. Синтезована модель враховує тип комп'ютерної системи, кількість процесів, на яких виконується завдання, пропускну здатність мережі передачі даних, параметри використовуваного математичного забезпечення (кількість можливих рішень, що обробляються в процесі роботи методу, частки рішень, що генеруються на кожній ітерації стохастичного пошуку за допомогою застосування операторів схрещування і мутації), а також параметри розв'язуваної прикладної задачі (кількість спостережень і кількість ознак в заданій множині даних, що описує результати спостережень за досліджуванним об'єктом або процесом). Розроблено програмне забезпечення, що реалізує синтезовану модель планування ресурсів. Виконано експерименти, що підтверджують адекватність запропонованої моделі. Результати експериментів дозволяють рекомендувати застосування розробленої моделі на практиці.

Ключові слова: вибірки даних, паралельні обчислення, планування ресурсів, нейро-нечітка модель, нейронна мережа.

НОМЕНКЛАТУРА

CPU – Central Processing Unit;

GPU – Graphical Processing Unit;

M – кількість ознак в множині спостережень S ;

N_{pr} – кількість процесів, на яких виконується завдання;

N_{χ} – кількість можливих рішень, з якими оперує система при синтезі нейро-нечітких мереж;

Q – кількість спостережень в заданій множині спостережень S ;

S – навчальна вибірка;

T_{NN} – значення виходу нейромережевої моделі;

V – пропускну здатність середовища передачі даних,

Гб / с;

w – матриця вагових коефіцієнтів;

w_0 – порогове значення функції $\Phi(w;x)$;

$|x|$ – кількість аргументів функції $\Phi(w;x)$;

β – частка рішень, що генеруються на кожній ітерації стохастичного пошуку за допомогою застосування оператора схрещування;

γ – частка рішень, що генеруються на кожній ітерації стохастичного пошуку за допомогою застосування оператора мутації;

$\Psi(\mu, \rho)$ – функція активації ρ -го нейрона μ -го шару;

$\Phi(\mu, \rho)$ – дискримінантна функція ρ -го нейрона μ -го шару.

ВСТУП

Нейро-нечіткі мережі (ННМ) широко застосовують-ся при вирішенні практичних задач розпізнавання, діаг-

ностування, контролю якості продукції та ін. [1–7]. У роботах [8–10] запропоновано метод синтезу ННМ на основі паралельних обчислень, що дозволяє вирішувати прикладні задачі, які передбачають необхідність побудови моделей на основі різних вибірок даних [8]. Виконання цього методу вимагає великих витрат часу та обчислювальних ресурсів, що обумовлює необхідність застосування паралельних комп'ютерних систем [11–14]. Використання ресурсів для паралельних обчислень на кластерах CPU і графічних процесорах (GPU) є досить дорогим і не доступним для більшості користувачів [11, 12]. Для ефективного застосування паралельних комп'ютерних систем виконують попереднє планування ресурсів, що може бути здійснено за допомогою методів математичного моделювання поведінки системи при вирішенні відповідних задач [14–17]. При цьому доцільно застосовувати методи нейромережевого моделювання [18–26], оскільки вони дозволяють виявляти нелінійні залежності між характеристиками паралельної комп'ютерної системи, параметрами використовуваного математичного забезпечення, аргументами розв'язуваної в системі прикладної задачі, та часом, затраченим системою на вирішення прикладної задачі.

Метою роботи є побудова моделі планування ресурсів паралельних комп'ютерних систем, що здійснюють вирішення прикладних задач на основі паралельного методу синтезу нейро-нечітких мереж [9].

1 ПОСТАНОВА ЗАВДАННЯ

При плануванні ресурсів паралельної системи важливими характеристиками, від яких залежить її ефек-

тивність і швидкість одержання результату, є наступні групи параметрів: технічні характеристики паралельної системи, параметри використовуваного програмного (математичного) забезпечення, особливості розв'язуваної прикладної задачі [8–10].

Основними характеристиками паралельної системи, що впливають на час вирішення практичної задачі, є:

- x_1 – тип системи *type* (кластер CPU або GPU);
- x_2 – кількість процесів, на яких виконується задача

N_{pr} ;

- x_3 – пропускна здатність середовища передачі даних V , Гб/с.

Серед основних параметрів використовуваного математичного забезпечення (у даному випадку, методу синтезу ННМ [8]) виділимо такі:

- x_4 – кількість можливих рішень N_χ [8], з якими оперує система на кожній ітерації методу;
- x_5 – частка рішень β , що генеруються на кожній ітерації стохастичного пошуку за допомогою застосування оператора схрещування [8–10];
- x_6 – частка рішень γ , що генеруються на кожній ітерації стохастичного пошуку за допомогою застосування оператора мутації [8–10].

Як параметри прикладної задачі, що істотно впливають на швидкість роботи паралельної системи при синтезі ННМ, доцільно використовувати:

- x_7 – кількість спостережень Q у заданій множині спостережень S ;
- x_8 – кількість ознак M у множині спостережень S .

Таким чином, для оцінювання часу роботи t паралельної комп'ютерної системи при синтезі ННМ необхідно побудувати модель виду (1):

$$t = t(\textit{type}, N_{pr}, V, N_\chi, \beta, \gamma, Q, M), \quad (1)$$

яка дозволяє виконувати прогнозування витраченого на виконання паралельного методу синтезу ННМ часу в залежності від характеристик системи, параметрів математичного забезпечення та особливостей розв'язуваної прикладної задачі.

2 ОГЛЯД ЛІТЕРАТУРИ

У роботах [27, 28] запропоновано моделі планування ресурсів паралельної обчислювальної системи, що дозволяють оцінити час її роботи при синтезі нейро-нечітких мереж на основі методу [8].

У синтезованій в роботі [27] моделі враховувалися лише 4 параметри (тип системи, кількість процесів, на яких виконується задача, пропускна здатність мережі, а також кількість можливих рішень N_χ). Крім того, важливо відзначити, що модель [27] була побудована за даними лише однієї навчальної вибірки (при вирішенні задачі побудови ННМ для індивідуального прогнозування стану хворого гіпертонічною хворобою [10]) і, відповідно, могла бути використана лише при вирішенні подібного класу задач.

У роботі [28] додатково враховувалися параметри прикладної задачі, що істотно впливають на швидкість роботи паралельної системи при синтезі ННМ (кількість спостережень і число ознак у заданій множині спостережень S). Однак така модель також характеризувалася деякими недоліками:

- не були враховані деякі важливі параметри паралельного методу синтезу ННМ [8], що істотно впливають на швидкість пошуку оптимального рішення (частки рішень, що генеруються на кожній ітерації стохастичного пошуку за допомогою застосування операторів схрещування і мутації);
- синтезована модель містила всі можливі прямонаправлені зв'язки між нейронами.

Наявність зазначених недоліків моделей планування ресурсів, синтезованих у роботах [27, 28], обумовлює необхідність побудови моделі, яка є досить простою для прийняття і враховує основні технічні характеристики паралельної системи, параметри використовуваного програмного (математичного) забезпечення та особливості розв'язуваної прикладної задачі.

3 МАТЕРІАЛИ ТА МЕТОДИ

Розглянутий метод [8–10] був застосований на кластері CPU та на GPU при обробці даних з публічного репозиторію [29], а також при вирішенні задачі індивідуального прогнозування стану здоров'я хворих артеріальною гіпертонією [10]. Характеристики оброблюваних наборів даних наведено в табл. 1.

У результаті обробки даних [10, 29] було сформовано навчальну вибірку (2), що містить 2064 результати виконання методу, кожний з яких характеризувався вісьмома ознаками:

$$D = \langle X, T \rangle, \quad (2)$$

где $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $x_i = \{x_{i1}, x_{i2}, \dots, x_{iN}\}$, $N = 2064$, $T = \{t_1, t_2, \dots, t_N\}$.

Таблиця 1 – Характеристики оброблюваних наборів даних

№ з/п	Назва задачі	Тип даних	Кількість спостережень Q	Кількість ознак M
1	Hypertensive patient condition [10]	цїлі, дійсні, якісні	2418	24
2	Auto MPG	дійсні, якісні	398	8
3	Automobile	цїлі, дійсні, якісні	205	26
4	Computer Hardware	цїлі	209	9
5	Housing	цїлі, дійсні, якісні	506	14
6	Servo	цїлі, якісні	167	4
7	Solar Flare	якісні	1389	10
8	Forest Fires	дійсні	517	13
9	Concrete Compressive Strength	дійсні	1030	9
10	Communities and Crime	дійсні	1994	128
11	Parkinsons Telemonitoring	цїлі, дійсні	5875	26
12	Energy efficiency	цїлі, дійсні	768	8

Таким чином навчальна вибірка являла собою таблицю чисел, що складається з 2064 рядків і дев'яти стовпців, які містять значення восьми вхідних ознак і одного вихідного (часу роботи методу) для кожного випадку застосування розглянутого методу в паралельній системі. Фрагмент навчальної вибірки наведено у табл. 2.

Для виключення впливу різного порядку значень ознак на синтезовану модель виконувалося нормування ознак, тобто приведення діапазону їхніх значень до єдиного інтервалу [0; 1] за формулою (3):

$$x_{ijn} = \frac{x_{ij} - x_{i \min}}{x_{i \max} - x_{i \min}}, \quad (2)$$

де $i = 1, 2, \dots, 8$; $j = 1, 2, \dots, 2064$.

Як базис для побудови моделі залежності виду (1) використовувалися нейронні мережі прямого поширення, що дозволяють апроксимувати складні нелінійні залеж-

ності з високою точністю. Модель (1) синтезувалася у виді тришарового перцептрону [18, 22], перший шар якого містив три нейрони, другий шар – чотири нейрони, третій шар – один нейрон. Усі нейрони мали сигмоїдну функцію активації (4):

$$\psi(\varphi) = \frac{1}{1 + e^{-\varphi}}, \quad (3)$$

де $\varphi = \Phi(w; x)$.

У якості дискримінантної функції нейронів використовувалася зважена сума (5):

$$\Phi(w; x) = w_0 + \sum_{i=1}^{|x|} w_i x_i, \quad (4)$$

де w_i визначає вагу i -го вхідного параметру x_i у функції $\Phi(w; x)$.

Таблиця 2 – Фрагмент навчальної вибірки

type	Значення ознак							T
	N_{pr}	V	N_{χ}	β	γ	Q	M	
0	1	20	50	0,85	0,1	2418	24	101,18
0	8	20	50	0,83	0,11	2418	24	18,09
0	32	1	100	0,58	0,3	2418	24	51,65
1	100	32	50	0,58	0,31	2418	24	30,33
1	180	64	100	0,51	0,43	2418	24	59,44
1	260	32	100	0,53	0,41	2418	24	67,54
0	16	1	50	0,74	0,15	209	9	14,90
0	2	20	100	0,7	0,18	209	9	106,42
1	60	32	50	0,58	0,31	209	9	33,83
1	100	64	100	0,52	0,42	209	9	48,41
0	4	1	50	0,77	0,14	506	14	47,59
0	18	20	100	0,67	0,2	506	14	18,87
1	200	32	50	0,57	0,33	506	14	28,52
1	240	64	100	0,51	0,44	506	14	56,43
0	8	1	50	0,76	0,14	1389	10	29,93
0	14	20	100	0,67	0,2	1389	10	24,79
1	140	32	50	0,57	0,32	1389	10	27,47
1	100	32	100	0,54	0,38	1389	10	61,56
0	16	20	50	0,81	0,11	1994	128	10,42
0	1	1	50	0,77	0,13	1994	128	176,78
1	80	32	50	0,58	0,31	1994	128	41,39
1	100	64	50	0,56	0,34	1994	128	28,07
0	6	20	50	0,84	0,1	5875	26	27,11
0	22	1	50	0,72	0,16	5875	26	19,98
0	32	20	100	0,64	0,23	5875	26	23,75
1	120	32	50	0,57	0,32	5875	26	33,78
1	180	32	100	0,53	0,39	5875	26	74,78
...
1	240	64	100	0,51	0,44	5875	26	71,77

Таким чином, структура синтезованої тришарової нейромоделі T_{NN} може бути представлена в такий спосіб (6):

$$\begin{cases} T_{NN} = \Psi(3,1)(\Phi(3,1)(w(3,1); \Psi(2))); \\ \Psi(2) = \{\Psi(2,1), \Psi(2,2), \Psi(2,3), \Psi(2,4)\}; \\ \Psi(2,k) = \Psi(2,k)(\Phi(2,k)(w(2,k); \Psi(1))), k = 1, 2, 3, 4; \\ \Psi(1) = \{\Psi(1,1); \Psi(1,2); \Psi(1,3)\}; \\ \Psi(1,l) = \Psi(1,l)(\Phi(1,l)(w(1,l); X)), l = 1, 2, 3. \end{cases} \quad (5)$$

Для побудови нейромоделі та визначення значень її параметрів (вагових коефіцієнтів і зсувів кожного нейрону) на її входи подавалися значення пронормованих ознак, на вихід – значення часу виконання методу синтезу нейро-нечітких мереж у паралельній системі. Як цільова функція при навчанні нейромоделі використовувався мінімум середньоквадратичної помилки E .

З метою виключення з моделі $t = t(type, N_{pr}, V, N_{\chi}, \beta, \gamma, Q, M)$ надлишкових зв'язків між нейроелементами, навчання проводилося за допомогою методу структурно-параметричного синтезу нейромоделі на основі еволюційного підходу [8, 9]. Прийнятним вважалося досягнення середньоквадратичної помилки порядку 10^{-4} .

Після підстановки значень вагових коефіцієнтів і зсувів у (6) з урахуванням функції активації (4) та дискримінантної функції (5) одержуємо математичний опис синтезованої нейромережевої моделі (7), що описує залежності між характеристиками системи, у якій здійснюється синтез нейро-нечітких мереж, параметрами досліджуваного методу та часом, витраченим системою на виконання синтезу мереж. Графічну інтерпретацію синтезованої нейромоделі наведено на рис. 1. Значення середньоквадратичної помилки синтезованої моделі склали $3,3 \times 10^{-4}$, що є прийнятним для подібного роду задач.

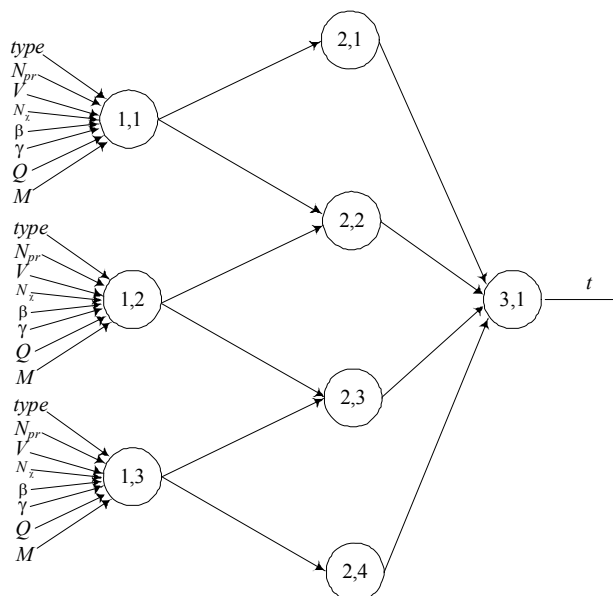


Рисунок 1 – Синтезована нейромережева модель

Таким чином, побудована нейромережева модель являє собою ієрархічну структуру, що містить нейро-подібні обчислювальні елементи, і дозволяє оцінювати час, необхідний паралельній комп'ютерній системі для синтезу ННМ при моделюванні складних об'єктів і процесів.

4 ЕКСПЕРИМЕНТИ

Для виконання експериментального дослідження запропонованої нейромережевої моделі використані такі комп'ютерні системи:

- кластер Інституту проблем моделювання в енергетиці імені Г. Е. Пухова НАН України (ІПМЕ) м. Київ: процесори Intel Xeon 5405, оперативна пам'ять – 4×2 ГБ DDR-2 на кожен вузол, комунікаційне середовище InfiniBand 20Гб/з, middleware Torque і OMPI;
- кластер Запорізького національного технічного університету (ЗНТУ) м.Запорозжя: процесори Intel E3200,

$$\begin{cases} Y_{NN} = \Psi(3,1) = \left(1 + e^{-(43,31 - 52,55\Psi(2,1) - 76,96\Psi(2,2) - 5,58\Psi(2,3) + 34,87\Psi(2,4))}\right)^{-1}; \\ \Psi(2,1) = \left(1 + e^{-(28,84 - 0,58\Psi(1,1))}\right)^{-1}; \\ \Psi(2,2) = \left(1 + e^{-(29,93 - 92,2\Psi(1,1) - 95,32\Psi(1,2))}\right)^{-1}; \\ \Psi(2,3) = \left(1 + e^{-(8,75 - 19,75\Psi(1,2) + 5,95\Psi(1,3))}\right)^{-1}; \\ \Psi(2,4) = \left(1 + e^{-(6,44 + 9,67\Psi(1,3))}\right)^{-1}; \\ \Psi(1,1) = \left(1 + e^{-(3,71 + 7,29type - 10,31N_{pr} - 6,46V + 3,97N_{\chi} + 4,3\beta - 8,59\gamma - 0,06Q - 0,01M)}\right)^{-1}; \\ \Psi(1,2) = \left(1 + e^{-(0,41 + 1,02type - 0,84N_{pr} - 0,02V + 0,11N_{\chi} + 1,29\beta + 0,83\gamma + 0,04Q + 0,01M)}\right)^{-1}; \\ \Psi(1,3) = \left(1 + e^{-(0,6 + 4,51type - 3,7N_{pr} + 0,89V + 0,48N_{\chi} + 5,09\beta - 0,2\gamma - 0,08Q - 0,02M)}\right)^{-1}. \end{cases} \quad (6)$$

оперативна пам'ять 1 Гб DDR-2 на кожен вузол, комунікаційне середовище Gigabit Ethernet 1 Гб/с, middleware Torque і MPICH;

- GPU NVIDIA GTX 285+ 240 ядер CUDA;
- GPU NVIDIA GTX 960 1024 ядра CUDA.

В експериментах кількість процесів x_2 , на яких виконувався метод, варіювалося від 1 до 32 для кластерів та від 60 до 260 для GPU. Пропускна здатність мережі x_3 – від 1 до 20 Гб/с, кількість можливих рішень на етапі ініціалізації методу N_χ – від 50 до 100. Частка рішень β , що генеруються на кожній ітерації стохастичного пошуку за

допомогою застосування оператора схрещування, x_5 – від 0,51 до 0,85, частка рішень γ , створюваних на кожній ітерації стохастичного пошуку за допомогою застосування оператора мутації, x_6 – від 0,1 до 0,45. Для проведення експериментів було розроблено програмне забезпечення мовою С з застосуванням бібліотеки MPI [30].

5 РЕЗУЛЬТАТИ

Результати експериментів на кластерах ЗНТУ (задача № 4 у табл. 1, $N_\chi=50$) та ІПМЕ (задача № 5 у табл. 1, $N_\chi=100$) наведено на рис. 2 і рис. 3, відповідно. Рис. 4 демонструє

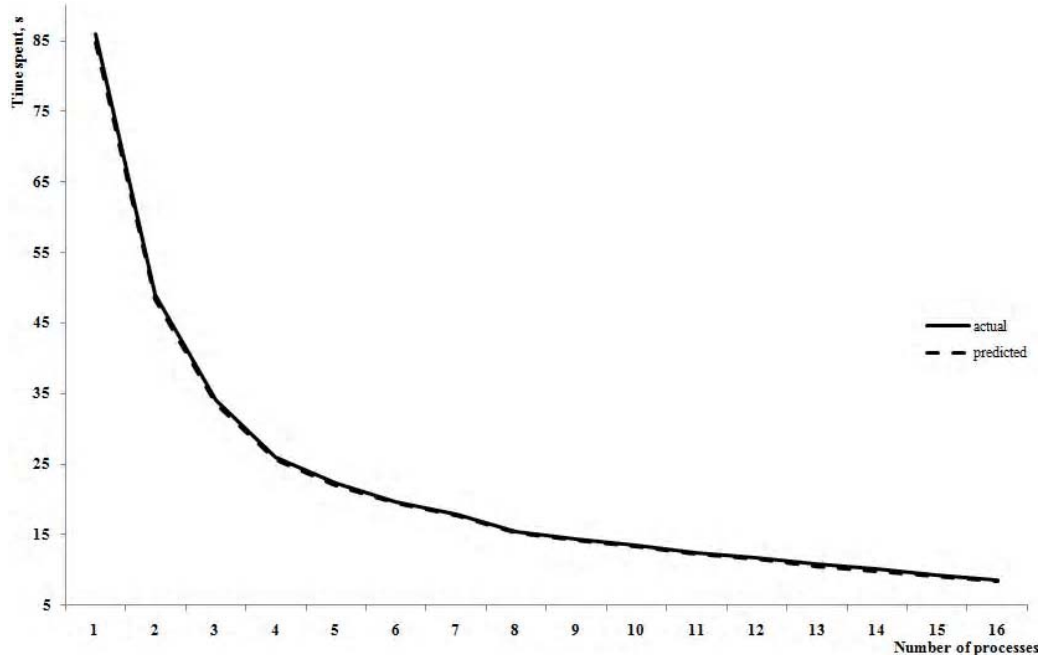


Рисунок 2 – Результати експериментів на кластері ІПМЕ (задача № 4 у табл. 1, $N_\chi=100$)

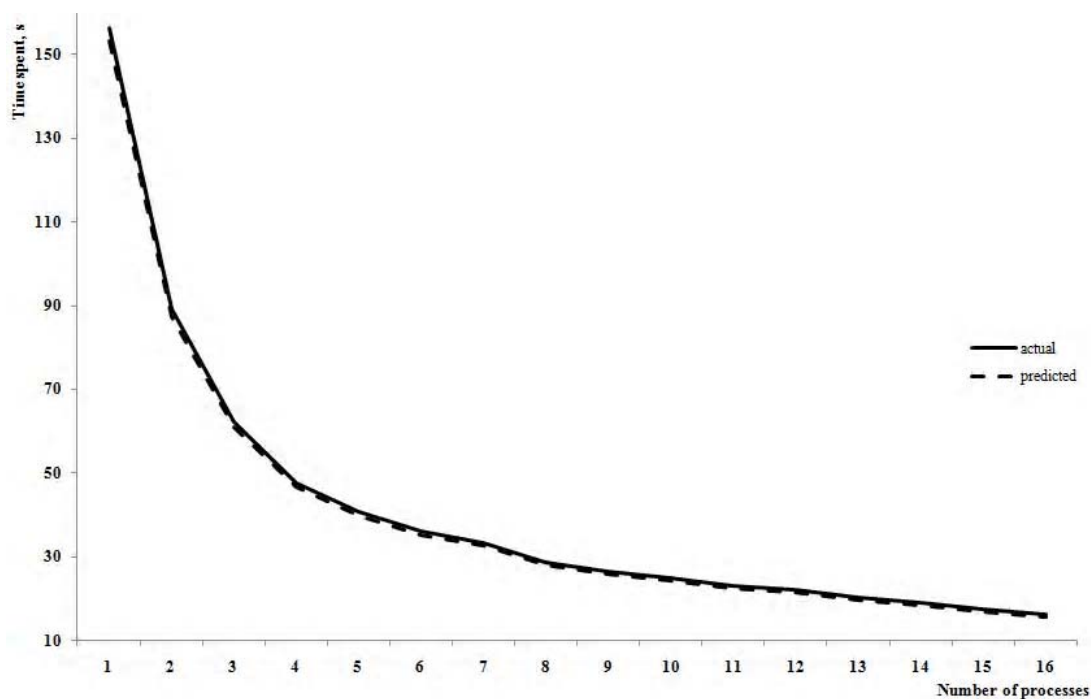


Рисунок 3 – Результати експериментів на кластері ЗНТУ (задача № 5 у табл. 1, $N_\chi=50$)

експериментальну перевірку запропонованої моделі на GPU NVIDIA GTX 960 у задачі № 11 з табл. 1, при цьому $N_{\chi}=100$. Рис. 5 демонструє результати експериментів на GPU NVIDIA GTX 285+ у вирішенні задачі № 10 з $N_{\chi}=50$. Суцільною лінією зображено час, фактично витрачений

системою на виконання паралельного методу синтезу нейро-нечітких мереж [8], а пунктирною лінією – розрахунковий час за допомогою запропонованої моделі.

У таблиці 3 наведено значення середньоквадратичної помилки (mse) запропонованої моделі в порівнянні з моделями [27] і [28].

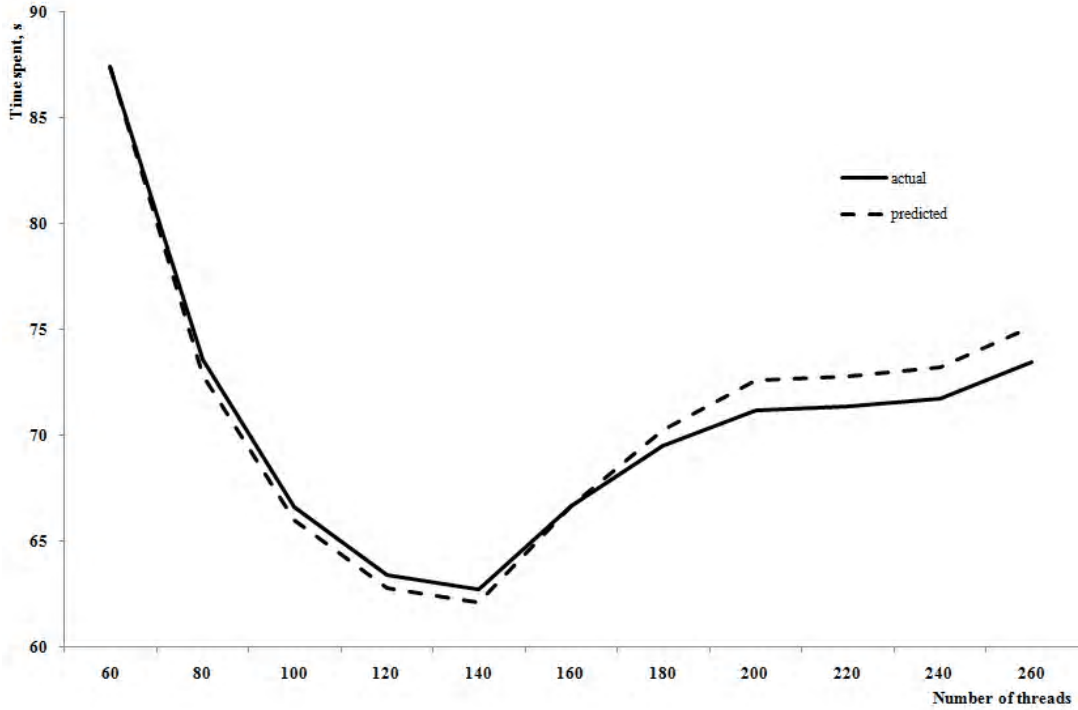


Рисунок 4 – Результати експериментів на GPU NVIDIA GTX 960 (задача № 11 у табл.1, $N_{\chi}=100$)

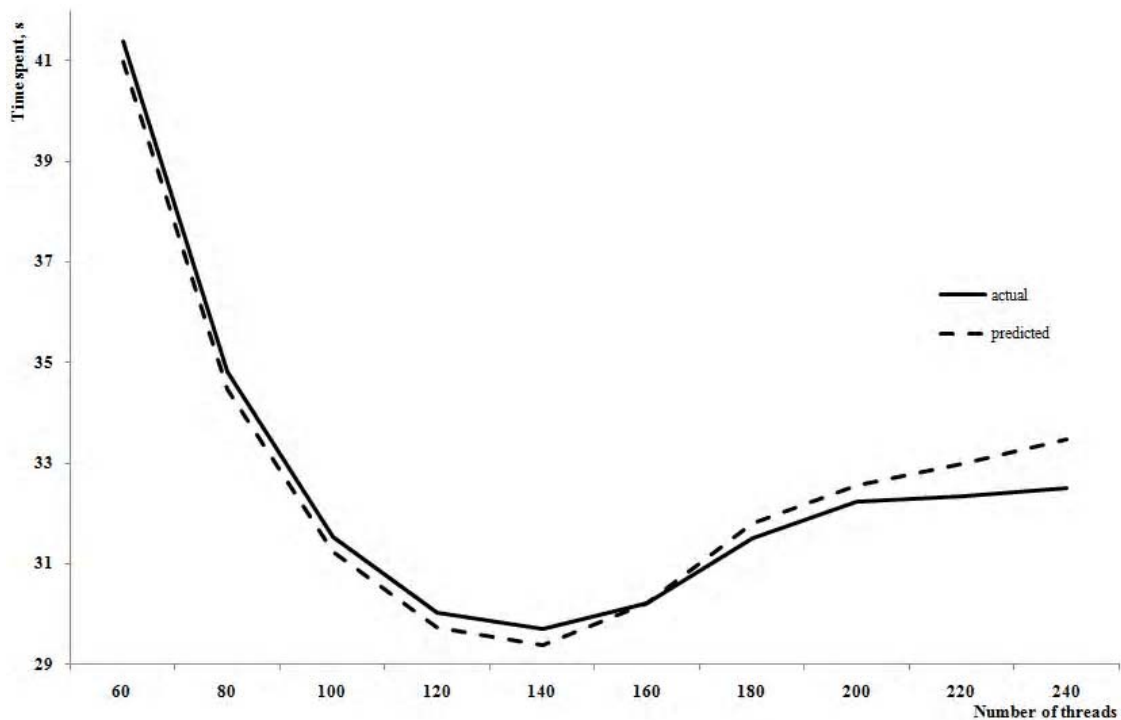


Рисунок 5 – Результати експериментів на GPU NVIDIA GTX 285+ (задача № 10 у табл. 1, $N_{\chi}=50$)

Таблиця 3 – Значення середньоквадратичних помилок розглянутих моделей

Задача	Модель	MSE запропонованої моделі (8 параметрів)	MSE моделі [27] (6 параметрів)	MSE моделі [28] (4 параметри)
Computer Hardware (на кластері)		$2,05 \times 10^{-6}$	$1,03 \times 10^{-5}$	$2,74 \times 10^{-5}$
Housing (на кластері)		$8,29 \times 10^{-7}$	$6,61 \times 10^{-5}$	$9,32 \times 10^{-5}$
Communities and Crime (на GPU)		$3,32 \times 10^{-5}$	$1,67 \times 10^{-5}$	$2,61 \times 10^{-5}$
Parkinsons Telemonitoring (на GPU)		$2,08 \times 10^{-4}$	$2,17 \times 10^{-5}$	$3,47 \times 10^{-4}$

6 ОБГОВОРЕННЯ

Тестова вибірка, що складається з результатів 69 експериментів, включала екземпляри рішень практичних задач у паралельній системі, що не входять у навчальну вибірку. При проведенні експериментів адекватність складності розв'язуваних задач відповідала продуктивності використовуваних паралельних систем.

З рис. 2 і 3 видно, що час рішення задачі на кластері, розрахований за допомогою запропонованої моделі, як правило, є трохи меншим у порівнянні з фактично витраченим часом. Це можна пояснити тим, що час, витрачений на синхронізацію та на пересилання даних між процесами кластера, значно варіюється в залежності від застосованого середовища передачі даних. При цьому, чим більше задіяно процесів кластера, тим істотнішим є вплив синхронізацій і пересилань та тим більшим є відхилення між фактичним і прогнозованим часом вирішення задачі.

Рисунки 4 і 5 дозволяють судити про те, що на GPU з ростом кількості задіяних потоків до 140 включно спостерігається зниження витраченого на виконання методу часу. Прогнозований за допомогою запропонованої моделі час поводить ся аналогічним чином. При подальшому збільшенні кількості задіяних потоків GPU вплив накладних витрат істотно зростає, частка пересилань і синхронізацій починає перевищувати обсяг цільових обчислень, тому фактично витрачений на виконання методу час починає зростати. Прогнозований за допомогою моделі час також зростає, але дещо швидшими темпами. Таким чином, запропонована модель враховує вплив накладних витрат обчислювального процесу в паралельній системі.

Таблиця 3 дозволяє порівняти середньоквадратичну помилку (основний критерій перевірки адекватності моделі) запропонованої моделі з моделями [27] і [28]. Як видно з таблиці 3 при використанні кластера збільшення параметрів у моделі (1) приводить до зменшення середньоквадратичної помилки. Отже, більш доцільним є застосування на практиці запропонованої моделі в порівнянні з моделями [27] і [28]. На графічних процесорах не спостерігається значного поліпшення значення критерію mse у порівнянні з розглянутими моделями [27] і [28], що обумовлено більш істотним впливом на час вирішення задачі накладних витрат (пересилань і синхронізацій) у порівнянні з кластерами CPU.

ВИСНОВКИ

У роботі вирішено завдання планування ресурсів паралельних комп'ютерних систем при синтезі нейро-нечітких мереж.

Наукова новизна полягає в тому, що запропоновано модель планування ресурсів паралельних комп'ютерних

систем при синтезі ННМ, що враховує тип комп'ютерної системи, кількість процесів, на яких виконується задача, пропускну здатність мережі передачі даних, параметри використовуваного математичного забезпечення (кількість можливих рішень, оброблюваних у процесі роботи методу, частки рішень, що генеруються на кожній ітерації стохастичного пошуку за допомогою застосування операторів схрещування та мутації), а також параметри розв'язуваної прикладної задачі (кількість спостережень і кількість ознак у заданій множині даних, що описує результати спостережень за досліджуванним об'єктом чи процесом). Синтезована модель дозволяє виконувати оцінювання часу, необхідного паралельній системі для виконання методу синтезу нейро-нечітких мереж.

Практична цінність отриманих результатів полягає в розробленому програмному забезпеченні, що реалізує запропоновану модель і дозволяє раціонально планувати вибір ресурсів комп'ютерної системи для побудови ННМ при вирішенні практичних задач розпізнавання образів, діагностування та прогнозування.

ПОДЯКИ

Роботу виконано в рамках науково-дослідної теми «Методи і засоби обчислювального інтелекту та паралельного комп'ютерного для оброблення великих даних в системах діагностування» (№ державної реєстрації 0116U007419) кафедри програмних засобів Запорізького національного технічного університету.

СПИСОК ЛІТЕРАТУРИ

1. Nauck D. Foundations of neuro-fuzzy systems / D. Nauck, F. Klawonn, R. Kruse. – Chichester : John Wiley & Sons, 1997. – 305 p.
2. Бодяньський С. В. Еволюційна каскадна система на основі нейро-фаззи вузлів / С. В. Бодяньський, О. К. Тищенко, О. О. Бойко // Радіоелектроніка, інформатика, управління. – 2016. – № 2. – С. 40–45.
3. Субботин С. А. Метод синтеза диагностических моделей на основе радиально-базисных нейронных сетей с поддержкой обобщающих свойств / С. А. Субботин // Радіоелектроніка, інформатика, управління. – 2016. – № 2. – С. 64–69.
4. Oliinyk A. O. Synthesis of Neuro-Fuzzy Networks on the Basis of Association Rules / A. O. Oliinyk, T. A. Zayko, S. A. Subbotin // Cybernetics and Systems Analysis. – 2014. – Vol. 50, Issue 3. – P. 348–357. DOI: 10.1007/s10559-014-9623-7.
5. Oliinyk A. Training Sample Reduction Based on Association Rules for Neuro-Fuzzy Networks Synthesis / A. Oliinyk, T. Zaiko, S. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2014. – Vol. 23, № 2. – P. 89–95. DOI: 10.3103/S1060992X14020039.
6. Oliinyk A. The decision tree construction based on a stochastic search for the neuro-fuzzy network synthesis / A. Oliinyk, S. A. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2015. – Vol. 24, № 1. – P. 18–27. DOI: 10.3103/S1060992X15010038.

7. Олійник А. О. Видобування продукційних правил на основі негативного відбору / А. О. Олійник // *Радіоелектроніка, інформатика, управління*. – 2016. – № 1. – С. 40–49.
8. Oliinyk A. O. Using Parallel Random Search to Train Fuzzy Neural Networks / A. O. Oliinyk, S. Yu. Skrupsky, S. A. Subbotin // *Automatic Control and Computer Sciences*. – 2014. – Vol. 48, Issue 6. – P. 313–323. DOI: 10.3103/S0146411614060078.
9. Oliinyk A. O. Experimental Investigation with Analyzing the Training Method Complexity of Neuro-Fuzzy Networks Based on Parallel Random Search / A. O. Oliinyk, S. Yu. Skrupsky, S. A. Subbotin // *Automatic Control and Computer Sciences*. – 2015. – Vol. 49, Issue 1. – P. 11–20. DOI: 10.3103/S0146411615010071.
10. Subbotin S. Individual prediction of the hypertensive patient condition based on computational intelligence / S. Subbotin, A. Oliinyk, S. Skrupsky // *Information and Digital Technologies : International Conference IDT'2015, Zilina, 7–9 July 2015 : proceedings of the conference*. – Zilina : Institute of Electrical and Electronics Engineers, 2015. – P. 336–344. DOI: 10.1109/DT.2015.7222996.
11. Sulistio A. Simulation of Parallel and Distributed Systems: A Taxonomy – and Survey of Tools / A. Sulistio, C.S. Yeo. R. Buyya // *International Journal of Software Practice and Experience*. Wiley Press. – 2002. – P. 1–19.
12. Методы и модели планирования ресурсов в GRID-системах : монография / [В. С. Пономаренко, С. В. Листровой, С. В. Минухин, С.В. Знахур]. – Х. : ИД «ИНЖЭК», 2008. – 408 с.
13. Introduction to GPUs. – Режим доступа: URL: <https://www.cs.utexas.edu/~pingali/CS378/2015sp/lectures/IntroGPUs.pdf>. – Загл. з екрану.
14. Gebali F. Algorithms and Parallel Computing / F. Gebali. – New Jersey : John Wiley & Sons, 2011. – 364 p. DOI: 10.1002/9780470932025.
15. Fokkink W. Distributed Algorithms: An Intuitive Approach / Wan Fokkink. – Cambridge : MIT Press, 2013. – 248 p.
16. Herlihy M. The Art of Multiprocessor Programming Revised Reprint / M. Herlihy, N. Shavit. – Boston : Morgan Kaufmann, 2012. – 536 p.
17. Roosta S. H. Parallel Processing and Parallel Algorithms: Theory and Computation / S. H. Roosta. – New York : Springer-Verlag, NY, 2000. – 566 p. DOI: 10.1007/978-1-4612-1220-1.
18. Хайкин С. Нейронные сети : полный курс / С. Хайкин. – СПб : Вильямс, 2005. – 1104 с.
19. Suzuki K. Artificial Neural Networks: Architectures and Applications / K. Suzuki. – New York : InTech, 2013. – 264 p. DOI: 10.5772/3409.
20. Hanrahan G. Artificial Neural Networks in Biological and Environmental Analysis / G. Hanrahan. – Boca Raton, Florida : CRC Press, 2011. – 214 p. DOI: 10.1201/b10515.
21. Bodyanskiy Ye. Hybrid adaptive wavelet-neuro-fuzzy system for chaotic time series identification / Ye. Bodyanskiy, O. Vynokurova // *Information Sciences*. – 2013. – Vol. 220. – P. 170–179. DOI: 10.1016/j.ins.2012.07.044.
22. Bishop C. Neural Networks for pattern recognition / C. Bishop. – New York : Oxford University Press, 1995. – 482 p.
23. Computational intelligence: collaboration, fusion and emergence / [ed. Ch. L. Mumford]. – New York : Springer, 2009. – 752 p. DOI: 10.1007/978-3-642-01799-5.
24. Bodyanskiy Ye. An evolving radial basis neural network with adaptive learning of its parameters and architecture / Ye. Bodyanskiy, A. K. Tyshchenko, A. Deineko // *Automatic Control and Computer Sciences*. – 2015. – Vol. 49, Issue 5. – P. 255–260. DOI: 10.3103/S0146411615050028.
25. Intelligent data analysis : an introduction / [eds. M. Berthold, D. J. Hand]. – New York : Springer Verlag, 2007. – 525 p. DOI: 10.1007/978-3-540-48625-1.
26. Tenne Y. Computational Intelligence in Expensive Optimization Problems / Y. Tenne, C.-K. Goh. – Berlin : Springer, 2010. – 800 p. DOI: 10.1007/978-3-642-10701-6.
27. Скрупский С. Ю. Экспериментальное исследование метода синтеза нейро-нечетких моделей в параллельной компьютерной системе / С. Ю. Скрупский // *Радиоелектроніка, інформатика, управління*. – 2016. – № 2. – С. 56–63.
28. Oliinyk A. Parallel Computer System Resource Planning for Synthesis of Neuro-Fuzzy Networks / A. Oliinyk, S. Skrupsky, S. Subbotin // *Recent Advances in Systems, Control and Information Technology. Advances in Intelligent Systems and Computing – 2017*. – Vol. 543. – P. 124–133. DOI: 10.1007/978-3-319-48923-0_12
29. UCI Machine Learning Repository. <http://archive.ics.uci.edu/ml/>.
30. Quinn M.J. Parallel Programming in C with MPI and OpenMP / M. J. Quinn. – New York, NY : McGraw-Hill, 2004. – 529 p.

Стаття надійшла до редакції 05.10.2016.
Після доробки 20.11.2016.

Олейник А. А.¹, Скрупский С. Ю.², Субботин С. А.³, Благодарев А. Ю.⁴, Гофман Е. А.⁵

¹Канд. техн. наук, доцент кафедры программных средств, Запорожский национальный технический университет, Запорожье, Украина

²Канд. техн. наук, доцент кафедры компьютерных систем и сетей, Запорожский национальный технический университет, Запорожье, Украина

³Д-р техн. наук, заведующий кафедры программных средств, Запорожский национальный технический университет, Запорожье, Украина

⁴Аспирант кафедры программных средств, Запорожский национальный технический университет, Запорожье, Украина

⁵Канд. техн. наук, с.н.с. кафедры программных средств, Запорожский национальный технический университет, Запорожье, Украина

ПЛАНИРОВАНИЕ РЕСУРСОВ ПАРАЛЛЕЛЬНОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ ПРИ СИНТЕЗЕ НЕЙРО-НЕЧЕТКИХ МОДЕЛЕЙ ДЛЯ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

Решена задача планирования ресурсов параллельных компьютерных систем при синтезе нейро-нечетких сетей. Объект исследования – процесс синтеза нейро-нечетких моделей. Предметом исследования являются методы планирования ресурсов параллельных компьютерных систем. Цель работы заключается в построении модели планирования ресурсов параллельных компьютерных систем, осуществляющих решение прикладных задач на основе параллельного метода синтеза нейро-нечетких сетей. Предложена модель планирования ресурсов параллельных компьютерных систем при синтезе нейро-нечетких сетей. Синтезированная модель учитывает тип компьютерной системы, число процессов, на которых выполняется задача, пропускную способность сети передачи данных, параметры используемого математического обеспечения (число возможных решений, обрабатываемых в процессе работы метода, доли решений, генерируемых на каждой итерации стохастического поиска с помощью применения операторов скрещивания и мутации), а также параметры решаемой прикладной задачи (число наблюдений и число признаков в заданном множестве данных, описывающем результаты наблюдений за исследуемым объектом или процессом). Разработано программное обеспечение, реализующее синтезированную модель планирования ресурсов. Выполнены эксперименты, подтверждающие адекватность предложенной модели. Результаты экспериментов позволяют рекомендовать применение разработанной модели на практике.

Ключевые слова: выборки данных, параллельные вычисления, планирование ресурсов, нейро-нечеткая модель, нейронная сеть.

Oliinyk A.¹, Skrupsky S.Yu.², Subbotin S.³, Blagodariov O.⁴, Gofman Ye.⁵

¹PhD, Associate Professor of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine

²PhD, Associate Professor of Computer Systems and Networks Department, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine

³Dr.Sc., Head of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine

⁴Postgraduate student of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine

⁵PhD, Senior Researcher of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine

PARALLEL COMPUTING SYSTEM RESOURCES PLANNING FOR NEURO-FUZZY MODELS SYNTHESIS AND BIG DATA PROCESSING

The article deals with the problem of planning resources of parallel computer systems for the synthesis of neuro-fuzzy networks. The object of research is a process of synthesis of neuro-fuzzy models. The subject of research are the methods of resource planning of parallel computer systems. The purpose of the work is to construct a model of parallel computing systems for resource planning, carrying out the decision of practical applications based on parallel method of neuro-fuzzy networks synthesis. A model of parallel computer systems resource planning for the synthesis of neuro-fuzzy networks is proposed. Synthesized model takes into account the type of computer system, the number of processes in which the task is executed, the capacity of data network, the parameters of the mathematical software (number of possible solutions to be processed during the operation of the method, the proportion of solutions generated in each iteration of stochastic search through the use of crossover and mutation operator), as well as parameters of the solved applied problem (the number of observations and the number of features in a given data sample, which describes the results of observing the researching object or process). The software that implements a synthesized model of resource planning is developed. Experiments confirming the adequacy of the proposed model are executed. The experimental results allow us to recommend the usage of the developed model in practice.

Keywords: data sample, parallel computing, resource planning, neuro-fuzzy models, neural network.

REFERENCES

1. Nauck D., Klawonn F., Kruse R. Foundations of neuro-fuzzy systems. Chichester, John Wiley & Sons, 1997, 305 p.
2. Bodyanskiy Ye. V., Tyshchenko O. K., Boiko O. O. An evolving cascade system based on neurofuzzy nodes, *Radio Electronics, Computer Science, Control*, 2016, No. 2, pp. 40–45.
3. Subbotin S. A. The method of diagnostic model synthesis based on radial basis neural networks with the support of generalization properties, *Radio Electronics, Computer Science, Control*, 2016, No. 2, pp. 64–69.
4. Oliinyk A. O., Zayko T. A., Subbotin S. A. Synthesis of Neuro-Fuzzy Networks on the Basis of Association Rules, *Cybernetics and Systems Analysis*, 2014, Vol. 50, Issue 3, pp. 348–357. DOI: 10.1007/s10559-014-9623-7.
5. Oliinyk A., Zaiko T., Subbotin S. Training Sample Reduction Based on Association Rules for Neuro-Fuzzy Networks Synthesis, *Optical Memory and Neural Networks (Information Optics)*, 2014, Vol. 23, No. 2, pp. 89–95. DOI: 10.3103/S1060992X14020039.
6. Oliinyk A., Subbotin S. A. The decision tree construction based on a stochastic search for the neuro-fuzzy network synthesis, *Optical Memory and Neural Networks (Information Optics)*, 2015, Vol. 24, No. 1, pp. 18–27. DOI: 10.3103/S1060992X15010038.
7. Oliinyk A. Production rules extraction based on negative selection, *Radio Electronics, Computer Science, Control*, 2016, No. 1, pp. 40–49.
8. Oliinyk A. O., Skrupsky S. Yu., Subbotin S. A. Using Parallel Random Search to Train Fuzzy Neural Networks, *Automatic Control and Computer Sciences*, 2014, Vol. 48, Issue 6, pp. 313–323. DOI: 10.3103/S0146411614060078.
9. Oliinyk A. O., Skrupsky S. Yu., Subbotin S. A. Experimental Investigation with Analyzing the Training Method Complexity of Neuro-Fuzzy Networks Based on Parallel Random Search, *Automatic Control and Computer Sciences*, 2015, Vol. 49, Issue 1, pp. 11–20. DOI: 10.3103/S0146411615010071.
10. Subbotin S., Oliinyk A., Skrupsky S. Individual prediction of the hypertensive patient condition based on computational intelligence, *Information and Digital Technologies : International Conference IDT'2015, Zilina, 7–9 July 2015 : proceedings of the conference*. Zilina, Institute of Electrical and Electronics Engineers, 2015, pp. 336–344. DOI: 10.1109/DT.2015.7222996.
11. Sullistio A., Yeo C. S., R. Buyya Simulation of Parallel and Distributed Systems: A Taxonomy – and Survey of Tools, *International Journal of Software Practice and Experience*. Wiley Press, 2002, pp. 1–19.
12. Ponomarenko V. S., Listrovoj S. V., Minuhin S. V., Znahur S. V. Metody i modeli planirovaniya resursov v GRID-sistemah: monografija. Xar'kov, ID «INZhJeK», 2008, 408 p.
13. Introduction to GPUs. Rezhim dostupa: URL: <https://www.cs.utexas.edu/~pingali/CS378/2015sp/lectures/IntroGPUs.pdf>. – Zagl. z ekranu.
14. Gebali F. Algorithms and Parallel Computing. New Jersey, John Wiley & Sons, 2011, 364 p. DOI: 10.1002/9780470932025.
15. Fokkink W. Distributed Algorithms: An Intuitive Approach. Cambridge, MIT Press, 2013, 248 p.
16. Herlihy M., Shavit N. The Art of Multiprocessor Programming Revised Reprint. Boston, Morgan Kaufmann, 2012, 536 p.
17. Roosta S. H. Parallel Processing and Parallel Algorithms: Theory and Computation. New York, Springer-Verlag, NY, 2000, 566 p. DOI: 10.1007/978-1-4612-1220-1.
18. Hajkin S. Nejrornyie seti: polnyj kurs. Sankt-Peterburg, Vil'jams, 2005, 1104 p.
19. Suzuki K. Artificial Neural Networks: Architectures and Applications. New York, InTech, 2013, 264 p. DOI: 10.5772/3409.
20. Hanrahan G. Artificial Neural Networks in Biological and Environmental Analysis. Boca Raton, Florida, CRC Press, 2011, 214 p. DOI: 10.1201/b10515.
21. Bodyanskiy Ye., Vynokurova O. Hybrid adaptive wavelet-neuro-fuzzy system for chaotic time series identification, *Information Sciences*, 2013, Vol. 220, pp. 170–179. DOI: 10.1016/j.ins.2012.07.044.
22. Bishop C. Neural Networks for pattern recognition. New York, Oxford University Press, 1995, 482 p.
23. Ch. L. ed. Mumford Computational intelligence: collaboration, fusion and emergence. New York, Springer, 2009, 752 p. DOI: 10.1007/978-3-642-01799-5.
24. Bodyanskiy Ye., Tyshchenko A. K., Deineko A. An evolving radial basis neural network with adaptive learning of its parameters and architecture, *Automatic Control and Computer Sciences*, 2015, Vol. 49, Issue 5, pp. 255–260. DOI: 10.3103/S0146411615050028.
25. Berthold M., Hand D. J. eds. Intelligent data analysis: an introduction. New York, Springer Verlag, 2007, 525 p. DOI: 10.1007/978-3-540-48625-1.
26. Tenne Y., Goh C.-K. Computational Intelligence in Expensive Optimization Problems. Berlin, Springer, 2010, 800 p. DOI: 10.1007/978-3-642-10701-6.
27. Skrupsky S. Yu. Experimental investigation of method for the synthesis of neurofuzzy models in a parallel computer system, *Radio Electronics, Computer Science, Control*, 2016, No. 2, pp. 56–63.
28. Oliinyk A., Skrupsky S., Subbotin S. Parallel Computer System Resource Planning for Synthesis of Neuro-Fuzzy Networks, *Recent Advances in Systems, Control and Information Technology. Advances in Intelligent Systems and Computing*, 2017, Vol. 543, pp. 124–133. DOI: 10.1007/978-3-319-48923-0_12
29. UCI Machine Learning Repository. <http://archive.ics.uci.edu/ml/>.
30. Quinn M. J. Parallel Programming in C with MPI and OpenMP. New York, NY, McGraw-Hill, 2004, 529 p.