

МОДЕЛЬ ПАРАЛЕЛЬНОЇ СОРТУВАЛЬНОЇ НЕЙРОННОЇ МЕРЕЖІ ДИСКРЕТНОГО ЧАСУ

© Тимошук П., 2020

Представлено модель паралельної сортувальної нейронної мережі дискретного часу. Модель описується системою різницевих і вихідних рівнянь. Мережа відзначається високою швидкістю, довільною скінченною роздільною здатністю вхідних даних і придатна для обробки невідомих вхідних даних зі скінченними значеннями, розміщених у довільному відомому скінченному діапазоні. Мережа характеризується незначною обчислювальною складністю і складністю схемотехнічної реалізації. Наведено результати комп'ютерного моделювання, які ілюструють ефективність мережі.

Ключові слова: Модель, різницева рівняння, нейронна мережа, паралельне сортування, роздільна здатність.

Постановка проблеми

Спроекуємо модель штучної нейронної мережі дискретного часу, призначеної для сортування вхідних даних у паралельному режимі. Модель повинна забезпечувати високу швидкість обробки даних, мати незначну обчислювальну складність, а також довільну скінченну роздільну здатність стосовно вхідних даних. Крім цього, модель має бути придатною для сортування невідомих вхідних даних, які знаходяться у будь-якому відомому скінченному діапазоні. Проілюструємо функціонування спроектованої моделі за допомогою чисельних експериментів.

Актуальність дослідження

Сортування – це процес упорядкування даних. Сортування є базовою операцією обробки даних, яка займає більше, ніж 25% часу обробки. Головна мета алгоритмів сортування – розміщення даних у необхідному порядку протягом мінімального часу з використанням мінімального об'єму пам'яті. Сортування використовується при проектуванні великих інтегральних схем, для обробки даних і сигналів, у комунікаційних мережах та в інших застосуваннях [1]. Тому проектування систем сортування з покращеними характеристиками є актуальною задачею.

Формулювання мети і завдань статті

Метою статті є проектування покращеної моделі штучної нейронної мережі дискретного часу, призначеної для сортування даних. Для досягнення сформульованої мети треба виконати такі завдання:

- проаналізувати останні дослідження та опубліковані результати за темою статті;
- спроектувати модель штучної нейронної схеми дискретного часу, яка має лежати в основі функціонування сортувальної мережі;
- сконструювати модель штучної нейронної мережі дискретного часу, призначено для паралельного сортування невідомих вхідних даних, що знаходяться у довільному відомому скінченному діапазоні;
- навести результати моделювання спроектованої мережі за допомогою чисельних експериментів, які демонструють ефективність моделі;
- сформулювати на базі одержаних результатів висновки і напрямки подальших досліджень.

Аналіз останніх досліджень і публікацій

Традиційні послідовні методи сортування вхідних даних описані, наприклад, в [1]. Такі підходи характеризуються квадратичною обчислювальною складністю. На відміну від них підходи, пов'язані з паралельним сортуванням, відзначаються лінійною обчислювальною складністю [2]. Алгоритми паралельного сортування можуть реалізуватись за допомогою нейронних мереж неперервного часу [3] – [10]. Так, наприклад, сортувальні мережі, призначені для паралельного сортування, наведені в [3] – [5], базуються на нейронних мережах Хопфілда. Паралельні сортувальні мережі, описані в [6], використовують нейронні мережі, призначені для пошуку максимальних за значеннями даних. В [7] для реалізації процедури сортування застосовується квадратичний перцептрон. У [8] запропоновано аналогову нейронну мережу, призначену для паралельного сортування, яка здійснює мінімізацію так звану функцію вартості глобального неупорядкування, придатну для реалізації за допомогою інтегральних схем. Модель паралельної сортувальної нейронної мережі, яка базується на так званій КВТА-моделі неперервного часу, запропоновано в [10]. Аналогову реалізацію систем сортування на основі інтегральних схем можна знайти, наприклад, в [8], [11] - [13].

Виклад основного матеріалу

Модель базової нейронної схеми дискретного часу

Нейронні мережі дискретного часу порівняно з аналогами неперервного часу мають такі переваги, як точність, надійність функціонування, а також вища якість обробки реальних неточних даних. Такі мережі можна реалізувати у паралельному як програмному, так і апаратному забезпеченні [13]. Спроєкуємо модель паралельної сортувальної нейронної мережі дискретного часу. Для цього розглянемо вектор $\mathbf{a} = (a_1, a_2, \dots, a_N)^T$, $1 < N < \infty$ дійсних вхідних даних з невідомими, але скінченними значеннями елементів. Нехай такі значення є різними, тобто

$$a_i \neq a_j, \quad -\infty < a_i, a_j < \infty, \quad (1)$$

де $i \neq j = 1, 2, \dots, N$. В [14] запропоновано просту швидкісну нейронну схему, яка може знаходити K найбільших з цих значень. Схема обробляє вектор \mathbf{a} вхідних даних у такий спосіб, що на її виході отримується вектор $\mathbf{b} = (b_1, b_2, \dots, b_N)$ відповідних вихідних даних, елементи якого задовольняють наступні нерівності:

$$b_i > 0, i = 1, 2, \dots, K; \quad b_j < 0, j = K+1, K+2, \dots, N. \quad (2)$$

Спростимо модель нейронної схеми, наведеної в [14], описавши її різницеvim рівнянням виду:

$$x(l+1) = \alpha x(l) + A \operatorname{sgn}(D(x)) \mathbf{a}^l, \quad (3)$$

де l - номер ітерації (крок дискретизації), $D(x) = \sum_{k=1}^N S_k(x) - K$ - функція різниці між необхідним і отриманим числами додатніх вихідних даних,

$$S_k(x) = \begin{cases} 1, & \text{якщо } a_k - a_{\min} - x > 0; \\ 0, & \text{інакше} \end{cases} \quad (4)$$

- ступінчаста функція,

$$\operatorname{sgn}(D(x)) = \begin{cases} 1, & \text{якщо } D(x) > 0; \\ 0, & \text{якщо } D(x) = 0; \\ -1, & \text{якщо } D(x) < 0 \end{cases} \quad (5)$$

- знакова функція, $x(l)$ - змінна стану дискретного часу, $0.5 < \alpha < 1$ - параметр, який гарантує збіжність $x(l)$ до коректного встановленого стану; $A = a_{\max} - a_{\min}$ - проміжок зміни вхідних даних, a_{\max} і a_{\min} - найбільше і найменше значення, яких можуть набувати вхідні дані, $0 \leq x(l) \leq A$ - початкове значення змінної стану.

Число ітерацій, які необхідні для збіжності пошукового процесу до встановленого стану, обмежується знизу наступною нерівністю:

$$m > \log_a(2\alpha - 1). \quad (6)$$

Слід відмітити, що нейронна схема, яка описується різницеvim рівнянням (3) і ступінчастими функціями (4), коректно обробляє вхідні дані, розміщені у відомому проміжку A .

Модель нейронної мережі паралельного сортування дискретного часу

Порядок паралельного сортування можна описати за допомогою матриці перестановок, у якій “1” у рядку a_i і стовпчику g_j вказує на i -й елемент вектора вхідних даних і на j -й елемент вектора вихідних даних [8]. Так, наприклад, при $i = 1, 2, 3, 4, 5, 6, 7$ матриця перестановок

$$\begin{array}{cccccccc}
 d_1 & d_2 & d_3 & d_4 & d_5 & d_6 & d_7 & \text{ранг} \\
 a_1 & 0 & 0 & 1 & 0 & 0 & 0 & 3 \\
 a_2 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
 a_3 & 0 & 0 & 0 & 1 & 0 & 0 & 4 \\
 a_4 & 0 & 1 & 0 & 0 & 0 & 0 & 2 \\
 a_5 & 0 & 0 & 0 & 0 & 1 & 0 & 5 \\
 a_6 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
 a_7 & 0 & 0 & 0 & 0 & 0 & 1 & 6
 \end{array} \tag{7}$$

відображує початкові вхідні дані $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$ і упорядковані вхідні дані $\{a_2, a_4, a_1, a_3, a_5, a_7, a_6\}$. Матриця (7) може бути перетворена до такої відсортованої матриці:

$$\begin{array}{cccccccc}
 s^1 & s^2 & s^3 & s^4 & s^5 & s^6 & s^7 & \text{ранг} \\
 a_1 & 0 & 0 & 1 & 1 & 1 & 1 & 3 \\
 a_2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 a_3 & 0 & 0 & 0 & 1 & 1 & 1 & 4 \\
 a_4 & 0 & 1 & 1 & 1 & 1 & 1 & 2 \\
 a_5 & 0 & 0 & 0 & 0 & 1 & 1 & 5 \\
 a_6 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
 a_7 & 0 & 0 & 0 & 0 & 0 & 1 & 6
 \end{array} \tag{8}$$

У загальному випадку результати сортування можна подати так:

$$d_{k+1} = a^T s^1, \quad d_{k+1} = a^T (s^{k+1} - s^k), \tag{9}$$

де $s^k = [s_1^k, s_2^k, \dots, s_N^k]^T$, $K = 1, 2, \dots, N-1$ - ступінчасті функції (4), а $s^{k+1} = [s_1^{k+1}, s_2^{k+1}, \dots, s_N^{k+1}]^T$ - ступінчасті функції

$$s_{k+1}(x) = \begin{cases} 1, & \text{якщо } a_{k+1} - a_{\min} - x > 0; \\ 0, & \text{інакше} \end{cases} \tag{10}$$

Відмітимо, що для виконання паралельного сортування необхідно тільки $N-1$ рівнянь (3) і виразів (4), (10), оскільки $s^N = [1, 1, \dots, 1]^T$. Тому, на відміну від інших мереж сортування, які потребують N^2 нейронів, описаній вище мережі достатня наявність тільки $N-1$ нейронів. Так, для знаходження елемента вектора вхідних даних з максимальним значенням необхідна одна 1WTA нейронна схема. Другий максимальний за значення елемент вектора вхідних даних може бути знайдений за допомогою однієї 1WTA- схеми і однієї 2WTA схеми у паралельному режимі без повторного знаходження першого елемента. Третій елемент вектора вхідних даних з максимальним значенням знаходиться за допомогою однієї 2WTA-схеми і однієї 3WTA-схеми у паралельному режимі без повторного знаходження другого елемента і так далі. Тому, оскільки $s^N = [1, 1, \dots, 1]^T$, увесь вектор вхідних даних, який містить N елементів, можна відсортувати за допомогою $N-1$ нейронних схем, які описуються рівняннями (3) і виразами (4), (9), (10) [3], [8].

Комп'ютерне моделювання мережі

Змоделюємо на комп'ютері функціонування нейронної мережі дискретного часу, призначеної для паралельного сортування, яка описується рівняннями (3) і виразами (4), (9), (10). З цією метою

посортуємо елементи такого вектора однорідно розподілених на інтервалі $(-5.0, 5.0)$ випадкових чисел: $a = (-0.8805, 2.4457, -2.3205, -0.6008, 4.3338, 1.8333, -2.8744, 3.3924, 1.2878)^T$, тобто $N = 9$. Використаємо для цього нейронні схеми дискретного часу, що описуються різницевиими рівняннями (3) та ступінчастими функціями (4), (10). Покладемо, що початкові значення змінних станів дискретного часу $x^k(1) = 5.0, k = 1, 2, \dots, N$. Відмітимо, що на протипагу 81 нейронам сортувальної мережі, запропонованій у [9], при використанні 8 нейронних схем, що описуються рівняннями (3) і виразами (4), (9), (10) нам знадобиться тільки 8 нейронів. На рис. 1 представлено траєкторії вхідних та вихідних даних описаної вище нейронної мережі дискретного часу, застосованої для паралельного сортування при $\alpha = 0.6, a_{min} = -5.0$ і $A = 10.0$. Коректні результати сортування вхідних даних $d = (4.3338, 3.3924, 2.4457, 1.8333, 1.2878, -0.6008, -0.8805, -2.3205, -2.8744)^T$ отримуються після $m = 6 > \log_{0.6}(2 \times 0.6 - 1) = 3.1507$ ітерацій згідно з оцінкою (6). Для завершення процесу сортування у середовищі Matlab необхідно приблизно 0.01 с.

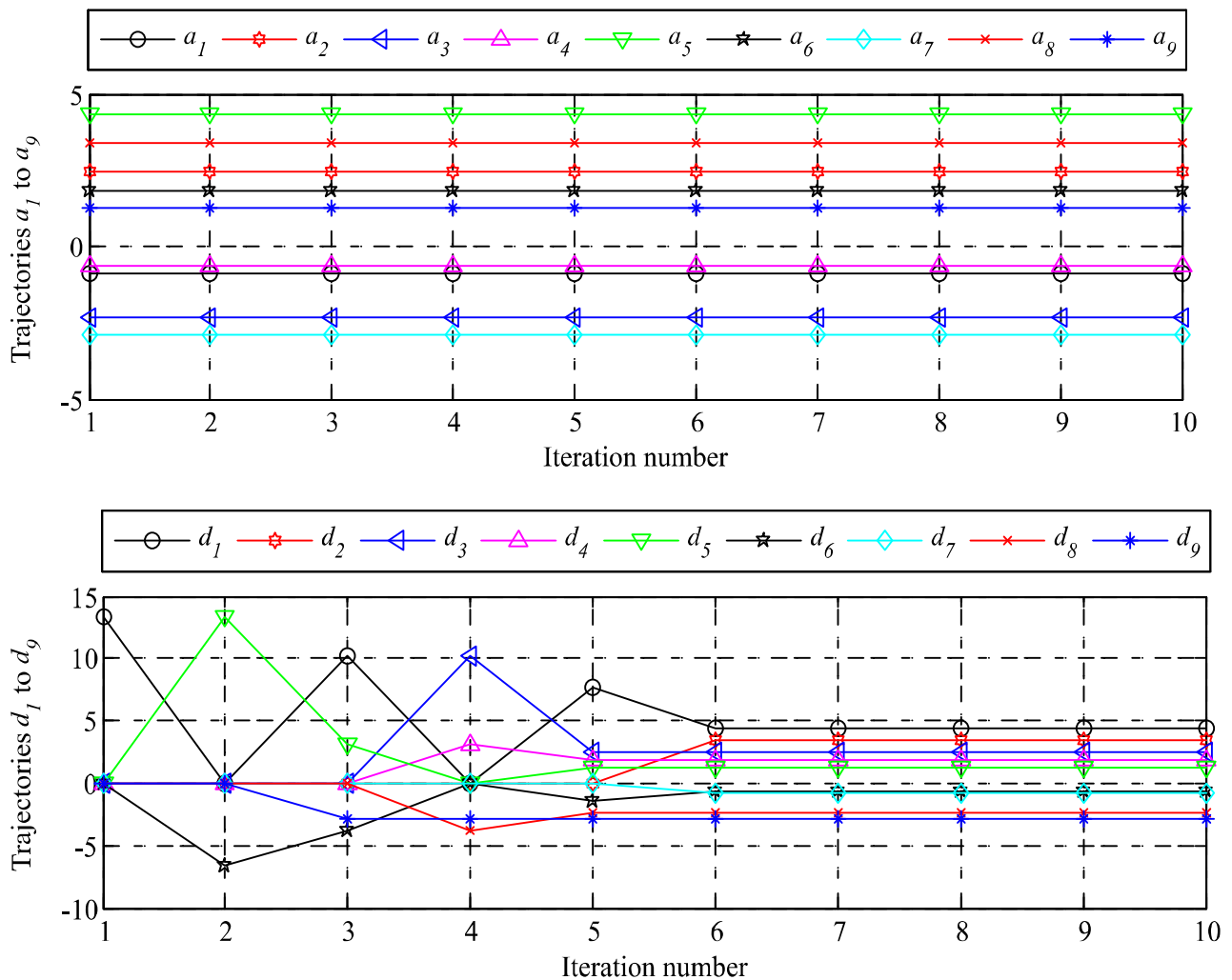


Рис. 1. Траєкторії вхідних і вихідних даних нейронної мережі дискретного часу, яка описується рівняннями (3) і виразами (4), (9), (10), використаної для здійснення паралельного сортування елементів вектора $a = (-0.8805, 2.4457, -2.3205, -0.6008, 4.3338, 1.8333, -2.8744, 3.3924, 1.2878)$ при $x^k(1) = 5.0, k = 1, 2, \dots, N$

На рис. 2 подано найбільші значення для числа ітерацій, необхідних для виконання сортування однорідно розподілених на інтервалі $(-5.0, 5.0)$ випадкових чисел при різних величинах параметра $0.5 < \alpha < 0.9$ і початкових значеннях змінних стану дискретного часу $x^k(1) = 5.0, k = 1, 2, \dots, N$. Як видно з рис. 2, найбільше число ітерацій не перевищує 50 для всіх значень

параметра $0.5 < \alpha < 0.9$. Найбільший час, необхідний для здійснення паралельного сортування у середовищі Matlab, дорівнює 0.016 с.

На рис. 3 показано найбільші числа ітерацій, необхідні для виконання паралельного сортування однорідно розподілених на інтервалі $(-5.0, 5.0)$ випадкових чисел при $\alpha = 0.6$ і різних величинах змінних станів дискретного часу $x^k(1), k = 1, 2, \dots, N$, які знаходяться на інтервалі $(-5.0, 5.0)$. Як можна побачити з рис. 3, найбільші числа ітерацій, необхідні для здійснення паралельного сортування даних у цьому випадку, не перевищують 25 для всіх величин $x^k(1), k = 1, 2, \dots, N$. Найбільший час, необхідний для виконання такого сортування у середовищі Matlab, є приблизно рівним 0.016 с.

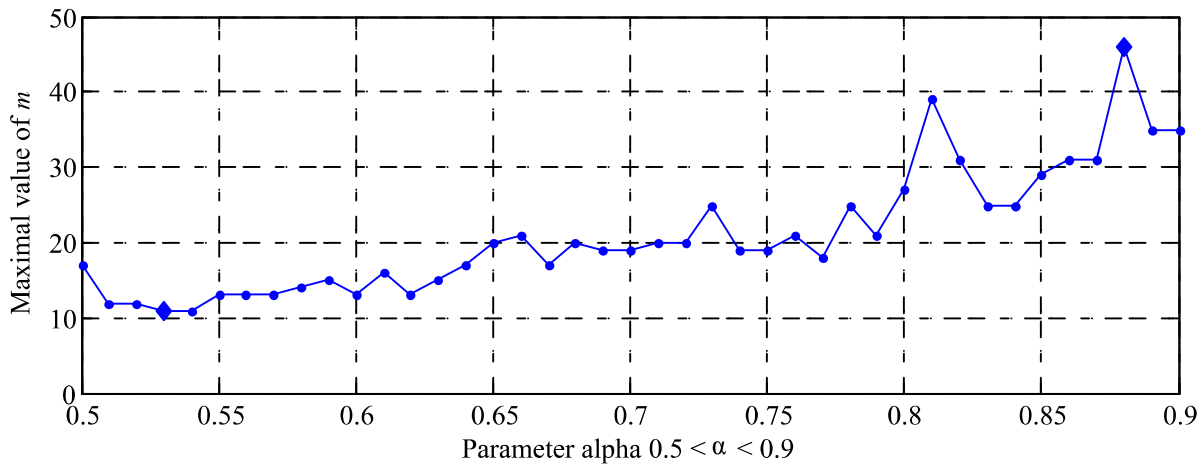


Рис. 2. Найбільші значення числа ітерацій, необхідних для виконання паралельного сортування однорідно розподілених на інтервалі $(-5.0, 5.0)$ випадкових чисел при різних величинах параметра $0.5 < \alpha < 0.9$ і значеннях змінних станів дискретного часу $x^k(1) = 5.0, k = 1, 2, \dots, N$

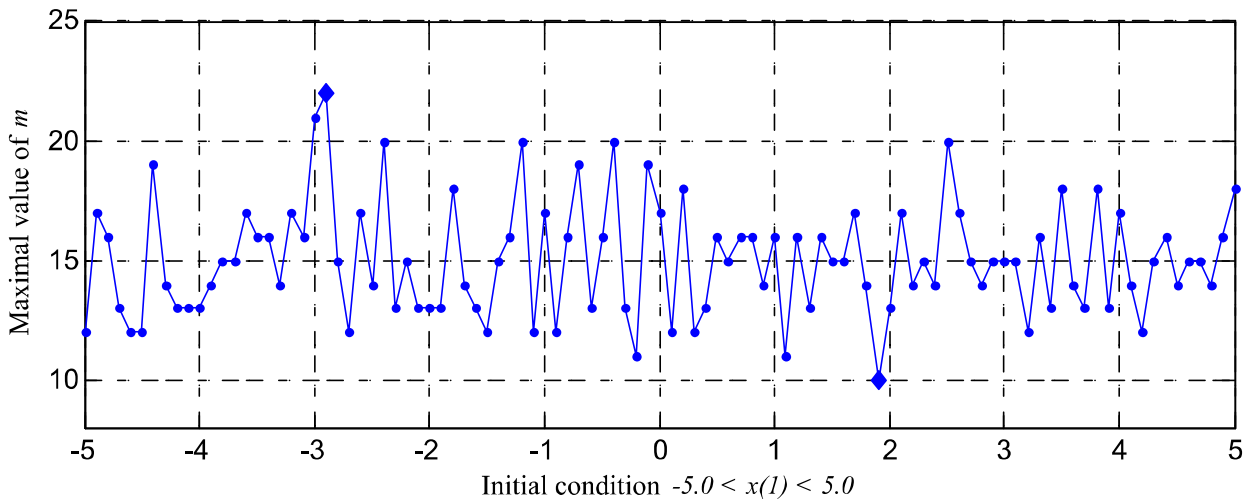


Рис. 3. Найбільші числа ітерацій, необхідні для здійснення паралельного сортування однорідно розподілених на інтервалі $(-5.0, 5.0)$ випадкових чисел при $\alpha = 0.6$ і різних початкових величинах змінних станів дискретного часу $x^k(1), k = 1, 2, \dots, N$, які знаходяться на інтервалі $(-5.0, 5.0)$

Висновки

Спроековано модель паралельної сортувальної мережі дискретного часу довільних невідомих скінченних вхідних даних, розміщених у відомому діапазоні. Мережа відзначається високою швидкодією і роздільною здатністю. Така мережа може бути реалізована у паралельному

програмному забезпеченні. Програмне забезпечення може бути вибране, якщо необхідна точна, проста і гнучка паралельна реалізація. Мережу можна реалізувати у цифровому апаратному забезпеченні за допомогою великих інтегральних схем, якщо необхідна швидкісна, надійна і завадостійка система сортування. Ефективність мережі продемонстрована за допомогою комп'ютерного моделювання.

Перспективи подальших досліджень

Подальші дослідження за тематикою статті пов'язані із застосуванням розробленої моделі паралельної сортувальної нейронної мережі дискретного часу для проектування ефективних систем, призначених для класифікації і фільтрування.

1. D. E. Knuth, *The Art of Computer Programming, Sorting, and Searching*. Reading, MA: Addison-Wesley, 1973.
2. S. G. Akl, *Parallel Sorting Algorithms*, Orlando, FL: Academic, 1985.
3. M. Atkins, "Sorting by Hopfield nets," in *Proc. Int. Joint Conf. Neural Netw.*, Washington, DC, USA, 1989, pp. 65-68.
4. Y. Takefuji and K.-S. Lee, "A super parallel sorting algorithm based on neural networks," *IEEE Trans. Circuits Syst.*, vol. CAS-37, no. 11, pp.1425-1429, 1990.
5. W. Chen and K. Hsieh, "A neural sorting network with $O(1)$ time complexity," in *Proc. Int. Joint Conf. Neural Networks*, vol. III, San Diego, CA, 1990, pp. 793-798.
6. T. M. Kwon and M. Zervakis, "A parallel sorting network without comparators: A neural network approach," in *Proc. Int. Joint Conf. Neural Networks*, vol. I, Baltimore, MD, 1992, pp. 701-706.
7. Y.-H. Tseng and J.-L. Wu, "Solving sorting and related problems by quadratic perceptrons." *Electron. Lett.*, vol. 28, no. 10, pp. 906-908, 1992.
8. J. Wang, "Analysis and design of an analog sorting network," *IEEE Trans. Neural Networks*, vol. 6, no. 4, pp. 962-971, Jul. 1995.
9. T. M. Kwon and M. Zervakis, "KwTA networks and their applications," *Multidimensional Syst. and Signal Processing*, vol. 6, no. 4, pp. 333-346, Oct. 1995.
10. J. Wang, "Analysis and design of a k-winners-take-all model with a single state variable and the Heaviside step activation function," *IEEE Trans. Neural Networks*, vol. 21, no. 9, pp. 1496-1506, Sept. 2010.
11. H. M. Alnuweiri and V. K. P. Kumar, "Optimal VLSI sorting with a reduced number of processors," *IEEE Trans. Comput.*, vol. C-40, pp. 105-110, 1991.
12. S. Rovetta and R. Zunino, "Minimal-connectivity programmable circuit for analog sorting," *IEE Proc. Circuits, Devices Syst.*, vol. 146, no. 3, pp. 108-110, Aug. 1999.
13. P. V. Tymoshchuk and S. V. Shatnyi, "Hardware implementation design of analog sorting neural network", in *Proc. XX Int. Seminar/Workshop "Direct and inverse problems of electromagnetic and acoustic wave theory"*, pp. 168-171, Sept. 2015.
14. P.V.Tymoshchuk, "A discrete-time dynamic K-winners-take-all neural circuit", *Neurocomputing*, vol. 72, 2009, pp. 3191-3202.

P. Tymoshchuk

Lviv Polytechnic National University

A MODEL OF PARALLEL SORTING NEURAL NETWORK OF DISCRETE-TIME

© Tymoshchuk P., 2020

A model of parallel sorting neural network of discrete-time is presented. The model is described by a system of differential equations and by step functions. The network has high speed, any finite resolution of input data and it can process unknown input data of finite values located in arbitrary finite range. The network is characterized by moderate computational complexity and complexity of hardware implementation. The results of computer simulation illustrating the efficiency of the network are provided.

Keywords: Model, difference equation, neural network, parallel sorting, resolution.