

СПРОЩЕНА МОДЕЛЬ НЕЙРОННОЇ МЕРЕЖІ ДИСКРЕТНОГО ЧАСУ ДЛЯ ПАРАЛЕЛЬНОГО СОРТУВАННЯ

П. В. Тимощук

Національний університет “Львівська політехніка”,
кафедра систем автоматизованого проектування

© Тимощук П. В., 2020

Запропоновано модель паралельної сортувальної нейронної мережі дискретного часу. Модель описується системою різницевих рівнянь і ступінчастими функціями. Модель базується на спрощеній нейронній схемі дискретного часу, призначеній для ідентифікації максимальних/мінімал за значеннями вхідних даних, яка описується різницеvim рівнянням і ступінчастими функціями. Визначається обмеження згори на кількість ітерацій, необхідних для досягнення пошуковим процесом збіжності до встановленого стану. Модель не потребує знання діапазону зміни вхідних даних. Для використання моделі має бути відомою мінімальна різниця між значеннями вхідних даних. Мережа придатна для обробки невідомих вхідних даних зі скінченними значеннями, розміщеними у довільному невідомому скінченному діапазоні. Мережа характеризується незначними обчислювальною складністю і складністю програмної реалізації, довільною скінченною роздільною здатністю вхідних даних, швидкодією. Наведено результати комп'ютерного моделювання, які ілюструють ефективність мережі.

Ключові слова: паралельне сортування, нейронна мережа, різницеve рівняння, обчислювальна складність, апаратна реалізація.

Постановка задачі

Необхідно спроектувати модель паралельної сортувальної нейронної мережі дискретного часу. Мережа має бути придатною для обробки будь-яких невідомих вхідних даних, розміщених у довільному невідомому скінченному діапазоні. Мережа повинна відзначатись довільною скінченною роздільною здатністю, високою швидкістю обробки даних і лінійною обчислювальною складністю $\mathcal{O}(N)$. Необхідно навести приклад комп'ютерного моделювання, який ілюструє ефективність функціонування мережі.

Актуальність дослідження

Сортування полягає у впорядкуванні даних. Сортування є базовою операцією обробки даних, що займає більше, ніж 25 % часу обробки. Головна мета сортування – розміщення даних у необхідному порядку протягом мінімального часу за допомогою якомога меншого обсягу пам'яті. Сортування використовується при обробці даних і сигналів, у комунікаційних мережах, при проектуванні великих інтегральних схем та в інших застосуваннях [1]. Тому проектування систем сортування з покращеними характеристиками є актуальною задачею.

Формулювання мети та завдань статті

Метою статті є проектування моделі паралельної сортувальної нейронної мережі дискретного часу. Для досягнення поставленої мети необхідно виконати такі завдання:

- проаналізувати останні дослідження і публікації за темою статті;
- розробити модель базової нейронної схеми дискретного часу, на основі якої має функціонувати модель сортувальної мережі;
- спроекувати модель нейронної мережі паралельного сортування дискретного часу;
- навести приклад комп'ютерного моделювання мережі, який підтверджує теоретичні положення та ілюструє функціонування мережі;
- на основі отриманих результатів проектування сформулювати висновки і тематику подальших досліджень.

Аналіз останніх досліджень та публікацій

У [1] описано послідовні підходи до розв'язання задачі сортування. На відміну від послідовних схем сортування, які мають квадратичну обчислювальну складність $\mathcal{O}(N^2)$, паралельні сортувальні системи забезпечують лінійну обчислювальну складність $\Theta(N)$. У [2] описано алгоритми та архітектури паралельного сортування. Паралельні сортувальні системи, які ґрунтуються на нейронних мережах неперервного часу, можна знайти в [3–10]. Зокрема, у паралельних сортувальних мережах, описаних у [3–5], використовують нейронні мережі Хопфілда. Паралельні сортувальні системи, наведені у [6], застосовують мережі, які визначають найбільші за значеннями вхідні дані. В [7] для виконання операції сортування використовується квадратичний персептрон. У [8] проектується аналогова сортувальна нейронна мережа, що мінімізує так звану функцію вартості глобального невпорядкування, яка реалізується за допомогою великих інтегральних схем. Нейронна мережа, призначена для паралельного сортування, що ґрунтується на нейронній схемі неперервного часу, призначеній для знаходження K найбільших за значеннями вхідних даних, де $1 \leq K < N$, описано в [10]. Схемотехнічну реалізацію систем сортування за допомогою великих інтегральних схем наведено, наприклад, у [8, 9, 11, 12].

Нейронні мережі дискретного часу відзначаються точністю, надійністю, краще обробляють реальні зашумлені дані. Такі типи мереж легко реалізуються у паралельному програмному забезпеченні. Мережі дискретного часу можна реалізувати у сучасному апаратному забезпеченні [13].

Модель базової KWTА-нейронної схеми дискретного часу

Розглянемо вектор $\mathbf{a} = (a_1, a_2, \dots, a_N)^T$, $1 < N < \infty$ дійсних вхідних даних з невідомими значеннями елементів. Нехай дані є такими, що їх можна розрізнити, і скінченними, тобто

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

де $i \neq j = 1, 2, \dots, N$. У [14] описано просту і швидкісну нейронну схему, здатну визначати K найбільших із цих даних. Така схема обробляє вектор \mathbf{a} вхідних даних так, що отримується вектор $\mathbf{b} = (b_1, b_2, \dots, b_N)$ відповідних вихідних даних, який задовольняє таку властивість:

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

Спростимо нейронну схему з [14], описавши її різницеvim рівнянням виду:

$$y(l+1) = y(l) + r \operatorname{sgn}(D(y)), \quad (3)$$

де l – номер ітерації (крок); $y(l)$ – змінна стану дискретного часу; $-\infty < y(1) < \infty$ – початкова умова; r – параметр, який повинен мати значення менше ніж мінімальна різниця між вхідними даними, тобто мають задовольнятися нерівності $0 < r \leq \min |a_i - a_j|$, $i \neq j = 1, 2, \dots, N$,

$$\operatorname{sgn}(D(y)) = \begin{cases} 1, & \text{якщо } D(y) > 0; \\ 0, & \text{якщо } D(y) = 0; \\ -1, & \text{якщо } D(y) < 0 \end{cases} \quad (4)$$

– сигнум – (жорстко обмежувальна) функція, $D(y) = \sum_{k=1}^N S_k(y) - K$ – функція різниці між отриманою і необхідною кількостями позитивних значень вихідних даних,

$$S_k(y) = \begin{cases} 1, & \text{якщо } a_k - y > 0; \\ 0, & \text{інакше} \end{cases} \quad (5)$$

– ступінчаста функція.

Кількість ітерацій, необхідних для досягнення збіжності пошукового процесу до встановленого стану, обмежується згори так:

$$m \leq |y(1) - y^*| / r, \quad (6)$$

де y^* – значення $y(l)$ у встановленому режимі. Як можна побачити, нейронна схема, яка описується різницеvim рівнянням (3) і ступінчастими функціями (5), потребує задавання значення параметра r , яке має бути меншим від мінімальної різниці між значеннями вхідних даних. Це означає, що мінімальна різниця між значеннями вхідних даних має бути відомою.

Модель нейронної схеми дискретного часу паралельного сортування

Під час паралельного сортування порядок сортування може бути представлений у вигляді матриці перестановок. У такій матриці “ I ” у рядку, позначеному через a_i і стовпчику, позначеному через g_j , може вказувати на i -й член у невідсортованому списку і на j -й член у відсортованому списку [8]. Наприклад, при $i = 1, 2, 3, 4, 5, 6, 7$ матриця перестановок

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	ранг
a_1	0	0	1	0	0	0	0	3
a_2	1	0	0	0	0	0	0	1
a_3	0	0	0	1	0	0	0	4
a_4	0	1	0	0	0	0	0	2
a_5	0	0	0	0	1	0	0	5
a_6	0	0	0	0	0	0	1	7
a_7	0	0	0	0	0	1	0	6

(7)

представляє невідсортований список $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$ і впорядкований список $\{a_2, a_4, a_1, a_3, a_5, a_7, a_6\}$. Матриця (7) може бути перетворена до відсортованої матриці виду:

S^1	S^2	S^3	S^4	S^5	S^6	S^7	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	6
a_7	0	0	0	0	0	1	6

(8)

У загальному випадку результати сортування можуть бути представлені у вигляді таких виразів:

$$d_1 = a^T S^1, \quad d_{K+1} = a^T (S^{K+1} - S^K), \quad (9)$$

де елементи K -го стовпчика $S^K = [S_1^K, S_2^K, \dots, S_N^K]^T$, $K = 1, 2, \dots, N-1$ матриці сортування визначаються за допомогою ступінчастих функцій (4), використовуючи різницеве рівняння (3) нейронної схеми, а елементи $K+1$ -го стовпчика $S^{K+1} = [S_1^{K+1}, S_2^{K+1}, \dots, S_N^{K+1}]^T$, $K=1, 2, \dots, N-1$ матриці сортування визначаються за допомогою ступінчастих функцій (5), використовуючи різницеві рівняння (3) так:

$$S_{k+1}(y) = \begin{cases} 1, & \text{якщо } a_{k+1} - y > 0; \\ 0, & \text{інакше} \end{cases} \quad (10)$$

Оскільки $S^N = [1, 1, \dots, 1]^T$, необхідно лише $N-1$ рівнянь (3), і кожне з цих рівнянь визначає один стовпчик матриці сортування зліва направо із значенням k , яке зростає від 1 до $N-1$. Тому порівняно з іншими системами сортування, які використовують N^2 нейронів [9, 10], необхідно тільки $N-1$ нейронів. Зокрема, для визначення найбільшого елемента списку необхідна одна 1WTA нейронна схема. Для обчислення другого елемента списку у паралельному режимі без перерахунку першого елемента необхідні 1WTA і 2WTA схеми. Для визначення третього елемента списку у паралельному режимі без перерахунку другого елемента треба 2WTA і 3WTA схеми і так далі. Увесь список N елементів можна відсортувати за допомогою KWTA, $K = 1, 2, \dots, N-1$ схем без необхідності обчислення останнього елемента встановленням $S^N = [1, 1, \dots, 1]^T$ [3, 8].

Комп'ютерне моделювання мережі

Розглянемо приклад комп'ютерного моделювання, який демонструє функціонування паралельної сортувальної нейронної мережі дискретного часу, яка описується рівняннями (3) і (9). Відсортуємо 9 елементів наступного 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 $N = 9$, за допомогою нейронної схеми дискретного часу, що описується рівняннями (3) і (9). Для цього задамо початкові значення змінних станів дискретного часу $x^k(l) = 5.0$, $k = 1, 2, \dots, N$. Використовуючи 8 нейронних схем, які описуються різницеvim

рівнянням (3) і ступінчастими функціями (5), необхідно лише 8 нейронів на відміну від 81 нейронів у сортувальній мережі з [9].

На рис. 1 показано динаміку вихідних і вихідних даних нейронної мережі дискретного часу, яка описується рівняннями (3) і (9), використаної для паралельного сортування при $r = 0.1$, $y^k(1) = 0.0$, $k = 1, 2, \dots, N$. Як можна побачити, коректне сортування вхідних даних досягається за $n < 40 < \frac{|0-5|}{0.01} = 500$ ітерацій згідно з оцінкою (6). Час, затрачений на таке сортування у середовищі Matlab, приблизно дорівнює to 0.07 s.

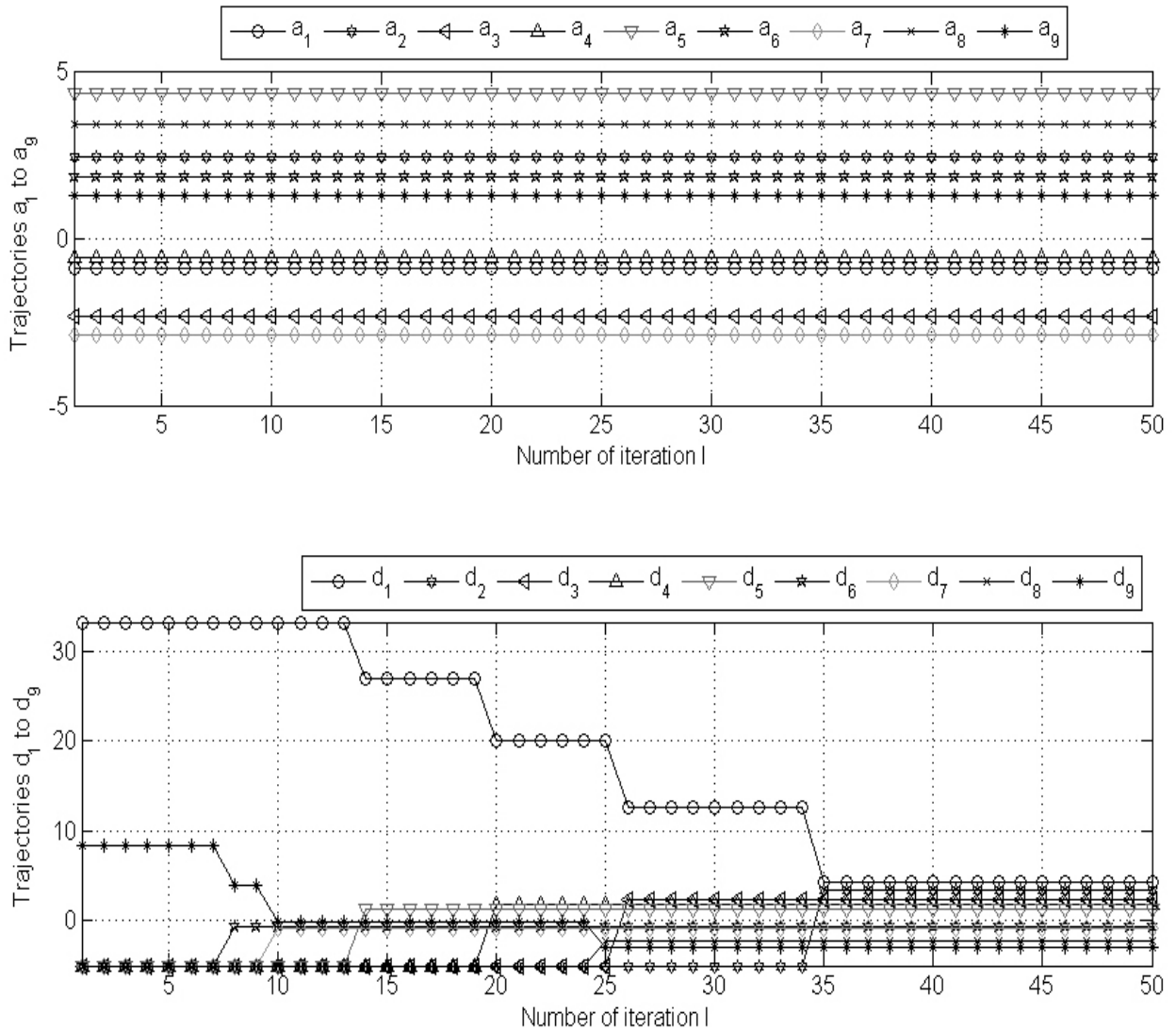


Рис. 1. Траєкторії вхідних і вихідних даних сортувальної нейронної мережі дискретного часу, яка описується рівняннями (3) і (9) при $r = 0.1$ і $y^k(1) = 0.0$, $k = 1, 2, \dots, N$.

На рис. 2 показано максимальні кількості ітерацій, необхідних для паралельного сортування за допомогою нейронної мережі дискретного часу, яка описується рівняннями (3) і (9)

при $r = 0.1$ для різних початкових значень змінної станів дискретного часу $y^k(1)$, $k = 1, 2, \dots, N$, вибраних на інтервалі $(-5.0, 5.0)$. Як можна побачити, максимальні кількості ітерацій, необхідних для сортування вхідних даних у цьому випадку, не перевищують 100 для всіх значень $y^k(1)$, $k = 1, 2, \dots, N$. Максимальний час, який необхідно затратити на таке сортування у середовищі Matlab, приблизно дорівнює 0.08 s.

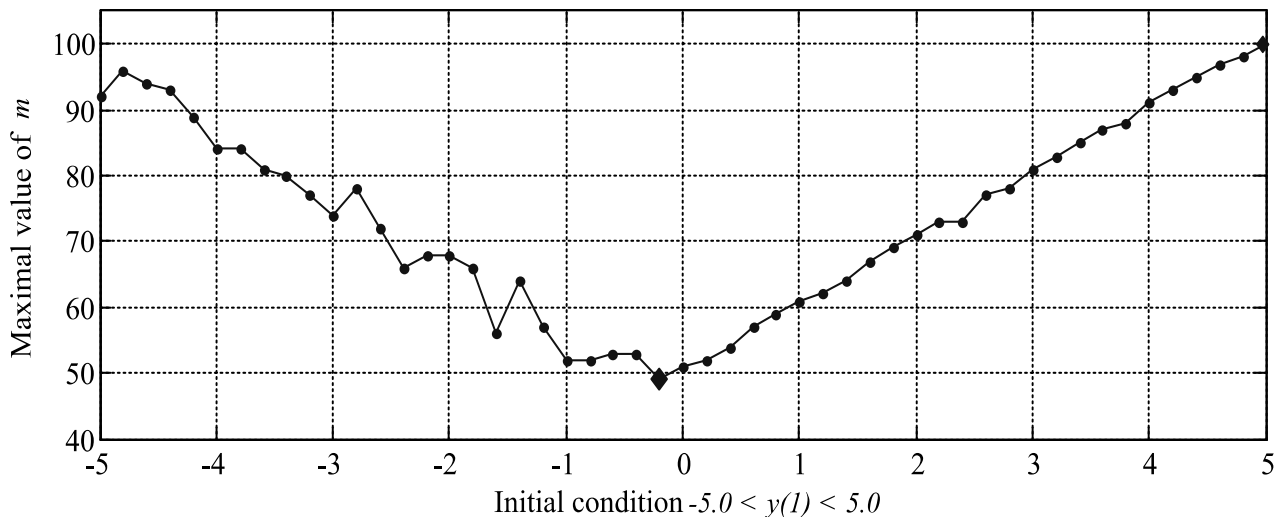


Рис. 2. Максимальні кількості ітерацій, необхідних для паралельного сортування за допомогою нейронної мережі дискретного часу, що описується рівняннями (3) і (9) при $r = 0.1$ і різних початкових значеннях змінних станів дискретного часу $y^k(1)$, $k = 1, 2, \dots, N$, вибраних на інтервалі $(-5.0, 5.0)$.

З результатів моделювання, бачимо, що паралельна сортувальна мережа, яка базується на нейронній схемі дискретного часу, й описується різницеvim рівнянням (3) і ступінчастими функціями (5), не потребує знання діапазону зміни вхідних даних. Для використання такої мережі необхідне знання мінімальної різниці між значеннями вхідними даними.

Висновки

Запропоновано модель паралельної сортувальної мережі дискретного часу вхідних даних із довільними невідомими скінченними значеннями, розміщеними у невідомому діапазоні. Мережа характеризується лінійною обчислювальною складністю, високою швидкістю і довільною скінченною роздільною здатністю. Практична швидкість функціонування мережі та її роздільна здатність обмежуються програмною і апаратною реалізацією мережі. Мережа може бути реалізована у паралельному програмному або апаратному забезпеченні. Програмне забезпечення може бути вибрано, якщо необхідна точна, проста і гнучка сортувальна система. Мережу можна реалізувати у цифровому апаратному забезпеченні, якщо необхідна швидкісна, надійна і завадостійка система сортування. Представлену мережу можна рекомендувати для використання при невідомому діапазоні зміни невідомих вхідних даних і відомій мінімальній

різниці між вхідними даними, якщо необхідно сортувати велику кількість вхідних даних із високою роздільною здатністю.

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.

Список літератури

1. Knuth D. E., *The Art of Computer Programming, Sorting and Searching*. Reading, MA: Addison-Wesley, 1973.
2. Akl S. G., *Parallel Sorting Algorithms*, Orlando, FL: Academic, 1985.
3. Atkins M., "Sorting by Hopfield nets", in *Proc. Int. Joint Conf. Neural Netw.*, Washington, DC, USA, 1989, pp. 65–68.
4. Takefuji Y. and Lee K.-S., "A super parallel sorting algorithm based on neural networks", *IEEE Trans. Circuits Syst.*, vol. CAS-37, 1990. no. 11, pp. 1425–1429.
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. Kwon T. M. and Zervakis M., „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. Tseng Y.-H. and Wu J.-L., "Solving sorting and related problems by quadratic perceptrons", *Electron. Lett.*, 1992. vol. 28, no. 10, pp. 906–908.
8. Wang J., "Analysis and design of an analog sorting network," *IEEE Trans. Neural Networks*, 1995. vol. 6, no. 4, pp. 962–971, Jul.
9. Kwon T. M. and Zervakis M., "KWTA networks and their applications", *Multidimensional Syst. and Signal Processing*, 1995. vol. 6, no. 4, pp. 333–346, Oct.
10. Wang J., "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*. Sept. 2010. vol. 21, no. 9, pp. 1496–1506.
11. Alnuweiri H. M. and Kumar V. K. P., "Optimal VLSI sorting with reduced number of processors", *IEEE Trans. Comput.*, 1991. vol. C-40, pp. 105–110.
12. Rovetta S. and Zunino R., "Minimal-connectivity programmable circuit for analog sorting", *IEE Proc. Circuits, Devices Syst.*, vol. 146, no. 3, pp. 108–110, Aug. 1999.
13. Cichocki A. and Unbehauen R., *Neural Networks for Optimization and Signal Processing*. New York, NY, USA: Wiley, 1993.
14. Tymoshchuk P. V., "A discrete-time dynamic K -winners-take-all neural circuit", *Neurocomputing*, vol. 72, 2009, pp. 3191–3202.

**SIMPLIFIED PARALLEL SORTING DISCRETE-TIME
NEURAL NETWORK MODEL**

P. Tymoshchuk

Lviv Polytechnic National University,
Department of Computer-Aided Design Systems

© *Tymoshchuk P., 2020*

A model of parallel sorting neural network of discrete-time has been proposed. The model is described by system of difference equations and by step functions. The model is based on simplified neural circuit of discrete-time that identifies maximal/minimal values of input data and is described by difference equation and by step functions. A bound from above on a number of iterations required for reaching convergence of search process to steady state is determined. The model does not need a knowledge of change range of input data. In order to use the model a minimal difference between values of input data should be known. The network can process unknown input data with finite values, located in arbitrary unknown finite range. The network is characterized by moderate computational complexity and complexity of software implementation, any finite resolution of input data, speed,. Computing simulation results illustrating efficiency of the network are given.

Keywords: — Parallel sorting, neural network, difference equation, computational complexity, hardware implementation.