

3. Розроблено структуру бази даних з підприємств, які потенційно можуть увійти до складу інноваційної моделі виробничого комплексу.
4. Визначено критерії ефективності роботи створеної структури.
5. Розроблено інформаційну технологію управління інноваційною моделлю на основі процесного підходу стандартів ISO серії 9000.

1. Закон України «Про затвердження Загальноцільової економічної програми розвитку промисловості до 2017 року». Міністерство промислової політики <http://www.industry.gov.ua>. 2. Ткаченко Н. Використання процесного підходу стандартів ISO для забезпечення регулярного виробничого циклу промислового підприємства // Вісник Нац. ун-ту «Львівська політехніка»: «Комп'ютерні науки та інформаційні технології». – № 604. – 2008. – С. 53–57. 3. Ткаченко Н. Аспекти процесного підходу при побудові інноваційної моделі виробничого комплексу // Вісник Нац. ун-ту «Львівська політехніка»: Комп'ютерні науки та інформаційні технології» № 616, 2008р.С.123-128. 4. Проектування структури інноваційної моделі вітчизняного виробничого комплексу // Вісник Нац. ун-ту «Львівська політехніка»: «Комп'ютерні науки та інформаційні технології» № 663. – 2010. – С.133–138. 5. Dragan Y., Tkachenko N. Production Complex Innovative Structure Project Proceeding of the VI International Conference MEMSTECH-2010. – Lviv-Polyana, 2010. – P. 150–151. 6. Федорчук Е., Ткаченко Н. Організація та дослідження інноваційної структури складного виробничого комплексу // Вісник Нац. ун-ту «Львівська політехніка» «Комп'ютерні науки та інформаційні технології». – № 651. – 2009. – С.10–14. 7. Федорчук Е., Ткаченко Н. Алгоритм управління інноваційною моделлю виробничого комплексу // Матеріали Міжнародної науково-практичної конференції «Інформаційні технології та комп'ютерна інженерія». – Вшниця, 2010. – С. 438–440. 8. Ткаченко Н., Федорчук Е. Формування показників якості технологічного процесу // Матеріали V міжнародної конференції CADSM 2009. – Львів-Поляна. – С. 34–36.

УДК 681.14

<sup>1</sup>В. Коцовський, <sup>1</sup>Ф. Гече, <sup>2</sup>А. Батюк

<sup>1</sup>Ужгородський національний університет,

<sup>2</sup> Національний університет “Львівська політехніка”,  
кафедра автоматизованих систем управління

## ОЦІНКИ ВЕЛИЧИНИ ЦІЛОЧИСЛОВИХ ВАГОВИХ КОЕФІЦІЄНТІВ ДВОПОРОГОВИХ НЕЙРОННИХ ЕЛЕМЕНТІВ

© Коцовський В., Гече Ф., Батюк А., 2011

Розглянуто питання, пов'язані з класифікацією елементів скінченних множин за допомогою двопорогових нейронних елементів (ДНЕ). Наведено оцінки величини цілочислових вагових коефіцієнтів ДНЕ. Показано, що середній об'єм пам'яті, необхідний для збереження вагових коефіцієнтів  $n$ -місної порогової (двопорогової) булевої функції, є величиною порядку  $\Omega(n^2)$ .

We study finite set dichotomies on bithreshold neurons. We also give bounds on size of the integral weight-vector of BN for Boolean threshold functions similar to the same for threshold neurons. We prove that upon the average one need  $\Omega(n^2)$  bits for store weight-vector of BN for arbitrary Boolean threshold function.

Нейромережі, побудовані з нейронних елементів (НЕ), успішно використовуються для розв'язування широкого кола практичних задач [1]. При цьому залежно від специфіки задач вибирають ту чи іншу активаційну функцію НЕ. З огляду на ефективну апаратну чи програмну реалізацію найбільший інтерес становить вивчення НЕ з цілочисловими ваговими коефіцієнтами. В

зв'язку з цим виникає питання про обмеження на значення цілочислових вагових коефіцієнтів. Для випадку НЕ з пороговою функцією активації Мурога [2] встановив, що для реалізації  $n$ -місних порогових булевих функцій (ПБФ) у базисі  $Z_2 = \{0,1\}$  досить обмежитися НЕ, цілочислові вагові коефіцієнти яких задовольняють умову  $|w_i| \leq (n+1)^{(n+1)/2}$ . Пізніше для базису  $E_2 = \{-1,1\}$  цю оцінку було покращено до

$$|w_i| \leq 2^{-n} (n+1)^{(n+1)/2}. \quad (1)$$

Також були наведені приклади ПБФ, для реалізації яких потрібні НЕ з цілочисловими вагами, не меншими за  $c^n$ , де  $c = (1 + \sqrt{5})/2$ . Довгий час не вдавалося усунути різницю між верхньою і нижньою оцінками. Нарешті, у 1994 р. Хастад [3] навів приклад  $n$ -місних ПБФ, усі цілочислові вагові коефіцієнти яких задовольняють нерівність

$$|w_i| \geq \frac{1}{2n} e^{-4nb} 2^{(n \log_2 n)/2 - n}, \quad (2)$$

де  $b = \log_2 \frac{3}{2}$ , а  $n$  — степінь двійки ( $n \geq 8$ ). Також у роботі [4] показано, що середнє (по усіх ПБФ) значення найбільшого за модулем вагового коефіцієнта оптимального цілочислового НЕ обмежене знизу величиною  $2^{n/2}$ .

У роботі розглянуто питання реалізації БФ на двопорогових нейронних елементах (ДНЕ) і показано, що для вагових коефіцієнтів двопорогових нейронних елементів (ДНЕ) справедливі оцінки, аналогічні до (1)-(2). Також ми доведемо, що для довільного  $a \in (0,1)$ , починаючи з деякого  $n$ , середнє значення цілочислових коефіцієнтів оптимального НЕ або ДНЕ не менше ніж  $2^{an}$ , що є покращенням результату роботи [4].

Двopоговим дійсним нейронним елементом з ваговим вектором  $\mathbf{w} \in \mathbb{R}^n$  і порогоми  $t_1, t_2 \in \mathbb{R}$  ( $t_1 < t_2$ ) називатимемо логічний елемент з  $n$  дійсними входами  $x_1, x_2, \mathbf{K}, x_n$  та одним виходом  $y \in \{a, b\}$ , ( $a < b$ ), поведінка якого описується співвідношеннями:

$$y = \begin{cases} a, & \text{якщо } t_1 < (\mathbf{w}, \mathbf{x}) < t_2, \\ b, & \text{якщо } (\mathbf{w}, \mathbf{x}) \leq t_1 \text{ або } (\mathbf{w}, \mathbf{x}) \geq t_2. \end{cases}$$

Надалі обмежимося випадками  $\{a, b\} = Z_2$  або  $\{a, b\} = E_2$ . Якщо вважати, що  $t_1 = -\infty$ , то ми отримуємо звичайний НЕ.

Нехай  $A$  — довільна скінченна множина у просторі  $\mathbb{R}^n$ . Тоді кожному ДНЕ із вектором структури  $(\mathbf{w}, t_1, t_2)$  можна поставити у відповідність таке розбиття  $(A^+, A^-)$  множини  $A$ :  $A^- = \{\mathbf{x} \in A \mid t_1 < (\mathbf{w}, \mathbf{x}) < t_2\}$ ,  $A^+ = A \setminus A^-$ , яке ми будемо називати д-розбиттями. Для практичних застосувань важливим є випадок  $A = Z_2^n$  або  $A = E_2^n$ . Вважатимемо, що булева функція  $f(x_1, \mathbf{K}, x_n)$  є двопороговою (ДПБФ), якщо знайдеться такий ДНЕ із структурою  $(\mathbf{w}, t_1, t_2)$ , що  $f(\mathbf{x}) = 0 \Leftrightarrow t_1 < (\mathbf{w}, \mathbf{x}) < t_2$ .

Добре відомо, що будь-яке д-розбиття скінченних множин можна виконати на ДНЕ, коефіцієнти вектора структури якого є цілими числами. Для практичних застосувань цікавим є питання про обмеження, які задовольняють ці коефіцієнти.

Нехай ДНЕ із цілочисловим вектором структури  $(\mathbf{w}, t_1, t_2) \in Z^{n+2}$  здійснює д-розбиття  $(A^+, A^-)$  скінченної множини  $A \subset Z^n$ . Тоді

$$\begin{aligned} t_1 + 1 &\leq (\mathbf{w}, \mathbf{x}) \leq t_2 - 1, & \mathbf{x} \in A^-, \\ (\mathbf{w}, \mathbf{x}) &\leq t_1 \text{ або } (\mathbf{w}, \mathbf{x}) \geq t_2, & \mathbf{x} \in A^+. \end{aligned}$$

За допомогою нескладних перетворень попередню систему можна звести до системи

$$(\mathbb{W}, \mathbb{G}) \leq b_k, \quad k = 1, \mathbf{K}, m, \quad (3)$$

де  $\mathbb{W} = (w_1, \mathbf{K}, w_m, t_1, t_2)$ ,  $\mathbb{G} = (a_k x_1^k, \mathbf{K}, a_k x_n^k, b_k, g_k)$ ,  $b_k, b_k, g_k \in \{-1, 0, 1\}$ ,  $b_k \cdot g_k = 0$ ,  $a_k \in E_2$ ,  $m = \text{Card } A$ ,  $k = 1, \mathbf{K}, m$ .

Теорема 1. Довільне д-розбиття  $(A^+, A^-)$  скінченної множини  $A \subset Z^n$  можна здійснити за допомогою ДНЕ з таким цілочисловим вектором структури  $(\mathbf{w}, t_1, t_2) \in Z^{n+2}$ , що

$$\|(\mathbf{w}, t_1, t_2)\|_\infty = \max\{w_1, \mathbf{K}, |w_n|, |t_1|, |t_2|\} \leq \left(\max_{x \in A} \|x\|^2 + 2\right)^{(n+1)/2}, \quad (4)$$

де  $\|\mathbf{x}\|$  — звичайна евклідова норма вектора  $x$ .

Доведення. Система нерівностей (3) має вузлову підсистему [5]. Тому знайдеться такий розв'язок системи (3), який перетворює на рівність  $n+2$  нерівності з лінійно незалежними правими частинами. Тобто знайдеться вектор  $\mathfrak{W} \in R^{n+2}$ , який задовольняє систему

$$C \mathfrak{W}^T = \mathbf{b}, \quad (5)$$

де

$$C = \begin{pmatrix} \mathfrak{W}_1^i & \mathbf{K} & \mathfrak{W}_{n+2}^i \\ \mathbf{M} & \mathbf{O} & \mathbf{M} \\ \mathfrak{W}_1^{k_{n+2}} & \mathbf{K} & \mathfrak{W}_{n+2}^{k_{n+2}} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_{k_1} \\ \mathbf{M} \\ b_{k_{n+2}} \end{pmatrix}$$

і ДНЕ із структурою  $\mathfrak{W}$  здійснює д-розбиття  $(A^+, A^-)$ . Розв'язок системи (5) з невідродженою квадратною матрицею  $C$  будемо шукати за правилом Крамера

$$\mathfrak{W}_i = \frac{\det C_i}{\det C}, \quad i = 1, \mathbf{K}, n+2,$$

де  $C_i$  — матриця, яку отримують з матриці  $C$  заміною  $i$ -го стовпчика стовпчиком вільних членів  $\mathbf{b}$ . Оскільки знаменники у формулах для  $\mathfrak{W}_i$  однакові, то, помноживши кожний  $\mathfrak{W}_i$  на  $\det C$ , одержимо цілочисловий вектор структури  $(\mathbf{w}, t_1, t_2)$  ДНЕ, який здійснює д-розбиття  $(A^+, A^-)$ .

Оцінимо значення визначника матриці  $C_i$ . Для цього скористаємося відомою нерівністю Адамара [6], згідно з якою для довільної квадратної комплексної матриці  $D = (d_{ij})_{i,j=1,\overline{n}}$

$$\begin{vmatrix} d_{11} & \mathbf{K} & d_{1n} \\ \mathbf{M} & \mathbf{O} & \mathbf{M} \\ d_{n1} & \mathbf{K} & d_{nn} \end{vmatrix}^2 \leq \sum_{i=1}^n |d_{1i}|^2 \sum_{i=1}^n |d_{2i}|^2 \mathbf{K} \sum_{i=1}^n |d_{ni}|^2,$$

причому рівність справедлива лише для унітарних матриць. Тоді

$$\det^2 C_i \leq \prod_{k=1}^{n+2} (\|\mathfrak{W}_k\|^2 + 1) \leq \left(\max_{x \in A} \|x\|^2 + 2\right)^{n+2}.$$

Тому

$$\max\{w_1, \mathbf{K}, |w_n|, |t_1|, |t_2|\} \leq \left(\max_{x \in A} \|x\|^2 + 2\right)^{(n+2)/2}.$$

Теорема доведена.

Наслідок 1. Для довільної  $n$ -місної ДПБФ в алфавіті  $Z_2$

$$\max\{w_1, \mathbf{K}, |w_n|\} \leq \sqrt{(n+1)n^{n+1}}, \quad \max\{|t_1|, |t_2|\} \leq \sqrt{(n+2)(n+1)^{(n+1)}}.$$

Доведення випливає з того, що при  $1 \leq i \leq n$  усі рядки матриці  $C_i$ , крім щонайбільше одного, містять не менше ніж два нулі, а у матрицях  $C_{n+1}$  і  $C_{n+2}$  може бути не більше від одного рядка, усі елементи якого відмінні від 0.

Наслідок 2. Для довільної  $n$ -місної ДПБФ в алфавіті  $E_2$

$$\max\{w_1, \mathbf{K}, |w_n|, |t_1|, |t_2|\} \leq 2^{-n} (n+2)^{(n+2)/2}.$$

Доведення випливає з (4) та з того, що визначник довільної  $\{-1,1\}$ -матриці  $n \times n$  націло ділиться на  $2^{n-1}$ , а тому чисельник і знаменник дробу у виразі для  $\mathfrak{W}_i$  можна скоротити на  $2^n$ . Отримана оцінка аналогічна до (1).

Попередня теорема показує, що для зберігання усіх коефіцієнтів вектора структури ДНЕ, на якому можна реалізувати довільну ДПБФ, потрібно не більше ніж  $O(n^2 \log n)$  бітів пам'яті.

Покажемо, що знайдеться ДПБФ, для збереження вектора структури якої знадобиться  $\Omega(n^2 \log n)$  біт пам'яті.

Лема 1. Якщо для довільної ПБФ  $f(x_1, \mathbf{K}, x_n)$  виконується умова  $\text{Card } A^- \geq 2^{n-1}$ , де  $A^- = \{\mathbf{x} \in E_2^n \mid f(\mathbf{x}) = -1\}$ , то для довільного ДНЕ із вектором структури  $(\mathbf{w}, t_1, t_2)$ , який здійснює д-розбиття  $(A^+, A^-)$  множини  $E_2^n$ , одна з множин  $\{\mathbf{x} \in A^+ \mid (\mathbf{w}, \mathbf{x}) \leq t_1\}$  або  $\{\mathbf{x} \in A^+ \mid (\mathbf{w}, \mathbf{x}) \geq t_2\}$  є порожньою.

Доведення. Нехай ДНЕ із структурою  $(\mathbf{w}, t_1, t_2)$  здійснює д-розбиття  $(A^+, A^-)$ . Припустимо, що  $\{\mathbf{x} \in A^+ \mid (\mathbf{w}, \mathbf{x}) \leq t_1\} \neq \emptyset$  і  $\{\mathbf{x} \in A^+ \mid (\mathbf{w}, \mathbf{x}) \geq t_2\} \neq \emptyset$ . Оскільки  $(\mathbf{w}, \mathbf{x}) > (\mathbf{w}, \mathbf{y}) \Leftrightarrow (\mathbf{w}, -\mathbf{x}) < (\mathbf{w}, -\mathbf{y})$ , то знайдеться такий  $\mathbf{x} \in A^+$ , що  $-\mathbf{x} \in A^+$ . Тоді  $\text{Card } A^+ > 2^{n-1}$ , що суперечить умовам леми.

Зауваження. Доведена лема вказує умови, виконання яких забезпечує відсутність «відмінностей» між лінійними і двопороговими розбиттями множини  $Z^n$ . Вона перестане бути правильною для множин, довільно лінійно розділених у просторі  $R^n$ . Нехай

$$A^- = \{(1,0), (2,0)\}, \quad A^+ = \{(1,-2), (0, n-2)\}, \quad n \in \mathbb{N}.$$

Тоді лінійне розбиття  $(A^+, A^-)$  можна задати за допомогою ДНЕ із вектором структури  $((0,1), -1, 1)$ , а для довільного НЕ з ваговим вектором  $\mathbf{w} \in Z^n$ , який задає це розбиття, виконується умова  $\|\mathbf{w}\|_\infty \geq n$ .

Теорема 2. Якщо  $n = 2^k, k \geq 3$  то для довільного ДНЕ із вектором структури  $(\mathbf{w}, t_1, t_2)$ , який реалізує ПБФ Хастада [3], виконується нерівність

$$\|\mathbf{w}\|_\infty \geq \frac{1}{2n} e^{-4nb} 2^{(n \log_2 n)/2-n},$$

де  $b = \log_2 \frac{3}{2}$ .

Доведення. Функція Хастада є непарною [3]. Тому для неї  $\text{Card } A^- = 2^{n-1}$ . За лемою 1 для всіх ДНЕ  $(\mathbf{w}, t_1, t_2)$ , які реалізують цю функцію, один з порогів є зайвим. Тоді твердження теореми випливає з (2).

Виникає питання про середній об'єм пам'яті, необхідний для збереження цілочислових вагових коефіцієнтів НЕ і ДНЕ. Покажемо, що в середньому для цього необхідно щонайменше  $\Omega(n^2)$  бітів. Для цього розглянемо сумарне та середнє значення коефіцієнтів цілочислового вектора структури НЕ:

$$S(\mathbf{w}, t) = \sum_{i=1}^n |w_i| + |t|, \quad E(\mathbf{w}, t) = \frac{S(\mathbf{w}, t)}{n+1}$$

і нехай  $E(LT_n) = \frac{1}{\text{Card } LT_n} \sum_{f \in LT_n} E(\mathbf{w}_f, t_f)$ , де  $LT_n$  — множина усіх  $n$ -місних ПБФ,  $(\mathbf{w}_f, t_f)$  — «мінімальний цілочисловий вектор структури»  $n$ -місної ПБФ  $f$  (тобто такий вектор структури, для якого  $S(\mathbf{w}, t)$  набуває найменшого значення). Величина  $E(LT_n)$  — математичне сподівання середнього арифметичного коефіцієнтів мінімальних цілочислових векторів структури  $n$ -місних ПБФ. Існує така теорема:

Теорема 3. Для довільного  $a \in (0, 1)$  знайдеться таке натуральне  $n_0(a)$ , що для всіх  $n \geq n_0(a)$  справджується нерівність  $E(LT_n) > 2^{an}$ .

Доведення. Візьмемо довільне  $b$ , яке задовольняє умову  $a < b < 1$ . Підрахуємо  $N(b)$  — кількість тих  $n$ -місних ПБФ  $f$ , мінімальні цілочислові вектори структури яких задовольняють умову

$S(\mathbf{w}_f, t_f) \leq 2^{bn} \cdot (n+1)$ . Ця кількість не перевищує  $2^{n+1} \sum_{m=1}^{\lfloor 2^{bn(n+1)} \rfloor} Q_m$ , де  $Q_m$  — кількість різних невід'ємних цілочислових векторів структур  $(\mathbf{w}, t)$ , які задовольняють умову  $S(\mathbf{w}, t) = m$ , а  $[x]$  —

позначення для цілої частини числа  $x$ . У попередній сумі множник  $2^{n+1}$  дорівнює кількості способів вибору знаків коефіцієнтів. За відомою комбінаторною властивістю [7]  $Q_m = C_{m+n}^n$ . Тому

$$\begin{aligned} \frac{N(\mathbf{b})}{2^{n+1}} &\leq \sum_{m=1}^{\lfloor 2^{bn}(n+1) \rfloor} Q_m < \sum_{m=n}^{\lfloor 2^{bn}(n+1)+n \rfloor} C_m^n = C_{\lfloor 2^{bn}(n+1) \rfloor + n + 1}^{n+1} = \frac{\left(\lfloor 2^{bn}(n+1) \rfloor + n + 1\right) \mathbf{K}\left(\lfloor 2^{bn}(n+1) \rfloor + 1\right)}{(n+1)!} < \\ &< \left(2^{bn}(n+1)\right)^{n+1} = 2^{bn^2 + bn + (n+1)\log_2(n+1)} = 2^{bn^2 + O(\log_2 n)}. \end{aligned}$$

Тоді і  $N(\mathbf{b}) = 2^{bn^2 + O(\log_2 n)}$ . Звідси

$$E(LT_n) = \frac{1}{\text{Card } LT_n} \sum_{f \in LT_n} E(\mathbf{w}_f, t_f) > \frac{N(\mathbf{b})}{(n+1)\text{Card } LT_n} + 2^{bn} \left(1 - \frac{N(\mathbf{b})}{\text{Card } LT_n}\right).$$

Відомо [8], що  $\text{Card } LT_n = 2^{n^2 - n \log n + O(n)}$ . Тоді для достатньо великих  $n$  (у разі потреби можна явно вказати вираз для функції  $n_0(a)$ ) виконується нерівність  $E(LT_n) > 2^{bn} \left(1 - 2^{-(1-b)n^2 + O(n \log_2 n)}\right) > 2^{an}$ .

Зауваження 1. Якщо ввести позначення  $S(\mathbf{w}, t_1, t_2) = \sum_{i=1}^n |w_i| + |t_1| + |t_2|$ ,

$$E(\mathbf{w}, t_1, t_2) = \frac{S(\mathbf{w}, t_1, t_2)}{n+2}, \quad E(LBT_n) = \frac{1}{\text{Card } LBT_n} \sum_{f \in LBT_n} E(\mathbf{w}_f, t_{1f}, t_{2f}),$$

де  $LBT_n$  — множина усіх  $n$ -місних ДПБФ, то майже аналогічно до доведення теореми 3 з урахуванням оцінки  $\text{Card } LBT_n = 2^{n^2 - n \log n + O(n)}$ , отримуємо, що для довільного  $a \in (0, 1)$ , починаючи з деякого номера  $n_0(a)$ ,  $E(LBT_n) > 2^{an}$ .

Зауваження 2. Теорема 3 є значним підсиленням відомого результату роботи [4], де було встановлено, що математичне сподівання найбільшого за модулем цілочислового вагового коефіцієнта не менше за  $2^{n/2}$ .

Зауваження 3. Під час доведення теореми 3 виявлено цікавий факт: для довільного  $a \in (0, 1)$ , починаючи з деякого  $n_0(a)$ , майже для всіх  $n$ -місних ПБФ  $f$  їхні мінімальні цілочислові вектори структури задовольняють умову  $E(\mathbf{w}_f, t_f) > 2^{an}$ .

1. Хайкин С. *Нейронные сети: полный курс* / С. Хайкин. — 2-е изд. — М.: Вильямс-Телеком, 2006. — 1104 с.
2. Muroga S. *Threshold Logic and its Applications* / S. Muroga. — New York: Wiley, 1971.
3. Hastad J. *On the size of weights for threshold gates* / J. Hastad // *SIAM Journal on Discrete Mathematics*. — 1994, 7(3). — PP. 484-492.
4. Hampson S. E. *Linear Function Neurons: Structure and Training* / S. E. Hampson & D. J. Volper // *Biol. Cyber.* — 1986 vol. 53. — PP. 203–217.
5. Черников С. Н. *Линейные неравенства* / С. Н. Черников. — М.: Наука, 1968. — 488 с.
6. Хорн Р. *Матричный анализ* / Р. Хорн, Ч. Джонсон. — М.: Мир, 1989. — 656 с.
7. Рыбников К. А. *Введение в комбинаторный анализ* / К. А. Рыбников. — М.: Изд-во МГУ, 1985. — 308 с.
8. Kahn J. *On the probability that a random  $\{\pm 1\}$   $n$ -matrix is singular* / Kahn J., Komlos J., Szemerédi E.J. // *Amer. Math. Soc.* — 1995 vol. 8, №1. — PP. 223–240.