УДК 681.14:517.11

Ф.Е. Гече, В.М. Коцовський, А.Є. Батюк

ПРО ДЕЯКІ КРИТЕРІЇ ПОРОГОВОСТІ БУЛЬОВИХ ФУНКЦІЙ.

В роботі, за допомогою аналізу спектральних характеристик досліджуваних функцій, встановлюються нові та спрощуються деякі відомі критерії реалізовності бульових функцій на одному нейронному елементі. Це приводить до суттєвого скорочення кількості операцій при перевірці пороговості бульових функцій.

In the paper with the help of the analysis of spectral characteristics of investigated function new criteria are established and some known criteria of realizability of Boolean function on one neuron element are simplified. This leads to a great decreasing of the quantity of operations while checking of threshold of Boolean function.

Нехай $f(x_1, \dots, x_n)$ – n-місна бульова функція, $f^{-1}(1) = \{\boldsymbol{X} \in Z_2^n \mid f(\boldsymbol{X}) = 1\}$ і $f^{-1}(0) = \{\boldsymbol{X} \in Z_2^n \mid f(\boldsymbol{X}) = 0\}$. Кажуть, що нейронний елемент з вектором структури $[\boldsymbol{w}; T]$ (де $\boldsymbol{w} = (\omega_1, \dots, \omega_n)$ - n-вимірний дійсний вектор, який називається ваговим вектором, T - дійсне число, яке називається порогом) реалізує бульову функцію $f(\boldsymbol{X})$, якщо для всіх $\boldsymbol{X} \in Z_2^n$ виконується умова $\boldsymbol{X} \in f^{-1}(0) \Leftrightarrow (\boldsymbol{w}, \boldsymbol{X}) < T$.

Бульова функція $f(\boldsymbol{X})$, яка реалізується на одному нейронному елементі називається пороговою(нейрофункцією). Для введення спектральних характеристик зручно розглядати бульові функції в алфавіті $\{-1, 1\}$. Нехай $G_n = \{-1, 1\}^n$. Бульовій функції $f(\boldsymbol{X})$ у алфавіті $\{0, 1\}$ поставимо у відповідність бульову функцію $g_f(\boldsymbol{Y})$ в алфавіті $\{-1, 1\}$, де $\boldsymbol{Y} = 2\boldsymbol{X} - 1$, $g_f(\boldsymbol{Y}) = 2f(\boldsymbol{X}) - 1$. Нехай $\boldsymbol{a} = (\alpha_1, \dots, \alpha_n) \in G_n$. Покладемо

$y_i(\boldsymbol{a}) = \alpha_i$, якщо $i \in \{1,2,\dots,n\}$ і $y(\boldsymbol{a}) = 1$, якщо $i = 0$. Визначимо характеристичний вектор $\boldsymbol{b}_f = (b_1, b_2, \dots, b_n; b_0)$ бульової функції f за формулами $b_i = \sum_{\boldsymbol{a} \in G_n} y_i(\boldsymbol{a}) g_f(\boldsymbol{a}), i = \overline{0, n}$.

Нехай відображення $\rho : R^n \times Z_2^n \times Z_2^n \to R$ визначається так: $\rho(\boldsymbol{w}, \boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{w}, \boldsymbol{X}) - (\boldsymbol{w}, \boldsymbol{Y})$. Введемо в розгляд множину $\Omega_n = \{\boldsymbol{w} \in R^n \mid \rho(\boldsymbol{w}, \boldsymbol{X}, \boldsymbol{Y}) = 0 \Leftrightarrow \boldsymbol{X} = \boldsymbol{Y}$ для всіх $\boldsymbol{X}, \boldsymbol{Y} \in Z_2^n\}$. В [1] показано, що розгляд нейронних елементів з ваговими векторами з множини Ω_n не звужує класу порогових функцій. Нехай $\boldsymbol{w} \in \Omega_n$, $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_{2^n}$ – послідовність бульових векторів, таких , що $\rho(\boldsymbol{w}, \boldsymbol{X}_i, \boldsymbol{X}_j) > 0$ при $i > j$. В [2] показано, що перші 2^{n-1} вектори, які входять у наведену послідовність є попарно толерантними відносно відношення толерантності, визначеного на множині Z_2^n так:

$(x_1, \dots, x_n)\ \tau\, (y_1, \dots, y_n) \Leftrightarrow$ знайдеться таке i, що $x_i = y_i$.

Поставимо вектору $\boldsymbol{w} \in \Omega_n$ у відповідність матрицю толерантності L_w розмірності $2^{n-1} \times n$, де

$L_w = \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_{2^{n-1}} \end{pmatrix}$. Позначимо $E_n = \bigcup_{w \in \Omega_n} L_w$. Вектори $\boldsymbol{v}, \boldsymbol{w} \in \Omega_n$ назвемо еквіалентними, якщо для всіх $\boldsymbol{X}, \boldsymbol{Y}$ виконується умова $sgn(\rho(\boldsymbol{v}, \boldsymbol{X}, \boldsymbol{Y})) = sgn(\rho(\boldsymbol{w}, \boldsymbol{X}, \boldsymbol{Y}))$. Якщо $\boldsymbol{v}, \boldsymbol{w}$ – еквіалентні, то очевидно $L_v = L_w$.

Нехай σ – елемент групи перестановок S_n, $\boldsymbol{g}$ – n-вимірний бульовий вектор. Позначимо результат дії елементів σ та $\boldsymbol{g}$ на вектор $\boldsymbol{w} \in \Omega_n$ і бульову функцію f відповідно як $\boldsymbol{w}^{\sigma}$, $\boldsymbol{gw}$, f^{σ}, $\boldsymbol{g}f$, де:

$\boldsymbol{w}^{\sigma} = (\omega_{\sigma(1)}, ..., \omega_{\sigma(n)})$, $f^{\sigma}(x_1, ..., x_n) = f(x_{\sigma(1)}, ..., x_{\sigma(n)})$, $\boldsymbol{g}f(x_1, ..., x_n) = f(x_1 \oplus g_1, ..., x_n \oplus g_n)$,

$(\boldsymbol{gw})_i = \omega_i$, якщо $g_i = 0$ і $(\boldsymbol{gw})_i = -\omega_i$, якщо $g_i = 1$. Нехай $\boldsymbol{v} = \boldsymbol{w}^{\sigma}$, $\boldsymbol{u} = \boldsymbol{gw}$. Тоді

$$\rho(\boldsymbol{v}, \boldsymbol{X}^{\sigma}, \boldsymbol{Y}^{\sigma}) = \rho(\boldsymbol{w}, \boldsymbol{X}, \boldsymbol{Y}), \tag{1}$$

$$\rho(\boldsymbol{u}, \boldsymbol{X} \oplus \boldsymbol{g}, \boldsymbol{Y} \oplus \boldsymbol{g}) = \rho(\boldsymbol{w}, \boldsymbol{X}, \boldsymbol{Y}). \tag{2}$$

Рівність (1) очевидна. Справедливість (2) випливає зі співвідношень

$\rho(\boldsymbol{u}, \boldsymbol{X} \oplus \boldsymbol{g}, \boldsymbol{Y} \oplus \boldsymbol{g}) = \sum_{i:g_i=0} \omega_i (x_i - y_i) + \sum_{i:g_i=1} -\omega_i (\overline{x_i} - \overline{y_i}) = \sum_{i=1}^{n} \omega_i (x_i - y_i) = \rho(\boldsymbol{w}, \boldsymbol{X}, \boldsymbol{Y})$. З (1), та (2) випливає, що $L_v = L_w^{\sigma} = \begin{pmatrix} \boldsymbol{X}_1^{\sigma} \\ \vdots \\ \boldsymbol{X}_{2^{n-1}}^{\sigma} \end{pmatrix}$, $L_u = \boldsymbol{g} L_w = \begin{pmatrix} \boldsymbol{X_1} \oplus \boldsymbol{g} \\ \vdots \\ \boldsymbol{X}_{2^{n-1}} \oplus \boldsymbol{g} \end{pmatrix}$. Якщо $h = (\boldsymbol{g}f)^{\sigma}$, то як показано в [3]:

$$\boldsymbol{b}_h = (\boldsymbol{g}\boldsymbol{b}_f)^{\sigma}. \tag{3}$$

Нехай $\boldsymbol{b}_f = (b_1, b_2, ..., b_n; b_0)$ – характеристичний вектор бульової функції $f(x_1, ..., x_n)$. Тоді якщо f – порогова, $\boldsymbol{w}$ – ваговий вектор нейронного елемента, який реалізує f, то справедливі співвідношення[3]:

$$sgn\, b_i = sgn\, \omega_i, \quad \text{якщо } b_i \neq 0, \tag{4}$$

$$b_i > b_j \Rightarrow \omega_i > \omega_j. \tag{5}$$

Якщо $b_i = 0$, то очевидно, що для порогової функції можна покласти $\omega_i = 0$. Тому задача перевірки реалізовності бульової функції $f(x_1, ..., x_n)$ зводиться до задачі встановлення пороговості деякої $(n - k)$-місної бульової функції, де k – кількість нульових координат серед перших n координат характеристичного вектора $\boldsymbol{b}_f$.

Розглянемо випадок, коли $b_i \neq 0$, $1 \le i \le n$. Побудуємо бульовий вектор $\boldsymbol{g}$ і бульову функцію h за такими правилами:

$$g_i = 0, \text{ якщо } b_i < 0 \text{ і } g_i = 1, \text{ якщо } b_i > 0, \tag{6}$$

$$h = (\boldsymbol{g}f)^{\sigma}. \tag{7}$$

Нехай $\boldsymbol{c}_h = (c_1, c_2, ..., c_n; c_0)$ – характеристичний вектор бульової функції h. Тоді існує

$\sigma \in S_n$, таке, що справедливі нерівності:

$$0 > c_1 \ge c_2 \ge ... \ge c_n. \tag{8}$$

Згідно властивостей інваріантних операцій[3] бульова функція f порогова тоді і тільки тоді, коли пороговою є бульова функція h. З (4), (5), (8) маємо, що бульова функція h порогова тоді і тільки тоді, коли знайдеться нейронний елемент з ваговим вектором з множини $\Omega_n^- = \{\boldsymbol{w} \in \Omega_n \mid 0 > \omega_1 > \omega_2 > ... > \omega_n\}$ і відповідним порогом T, який реалізує дану функцію.

Нехай K і L – матриці, рядками яких є n-вимірні бульові вектори, причому q – кількість рядків матриці K, не перевищує кількості рядків матриці L. Матрицю K називають передматрицею матриці L, якщо в ній можна так переставити рядки, що отримана матриця буде співпадати з матрицею, яка складається з перших q рядків матриці L. Той факт, що K є передматрицею матриці L, будемо позначати $K \Delta L$. Введемо поняття ядра $K(f)$ для бульової функції f, покладаючи $K(f) = f^{-1}(1)$, якщо $|f^{-1}(1)| \leq |f^{-1}(0)|$ і $K(f) =$ $= f^{-1}(0)$ в противному випадку. Запишемо елементи ядра $K(f)$ в деякому порядку в рядки матриці, зберігаючи для отриманої матриці позначення $K(f)$. Якщо бульова функція h визначається з (7), то очевидно

$$K(h) = (\boldsymbol{g}K(f))^{\sigma}. \tag{9}$$

Як показано в [2] бульова функція $f(x_1, ..., x_n)$ з ядром $K(f)$ є пороговою тоді і тільки тоді, коли існує така матриця толерантності $L \in E_n$, що $K(f) \,\Delta\, L$, причому $L = L_w$ у випадку $K(f) = f^{-1}(1)$ і $L = \overline{L}_w$ в протилежному випадку, де L_w – матриця, яка відповідає ваговому вектору $\boldsymbol{w} \in \Omega_n$ деякого нейронного елемента, на якому реалізується бульова функція f, $\overline{L}_w$ – матриця, яка утворилася з матриці L_w шляхом бульового заперечення її векторів-рядків. З (7) та (9) випливає, що $f(x_1, ..., x_n)$ з ядром $K(f)$ є пороговою тоді і тільки тоді, коли знайдуться $\sigma \in S_n$, $L_w \in E_n^-$, такі, що

$$(\boldsymbol{g}K(f))^{\sigma} \,\Delta\, L_w \text{ у випадку } K(f) = f^{-1}(1), \tag{10}$$

$$(\overline{\boldsymbol{g}}K(f))^{\sigma} \,\Delta\, L_w \text{ у випадку } K(f) = f^{-1}(0), \tag{11}$$

де $\boldsymbol{g}$ визначається згідно (6), σ задовольняє (8), а множина матриць толерантності E_n^- визначається так: $E_n^- = \bigcup_{w \in \Omega_n^-} L_w$. Оскільки для всіх $\boldsymbol{w} \in \Omega_n^-, \boldsymbol{X} \in Z_2^n$ $\rho(\boldsymbol{w}, \boldsymbol{0}, \boldsymbol{X}) \geq 0$, то перший рядок матриці толерантності L_w складається з нулів. Тому $\boldsymbol{g} \in K(f)$ у випадку $K(f) = f^{-1}(1)$ і $\overline{\boldsymbol{g}} \in K(f)$ у випадку $K(f) = f^{-1}(0)$. Тому існування у ядрі вектора $\boldsymbol{g}$ (або $\overline{\boldsymbol{g}}$ у випадку $K(f) = f^{-1}(0)$) є необхідною умовою пороговості бульової функції f. Конкретизуємо тепер вигляд перестановки σ, яка фігурує в (7). Очевидно, що у випадку , коли для $i \neq j$, $1 \leq i, j \leq n$ відповідні координати характеристичного вектора співпадають, то перестановка σ, яка задовольняє нерівності (8) визначається неоднозначно. Покажемо, що тоді для перевірок умов (10) або (11) можна взяти довільне $\sigma \in S_n$, яке задовольняє умову (8). Справді, нехай перестановки σ, $\gamma \in S_n$ забезпечують виконання умови (8). Тоді функції $h_1 = (\boldsymbol{g}f)^{\sigma}$ і $h_2 = (\boldsymbol{g}f)^{\gamma}$ мають однакові характеристичні вектори, а тому згідно першої теореми Чоу[3] одночасно порогові або непорогові. Тому перевірка умов (10) або (11) для перестановок σ або γ приведе до аналогічних результатів. Тому вибір перестановки σ з множини перестановок, що задовольняють (8) є довільним. Наведемо остаточне формулювання отриманих результатів:

Теорема 1. *Бульова функція $f(x_1, ..., x_n)$ з ядром $K(f)$ є пороговою тоді і тільки тоді, коли в ядрі знайдеться вектор $\boldsymbol{g}$($\overline{\boldsymbol{g}}$ у випадку $K(f) = f^{-1}(0)$), що визначається за правилом (7), $\sigma \in S_n$, що задовольняє (8) та $L_w \in E_n^-$, такі, що в залежності від вигляду ядра $K(f)$ (прообраз 1 чи 0) виконуються співвідношення (10) або (11).*

***Зауваження* 1**. Аналогічні результати були отримані в [2]. Перевагою наведеного при доведені теореми 1 підходу, який базується на властивостях характеристичних векторів, є той факт, що побудова введеної в [2] так званої множини зведених ядер звелася до побудови одного цілком конкретного зведеного ядра за правилом (7). Крім того вдалося обгрунтувати вибір перестановки σ, яка фігурує в умовах (10)–(11).

Зауваження **2**. Отримані результати можна поширити на випадок останньої координати b_0 характеристичного вектора $\boldsymbol{b}_f$ бульової функції f, якщо використовувати більш широке коло інваріантних операцій, описаних в [3].

Розглянемо випадки, коли деякі особливості у будові характеристичного вектора дозволяють дещо спростити перевірку можливості реалізувати n-місну бульову функцію f на одному нейронному елементі. Під особливостями будемо розуміти рівність деяких координат характеристичного вектора нулю або співпадання деяких координат. Розгляд таких випадків є важливим, оскільки у багатьох бульових функцій наявні згадані властивості. Вище вже було показано, що у випадку нульових координат характеристичного вектора задача перевірки пороговості зводиться до еквіалентної задачі меншої розмірності. Розглянемо випадок, коли k з перших n компонент характеристичного вектора $\boldsymbol{b}_f$ бульової функції $f(x_1, \ldots, x_n)$ співпадають. Будемо вважати, що

$$b_{n-k+1} = b_{n-k+2} = \ldots = b_n \,. \tag{12}$$

Це припущення не обмежує загальності, оскільки в противному випадку можна побудувати перестановку $\sigma \in S_n$, таку, що для характеристичного вектора функції $f^{\sigma}(x_1, \ldots, x_n)$ умова (12) виконується і функції f і f^{σ} одночасно порогові або непорогові.

Нехай $\boldsymbol{a} = (\alpha_1, \ldots, \alpha_k)$, де $1 \le k < n$, $\alpha_i \in Z_2$. Введемо позначення

$f_a(x_1, \ldots, x_{n-k}) = f(x_1, \ldots, x_{n-k}, \alpha_1, \ldots, \alpha_k)$. Якщо бульова функція f порогова, то можна показати, що знайдеться нейронний елемент з таким ваговим вектором, для якого

$$\omega_{n-k+1} = \omega_{n-k+2} = \ldots = \omega_n = \omega \tag{13}$$

і який реалізує бульову функцію f. Визначимо норму Хеммінга на множині $Z_2^n : \| \boldsymbol{a} \| = \sum_{i=1}^{n} \alpha_i$. Тоді з (13) одержується необхідна умова пороговості бульової функції $f(x_1, \ldots, x_n)$ у випадку виконання умови (12):

$$\text{для всіх } \boldsymbol{a}\ \boldsymbol{c} \in Z_2^k : \|\boldsymbol{a}\| = \|\boldsymbol{c}\| \Rightarrow f_a(x_1, \ldots, x_{n-k}) = f_c(x_1, \ldots, x_{n-k}) \,. \tag{14}$$

Крім того, якщо виконується умова (12), то можна дещо послабити вимоги, які накладені в теоремі 4[4] при переході від задачі перевірки пороговості n-місної бульової функції до послідовністі задач розмірності $n - k$:

Теорема 2. *Бульова функція $f(x_1, \ldots, x_n)$ з характеристичним вектором $\boldsymbol{b}_f$, для якого виконується умова (12) реалізується на одному нейронному елементі тоді і тільки тоді, коли:*

1) виконується умова (14),

2) знайдуться такі дійсні числа T і ω і такий вектор $(\omega_1, \ldots, \omega_{n-k})$ з дійсними компонентами, що бульові функції f_s , де $f_s = f(x_1, \ldots, x_{n-k}, 0, \ldots, 0, \underbrace{1, \ldots, 1}_{s})$, $s = \overline{0, k}$ реалізуються на нейронних елементах з одним і тим самим ваговим вектором $(\omega_1, \ldots, \omega_{n-k})$ і відповідними порогами $T_s = T - s\omega$.

Доведення. Необхідна умова теореми випливає з необхідної умови теореми 4[4] та із (12), (13) та (14). Доведемо достатність умов 1–2. Для цього за достатньою умовою теореми 4[4] досить показати існування такого n–вимірного дійсного вектора $\boldsymbol{v}$ і такого дійсного числа T, що виконуються наступні умови:

бульові функції $f_{\boldsymbol{a}}$, де $\boldsymbol{a}$ приймає послідовно всі можливі значення з множини Z_2^k, реалізуються на нейронних елементах з одним і тим самим ваговим вектором (v_1, ..., v_{n-k}) і відповідними порогами $T_{\boldsymbol{a}} = T - \sum_{j=1}^{k} \alpha_j v_{n-k+j}$.Якщо покласти $\boldsymbol{v} = (\omega_1, \ldots, \omega_{n-k}, \omega, \ldots, \omega)$, то виконання згаданих умов забезпечується виконанням умов 1, 2 у формулюванні теореми. Теорема доведена.

***Зауваження* 3**. Теорема 2 дозволяє при виконанні умов (12) звести задачу перевірки пороговості n–місної бульової функції до послідовністі, яка складається з k + 1 задачі розмірності n – k, які пов'язані між собою співідношеннями для порогів, тоді як в загальному випадку згідно теореми 4[4] отримуємо 2^k задач. Крім того справедливим залишається зауваження 2 до теореми 1.

1. *Яджима С., Ибараки Т*. Нижняя оценка числа пороговых функций // Киберн. сб. Нов. сер.–1969. – Вып.6. - С. 72–81.
2. Параллельная обработка информации // Проблемно–ориентированные и специализированные средства обработки информации в 5–ти т./ Под ред *Б. Н. Малиновского и В. В. Грицыка*.- Киев: Наукова думка, 1990.-т. 5.-504 С.
3. *Дертоузос М*. Пороговая логика. – М.: Мир, 1967.-342 С.
4. *Ф. Е. Гече, В. М. Коцовський*. Про деякі властивості бульових функцій, які реалізуються на одному нейронному елементі // Наук. вісник Ужгород. ун-ту, сер. мат. вип. 4, -1999. – С. 25-29.