

## СИНТЕЗ ВИСОКОЕФЕКТИВНИХ БАГАТОШАРОВИХ ПЕРЦЕПТРОНІВ З НЕІТЕРАЦІЙНИМ НАВЧАННЯМ

© Ткаченко Р., Цмоць І., Скорохода О., 2009

Розглянуто модель багатошарових перцептронів з неітераційним навчанням на основі моделі геометричних перетворень, виділено базові операції, вдосконалено метод їх обчислення, розроблено базову структуру пристрою для обчислення оператора сум парних добутків і синтезовано проблемно-орієнтовану систему для реалізації багатошарових перцептронів з неітераційним навчанням.

The model of multilayer perceptron with noniterative learning on the basis of geometrical transformation model was considered, basic operations were emphasized, method of their calculation was improved, the basic structure of device for calculation of even products sums operator was developed and problem-oriented system for realization of multilayer perceptron with noniterative learning was synthesized.

### Вступ

Штучні нейронмережі (ШНМ) є ефективним засобом інформаційного моделювання в умовах невизначеностей, включаючи параметричні та непараметричні. Це зумовило достатнє поширення ШНМ, якими найчастіше є багатошарові перцептрони (БШП), для передбачень, прогнозування, класифікації, ущільнення та інших завдань опрацювання даних. У багатьох випадках БШП не забезпечують розв'язків великорозмірних завдань, яким властиві істотні нелінійності, за прийнятний час, оскільки їхнє навчання реалізується шляхом виконання процедур багатовимірної і багатокритеріальної оптимізації. Подолати вказаний недолік можна, як шляхом розроблення нових нейропарадигм ШНМ з неітераційним навчанням, так і створенням апаратних засобів для їх реалізації [1,2]. У зв'язку з цим актуальною задачею є розробка високопродуктивних ШНМ, які ґрунтуються на нових нейропарадигмах з неітераційним навчанням.

### Постановка завдання

Архітектура БШП повинна бути орієнтована на НВІС-реалізацію та конвеєрно-паралельну обробку інтенсивних потоків даних у реальному часі. Для переходу до НВІС-архітектури доцільно використовувати метод просторово-часового відображення алгоритмів навчання та функціонування БШП на паралельних однорідних структурах з високою ефективністю використання обладнання [3–7]. Схема графу БШП з одним прихованим шаром наведено на рис. 1.

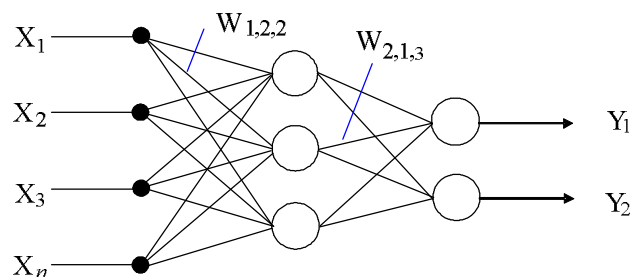


Рис.1. Схема графу БШП з одним прихованим шаром

Апаратна реалізація графів такого типу, включаючи відтворення процесорних функцій нейронних елементів (НЕ), значно прискорює роботу БШП в режимі функціонування, чого не

можна сказати про режим навчання. Останній передбачає почергове пряме та зворотне проходження сигналів відповідно до обраної процедури мінімізації відхилень на виходах мережі від еталонних значень. Отже, апаратне відтворення графу ШНМ не розв'язує задачі кардинального покращання швидкісних характеристик і одночасно може виявитися надто громіздким. У результаті, враховуючи різнотипність базових операцій, відмінності обчислювальних процедур функціонування та навчання БШП зі зворотним поширенням похибок вихідних сигналів, апаратні нейрокомп'ютери подібного типу виявилися малоефективними, дорогими і не отримали поширення.

Тому *мета роботи* полягає у вдосконаленні моделі навчання та функціонування БШП, виділенні базових операцій, розробленні нових алгоритмів і структур для їх реалізації та синтезі високоефективних БШП з неітераційним навчанням.

### Розв'язання задачі

**Модель БШП з неітераційним навчанням.** БШП на основі моделі геометричних перетворень (МГП) призначені для реалізації концепції інформаційного моделювання ("чорна скринька" та "сіра скринька") на принципово новій основі, яка забезпечує розв'язання завдань розпізнавання образів, прогнозування, класифікації, виділення головних (незалежних) компонент, оптимізації, відновлення втрачених даних або їх ущільнення, реалізації методів захисту та конфіденційності інформації, розв'язування систем лінійних алгебраїчних рівнянь, включаючи недовизначені та перевизначені, візуалізації багатовимірних даних тощо. Нейромережі МГП як універсальні апроксиматори реалізують принципи навчання та самонавчання, будуються на підставі алгоритмічних або апаратних варіантів виконання з використанням принципів просторового та часового розпаралелювання. Важливою особливістю БШП на основі МГП потрібно вважати також можливість виконання всіх обчислювальних операцій у форматі чисел з фіксованою комою, що істотно спрощує побудову функціональних блоків апаратної ШНМ [3,4].

Опишемо особливості навчання і функціонування БШП на основі МГП. Базовим елементом моделі МГП є матриця реалізацій, рядками якої є вектори реалізацій (спостереження) об'єкта моделювання, розпізнавання чи передбачення. Частина рядків матриці реалізацій становить тренувальну вибірку. Вважаємо, що кожному з рядків матриці реалізацій, включаючи і ті, що не входять до тренувальної вибірки, можна поставити унікальне число  $N$  – ідентифікатор реалізації, що дорівнює значенню певної функції від вхідних змінних:

$$N = \Psi(x_1, x_2, \mathbf{K}, x_n). \quad (1)$$

Своєю чергою, матрицю реалізацій можна розглядати як таблицю обчислених значень функції двох змінних  $F(N, i)$ , де перша змінна  $N$  задає рядок через ідентифікатор реалізації, а друга – номер  $i$  стовпця (ознаку реалізації). цю функцію двох змінних можна наблизити комбінацією функцій однієї змінної

$$F(N, i) = \sum f(N) \cdot j(i). \quad (2)$$

Із врахуванням (1) отримуємо:

$$F(N, i) = \sum f(\Psi(x_1, x_2, \mathbf{K}, x_n)) \cdot j(i) = \sum f_1(x_1, x_2, \mathbf{K}, x_n) \cdot j(i). \quad (3)$$

Якщо для простоти міркувань розглянути структуру перцептрона з одним прихованим шаром та лінійними НЕ у вихідному шарі, то вихідний сигнал БШП може бути описаний залежністю (3), де  $f_1$  – сигнали на виході НЕ прихованого шару,  $\varphi$  – величини відповідних синаптичних ваг.

Як виявилось [3], залежність (1) з достатньою точністю апроксимується функцією однієї змінної:

$$f_1(x_1, x_2, \mathbf{K}, x_n) \approx f_2(KS_j), \quad (4)$$

$$KS_j = \frac{\sum_{i=1}^n x_{ji} \cdot x_{Mi}}{\sum_{i=1}^n x_{Mi}^2}, \quad (5)$$

де  $i$  – номер вхідної ознаки;  $j$  – номер рядка таблиці реалізацій;  $M$  – номер рядка-вектора матриці реалізацій з найбільшою віддалю (за Евклідом) від початку координат.

Процедура навчання ШНМ, метою якого є отримання наближення елементів розкладу (3) з врахуванням (4,5), виконується шляхом покрокової ортогоналізації Грамма–Шмітта, де на кожному кроці формується передатна функція одного з НЕ прихованого шару. Режим функціонування попередньо навченої ШНМ реалізується також на підставі процедури ортогоналізації Грамма–Шмітта, причому для стовпців вхідних ознак справедливий повний збіг кроків обчислень як для навчання, так і для функціонування ШНМ.

Отже, можна дійти висновку про однотипність алгоритмів навчання та функціонування БШП даного типу, а також про обмежений набір базових операцій, покладених в їх основу – оператор суми парних добутоків та поліноміальна регресія. Якщо прийняти до уваги, що розширення вимірності вхідних векторів одним з відомих методів, зокрема радіальних базових функцій [4,6], що добре поєднується з нейропарадигмою цього типу, отримуємо лінеаризований варіант ШНМ, де єдиною базовою обчислювальною операцією як для навчання, так і для функціонування є оператор суми парних добутоків.

**Вимоги та принципи реалізації БШП.** При апаратній реалізації висока ефективність використання обладнання досягається узгодженням інтенсивності надходження даних  $P_d = hknF_d$ , де  $h$  – кількість трактів обробки,  $k$  – кількість каналів надходження даних;  $n$  – розрядність каналів надходження даних;  $F_d$  – частота надходження даних, із інтенсивністю обчислень (обчислювальною здатністю) апаратних засобів, яку визначають так:

$$D_k = \frac{hm_m n_m}{T_k}, \quad (6)$$

де  $m_m$  – кількість каналів надходження даних у сходінках конвеєра;  $n_m$  – розрядність каналів надходження даних у сходінках конвеєра;  $T_k$  – такт конвеєра. Кількість трактів обробки  $h$  в БШП дорівнює кількості нейронів в кожному шарі.

Одним із основних інтегральних параметрів оцінки структур БШП орієнтованих на НВІС-реалізацію є ефективність використання обладнання, який враховує кількість виводів інтерфейсу, однорідність структури, кількість і локальність зв'язків, пов'язує продуктивність з витратами обладнання та дає оцінку елементам пристрою за продуктивністю [7]. Кількісна величина ефективності використання обладнання визначається так:

$$E = \frac{R}{t_o (k_1 \sum_{i=1}^s W_{\phi_{y_i}} d_i + k_2 Q + k_3 Y)}, \quad (7)$$

де  $R$  – складність алгоритму, яка визначається кількістю елементарних арифметичних операцій, необхідних для його реалізації;  $t_o$  – час реалізації алгоритмів навчання та функціонування БШП;  $W_{\phi_{y_i}}$  – витрати обладнання на реалізацію  $i$ -го функціонального вузла;  $d_i$  – кількість функціональних вузлів  $i$ -го типу;  $k_1$  – коефіцієнт врахування однорідності  $k_1 = f(s)$ ;  $s$  – кількість видів функціональних вузлів;  $Q$  – загальна кількість зв'язків;  $k_2$  – коефіцієнт врахування регулярності зв'язків  $k_2 = f(\Delta j)$ ;  $\Delta j$  – просторова зв'язкова віддаль;  $Y$  – кількість виводів інтерфейсу;  $k_3$  – коефіцієнт врахування кількості виводів інтерфейсу зв'язку  $k_3 = f(Y)$ .

Задача проектування спеціалізованих і проблемно-орієнтованих структур БШП реального часу, орієнтованих на НВІС-реалізацію з високою ефективністю використання обладнання зводиться до мінімізації апаратних затрат, кількості виводів інтерфейсу, збільшення однорідності структури та регулярності зв'язків із забезпеченням режиму реального часу. Для найповнішого використання переваг сучасної НВІС-технології та забезпечення цих вимог в основу розробки БШП доцільно покласти такі принципи [6,7]:

- узгодженості інтенсивності надходження даних  $P_d$  з обчислювальною інтенсивністю  $D_k$  пристрою;

- конвеєризації, просторового паралелізму та використання базису елементарних арифметичних операцій;
- однорідності, регулярності та модульності;
- локалізації та зменшення кількості зв'язків між елементами пристрою.

**Методи та алгоритми обчислення базових операцій БШП.** При реалізації БШП з неітераційним навчанням на базі МГП єдиною базовою обчислювальною операцією як для навчання, так і для функціонування є оператор суми парних добутоків:

$$Y = \sum_{j=1}^m C_j X_j. \quad (8)$$

Особливістю цієї нейропарадигми є можливість попереднього обчислення та зберігання всіх множених  $C_j$ , які використовуються при обчисленні оператора суми парних добутоків. Існують два підходи до апаратної реалізації оператора суми парних добутоків [7–11]. Перший з них ґрунтується на операціях множення, додавання, другий – на операціях додавання, інверсії та зсуву. Перший підхід переважно, використовують при синтезі пристроїв обчислення оператора суми парних добутоків на базі окремих мікросхем (помножувачів, суматорів), а другий орієнтований на НВІС-реалізацію. Використання для НВІС-реалізацій алгоритмів на базі операцій додавання, інверсії та зсуву збільшують однорідність, регулярність структури та забезпечують оптимізацію пристрою за швидкістю і апаратними витратами.

Існуючі алгоритми обчислення оператора суми парних добутоків, що ґрунтуються на базисі елементарних арифметичних операцій, зводяться до формування і підсумовування часткових добутоків для кожної пари операндів [7]. Апаратна реалізація таких алгоритмів вимагає великих затрат обладнання і значної кількості виводів інтерфейсу, які залежать як від кількості операндів, так і від їх розрядності. Зменшити затрати обладнання та кількість виводів можна шляхом використання вертикального багатооперандного методу обчислення оператора суми парних добутоків, який передбачає одночасне послідовно-порозрядне надходження всіх операндів, які представлені числами в додатковому коді з фіксованою комою та формування в кожному такті відповідних частин макрочасткових добутоків. Для реалізації цього методу обчислення оператора суми парних добутоків необхідно, щоб множники  $X_j$  були представлені в порозрядному вигляді відповідно до формули [10,12]:

$$X_{j_i} = \sum_{i=0}^{n-1} (-1)^{2^i} 2^{-i} x_{ji}. \quad (9)$$

де  $x_{ji}$  – цифра  $i$ -го розряду множника  $X_j$ . Підставляємо значення формули (9) у формулу (8) і отримаємо:

$$Y = \sum_{j=1}^m \sum_{i=0}^{n-1} (-1)^{2^i} 2^{-i} C_j x_{ji}. \quad (10)$$

Об'єднавши часткові добуток з однаковими вагами, отримаємо макрочастковий добуток:

$$P_i = \sum_{j=1}^m C_j x_{ji}. \quad (11)$$

Підставивши значення формули (11) до формули (10), отримаємо формулу для обчислення оператора суми парних добутоків:

$$Y = \sum_{i=0}^{n-1} (-1)^{2^i} 2^{-i} P_i \quad (12)$$

Кількість макрочасткових добутоків залежить від кількості розрядів множників  $X_j$ , які аналізуються для їх отримання. Визначається кількість макрочасткових добутоків за формулою:

$$r = \left\lceil \frac{n+1}{g} \right\rceil, \quad (13)$$

де  $\lceil \cdot \rceil$  – знак округлення до більшого цілого;  $r$  – кількість груп, на які розбиваються множники  $X_j$ ;  $n$

– розрядність множників  $X_j$ ;  $g$  – кількість розрядів у групі, які аналізуються для отримання макрочасткових добутоків.

Формувати макрочасткові добуток доцільно з молодших розрядів, що забезпечить зменшення розрядності накопичувального суматора. Вираз, за яким формують макрочастковий добуток  $P_l$  має такий вигляд:

$$P_l = \sum_{j=1}^m P_{jl}, \quad (14)$$

де  $l=1, \dots, r$ .

Сам  $P_{jl}$  частковий добуток обчислюється так:

$$P_{jl} = C_j K_{jl}, \quad (15)$$

де  $K_{jl}$  –  $l$ -й коефіцієнт при множенні  $C_j$ .

Кількість коефіцієнтів  $Z$ , використовуваних для формування часткового добутку  $P_{jl}$ , визначається за такою формулою [9–11]:

$$Z = 2^g + 1.$$

Для визначення коефіцієнтів  $K_{jl}$  необхідно множники  $X_j$  розбити на групи із трьох і більше розрядів з розшифровкою кожної групи разом із старшим розрядом сусідньої групи, який розглядається як додатковий молодший розряд. При розшифровці кожної  $l$ -ої групи множника  $X_j$  визначається коефіцієнт  $K_{jl}$  шляхом додавання ваг ненульових розрядів цієї групи [7,8]. Ваги розрядів у кожній  $l$ -й групі множника  $X_j$  є такими:

$$\frac{x_{jl}g}{2^{g-1}} \quad \frac{x_{jl(g-1)}}{2^{g-2}} \quad \frac{x_{jl(g-2)}}{2^{g-3}} \quad \mathbf{L} \quad \frac{x_{jl2}}{2^1} \quad \frac{x_{jl1}}{2^0} \quad \frac{x_{j(l+1)g}}{2^0}. \quad (16)$$

Коефіцієнт  $K_{jl}$  при цьому може набувати одного із таких наступних значень:

$$0, \pm 1, \pm 2, \pm 3, \pm 4, \dots, \pm 2^{g-1}.$$

У загальному випадку обчислення оператора суми парних добутоків з аналізом  $g$  розрядів запишеться так:

$$Y = \sum_{l=1}^r 2^{-g(l-1)} P_l. \quad (17)$$

Із формули (17) видно, що обчислення оператора суми парних добутоків  $Y$  зводиться до підсумовування макрочасткових добутоків  $P_l$ , де кожний  $P_l$  макрочастковий добуток зсунутий відносно  $P_{l-1}$  на  $g$  розрядів праворуч.

Нейропарадигма, використовувана для реалізації БШП з неітеративним навчанням, забезпечує можливість попереднього обчислення всіх операндів  $C_j$ . Знаючи значення всіх операндів  $C_j$ , можна обчислити таблицю макрочасткових добутоків для всіх можливих комбінацій  $l$ -х груп розрядних зрізів множників  $X_j$  [12,13]. Кількість можливих варіантів макрочасткових добутоків  $P_l$  (обсяг таблиці) залежить як від кількості пар операндів  $m$ , так і від кількості розрядів у групі  $g$ , які аналізуються для формування часткових добутоків  $P_{jl}$ . Обсяг таблиці макрочасткових добутоків  $P_l$  визначається так:

$$Q = 2^{mg}. \quad (18)$$

Вдосконалення вертикального багатооперандного методу обчислення оператора суми парних добутоків [12,13] пов'язано з використанням таблиць макрочасткових добутоків  $P_l$  і з узгодженням інтенсивності надходження даних з інтенсивністю обчислень. Для такої реалізації необхідно обчислити всі можливі комбінації макрочасткових добутоків  $P_l$  і записати їх у пам'ять так, щоб комбінація значень розрядних зрізів множників  $X_j$  була адресою комірки пам'яті, в якій зберігається відповідне значення макрочасткового добутку  $P_l$ .

Найчастіше для запису в пам'ять макрочасткові добутоків  $P_l$  обчислюються за алгоритмом з аналізом одного розряду множників  $x_{ji}$  ( $g=1$ ) [12,13]. Вибір макрочасткового добутку  $P_l$  з пам'яті здійснюється за адресою, яка є розрядним зрізом множників  $X_j$ , відповідно до виразу:

$$P_l = \begin{cases} 0, & \text{якщо } x_{1l} = x_{2l} = x_{3l} = \dots = x_{ml} = 0 \\ C_1, & \text{якщо } x_{1l} = 1, x_{2l} = x_{3l} = \dots = x_{ml} = 0 \\ C_2, & \text{якщо } x_{1l} = 0, x_{2l} = 1, x_{3l} = \dots = x_{ml} = 0 \\ C_1 + C_2, & \text{якщо } x_{1l} = 1, x_{2l} = 1, x_{3l} = \dots = x_{ml} = 0 \\ \mathbf{M} \\ C_2 + C_3 + \dots + C_m, & \text{якщо } x_{1l} = 0, x_{2l} = x_{3l} = \dots = x_{ml} = 1 \\ C_1 + C_2 + \dots + C_m, & \text{якщо } x_{1l} = x_{2l} = x_{3l} = \dots = x_{ml} = 1 \end{cases} . \quad (19)$$

Основними етапами реалізації вдосконаленого вертикального багатооперандного методу обчислення оператора суми парних добутоків є [12,13]:

- обчислення таблиці всіх можливих макрочасткових добутоків  $P_l$ ;
- запис в пам'ять обчислених значень макрочасткових добутоків  $P_l$ ;
- зчитування макрочасткових добутоків  $P_l$  з комірок пам'яті за адресою, яка відповідає значенню розрядних зрізів множників  $X_j$ ;
- підсумовування макрочасткових добутоків  $P_l$  із зсувом праворуч на кількість розрядів, які аналізуються для їх отримання.

**Структури пристроїв обчислення базових операцій БШП.** Основними вимогами, що висуваються до пристроїв обчислення базових операцій БШП, є забезпечення режиму реального часу та високої ефективності використання обладнання. На етапі переходу від алгоритму обчислення базових операцій БШП до структури пристрою, що працює у реальному часі, високої ефективності використання обладнання досягається мінімізацією апаратних затрат. Основними шляхами такої мінімізації є [10,13]:

- узгодження інтенсивності надходження даних  $P_d$  із інтенсивністю обчислень  $D_k$  апаратних засобів;
- вибір ефективних алгоритмів формування макрочасткових добутоків, які забезпечать зменшення їх кількості;
- зменшення розрядності суматорів і каналів передачі даних;
- зменшення кількості трактів обробки.

Узгодження інтенсивності надходження даних  $P_d$  із інтенсивністю обчислень  $D_k$  апаратних засобів можна досягти вибором:

- тривалості конвеєрного такту  $T_k$ ;
- кількості трактів обробки  $h$ ;
- кількості каналів надходження даних у тракті обробки  $m_m$ ;
- розрядності каналів надходження даних  $n_m$ .

Для узгодження інтенсивності надходження даних  $P_d$  із інтенсивністю обчислень  $D_k$  апаратних засобів, які реалізують вдосконалений вертикальний багатооперандний метод обчислення оператора суми парних добутоків, насамперед необхідно мінімізувати тривалість конвеєрного такту  $T_k$ , який визначається швидкодією пам'яті та суматорів, а по-друге, змінювати кількості трактів обробки  $h$ , каналів надходження даних у тракті обробки  $m_m$  та їх розрядність  $n_m$ . Залежно від інтенсивності надходження даних  $P_d$  та вимог конкретного застосування можна синтезувати багато структур пристроїв для обчислення оператора суми парних добутоків, які відрізнятимуться як організацією обчислень, так і технічними характеристиками. Тому доцільно розробити та дослідити базову структуру пристрою обчислення оператора суми парних добутоків, яка буде основою для синтезу конкретного пристрою із заданими характеристиками. Базова структура такого пристрою, який використовує таблиці макрочасткових добутоків  $P_l$  для обчислення оператора суми парних добутоків, наведена на рис.2, де Рг – регістр; ПерК –перетворювач кодів; ФМЧД – формувач макрочасткових добутоків; ОЗП – оперативний запам'ятовуючий пристрій; БК – блок керування; СВ – суматор-віднімач; ША –шина адреси; ШД – шина даних; ШУ – шина управління; ШФ – шинний формувач; У1, У2, У3, У4, У5, У6, У7 – сигнали управління відповідно режимом роботи ПерК, режимом роботи ШФ1, записом/читанням в ОЗП, вибіркою ОЗП, режимом роботи СВ, скидом РгСм та режимом роботи ШФ2.

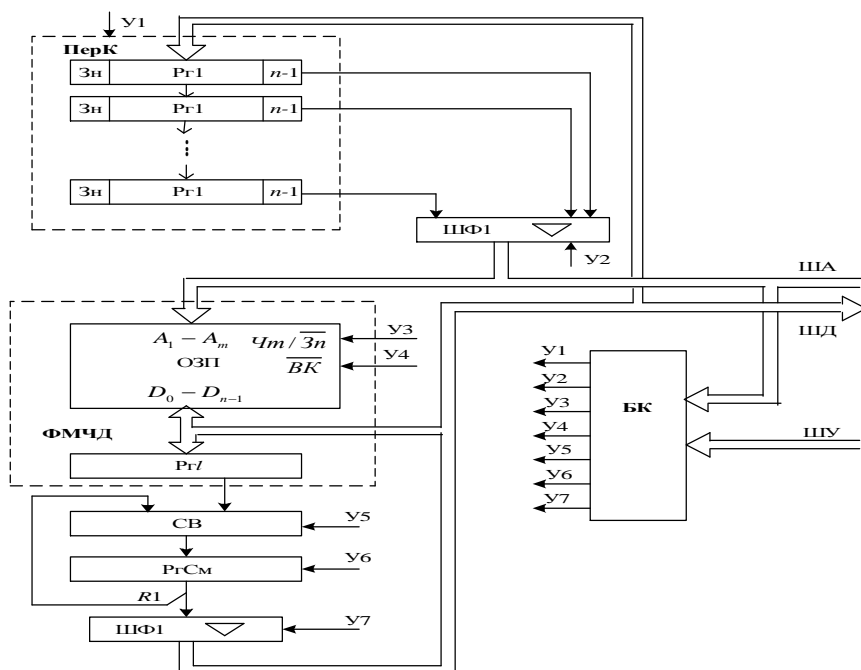


Рис. 2. Базова структура пристрою для обчислення оператора суми парних добутків

Апаратна реалізація вдосконаленого вертикального багатооперандного методу обчислення оператора суми парних добутків  $Y$  передбачає попереднє обчислення всіх можливих варіантів макрочасткових добутків  $P_l$  з подальшим записом їх в ОЗП. Крім того, перед початком обчислення оператора суми парних добутків  $Y$  значення множників  $X_j$  в паралельному коді зберігаються в регістрах перетворювача ПерК. Сигналом лог.1 на вході  $Y1$  перетворювача ПерК установлює регістри  $Pr1$ - $Prm$  в режим зсуву у бік молодших розрядів, а сигнал лог.1 на вході  $Y2$  шинного формувача ШФ1 дозволяє передачу розрядних зрізів множників  $X_j$  з виходів регістрів  $Pr1$ - $Prm$  на адресні входи ОЗП.

Обчислення оператора суми парних добутків  $Y$  в запропонованому пристрої здійснюється за тактами. В кожному такті роботи пристрою здійснюється зчитування з ОЗП макрочасткового добутку  $P_l$ , який зберігається за адресою, що відповідає значенню розрядного зрізу  $x_{1l}, \dots, x_{ml}$  відповідно до формули (19). Кожний  $l$ -й макрочастковий добуток  $P_l$  додається до зсунутої на один розряд праворуч відносно раніше накопиченої суми, що надходить з виходів регістра  $PrCm$ . Оскільки пристрій працює з числами в додатковому коді, то макрочастковий добуток  $P_l$ , який відповідає значенню знакового розрядного зрізу  $x_{13l}, \dots, x_{m3l}$ , віднімається від раніше накопиченої суми.

Період конвеєрного такту, з яким працює пристрій, визначається за формулою:

$$T_k = t_{ШФ} + t_{ОЗП} + t_{Pr}, \quad (20)$$

де  $t_{ШФ}$ ,  $t_{Pr}$ ,  $t_{ОЗП}$  – час спрацювання відповідно шинного формувача ШФ, регістра  $Pr$  і ОЗП.

Зменшити цей час можна шляхом вибору швидкодіючої елементної бази. Кількість тактів, необхідних для отримання результату обчислення оператора суми парних добутків  $Y$  за алгоритмом з аналізом одного розряду ( $g=1$ ) множників  $x_{1l}, \dots, x_{ml}$ , визначається розрядністю множників  $X_j$  і початковою затримкою конвеєра.

Затрати обладнання на реалізацію базової структури пристрою обчислення оператора суми парних добутків визначаються за таким виразом:

$$W_{Pr} = (m + 2)W_{Pr} + 2W_{ШФ} + W_{СВ} + W_{ОЗП} + W_{БК}, \quad (21)$$

де  $W_{Pr}$ ,  $W_{ШФ}$ ,  $W_{СВ}$ ,  $W_{ОЗП}$ ,  $W_{БК}$  –затрати обладнання відповідно на регістр, суматор-віднімач, ОЗП і блок керування.

Підвищити інтенсивність обчислень  $D_k$  базового пристрою можна збільшенням кількості розрядів  $g$ , які аналізуються для отримання макрочасткового добутку  $P_l$ . Існують два підходи до розв'язання цієї задачі. Перший з них ґрунтується на використанні одного ОЗП для зберігання макрочасткових добутків  $P_l$ , другий – на використанні декількох ОЗП. При використанні першого підходу обсяг пам'яті зростає за показниковим законом за формулою:

$$Q = 2^{mg}, \quad (22)$$

а за другим підходом – обсяг пам'яті зростає лінійно за формулою:

$$Q = g2^m. \quad (23)$$

З формул (22) і (23) видно, що другий підхід вимагає значно менших обсягів пам'яті. При використанні другого підходу максимальної інтенсивності обчислень  $D_k$  буде досягнуто при  $n$  ОЗП, тобто коли кількість ОЗП відповідає розрядності множників  $X_j$ . Структура блоку формування результату обчислення оператора суми парних добутків  $Y$  для цього випадку наведена на рис.3, де БСМ – багатовходовий суматор.

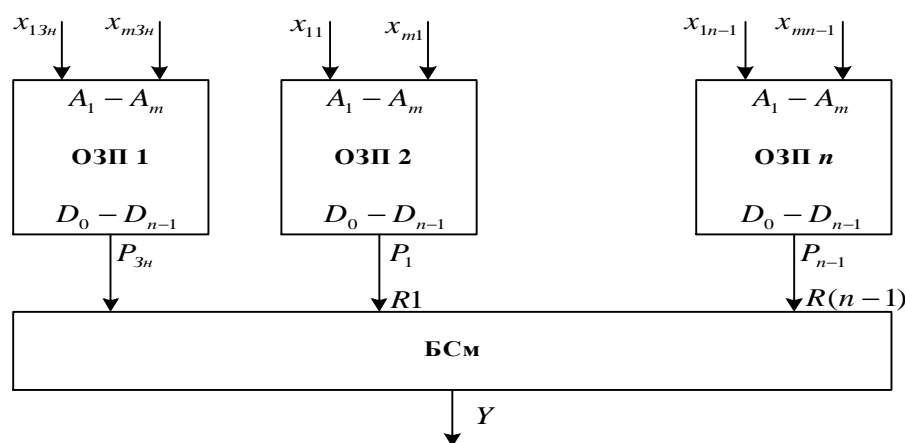


Рис. 3. Блок формування результату обчислення оператора суми парних добутків на базі  $n$  ОЗП

Формування результату обчислення оператора суми парних добутків у цьому блоці здійснюється за один такт, який дорівнює  $T_k = t_{\text{ОЗП}} + t_{\text{БСМ}}$ , де  $t_{\text{БСМ}}$  – час затримки проходження даних через БСМ.

Зменшити інтенсивність обчислень  $D_k$  базового пристрою можна зменшенням кількості каналів надходження даних  $m_m$ . Мінімальна кількість каналів надходження даних  $m_m$  може дорівнювати одиниці. Зменшення кількості каналів надходження даних  $m_m$  призводить до збільшення тактів, необхідних для формування макрочасткового добутку  $P_l$ . Структура формувача макрочасткових добутків  $P_l$  для випадку коли  $m_m=1$ , наведена на рис.4, де  $q = \lceil \log_2 m \rceil$ .

З допомогою даних на адресних входах  $A_1, \dots, A_q$  і  $A_{q+1}$  вибирається комірка ОЗП, з якої зчитується  $jl$ -й частковий добуток  $P_{jl}$  для пари операндів  $C_j X_j$ . Формування макрочасткового добутку  $P_l$  для  $l$ -го розрядного зрізу  $x_{1l}, \dots, x_{ml}$  в даному вузлі здійснюється за  $m$  тактів. Обсяг пам'яті, необхідний для зберігання часткових добутків  $P_{jl}$ , визначається за формулою:

$$Q = 2^{q+1}. \quad (24)$$

Зміна інтенсивності обчислень  $D_k$  базової структури пов'язана з модифікацією перетворювача ПерК та формувача ФМЧД. Збільшення інтенсивності обчислень  $D_k$  вимагає збільшення кількості регістрів, обсягів ОЗП, ускладнення структур формувача ФМЧД і перетворювача ПерК, а зменшення інтенсивності обчислень  $D_k$  вимагає менших обсягів ОЗП і спрощує структури формувача ФМЧД і перетворювача ПерК.



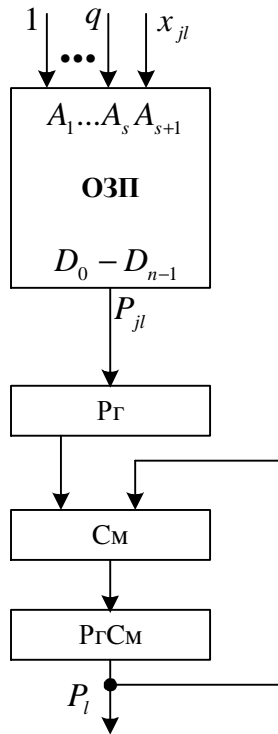


Рис.4. Формувач макрочасткових добутків  $P_l$

**Синтез структури БШП з неітераційним навчанням.** Задача синтезу БШП з неітераційним навчанням, що працює у реальному часі, зводиться до формування множин вимог  $\mathbf{R}=\{R_1, R_2, \dots, R_k\}$ , характеристик  $\mathbf{H}=\{H_1, H_2, \dots, H_m\}$  і обмежень  $\mathbf{B}=\{B_1, B_2, \dots, B_k\}$  та знаходження такого вектора  $\mathbf{H}^*=[H_1^*, H_2^*, \dots, H_m^*]$ ,  $H_i^*=f_i(R, H, B)$ ,  $i=1, \dots, m$ , який забезпечить максимальне значення ефективності використання обладнання  $E=\max f(\mathbf{R}, \mathbf{H}^*, \mathbf{B})$  [7,13].

Множина вимог  $\mathbf{R}$  складається з:  $R_1$  – кількість каналів надходження даних  $m$ ;  $R_2$  – розрядність каналів надходження даних  $n_m$ ;  $R_3$  – частота надходження даних  $F_d$ ;  $R_4$  – швидкодія елементної бази, яка визначається часом затримки вентиля  $t_6$ ;  $R_5$  – кількість трактів обробки  $h$ ;  $R_6$  – розрядність вхідного слова  $n$ . Множину характеристик  $\mathbf{H}$  становлять:  $H_1$  – загальна кількість зв'язків  $Z$ ;  $H_2$  – просторова зв'язкова віддаль  $\Delta j$ ;  $H_3$  – конвеєрний такт  $t_k$ ;  $H_4$  – витати обладнання  $W$ ;  $H_5$  – кількість видів функціональних вузлів  $s$ ;  $H_6$  – кількість каналів введення  $m_{66}$ ;  $H_7$  – розрядність каналів введення  $n_{66}$ ;  $H_8$  – кількість виводів інтерфейсу зв'язку  $Y$ . Обмеження  $\mathbf{B}$ , які необхідно враховувати при синтезі БШП з неітераційним навчанням, є такими:  $B_1$  – точність обчислення, яка визначається розрядністю результату  $n_p$ ;  $B_2$  – час обчислення  $t_o$  і витати обладнання  $W$ .

Для реалізації розроблених алгоритмів навчання та функціонування БШП будемо використовувати універсальні та спеціалізовані обчислювальні засоби, які об'єднуються в проблемно-орієнтовану обчислювальну систему. Як універсальні засоби будемо використовувати функціонально-орієнтовані процесори, а як спеціалізовані – алгоритмічну систему, архітектуру та організацію обчислювального процесу, в яких відображено структуру алгоритму розв'язання задачі [7]. Побудову такої алгоритмічної системи доцільно здійснювати на базі обчислювального поля, елементною базою якого є ПЛІС з динамічним репрограмуванням. Використання такої елементної бази відкриває нові можливості, які пов'язані з оперативним переналаштуванням ПЛІС на реалізацію конкретного алгоритму. Структура проблемно-орієнтованої обчислювальної системи, яка реалізує алгоритми навчання та функціонування БШП з неітераційним навчанням, наведена на рис.5, де ПП – паралельна пам'ять, БПП – буферна паралельна пам'ять, ФОП – функціонально-орієнтований процесор.

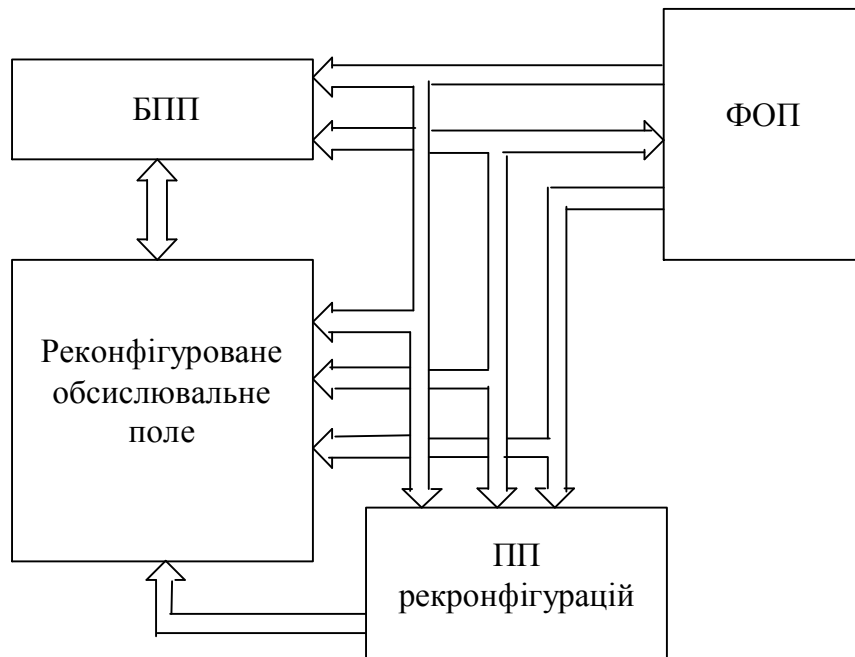


Рис. 5. Структура проблемно-орієнтованої обчислювальної системи для реалізації БШП з неітераційним навчанням

Основними компонентами розробленої системи є: буферна паралельна пам'ять і паралельна пам'ять реконфігурацій, функціонально-орієнтований процесор ФОП і реконфігуроване обчислювальне поле. Буферна паралельна пам'ять призначена для зберігання, упорядкування та комутації вхідних даних, а паралельна пам'ять реконфігурацій призначена для зберігання даних налаштування реконфігурованого обчислювального поля на реалізацію компонентів узгодженого графа БШП.

Архітектура ФОП повинна бути орієнтована на реалізацію алгоритмів навчання та функціонування БШП. В запропонованій проблемно-орієнтованій обчислювальній системі ФОП призначений для виконання попередніх обчислень, пов'язаних з формуванням макрочасткових добутків, здійснення налаштування реконфігурованої обчислювальної системи на реалізацію конкретних алгоритмів та управління обчислювальним процесом.

Реконфігуроване обчислювальне поле програмно налаштовується на апаратне відображення компонентів узгодженого потокового графу алгоритму навчання та функціонування БШП, де кожному НЕ ставиться у відповідність операційні пристрої, які забезпечують обчислення оператора суми парних добутків і виконання активаційної функції. Для апаратного відображення оператора суми парних добутків використовуються розроблені структури, інтенсивність обчислення в яких узгоджена з інтенсивністю надходження даних. Використання реконфігурованого обчислювального поля в проблемно-орієнтованій обчислювальній системі забезпечує оперативне налаштування програмним шляхом даного поля на реалізацію компонентів узгодженого потокового графу алгоритму навчання та функціонування БШП.

### Висновки

1. Розв'язок сучасних багатовимірних задач видобутку даних за сприятливий для реальних застосувань час доцільно здійснювати на основі БШП з неітераційним навчанням.

2. Ефективність апаратної реалізації БШП на основі МГП досягається за рахунок однотипності алгоритмів навчання і функціонування, базові операції яких ґрунтуються на обчисленні оператора суми парних добутків.

3. Вдосконалено вертикальний багатооперандний метод обчислення оператора суми парних добутоків в реальному часі шляхом орієнтації його на використання таблиць макрочасткових добутоків і на узгодження інтенсивності надходження даних з інтенсивністю обчислень пристрою, що забезпечило підвищення ефективності використання обладнання.

4. Використання таблиць макрочасткових добутоків і представлення алгоритмів обчислення оператора суми парних добутоків у базисі елементарних арифметичних операцій дало змогу оптимізувати структуру і повною мірою використати можливості НВІС-технології.

5. Об'єднання універсальних і спеціалізованих підходів, апаратних і програмних засобів при реалізації БШП з неітераційним навчанням для обробки даних в реальному часі забезпечило підвищення ефективності використання обладнання.

1. Круглов В.В., Борисов В.В. *Искусственные нейронные сети. Теория и практика.* – 2-е изд., стереотип. – М.: Горячая линия-Телеком, 2002. – 382 с. 2. Галушкин А.И. *Нейрокомпьютеры.* Кн.3. – М.; ИПРЖР, 2000. – 528 с. 3. Ткаченко Р.О. *Модель нейронных сетей* // *Вісник Держ. ун-ту “Львівська політехніка”*: “Комп’ютерна інженерія та інформаційні технології”. – 1998. – №349. – С. 83–86. 4. Ткаченко Р.О. *Нейронні мережі з нелінійними синаптичними зв’язками* // *Вісник Держ. ун-ту “Львівська політехніка”*: *Комп’ютерні системи проектування. Теорія і практика.* – 1999. – №373. – С. 20–22. 5. Осовский С. *Нейронные сети для обработки информации* / Пер. с польс. – М.: Финансы и статистика, 2002. – 344 с. 6. Ткаченко Р.О., Ткаченко П.Р., Цмоць І.Г. *Апаратна реалізація багатошарових перцептронів з неітераційним навчанням* / *Збірник наукових праць. Інститут проблем моделювання в енергетиці НАН України. Вип. 29.* – К., 2005. – С.103–113. 7. Цмоць І.Г. *Інформаційні технології та спеціалізовані засоби обробки сигналів і зображень у реальному часі.* Львів: УАД, 2005. – 227с. 8. Цмоць І.Г. *Паралельні алгоритми та матричні НВІС-структури пристроїв множення для комп’ютерних систем реального часу* // *Науково-технічний журнал “Інформаційні технології і системи”.* Т. 7.– Львів, 2004. – №1. – С. 5–16. 9. Стрямець С.П., Цмоць І.Г. *Паралельні алгоритми та НВІС-структури обчислення суми парних добутоків* // *Вісник Нац. ун-ту “Львівська політехніка” “Комп’ютерна інженерія та інформаційні технології”.* – Львів, 2003. – № 496. – С.255–263. 10. Цмоць І.Г., Іванців В.Р. *Паралельно-вертикальний метод і базова структура пристрою обчислення сум парних добутоків.* 11. Березький О.М., Цмоць І.Г. *Методи, алгоритми та НВІС-структури для множення матриці на вектор у реальному часі* // *Вісник Хмельницького національного університету.* – 2007. – №3. – Т1. – С.134–140. 12. Боюн В.П. *Динамическая теория информации. Основы и приложения.* – К.: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2001. – 326с. 13. Малиновський Б.Н., Боюн В.П., Козлов Л.Г. *Введение в кибернетическую технику. Параллельные структуры и методы.* – К.: Наук. думка, 1989. – 248с.