

## ДОСЛІДЖЕННЯ СТРУКТУРИ ГІБРИДНОЇ НЕЙРОМЕРЕЖІ ДЛЯ СИНТЕЗУ ІСППР В ЗАДАЧАХ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

**Вступ.** Керамічне виробництво традиційно характеризується значною часткою зворотного браку на різних етапах виконання технологічного процесу. Така ситуація обумовлена недостатньою повнотою теоретичних обґрунтувань фізико-хімічних процесів, а також низьким рівнем автоматизації технологічних процесів, що створює передумови для виникнення помилок при прийнятті управлінських рішень, погіршує їх якість і оперативність.

**Мета роботи.** Для вирішення задач прогнозу значень фізико-хімічних параметрів водної суспензії (керамічного шлікера) вивчити ефективність використання в якості компоненти гібридної структури з властивостями апроксимації динамічних нейронних мереж Елмана.

**Результати досліджень.** Як свідчать матеріали публікацій в галузі теорії та прикладного використання штучних нейронних мереж [1], гібридні структури при їх практичній реалізації, як правило, є більш ефективними у порівнянні з мережами, що базуються на основі окремих нейромережевих парадигм. Даний факт пояснюється тим, що результуюча структура набуває нових властивостей, які не притаманні кожній парадигмі окремо. В статті [2] розглянуті питання синтезу і моделювання процесів навчання та отримані на їх основі результати для компоненти гібридної структури з властивостями самоорганізації (мапа Кохонена). Вагові коефіцієнти нейронів мапи Кохонена по закінченню процесу навчання були заморожені. Таким чином, вхідними сигналами для компоненти гібридної структури з властивостями апроксимації буде виступати множина пар векторів  $(u_i, y_i)$ , де  $u_i$  – вектор, складений з вихідних сигналів синтезованої мапи Кохонена, а  $y_i$  – вектор цільових значень оригінального відображення  $(z_i, y_i)$   $y = [y_1, y_2, \dots, y_K]^T$ ;  $K = 9$  (фізико-хімічні параметри водної суспензії), якому відповідає  $u_i$ ,  $z = [z_1, z_2, \dots, z_N]^T$  – вхідний вектор для мапи Кохонена, що містить 27 значень

технологічних параметрів процесу структуроутворення водної суспензії (керамічного шлікера).

В даній статті викладені результати досліджень щодо можливості застосування в якості компоненти гібридної структури з властивостями апроксимації динамічної рекурсивної мережі Елмана. Наявність зворотних зв'язків між нейронами різних шарів (включаючи нейрони вихідного шару) надає меражам даного типу позитивних властивостей, які не можуть бути досягнуті в статичних багат шарових мережах [3]. Типовими прикладами застосування зазначених мереж є імітація нелінійних динамічних об'єктів, зокрема – систем управління машинами, пристроями і динамічними процесами, для яких динамічна мережа виступає в ролі моделі, а алгоритм уточнення вагових коефіцієнтів – у ролі процедури ідентифікації параметрів даної моделі [4]. Мережа Елмана характеризується частковою рекурентністю в формі зворотнього зв'язку між прихованим і вихідним шаром, що реалізується за допомогою одиничних елементів запізнювання [5]. Це дозволяє формувати правильну стратегію управління за рахунок врахування передісторії процесів, за якими спостерігають.

Вхідними сигналами моделі будуть являтися виходи компоненти з самоорганізацією. Вихід моделі з самоорганізацією при подачі на її вхід вектора  $z$  являє собою розріджену матрицю зв'язності розмірності [27x27]. З метою усунення несуттєвої інформації, яка буде привносити зайвий шум у вхідні дані  $i$ , в принципі, може впливати на вихідний сигнал, виконано перетворення матриці зв'язності у вектор індексів класів (кластерів). Після перетворення номер кластеру відповідає номеру строки матриці зв'язності. Зв'язаними з вектором  $U$  є два вихідних вектори: вектор - фактичних вихідних сигналів  $D$  і вектор цільових сигналів  $Y$ :

$$D = [d_0, d_{i,1}, d_{i,2}, \dots, d_{i,m}]^T \quad (1)$$

$$Y = [y_0, y_{i,1}, y_{i,2}, \dots, y_{i,m}]^T \quad (2)$$

Фактичним статистичним матеріалом при виконанні досліджень були дані виробничої діяльності, що накопичені в базі даних підприємства “Дніпрокераміка” за 5 останніх років.

Як відомо, рішення щодо остаточної структури моделі може бути прийнято тільки після виконання певної кількості циклів (epoch) її навчання і досягненні похибки узагальнення на рівні, який

признається задовільним. Однак однозначних підходів щодо визначення значення похибки узагальнення штучних нейронних мереж на даний час не сформовано. В роботі [5] запропоновано характеризувати її співвідношенням:

$$E(w) \approx E_L(w) + \varepsilon \left( \frac{p}{h}, E_L \right) \quad (3)$$

де  $E(w)$  – похибка узагальнення нейромережі;  $E_L(w)$  – похибка навчання;  $\varepsilon$  – довірчий інтервал;  $p$  – кількість навчальних вибірок.

При формуванні структури моделі також враховували, що мінімізація похибки узагальнення на малій кількості навчальних послідовностей відносно кількості вагових коефіцієнтів може спровокувати випадковий характер значень більшості вагових коефіцієнтів, що при переході від навчальних до тестових множин даних може стати причиною суттєвих відхилень фактичних значень  $D$  від очікуваних  $Y$ .

Найбільш широке застосування (у тому числі і при вирішенні прикладних задач теорії управління) знайшли багат шарові нейронні мережі, з функціями активації сигмоїдного типу. Останні дозволяють забезпечити задану точність апроксимації при меншій розмірності нейронної мережі. У нашому випадку на відміну від рекомендацій, викладених в роботі Наренди [6], досліджувались моделі з нелінійними вихідними нейронами, що описуються функцією гіперболічного тангенсу. Це дозволяє значно скоротити розмірність мережі, час її навчання і при певній обмеженості в кількості навчальних даних – покращити її узагальнюючі властивості.

При формуванні структури мережі серйозною проблемою є визначення кількості прихованих шарів і контекстних шарів та кількості нейронів у них [1]. З огляду на препроцесорну обробку вхідних даних компонентою з самоорганізацією, в дослідженнях вивчалась можливість застосування мережі Елмана, що складається з одного прихованого і одного контекстного шару нейронів (в які при необхідності будуть додаватися додаткові нейрони). Як відомо, основним завданням нейронів контекстного шару є генерація бажаного вихідного сигналу на підставі порівняння образу, що надійшов, і образу, що запам'ятовувався в контекстному шарі.

Мета навчання компоненти гібридної структури полягала в підборі таких значень вагових коефіцієнтів синаптичних зв'язків

нейронів, щоб при заданому вхідному векторі  $U$  відгук моделі відповідав значенням фактичних сигналів  $d_i$ , які з 94% точністю і більше будуть дорівнювати очікуваним значенням  $y_i$ , для  $i=1, 2, \dots, m$ . В якості функціоналу  $E$ , що характеризує якість навчання моделі, використовували:

$$E = \frac{1}{2} \sum_{t=1}^{T_H} \sum_{i=1}^{S^l} \left( y_i^t - d_i^{tS^l} \right)^2 \quad (4)$$

де  $T_H$  – об’єм навчальної вибірки;  $l$  – кількість шарів мережі;  $t$  – номер вибірки;  $S^l$  – кількість нейронів вихідного шару;  $d_i^{tS^l}$  – вектор сигналу на виході мережі;  $y_i^t$  – вектор цільових значень сигналу на виході мережі для вибірки з номером  $t$ .

Градиент цільової функції обчислювали за алгоритмом Левенберга-Марквардта [7]. Корегування значень вагових коефіцієнтів і зміщень здійснювали за методом зворотного розповсюдження похибки.

Відомо, що на результати навчання моделі значний вплив спричиняє вибір початкових значень вагових коефіцієнтів. Ідеальною є ситуація, коли початкові значення вагових коефіцієнтів є достатньо близькими до оптимальних. У цьому випадку будуть усунуті затримки в точках локальних мінімумів, що, відповідно, призведе до значного прискорення процесу навчання моделі. Однак універсального методу вибору початкових значень вагових коефіцієнтів, який би гарантував знаходження найкращої початкової точки для вирішення задач даного класу не існує, тому ініціалізація моделі виконувалась випадковими значеннями розподіленими рівномірно в діапазоні  $[-1 +1]$ .

Виконані експериментальні дослідження з моделювання процесу навчання мереж Елмана різної структури показали, що оптимальною з точки зору узагальюючих властивостей є модель структури 3-9, що вміщує 3 нейрони у прихованому шарі і, відповідно, 3 контекстних нейрони. При збільшенні числа нейронів у прихованому шарі спостерігається явище перенавчання мережі. Вихідний шар вміщує 9 нейронів, що відповідає кількості фізико-хімічних параметрів водної суспензії, які необхідно прогнозувати. На рис 1. показані результати моделювання процесу навчання мережі.

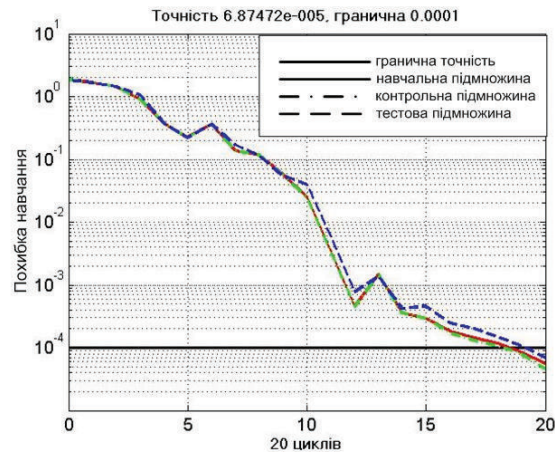


Рисунок 1 - Графік зміни похибки навчання моделі

Аналіз рис.1 свідчить, що необхідна відповідність між входом і виходом моделі забезпечується з середньоквадратичною похибкою 0,0001 на 18-му циклі навчання, при цьому похибки перевірки на контрольній і тестовій множині даних ведуть себе приблизно однаково, помітних тенденцій до перенавчання моделі не спостерігається.

Результати моделювання підсумкової нейромережі гібридної структури, що вирішує задачу прогнозу значень фізико-хімічних параметрів водної суспензії, наведено на рис. 2. Аналіз графіків свідчить, що апроксимаційні можливості мережі можна вважати задовільними: максимальне значення відносної похибки для контрольної підмножини складає 16,7% (середня відносна похибка моделі – 3,4%), для тестової підмножини – 5.7 % (середня відносна похибка моделі – 3,99%).

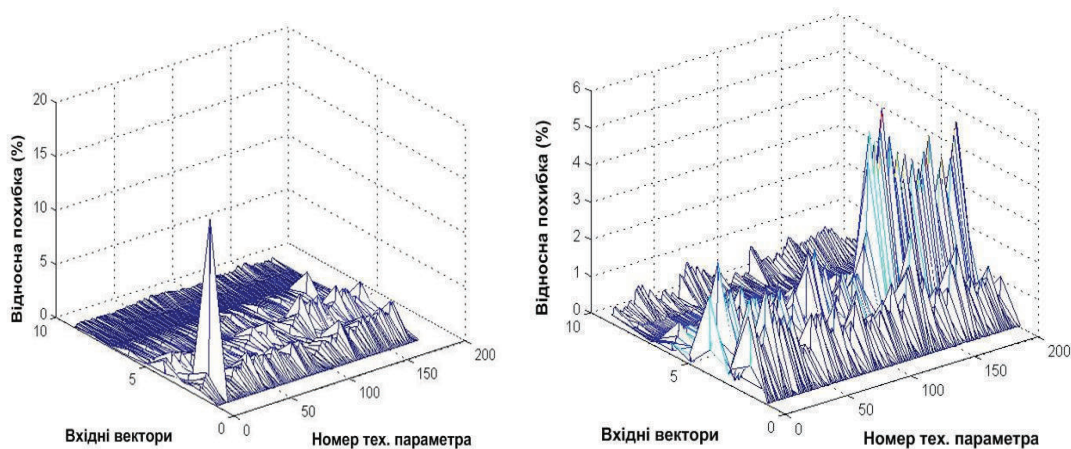


Рисунок 2 - Відносні похибки синтезованої моделі при пред'явленні на її вхід контрольної і тестової підмножин даних: а – контрольна; б – тестова підмножини вхідних даних

На заключеному етапі досліджень виконували порівняльний аналіз ефективності використання в якості компоненти гібридної структури з властивостями апроксимації мереж: прямого розповсюдження з повними послідовними зв'язками структури 3-9 та рекурентної мережі Елмана аналогічної структури. В обох випадках використовувались нейрони з функцією активації – гіперболічний тангенс, градієнт цільової функції обчислювали за алгоритмом Левенберга-Марквардта, корегування значень вагових коефіцієнтів і зміщень нейронів виконували за методом зворотного розповсюдження похибки.

**Висновки.** Аналіз отриманих результатів свідчить, що не дивлячись на певні переваги динамічних нейронних мереж, зокрема – їх можливості щодо прогнозування часових рядів, при вирішенні задачі прогнозування фізико-хімічних параметрів водних суспензій (керамічних шлікерів) в якості компоненти гібридної структури з властивостями апроксимації більш доцільно використовувати статичну мережу прямого розповсюдження з повними послідовними зв'язками: поперше – вона має кращі узагальнюючі властивості, по-друге – знімається проблема можливого виникнення її нестійких станів.

## ЛИТЕРАТУРА

1. Осовский С. Нейронные сети для обработки информации. – М.: Финансы и статистика, 2004. – 344 с.
2. Бешта О.С., Бабенко Т.В. Синтез компоненти гібридної нейромережі з властивостями самоорганізації // Науковий вісник НГУ. – 2007. – №3. – С. 61-63.
3. Руденко О.Г., Бодянський Є.В. Штучні нейронні мережі. – Харків: ТОВ «Компанія СМІТ», 2006. – 404с.
4. Tran Hoai L. Sieci neuronowe rekurencyjne - gradientowe algorytmy uczone i zastosowania w identyfikacji nieliniowych obiektow dynamicznych. Praca dyplomowa. – Warszawa: Politechnika Warszawska, 1997.
5. Haykin S. Neural networks, a comprehensive foundation. – N.Y: Macmillan Colege Publishing Company, 1994.
6. Narendra K. S., Parthasarathy K. Identification and control of dynamical systems using neural networks // IEEE Trans. Neural Networks, 1990. – Vol. 1. - P. 4-27.
7. Marquardt, D., "An Algorithm for Least Squares Estimation of Nonlinear Parameters," SIAM J. Appl. Math, 1963. – Vol. 11. - P. 431-441.

Получено 03.04.2007 г.