

М.В. Козлов, к.т.н., доц.*Житомирський державний технологічний університет***СПОСІБ ФОРМУВАННЯ ПАМ'ЯТІ У СИСТЕМАХ ШТУЧНОГО ІНТЕЛЕКТУ**

У статті розглянуто можливості фільтрації і впорядкування великого потоку інформації у пристроях збору і зберігання даних систем штучного інтелекту за рахунок введення в ці системи пристроїв проміжної пам'яті і фільтрації. Наведено приклад побудови подібної системи.

Вступ. За останні декілька десятиліть різко зріс потік інформації. Так збільшення в три рази нових наукових даних – це десятиразове зростання інформації [1]. Велику роль зіграла в цьому поява Інтернету. Багато інформації, що при цьому повторюється і квазіінформації, ставали серйозною проблемою виділення дійсно важливого.

Переробкою інформації, що поступає, у все більшому обсязі, останнім часом займаються системи штучного інтелекту (СШ) і це стає однією з основних їх функцій і від того, наскільки ефективно вони виконують цю роботу, залежить якість виконуваних ними завдань. У зв'язку з цим дослідження в області вдосконалення методів фільтрації і впорядкування інформації стосовно СШ актуальні.

Аналіз останніх досліджень і публікацій. У динамічних СШ, використовуваних для обробки розподіленої в часі вхідної інформації, використовується пам'ять, яку ділять на короткочасну (КП) і довготривалу (ДП) [2, 3]. У КП відображається поточний стан вхідної інформації і кожен новий образ (патерн), що поступає, призводить до оновлення стану цієї пам'яті. У ДП зберігаються дані, призначені для тривалого або постійного використання.

Таке ділення пам'яті аналогічно традиційній концепції поділу пам'яті кори головного мозку (неокортексу) людини на КП і ДП [4, 5]. Однак в останніх роботах по нейрофізіології [6] при аналізі діяльності центральної нервової системи людини було виявлено, що при формуванні ДП відбувається затримка в часі і, умовно, можна говорити про формування в неокортексі за участю гіпокампу і мигдалини лімбічної системи головного мозку проміжної (вторинної) пам'яті (ПП), в якій може накопичуватися інформація, що надходить з КП, у течії від 17 годин до двох років, частково стиратися і перетворюватися перед зберіганням її у ДП. При цьому можливість перенесення інформації з ПП в ДП буде, зокрема, залежати від частоти повторення цієї інформації. Однією з особливостей ПП є те, що година звернення до неї буде значно більша години звернення до ДП [6].

Окрім функцій запам'ятовування, зберігання і витягання інформації, СШ для збирання даних повинна володіти властивістю переробки і впорядкування інформації. Для цих цілей в [2] пропонується для впорядкування інформації проводити її ранжування за частотою появи.

У [7] для корпоративних баз даних пропонується використовувати для оцінки цінності отриманої інформації статистику (x, tx, T) , де x число транзакцій спостережуваних у напіввідкритому інтервалі часу $(0, T]$ і tx $(0 \leq tx < T)$ – час останньої транзакції.

Для отримання впорядкованої множини інформації, що надходить, можна використовувати і спосіб фільтрації даних, розглянутий у [8]. Стосовно обробки будь-яких вхідних даних, запропоновані в способі два критерії оцінки інформації можна сформулювати таким чином:

1) оцінка за годинною надходження інформації. При цьому для всіх однакових патернів вхідної інформації, годині надходження кожного з них привласнюватиметься деякий ваговий коефіцієнт, що убуває із збільшенням інтервалу між поточним часом і часом надходження;

2) оцінка кількості повторення у вхідних даних однакових патернів за певний проміжок години.

Розглянуті методи впорядкування даних, залежно від поставленого завдання, можуть бути використані під час виділення корисної інформації при обробці її у СШ.

Основною метою даної роботи є розглядання можливостей використання проміжної пам'яті у пристроях збирання і зберігання інформації систем штучного інтелекту.

Викладення основного матеріалу. Сучасні системи штучного інтелекту не є досконалими і їх взаємодія з інформацією, що знов поступає, далека від оптимальної. Виходячи з цього і враховуючи проведений вище аналіз, розглянемо побудову СШ, призначену для впорядкування великого потоку інформації, на користь широкого класу користувачів.

На схемі, представленій на рисунку 1, показано просторово-часове переміщення патернів від входу через КП і проміжну пам'ять до запису їх у ДП.

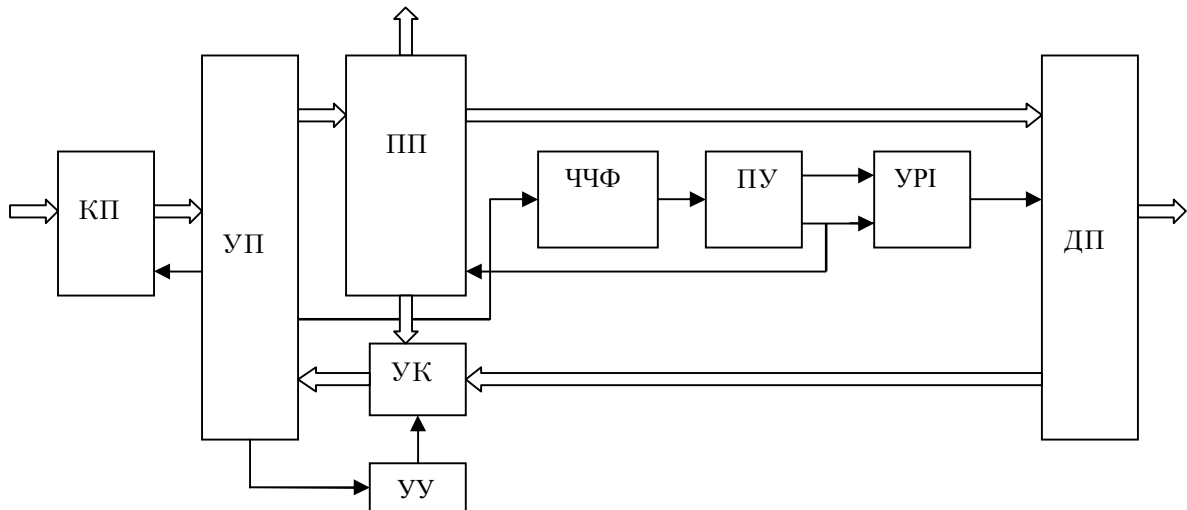


Рис. 1. Схема СШ з проміжною пам'яттю

Вхідні дані СШ послідовно через КП у вигляді патернів поступають на устрій порівняння (УП). На другий вхід УП через устрій комутації (УК), який перемикається устроєм управління (УУ), поступає інформація з ДП. У разі відсутності в ДП подібного патерну, УП, після перемикавання УК за допомогою УУ, порівнює інформацію, що зберігається в ПП, з даними в КП і за відсутності в ПП подібного патерну пропускає патерн, що поступив у КП в проміжну пам'ять. Одночасно на частотно-часовий фільтр (ЧЧФ) з УП поступає імпульс (ознака патерну) з фіксованою величиною амплітуди. Після закінчення циклу порівняння УП дозволяє подати на свій вхід з КП новий патерн. У подальшому, під час вступу на УП подібного патерну в момент порівняння його з патерном, що зберігається в ПП, на ЧЧФ також приходить з УП імпульс ознаки патерну. У ЧЧФ проводиться згортка цієї послідовності імпульсів з ваговою функцією, в результаті використання якої у вихідному сигналі ЧЧФ внесок кожного з цих імпульсів убавляється у міру збільшення тимчасового інтервалу до моменту їх надходження з УП.

У процесі надходження ознак подібних патернів, при досягненні на виході ЧЧФ верхнього порогового рівня U_v , пороговий устрій (ПУ) дає дозвіл на перезапис цього патерну з ПП в ДП через устрій ранжування інформації (УРІ). Сигнал на другому виході ПУ призначений для стирання інформації у ПП при зниженні на виході ЧЧФ сигналу до рівня нижнього порогового рівня U_n , після приходу на ЧЧФ першої ознаки подібних патернів.

Верхній поріг ПУ U_v у міру збільшення тривалості інтервалу роботи ЧЧФ з подібними патернами (T_x) поступово зменшуватиметься, що дозволяє, зокрема, проводити захист від первинних інтенсивних інформаційних атак на СШ. Нижній поріг U_n , також із збільшенням інтервалу T_x змінюється, але його зміна відбувається у бік збільшення значення U_n . Така робота ПУ, із зміною верхнього і нижнього порогових рівнів, підвищує робастність інформації, що зберігається в ДП.

Аналогічна обробка інформації у ЧЧФ проводиться паралельно за всіма ознаками патернів, що зберігаються в ПП.

Виходячи з цілей користувача СШ, можна ввести поділ інформації, що зберігається в ДП, на основі застосування критерію тривалого стійкого надходження однотипних патернів на вхід СШ. Для цього ДП розбивається на M сегментів і у процесі надходження на вхід СШ даних патерни переводяться в більш верхній сегмент ДП, що має більший показник якості за цим критерієм.

Для реалізації цього режиму роботи, УУ через УК з періодом T_k , який береться значно великим максимальної величини T_x , переводить УП у режим порівняння патерну з КП тільки за даними, що зберігаються в ДП. В цьому випадку, за наявності в ДП патерну подібного тому, що є на вході УП, з виходу УП на ЧЧФ поступає імпульс ознаки наявності патерну. Далі ЧЧФ працює аналогічно розглянутій вище процедурі і при досягненні на виході ЧЧФ верхнього порогового рівня U_v , ПУ видає сигнал на УРІ, яке враховує кількість вирішуючих сигналів з ПУ і дозволяє переформувати адресацію ДП для перекладу патерну у вищий за рівнем сегмент.

При зниженні на виході ЧЧФ сигналу до значення, рівного нижньому пороговому рівню U_n , сигнал з другого виходу ПУ також поступає на УРІ, який у цьому випадку дозволяє перенесення занесеного у ДП патерну в нижчий за рівнем сегмент, що відповідатиме зниженню активності надходження подібних патернів на вхід СШ.

При такій роботі УРІ, відповідно до критерію тривалого стійкого надходження на вхід СШ подібних патернів, в більш верхньому за номером сегменті будуть зберігатися патерни, які частіше поступають і досвід ДП для подачі інформації на УП можна починати з самого верхнього сегмента. Така

послідовність опиту ДП дозволить скоротити процедуру порівняння даних, що зберігаються в ДП, з тими, що поступають на вхід СШ.

Додатково до тієї, що знімається з ДП відфільтрованої і впорядкованої інформації з ПП можна знімати поточну інформацію, що поступає.

Обробка інформації в ЧЧФ може проводитися аналогічно наданому вище способу фільтрації даних [8]. Для можливості оцінки важливості якого-небудь патерну з множини подібних патернів V , введемо ціну ознаки цих патернів C_v . У такому разі оцінку активності патернів з множини V на момент проведення аналізу можна визначити на основі виразу, який є дискретною лінійною згортою [9] ціни ознаки патерну $C_v(n)$ з дискретно-тимчасовою ваговою функцією $W(n)$ з довжинами N_c і N_w відповідно.

$$P_v(n) = \sum_{m=0}^{L-1} W(m) \cdot \tilde{N}_v(n-m), \quad (1)$$

де $L = N_c + N_w - 1$.

Величина $C_v(n)$, залежно від надходження на вхід ЧЧФ ознаки патерну, що повторюється, або її відсутності, набуватиме дискретних значень C_v і 0.

Оскільки довжина послідовності подібних патернів N_c заздалегідь невідома, то умовно її можна вважати за нескінченну і лінійну згортку можна записати у вигляді:

$$P_v(n) = \sum_{m=0}^{\infty} W(m) \cdot C_v(n-m). \quad (2)$$

Для обчислення згортки вигляду (2) скористаємося методом перекриття з підсумовуванням, що достатньо легко реалізується на мові MATLAB [10]. При реалізації цього методу послідовність $C_v(n)$ розбивається на суміжні секції довжиною N_{ck} , порівняною з довжиною N_w і має вигляд:

$$\tilde{N}_v(n) = \sum_{k=0}^{\infty} C_{vk}(n). \quad (3)$$

У такому випадку згортка (2) записується так:

$$P_v(n) = \sum_{m=0}^{\infty} W(m) \sum_{k=0}^{L-1} C_{vk}(n-m), \quad (4)$$

де $L = N_{ck} + N_w - 1$.

Дискретно-тимчасову вагову функцію $W(n)$ можна записати у вигляді експоненціальної функції:

$$W(n) = \ell^{-(kn)^\beta}, \quad (5)$$

де k – коефіцієнт крутизни вагової функції; β – показник степеня.

Показник степеня β визначає форму вагової функції $W(n)$ і може теоретично набувати значень від 0 до ∞ . Практично можна обирати значення β від 1 до 5.

Як вагову функцію $W(n)$ можна також використовувати, часто вживану в нейроподібних елементах динамічних нейронних мереж, сигмоїдну функцію [11] вигляду:

$$W(n) = \frac{A}{1 + \ell^{(kn)^\beta}}, \quad (6)$$

де A – нормуючий множник.

Природно, можливе застосування у будь-якій іншій зручній для користувача ваговій функції.

На рисунку 2 представлена змодельована з використанням мови MATLAB [10] процедура проходження через ЧЧФ ознак подібних патернів.

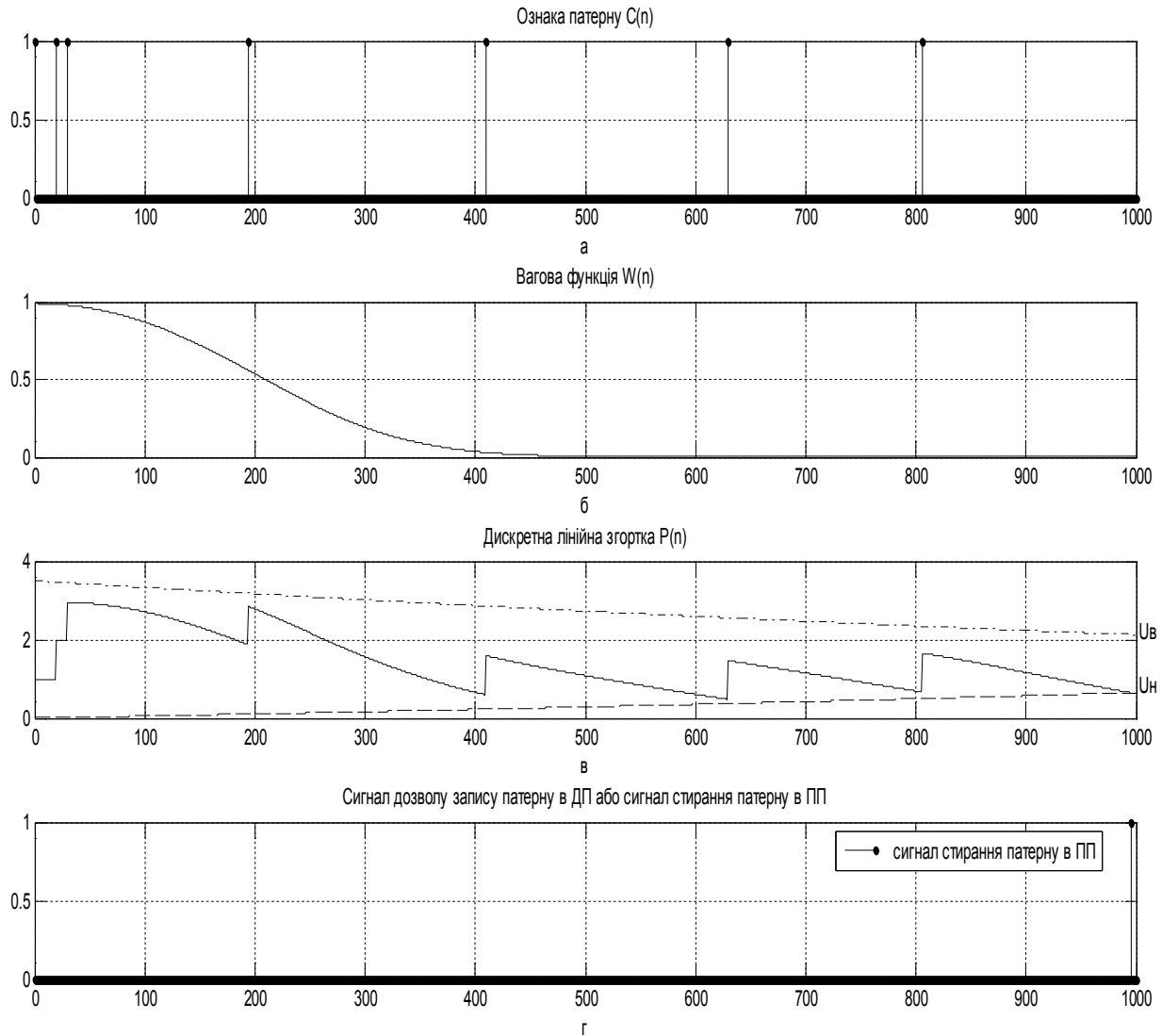


Рис. 2. Фільтрація послідовності ознак подібних патернів

З рисунка 2 видно, що робота ЧЧФ залежить від кількості і часу надходжень на вхід СШП подібних патернів. Як вагова функція використовувалася сигмоїдна функція вигляду (6) з $\beta = 2$.

Устрої ЧЧФ, ПУ і УРІ в даній СШП деякою мірою виконують функції гіпокампу і мигдалини лімбічної системи центральної нервової системи людини [6].

При побудові СШП на основі тимчасової ієрархічної пам'яті НТМ (Hierarchical Temporal Memory) [12] блоки УП і ПП об'єднуються в один блок ПП за рахунок наявності в системі НТМ великої кількості зворотних зв'язків, що значно перевищують кількість вхідних сигналів. У такому разі, розширюються можливості СШП за рахунок того, що НТМ дозволяє обробляти паралельно декілька патернів. Так за даними [13] мозок людини, значно спрощеною моделлю неокортексу якого є система НТМ, здатний обробляти одночасно 7 ± 2 вхідних патернів.

Для обробки вхідних даних, представлених у вигляді нечітких множин, можна при побудові устроїв СШП скористатися нейронечіткими мережами, розглянутими в [14].

Висновки. Застосування СШП, з розглянутою організацією пам'яті, дозволить їх ефективно використовувати в таких системах обробки даних, як формування основного фонду за різними напрямками потоку інформації, таких як: бібліотечний, патентний, формування корпоративних баз даних клієнтів і продукції, створення систем автоматизації маркетингу, таких як: CRM [15], виявлення профілів стійких інтересів користувачів Internet, що дозволяють Internet-сервісам адаптуватися до цих інтересів і в багатьох інших випадках, а також у процесі навчання СШП.

Список використаної літератури:

1. Лудченко А.А. Основы научных исследований : учеб. пособие / А.А. Лудченко, Я.А. Лудченко, Т.А. Примак. – К. : Знання, КОО, 2001. – 113 с.
2. Башмаков А.И. Интеллектуальные информационные технологии : учеб. пособие / А.И. Башмаков, И.А. Башмаков. – М. : Изд-во МГТУ им. Н.Э. Баумана, 2005. – 304 с.
3. Хайкин С. Нейронные сети: полный курс : пер. с англ. / С.Хайкин. – М. : ООО «И.Д. Вильямс», 2006. – 1104 с.
4. Блум Ф. Мозг, разум и поведение : пер. с англ. / Ф.Блум, А.Лейзерсон, Л.Хофстедтер. – М. : Мир, 1988. – 248 с.
5. Ларичев О.И. Теория и методы принятия решений, а также Хроника событий в Волшебных странах : учебник / О.И. Ларичев. – М. : Логос, 2002. – 392 с.
6. Физиология человека : В 3-х т. Т. 1. : пер. с англ. / под ред. Р.Шмидта, Г.Тевса. – М. : Мир, 2005. – 323 с.
7. Fader P.S. Probability Models for Customer-Base Analysis / P.S. Fader, B.G.S. Hardie // J. Interactive Marketing. – 2009. – Vol. 23, № 1. – Pp. 61–69.
8. Спосіб автоматичного впорядкування даних в пам'яті телефонних апаратів : пат. на винахід / М.В. Козлов, Ю.О. Подчащинський. – № 82718. – 12.05.08, Бюл. № 9.
9. Лайонс Р. Цифровая обработка сигналов : пер. с англ. / Р.Лайонс. – М. : ООО «Бином-Пресс», 2006. – 656 с.
10. Солонина А.И. Цифровая обработка сигналов. Моделирование в MATLAB / А.И. Солонина, С.М. Арбузов. – СПб. : БХВ-Петербург, 2008. – 816 с.
11. Харламов А.А. Использование сигмоидной функции в нейроподобных элементах с временной суммацией входных сигналов в составе динамической нейронной сети / А.А. Харламов, Г.Н. Рябов // Наука и образование: электронное научно-техническое издание. – М. : МГТУ им. Н.Э. Баумана, 2008. – № 2.
12. Hierarchical Temporal Memory including HTM Cortical Learning Algorithms, Version 0.2.1, Septemta 12, 2011. – © Numenta Inc., 2011. – P. 68.
13. Miller G.A. The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information / G.A. Miller // J. Psychological Review. – 1956. – Vol. 63. – Pp. 81–97.
14. Субботін С.О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : навч. посібник / С.О. Субботін. – Запоріжжя : ЗНТУ, 2008. – 341 с.
15. Ковалев А.М. Microsoft Dynamics CRM: первые шаги. Серия «Информационные бизнес технологии» / А.М. Ковалев. – М. : ЭКОМ Паблишерз, 2007. – 232 с.

КОЗЛОВ Михайло Венедиктович – кандидат технічних наук, доцент кафедри автоматизації і управління в технічних системах Житомирського державного технологічного університету.

Наукові інтереси:

– прилади та методи вимірювання електричних величин;

– методи прийняття рішень;

– системи штучного інтелекту.

Тел.: (0412) 44–59–26.

E-mail: mike21k@rambler.ru

Стаття надійшла до редакції 14.05.2012