

ДИНАМІЧНА АСОЦІАТИВНА ПАМ'ЯТЬ НА ОСНОВІ ВІДКРИТОЇ РЕКУРЕНТНОЇ НЕЙРОМЕРЕЖІ

Abstract. *Mathematical model of open dynamic recurrent neural network, that hasn't got hidden neurons, is described, and concept of "freezing" of its dynamic attractors is proposed. Solution of generalized stability equation is found and non-iterative method for training of this type of networks is developed. On the basis of this method a new model of neural associative memory is created, estimations of attraction radius and memory capacity are obtained. Size of dynamic attractor's trajectory is estimated and possibility of managing of dynamic attractors by changing level of influence of different feedback components is shown. Software model of dynamic associative memory is developed, and experimental confirmation of theoretical analysis is obtained. An analogy between dynamic attractors and neural activity patterns is remarked.*

Key words: *neural networks, associative memory, recurrent networks, non-iterative learning.*

Анотація. *Розглянуто модель відкритої рекурентної нейромережі, що відрізняється відсутністю прихованих нейронів, та запропоновано метод "заморожування" її динамічних аттракторів. Одержано вирішення рівняння стабільності й розроблено метод неітеративного навчання таких нейромереж. На його основі створено нову модель нейронної динамічної асоціативної пам'яті, одержано оцінки аттракторного радіуса та обсягу пам'яті. Оцінено розміри траєкторії динамічних аттракторів і виявлено можливість управління динамічними аттракторами шляхом зміни впливу різних складових зворотного зв'язку. Створено програмну модель динамічної асоціативної пам'яті та одержано експериментальне підтвердження результатів теоретичного аналізу. Відзначено аналогію між динамічними аттракторами та цугами нервової активності мозку.*

Ключові слова: *нейронні мережі, асоціативна пам'ять, рекурентні мережі, неітеративне навчання.*

Аннотация. *Рассмотрена модель открытой рекуррентной нейросети, отличающейся отсутствием скрытых нейронов, и предложен метод "замораживания" ее динамических аттракторов. Получено решение уравнения устойчивости и разработан метод неитеративного обучения таких нейронных сетей. На его основе разработана новая модель нейронной динамической ассоциативной памяти, получены оценки величины аттракторного радиуса и объема памяти. Оценены размеры траектории динамических аттракторов и показана возможность управления динамическими аттракторами путем изменения влияния различных составляющих обратной связи. Создана программная модель динамической ассоциативной памяти и получено экспериментальное подтверждение результатов теоретического анализа. Отмечена аналогия между динамическими аттракторами и цугами нервной активности мозга.*

Ключевые слова: *нейронные сети, ассоциативная память, рекуррентные сети, неитеративное обучение.*

1. Вступ

Асоціативні властивості рекурентних нейронних мереж, зокрема, мережі Хопфілда, базуються на їх мультистабільності [1, 2]. Стабільні стани (аттрактори) такої нейромережі можна розглядати як запам'ятовані образи, а процес конвергенції, тобто послідовних змін стану в напрямку найближчого аттрактора, – як пошук рішення за асоціацією з образом, заданим початковим станом нейромережі. У динамічних рекурентних нейромережах, міжнейронні зв'язки яких мають елементи затримки, аттракторами є стійкі послідовності станів-реакцій на відповідні послідовності зовнішніх стимулів [3–5]. Такі послідовності можна розглядати як динамічні образи, а саму динамічну рекурентну нейромережу – як динамічну асоціативну пам'ять, що здатна розпізнавати та відтворювати запам'ятовані динамічні образи.

Навчання динамічних рекурентних нейромереж здійснюють за ітеративним методом зворотного поширення похибки в часі [6–8]. Цей процес вимагає значних витрат часу та обчислювальних ресурсів, оскільки формування рекурентних зв'язків часто не є монотонним, а його результати не завжди стабільні. Складності виникають і при інтерпретації результатів навчання динамічних рекурентних нейромереж, оскільки методи та критерії, розроблені для статичних нейромереж, не завжди придатні для оцінки поведінки.

Це дає підстави для детального вивчення відносно простих відкритих динамічних рекурентних нейромереж, стан яких доступний для постійного контролю, оскільки вони не мають прихованих нейронів. До них можуть бути використані відомі аналітичні методи дослідження та неітеративні методи навчання [9], які не потребують значних обчислювальних ресурсів. Прозорість відкритих нейромереж робить їх перспективними для моделювання елементів нервової системи, зокрема, асоціативних полів та рекурентних структур зорової кори мозку [8]. Вони можуть скласти конкуренцію іншим, більш складним, динамічним нейромережам при створенні прикладних систем, зокрема, динамічної асоціативної пам'яті.

У даній роботі пропонується нова модель динамічної асоціативної пам'яті на основі відкритої динамічної рекурентної нейромережі. В розд. 2 розглянуто структуру та поведінку таких нейромереж, запропоновано концепцію уявної статичної нейромережі, яка дозволяє застосувати до них неітеративні методи навчання. Розділ 3 містить оцінку асоціативних властивостей відкритих динамічних рекурентних нейромереж. В розд. 4 розглянуто роботу рекурентної нейромережі як динамічної асоціативної пам'яті. Розділ 5 містить результати експериментів, виконаних на створеній програмній моделі відкритої рекурентної нейромережі, що демонструють функціонування механізму динамічної асоціативної пам'яті і дозволяють припускати існування аналогічних механізмів у латеральних нейроструктурах мозку.

2. Відкрита динамічна рекурентна нейромережа

Відкритою називатимемо рекурентну нейромережу без прихованих нейронів, в якій значення всіх сигналів на входах нейронів є доступними для зовнішнього спостереження. Прикладами відкритих нейромереж можуть служити мережа Хопфілда та двостороння асоціативна пам'ять [1, 10]. Якщо зворотні зв'язки такої нейромережі містять елементи затримки, вона має динамічні властивості.

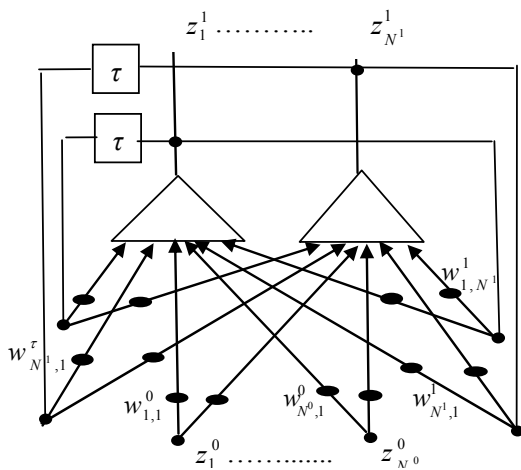


Рис. 1. Відкрита динамічна рекурентна нейромережа (ВДРН)

На рис. 1 представлено схематичне зображення відкритої динамічної рекурентної нейромережі, яка має N^1 нейронів і відповідно N^1 бінарних виходів. Число входів нейромережі N^0 може відрізнятись від числа нейронів. Вважатимемо, що сигнали на входах також є бінарними. Кількість затриманих зв'язків нейромережі N^τ може відрізнятись від числа нейронів ($N^\tau \leq N^1$). Також вважатимемо, що величини затримки сигналів τ однакові.

Сукупність прямих та затриманих сигналів, що одночасно діють на входи всіх

нейронів, визначає N^* – вимірний вектор стану нейромережі $Z^*(t) = \{z_n^*\}_{n=1}^{N^*}$, де $N^* = N^1 + N^\tau + N^0$. Розглядаючи послідовності станів у дискретні моменти часу $\dots t-1, t, t+1 \dots$, поточне значення реакції нейромережі представимо вектором

$$Z^1(t) = F[S^1(t-1)] = \left\{ f(s_n^1(t-1)) \right\}_{n=1}^{N^1}, \quad (1)$$

де $f(\cdot)$ – активаційна функція нейрона;

$S^1(t)$ – вектор постсинаптичних потенціалів (ПСП):

$$S^1(t) = W^{11}Z^1(t) + W^{1\tau}Z^1(t-\tau) + W^{10}Z^0(t), \quad (2)$$

де $Z^1(t)$, $Z^1(t-\tau)$ – вектори поточних та затриманих значень реакції нейронів;

$Z^0(t)$ – вектор входу нейромережі;

W^{11} , $W^{1\tau}$ та W^{10} – матриці ваги зв'язків між входами та виходами нейронів, входами та затриманими виходами, а також зовнішніми входами динамічної рекурентної нейромережі.

Умовою перебування нейромережі у стані атрактора є виконання рівності

$$Z^1(t) = F[S^1(t)] = \left\{ f(s_n^1(t)) \right\}_{n=1}^{N^1}. \quad (3)$$

Для статичної нейромережі це означає, що значення виходу, а відповідно і постсинаптичного потенціалу нейронів, перестає залежати від часу. У динамічної нейромережі, яка має затримані зворотні зв'язки, це неможливо. Її динамічні атрактори є послідовностями атракторних станів нейромережі, що є реакціями на відповідні послідовності зовнішніх стимулів. Позначимо m -й вектор такої послідовності, як R^{*m} . Його компонентами є три вектори: $R^{1,m}$ – поточних значень виходів нейронів, $R^{1,m-1}$ – попереднього $(m-1)$ -го атрактора, який відповідає затриманим виходам нейронів, $R^{0,m}$ – поточних значень зовнішнього стимулу.

Враховуючи монотонність активаційної функції, умову (3) можна записати як

$$R^{1,m1} = \Lambda[W^{11}R^{1,m} + W^{1\tau}R^{1,m-1} + W^{1,0}R^{0,m}], \quad (4)$$

де Λ – деяка позитивно визначена діагональна матриця $N^1 \times N^1$ ($\lambda_{ii} \geq 0$).

Якщо вектори $R^{1,m}$, $R^{1,m-1}$ та $R^{0,m}$ відомі, то це рівняння можна вирішити відносно матриць W^{11} , $W^{1\tau}$ та $W^{1,0}$. Для цього уявімо статичну симетричну нейромережу, що має N^* нейронів. Вектор стану такої нейромережі $Z^*(t)$ має три компоненти: $Z^1(t)$, $Z^1(t-\tau)$ та $Z^0(t)$. ПСП уявної нейромережі можна представити матричним рівнянням:

$$S^*(t) = \begin{vmatrix} S^1(t) \\ S^1(t-\tau) \\ S^0(t) \end{vmatrix} = \begin{vmatrix} W^{11} & W^{1\tau} & W^{10} \\ W^{\tau 1} & W^{\tau\tau} & W^{\tau 0} \\ W^{01} & W^{0\tau} & W^{00} \end{vmatrix} \begin{vmatrix} Z^1(t) \\ Z^1(t-\tau) \\ Z^0(t) \end{vmatrix} = \begin{vmatrix} W^{11}Z^1(t) + W^{1\tau}Z^1(t-\tau) + W^{10}Z^0(t) \\ W^{\tau 1}Z^1(t) + W^{\tau\tau}Z^1(t-\tau) + W^{\tau 0}Z^0(t) \\ W^{01}Z^1(t) + W^{0\tau}Z^1(t-\tau) + W^{00}Z^0(t) \end{vmatrix}.$$

Припустимо, що вектори її атракторних станів тотожні тим, що утворюють динамічні атрактори динамічної рекурентної нейромережі. Розглядаючи динамічні атрактори як послідовності векторів, що змінюють один одного в дискретні моменти часу $\dots t-\tau, t, t+\tau \dots$, встановимо відповідність між станами динамічної нейромережі та її атракторами: $Z^1(t) \Rightarrow R^{1,m}$; $Z^1(t-\tau) \Rightarrow R^{1,m-1}$. Представимо послідовність атракторних станів у вигляді матриці:

$$\mathfrak{R}^* = \begin{vmatrix} R^{1,1} \dots & R^{1,m} \dots & R^{1,M} \\ R^{1,0} \dots & R^{1,m-1} \dots & R^{1,M-1} \\ R^{0,1} \dots & R^{0,m} \dots & R^{0,M} \end{vmatrix},$$

стовпчиками якої є вектори статичних атракторів уявної нейромережі. Користуючись аналогією з рівнянням (4), запишемо узагальнене рівняння атракторного стану уявної нейромережі:

$$\mathfrak{R}^* = \Lambda W^* \mathfrak{R}^*, \quad (5)$$

Вирішуючи його відносно W^* , одержимо

$$W^* = \Lambda^{-1} \mathfrak{R}^* (\mathfrak{R}^*)^+, \quad (6)$$

де $(\mathfrak{R}^*)^+$ – псевдообернена матриця \mathfrak{R}^* [11]. Оскільки вектори $R^{1,m}$ та $R^{0,m}$ є бінарними, то можна вважати $\Lambda = I$.

Для знаходження елементів матриці W^* застосуємо псевдоінверсний алгоритм [12, 13]:

$$w_{i,j}^{*m+1} = w_{i,j}^{*m} + (r_i^{m+1} - s_i^{m+1})(r_j^{m+1} - s_j^{m+1}) / d^{m+1}, \quad (7)$$

$$s_i^{m+1} = \sum_{k=1}^{N^*} w_{ik}^{*m} r_k^{m+1}, \quad d^{m+1} = \sum_{k=1}^{N^*} r_{ik}^{m+1} (r_k^{m+1} - s_k^{m+1}),$$

де r_k^{m+1} – компонента вектора R^{*m+1} .

Матриця $W^* = \mathfrak{R}^* (\mathfrak{R}^*)^+$ є проекційною в лінійному просторі \mathfrak{Z} , напнутому на M векторів з \mathfrak{R}^* . Вона має такі властивості:

$$W^* = (W^*)^2, \quad w_{i,i}^* = \sum_{j=1}^{N^*} (w_{i,j}^*)^2, \quad Tr W^* = \sum_{i=1}^{N^*} w_{i,i}^* = M, \\ \overline{w_{i,i}^*} = \frac{M}{N^*}, \quad \frac{M}{N^*} \left[1 - \frac{M}{N^*} \right] \geq (w_{i,j}^*)^2 \geq \frac{M}{N^* (N^* - 1)} \left[1 - \frac{M}{N^*} \right], \quad i \neq j. \quad (8)$$

Величина недиагональних елементів знаходиться в межах між верхньою оцінкою, що відповідає розрідженій матриці, більшість елементів якої мають значення, близькі до нуля, та нижньою, що відповідає відносно рівномірному розподілу значень її елементів. Діагональні елементи визначають вагу позитивних зворотних зв'язків нейронів. При збільшенні відношення M/N^* нейрони втрачають чутливість до зовнішніх збуджень, що приводить до зупинки процесу конвергенції і появи хибних атракторів.

При відсутності затримок зворотних зв'язків ($\tau = 0$) розрахована за формулами (7) симетрична нейромережа відповідає моделі асоціативної пам'яті, керованої порогом [13], яка має $M < N^1$ головних атракторів, представлених парами векторів $\{R^{1,m}, R^{0,m}\}_{m=1}^M$. Якщо така нейромережа знаходиться в нестабільному стані, відбувається процес конвергенції – послідовних змін стану в напрямку найближчого головного атрактора.

При наявності затриманих зворотних зв'язків у матриці W^* залишаються блоки W^{11} , $W^{1\tau}$ та W^{10} , які відповідають реальним зв'язкам. У такій нейромережі процес конвергенції обмежений

першим кроком переходу нейронів з поточного в наступний стан. Таке обмеження не заважає існуванню динамічних атракторів рекурентної нейромережі, для яких наступний стан є іншим атрактором. Динамічний атрактор може починатись з будь-якого вектора послідовності, представлені матрицею \mathfrak{R}^* , і закінчуватись останнім її вектором $R^{*,M}$. Якщо початковий стан не є атрактором, ситуація виявляється більш складною, оскільки, на відміну від симетричної статичної нейромережі, наближається до найближчого атрактора монотонно, поведінка динамічної рекурентної нейромережі поза динамічним атрактором може бути не монотонною.

3. Атракторний радіус динамічної рекурентної нейромережі

Поведінка статичної рекурентної нейромережі залежить від заповнення її асоціативної пам'яті, визначеної як відношення числа атракторів до числа нейронів M/N . Процес конвергенції в ній може зупинитись, не досягнувши головного атрактора, причому зі збільшенням співвідношення M/N імовірність такої зупинки зростає. В мережі Хопфілда конвергенція припиняється при $M/N > 0,14$ [1, 2, 7]. При застосуванні псевдоінверсного алгоритму ця межа наближається до 0,25 [14]. Найбільше значення $M/N \approx 0,7$ досягнуто при застосуванні спеціального методу рознасичення синаптичної матриці [15].

Для оцінки атракторних властивостей статичної нейромережі в [12] запропоновано поняття атракторного радіуса, визначеного як число нейронів, що змінюють свій вихід на останньому кроці конвергенції у стан головного атрактора. Цей критерій відповідає моделі синхронного обчислення стану нейромережі, коли спочатку обчислюються ПСП всіх нейронів, а потім одночасно – значення їхніх виходів. В [2] показано, що він є ефективним і при асинхронному обчисленні стану нейромережі.

Спробуємо застосувати цей критерій для динамічної рекурентної нейромережі. Для цього скористаємось введеною нами моделлю уявної нейромережі і розглянемо її поведінку на останньому кроці конвергенції при досягненні стану головного атрактора. Нехай z_j^1 та z_j^0 – значення виходу та входу j -го нейрона уявної нейромережі перед переходом у стан атрактора, а r_j^1 та r_j^0 – після переходу. Вважатимемо, що $|r_j^1| = 1$. Користуючись рівністю для стану атрактора

$$s_i(t) = \sum_{j=1}^{N^*} w_{i,j}^* r_j^1 + \sum_{j=1}^{N^*} w_{i,j}^* r_j^0 = r_i$$

і враховуючи, що при переході у стан атрактора знаки ПСП та виходу кожного нейрона збігаються, представимо умову досягнення атрактора в такому вигляді:

$$r_i s_i(t) = r_i \left[\sum_{j=1}^{N^*} w_{i,j}^* r_j - \sum_{j=1}^{N^*} w_{i,j}^* [r_j - z_j(t)] \right] = (r_i)^2 - 2 \sum_{h=1}^H w_{i,j}^* r_i r_j > 0, \quad i = 1, 2, \dots, N^*, \quad (9)$$

де H – число елементів векторів $Z^*(t)$ і R^* , знаки яких є протилежними.

Для нейронів, вихід яких змінює знак при досягненні нейромережею стану атрактора, це співвідношення можна записати так:

$$2 \sum_{h=1}^{H-1} w_{i,j}^* r_i r_j < 1 - 2w_{i,i}^*.$$

Підсилюючи цю нерівність шляхом заміни членів суми їх середніми абсолютними значеннями, знайдемо верхню межу для величини атракторного радіуса:

$$H < 1 + (1 - 2\overline{w_{i,i}^*}) / 2\overline{w_{i,j}^*}. \quad (10)$$

Цей результат ми одержали на основі аналізу лише одного останнього кроку переходу уявної нейромережі у стан головного атрактора, в якому брали участь усі N^* уявних нейронів цієї нейромережі. Виокремлюючи з них N^1 реальних нейронів, для яких ця нерівність, безумовно, виконується, ми одержимо ту саму нерівність, яка відповідає кроку переходу динамічної рекурентної нейромережі з будь-якого стану у стан атрактора.

Підставляючи в нерівність (10) значення $\overline{w_{i,j}^*}$ з (8), одержимо оцінку середньої величини атракторного радіуса:

$$H < 1 + \left[\frac{N^*}{2} - M \right] \sqrt{\frac{N^* - 1}{M(N^* - M)}} \xrightarrow{M \ll N} 1 + \frac{N^*}{2\sqrt{M}}. \quad (11)$$

Для відкритої динамічної рекурентної нейромережі ця нерівність визначає максимальну відстань між сусідніми векторами навчальної послідовності, при якій ще можливе існування динамічних атракторів. Оскільки кожен з цих векторів містить компоненту зовнішнього стимулу, то динамічний атрактор можна розглядати просто як реакцію нейромережі на відповідну послідовність стимулів. Величина атракторного радіуса визначає максимальне відхилення векторів стимулу від їх еталонних значень, за яких динамічний атрактор відповідатиме навчальній послідовності.

4. Динамічна нейронна асоціативна пам'ять

Розглядаючи рекурентну нейромережу як динамічну асоціативну пам'ять, спробуємо оцінити її здатність запам'ятовувати послідовності векторів та відтворювати їх за асоціацією із послідовностями стимулів, що діють на її входи.

Представлятимемо поведінку такої нейромережі послідовністю кадрів, кожен з яких містить N^* – вимірний вектор її поточного стану $Z^*(t)$. Кадри надходять з інтервалом τ , який відповідає часу затримки зворотних зв'язків нейромережі. Навчальна послідовність включає M кадрів, які містять вектори атракторів $R^{*m} = [R^{1m}, R^{1m-1}, R^{0m}]$. В кожному кадрі вектор R^{0m} є ключем до наступного атракторного стану нейромережі, тому у випадку автоасоціативної динамічної пам'яті $R^{*m} = [R^{1m}, R^{1m-1}, R^{1m+1}]$.

Розглянемо поведінку відкритої динамічної рекурентної нейромережі, яка в початковий момент t знаходилась у стані атрактора $R^{*m} = [R^{1m}, R^{1m-1}, R^{0m+1}]$. Цей стан нейромережа зберігала до моменту $t + \tau$, коли затримане значення вектора $R^{1,m}$ надійшло на входи нейронів. За цей момент зовнішній стимул також змінив значення з R^{0m} на $Z^0(t + \tau)$, внаслідок чого ПСП нейромережі набуло нового значення:

$$S^1(t + \tau) = W^{11}R^{1,m} + W^{1\tau}R^{1,m} + W^{10}Z^0(t + \tau) = R^{1m+1} + \Delta^{11} + \Delta^0, \quad (12)$$

$$\Delta^{11} = W^{11}(R^{1m1} - R^{1m+1}), \quad \Delta^0 = W^{10}[Z(t + \tau) - R^{0,m+1}].$$

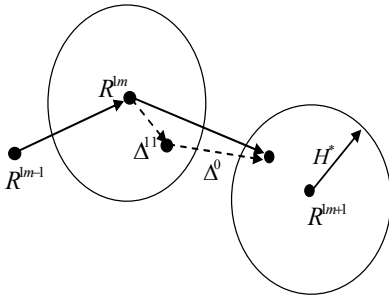


Рис. 2. Перехід ВРДН з атракторного стану R^{1m} в R^{1m+1}

Цю ситуацію зображено на рис. 2, де позначено атракторні стани нейромережі та виділено їхні області притягіння, обмежені величиною атракторного радіуса H^* . Вектори Δ^{11} та Δ^0 визначають переміщення вектора ПСП, який перед тим мав значення атрактора R^{1m} .

В зображеній ситуації вектор ПСП в момент $t + \tau$ опиняється в області притягіння атрактора R^{1m+1} , який стає наступним станом

нейромережі при завершенні ітерації. Якщо в зображеній ситуації зовнішній стимул залишити незмінним: $\Delta^{0,m}(t + \tau) = 0$, то при завершенні ітерації динамічна нейромережа повернеться у стан попереднього атрактора R^{1m} .

При меншій величині атракторного радіуса вектор ПСП може опинитись за межами областей притягіння головних атракторів. Тоді подальша поведінка нейромережі залежатиме від співвідношення між складовими $W^{11}R^{1,m}$, $W^{1\tau}R^{1,m}$ та $W^{10}Z^0(t + \tau)$. Змінюючи їх вагу, можна у значних межах коригувати динамічні властивості нейромережі. Збільшення ваги складової $W^{11}R^{1,m}$ наближує вектор ПСП до попереднього атрактора, що робить імовірним повернення нейромережі в попередній стан $R^{1,m}$. При зростанні складової $W^{1\tau}R^{1,m}$ більш імовірним стає перехід у наступний стан $R^{1,m+1}$. Динамічна нейромережа і далі намагатиметься переходити в наступні стани відповідно до порядку їх розташування в навчальній послідовності. При домінуванні цієї складової нейромережа буде постійно перебувати у стані динамічного атрактора, повторюючи фрагменти навчальної послідовності. Підсилення складової $W^{10}Z^0(t + \tau)$ веде до повністю керованої поведінки, коли динамічна нейромережа повторює стани, задані зовнішнім стимулом, незалежно від їх розташування в навчальній послідовності.

Розглянемо більш детально співвідношення (12) для пересічного нейрона

$$s_i^1(t + \tau) = r_i^{1,m+1} - 2 \sum_{h=1}^{H^{11}} w_{i,h}^{11} r_h^{1,m+1} - \sum_{j=0}^{N^0} w_{i,j}^0 [r_j^{0,m+1} - z_j^0(t + \tau)], \quad (13)$$

де H^{11} – відстань за Хеммінгом між векторами R^{1m} та R^{1m+1} .

Наступним станом нейромережі буде R^{1m+1} лише за умови $s_i^1(t)r_i^{1,m+1} > 0$, $i = 1.2...N^1$ або

$$(r_i^{1,m+1})^2 > r_i^{1,m+1} \left\{ 2 \sum_{h=1}^{H^{11}} w_{i,h}^{11} r_h^{1,m+1} + \sum_{j=0}^{N^0} w_{i,j}^0 [r_j^{0,m+1} - z_j^0(t + \tau)] \right\}.$$

Підсилимо цю нерівність, беручи середні абсолютні значення складових і виділяючи складову зворотного зв'язку нейрона $w_{i,i}^{11}$:

$$2 \sum_{h \neq i}^{H^{11}-1} |w_{i,h}^{11}| + \sum_{j=0}^{N^0} |w_{i,j}^0 [r_j^{0,m+1} - z_j^0(t + \tau)]| < 1 - 2|w_{i,i}^{11}|.$$

Найбільш сприятливою для виконання цієї нерівності є умова, що зовнішній стимул має однакове значення з очікуваним атрактором

$$z_j^0(t + \tau) = r_j^{0,m+1}. \quad (14)$$

Вважаючи цю умову виконаною, вирішимо нерівність (18) відносно H^{11} :

$$H^{11} < 1 + (1 - 2\overline{w_{i,i}^{11}}) / 2\overline{w_{i,j}^{11}}.$$

Одержане співвідношення повторює формулу (10) для атракторного радіуса уявної нейромережі, що є очікуваним, оскільки матриця W^{11} є діагональним блоком W^* .

Величина H^{11} є верхньою межею відстані по Хеммінгу між сусідніми кадрами навчальної послідовності. Оскільки H^{11} залежить від числа векторів цієї послідовності, то можна оцінити допустиме число таких векторів, а відтак і можливу тривалість динамічного атрактора. Вирішуючи нерівність (11) відносно M , одержимо

$$M < \frac{N^*}{2} \left(1 - \frac{H^{11} - 1}{\sqrt{N^* - 1}} \right). \quad (15)$$

З цього видно, що число векторів навчальної послідовності динамічної рекурентної нейромережі зменшується при зростанні атракторного радіуса і не може перевищувати половини числа нейронів уявної нейромережі. До цього висновку приводить і аналогія із статичною нейромережею, яка при $M > N^*/2$ втрачає здатність до конвергенції.

Розглянуті співвідношення характеризують поведінку відкритої динамічної рекурентної нейромережі, в якій присутні як затримані, так незатримані зв'язки. Якщо вилучити незатримані зв'язки, то формула (12) набуває вигляду

$$\begin{aligned} S^1(t + \tau) &= W^{1\tau} R^{1,m} + W^{10} Z^0(t + \tau) = S^1(t) + \Delta^{1,\tau} + \Delta^0, \\ \Delta^{1,\tau} &= W^{1\tau} (R^{1,m} - R^{1,m-1}), \quad \Delta^0 = W^{10} [Z(t + \tau) - R^{0,m-1}]. \end{aligned} \quad (16)$$

Тут $R^{1,m-1}$ – вектор стану нейромережі, який надходив на вхід нейронів у момент t ; $R^{1,m}$ – вектор, що надходить у момент $t + \tau$.

Для пересічного нейрона

$$s_i^1(t + \tau) = s_i^1(t) - 2 \sum_{h=1}^{H^{1\tau}} w_{i,h}^{1\tau} r_h^{1,m-1} - \sum_{j=0}^{N^0} w_{i,j}^0 [r_j^{0,m} - z_j^0(t + \tau)],$$

де $H^{1\tau}$ – відстань за Хеммінгом між атракторами $R^{1,m}$ та $R^{1,m-1}$.

Умовою для повернення в попередній стан є $s_i^1(t + \tau) r_i^{1,m} > 0$, $i = 1, 2, \dots, N^1$:

$$s_i^1(t) r_i^{1,m} > 2 r_i^{1,m} \sum_{h=1}^{H^{1\tau}} w_{i,h}^{1\tau} r_h^{1,m-1} + r_i^{1,m} \sum_{j=0}^{N^0} w_{i,j}^0 [r_j^{0,m} - z_j^0(t + \tau)].$$

Підсилимо цю нерівність, замінюючи члени суми їх абсолютними значеннями:

$$s_i^1(t) r_i^{1,m} > 2 H^{1\tau} \overline{w_{i,h}^{1\tau}} + \sum_{j=0}^{N^0} |w_{i,j}^0| [r_j^{0,m} - z_j^0(t + \tau)]$$

Найсприятливішим для її виконання є рівність значень зовнішнього стимулу і попереднього атрактора: $z_j^0(t + \tau) = r_j^{0,m}$. За цієї умови

$$H^{1\tau} < \overline{s_i^1(t)r_i^{1,m}} / 2\overline{w_{i,h}^{1,\tau}}, \quad i \neq h. \quad (17)$$

Оскільки в даному випадку кількість зворотних зв'язків скорочено наполовину порівняно з уявною нейромережею, то добуток $\overline{s_i^1(t)r_i^{1,m}} \approx 0,5$. Підставляючи в (17) оцінку величини $\overline{w_{i,h}^{1,\tau}}$ з (8), знайдемо

$$H^{1\tau} < \frac{N^* \sqrt{N^* - 1}}{4\sqrt{M(N^* - M)}}. \quad (18)$$

Порівнюючи з оцінкою (11), одержаною для уявної нейромережі, можна відзначити практично однаковий характер залежності від числа нейронів та розмірів навчальної послідовності. В даному випадку атракторний радіус є вдвічі меншим через скорочення кількості зв'язків нейромережі. Він визначає максимальне відхилення вектора стану нейромережі або зовнішнього стимулу від попереднього атрактора R^{1m} , при якому вона ще повертається в попередній стан. При перевищенні цієї межі динамічна нейромережа переходить у наступний стан, заданий зовнішнім стимулом. При відсутності зовнішнього стимулу вона переходить у стан динамічного атрактора, відтворюючи решту ланцюга станів навчальної послідовності.

5. Експериментальні результати

Експериментальне дослідження динамічної асоціативної пам'яті проводилось за допомогою створеної авторами програмної моделі відкритої рекурентної нейромережі. Нейромережа за архітектурою відповідала рис. 1 і включала 100 нейронів зі знаковою активаційною функцією. Обчислення стану нейронів виконувалось у синхронному режимі, тобто спочатку обчислювались значення постсинаптичних потенціалів, а потім одночасно встановлювались нові значення їх реакції. Величину затримки зворотного зв'язку τ можна було встановлювати в 2 – 6 разів більшою, ніж період обчислення стану нейромережі, що дозволяло виконувати за час затримки кілька локальних ітерацій конвергенції. Навчальна послідовність включала кілька динамічних атракторів, що представляли послідовності кадрів у вигляді триплетів $R^{*m} = [R^{1m}, R^{1m-1}, R^{1m+1}]$ і відрізнялись

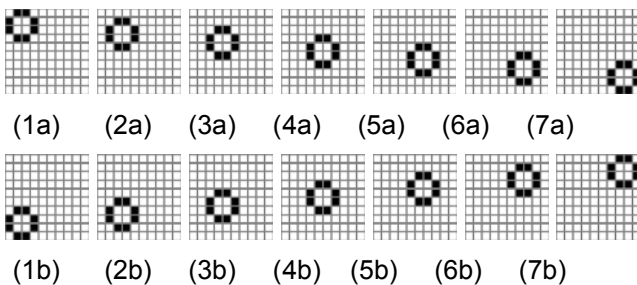


Рис. 3. Динамічний атрактор з одним спільним кадром ($4a = 4b$)

розмірами та формою зображення, ступенем схожості складових триплету. Використовувались зображення в формі прямих ліній та кілець, що переміщувались у кадри, а також послідовності латинських літер.

На рис. 3 представлено два динамічні атрактори у вигляді кілець, що переміщуються по схрещених траєкторіях.

Такі перехрещення приводять до інтерференції відбитків у пам'яті рекурентної нейромережі. Це видно на рис. 4, де верхній ряд представляє зовнішній стимул, а нижній – розподіл ПСП нейронів при проходженні пункту перетину. Залежно від напрямку руху стимулу розподіл ПСП в околі пункту перетину змінюється.

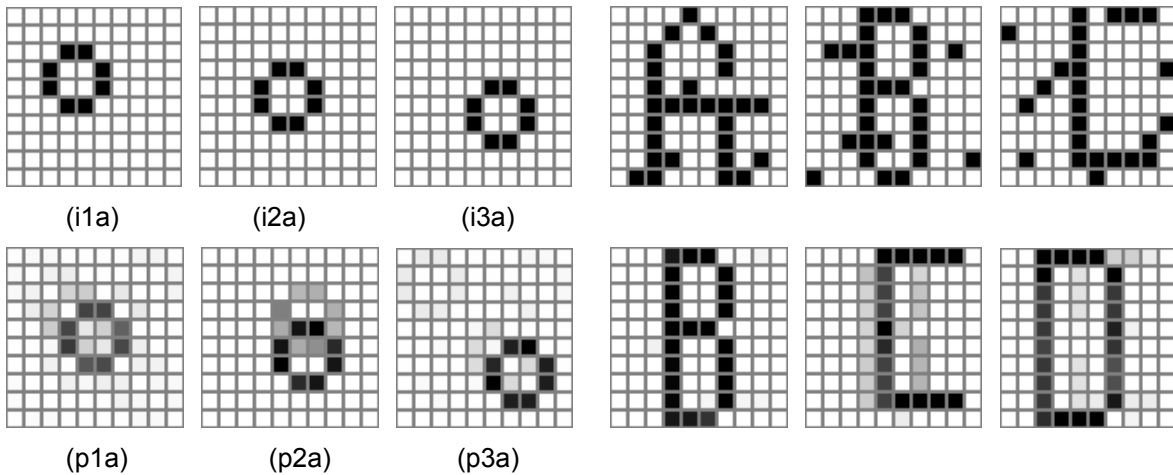


Рис. 4. Входи мережі (зверху) і відповідні постсинапси (знизу) при відтворенні динамічного атрактора із серії (а)

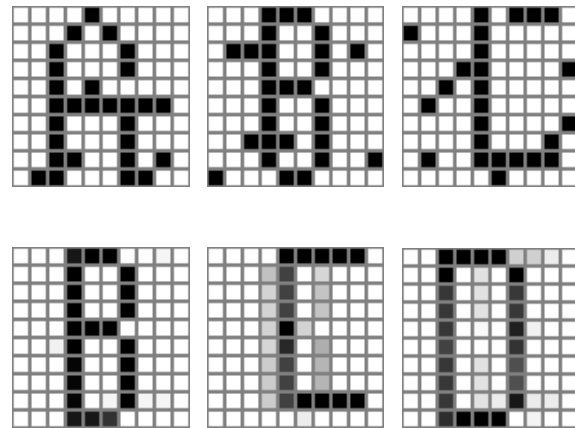


Рис. 5. Збурені входи (зверху) і відповіді мережі (знизу) для рівня шуму $H = 10$

Метою експериментів було оцінити здатність динамічної асоціативної пам'яті нейтралізувати збурення зовнішніх стимулів. Збурення вводилось шляхом зміни полярності H випадково обраних елементів вектора входу. На рис. 5 наведено приклад нейтралізації збурення для динамічного атрактора у вигляді послідовності з 8 літер при $H = 10$. Верхній ряд представляє діючі стимули, а нижній – відповідні реакції у вигляді розподілу ПСП нейронів. На окремих зображеннях простежується вплив збурення.

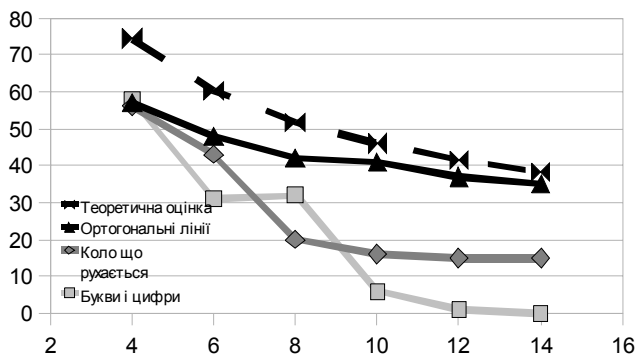


Рис. 6. Залежність атракторного радіуса від кількості запам'ятовуваних векторів і їх типу

На рис. 6 наведено експериментальні оцінки залежності величини атракторного радіуса H від числа запам'ятовуваних кадрів M , одержаних для різних типів динамічних атракторів при $\tau = 2$. Величина атракторного радіуса визначалась у серіях з 10 експериментів як найбільший рівень збурень H , при якому всі кадри динамічного атрактора відтворювались безпомилково. Для порівняння наведено теоретичну оцінку

атракторного радіуса за формулою (11). Можна відзначити, що теоретична та експериментальні залежності мають однаковий характер, але експериментальні дані істотно відрізняються для різних типів даних. Атракторний радіус виявився найбільшим для ортогональних векторів у вигляді прямих ліній, а найменшим – у зображень літер та цифр, що значною мірою перекривались.

6. Висновки

Проведені експерименти показали, що величина атракторного радіуса зростає при збільшенні величини τ . Це можна пояснити впливом локальної конвергенції нейромережі. Крім того, на поведінку нейромережі суттєво впливає співвідношення між зовнішньою та затриманою

компонентами ПСП. Відключення зовнішнього входу та підвищення компоненти зворотного зв'язку часто приводило до переходу мережі у стан динамічного атрактора, причому іноді спостерігались комбінації кількох динамічних атракторів, що нагадувало хибні атрактори в мережі Хопфілда [1]. Ці експерименти вказують на можливість провокувати появу фрагментів динамічних атракторів, які повторюють елементи навчальної послідовності. Вони нагадують цуги нервової активності, що спостерігаються при подразненні нервової тканини. Аналогії, які виникають при порівнянні поведінки динамічної асоціативної пам'яті з нервовою системою, підтверджують висловлену в [16] гіпотезу, згідно з якою щільні шари нервових клітин кори мозку складаються з локальних нейронних ансамблів, структура яких відповідає моделі динамічної рекурентної нейромережі. Подібні ансамблі, здатні не тільки підтримувати, але і відновлювати втрачені атракторні стани, становлять основу довготривалої пам'яті, забезпечуючи тривале збереження фрагментів нервової активності [17].

Роботу виконано за підтримки міжнародного гранту УНТЦ P-357 (EOARD 088008).

СПИСОК ЛІТЕРАТУРИ

1. Horfield J.J. Neural networks and physical systems with emergent collective computational abilities / J.J. Horfield // Proc. of the National Academy of Science 79:2554-58. – 1982. – P. 2554 – 2558.
2. Gupta M.M. Static and Dynamic Neural Networks / Gupta M.M., Jin L., Homma N. // From Fundamentals to Advanced Theory. – John Willey&Sons inc. Hoboken NJ, 2003. – 751 p.
3. Cao J. Absolute exponential stability of recurrent neural networks with Lipschitz-continuous activation functions and time delays / J. Cao, J. Wang // Neural Networks. – 2004. – Vol.17, N 3. – P. 379 – 390.
4. Lu H. Global exponential stability of delayed competitive neural networks with different time scales / H. Lu, Z. He // Neural Networks. – 2005. – Vol. 18, N 3. – P. 243 – 250.
5. Zhao H. Global asymptotic stability of Hopfield neural network involving distributed delays / H. Zhao // Neural Networks. – 2004. – Vol. 17, N 1. – P. 48 – 53.
6. Pineda F.J. Generalization of Back-propagation to Recurrent Neural Networks / F.J. Pineda // Physics Review Letters. – 1987. – Vol. 18. – P. 2229 – 2232.
7. Haykin S. Neural Networks. A comprehensive Foundation / S. Haykin // Prentice Hall, Inc. N.J. – 2ed. – 1999. – P. 690.
8. Yazdanbakhsh A. Fast synchronization of perceptual grouping in laminar visual cortical circuits / A. Yazdanbakhsh, S. Grossberg // Neural Networks. – 2004. – Vol. 17, N 5–6. – P. 707 – 718.
9. Reznik A.M. Non-Iterative Learning for Neural Networks / A.M. Reznik // Proc. International Joint Conference on Neural Networks (Washington DC, July 10–16 1999). – Washington, 1999. – Vol. 2. – P. 1374 – 1379.
10. Kosko B. Bi-directional associative memories / B. Kosko // IEEE Transactions on Systems, Man and Cybernetics. – 1987. – Vol. 18, N 1. – P. 49 – 60.
11. Albert A. Regression and the Moore-Penrose Pseudoinverse / A. Albert // Academic Press. – N-Y, 1972. – P. 97.
12. Personnaz L. Collective computational properties of neural networks: new learning mechanisms / L. Personnaz, I. Guyon, G. Dreyfus // Physical Review A. – 1986. – Vol. 34. – P. 4217 – 4228.
13. Reznik A.M. Projection learning algorithm for threshold-controlled neural networks / A.M. Reznik // Cybernetics and System Analysis. – 1994. – Vol. 30, N 4. – P. 603 – 613.
14. Weinfield W. A fully digital integrated CMOS Hopfield network including the learning algorithm / W. Weinfield // Proc. of Intern. Workshop on VLSI for Art. Intel. (Univ. Oxford, July 20–22, 1988). – Oxford, 1988. – E1. – P.1 – 11.
15. Gorodnichi D.O. Increasing Attraction of Pseudo- Inverse Autoassociative Networks / D.O. Gorodnichi, A.M. Reznik // Neural Processing Letters. – 1997. – Vol. 5, N 2. – P.123 – 127.
16. Reznik A.M. Hopfield ensembles in Lateral neurostructures of the cerebral cortex / A.M. Reznik // Mathematical Machines and Systems. – 2006. – N 1. – P. 3 – 12.
17. Reznik A.M. Associative Memories with "Killed" Neurons: the Methods of Recovery / A.M. Reznik, A.S. Sitchoy, O.K. Dekhtyarenko [et al.] // Proc. of the International Joint Conference on Neural Networks (Portland, Oregon, July 20–24, 2003). – Portland, Oregon, 2003. – Vol. 1. – P. 634 – 639.

Стаття надійшла до редакції 17.02.2010