13. Морелос-Сарагоса Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение / Р. Морелос-Сарагоса. – Москва: Техносфера, 2005. – 320 с.

14. Вернер М. Основы кодирования: учебник для ВУЗов / М. Вернер. – Москва: Техносфера, 2004. – 288 с.

15. Корченко О.Г. Оцінка корегувальної здатності завадостійких трійкових РС-кодів при передачі інформації повністю переплутаними станами кутритів квантовим каналом із шумом / О.Г. Корченко, Є.В. Васіліу, С.О. Гнатюк, В.М. Кінзерявий, А.М. Горчинська // Науково-технічний журнал «Захист інформації». – 2010. – №4. – С. 44–53.

16. Прескилл Дж. Квантовая информация и квантовые вычисления / Дж. Прескилл. – Т. 1. – Ижевск: "Регулярная и хаотическая динамика", 2008. – 464 с.

Надійшла: 21.07.12 р.

Рецензент: д.т.н., професор Корченко О.Г.

УДК 004.8.565.5

Терейковський І. А.

ОПТИМІЗАЦІЯ СТРУКТУРИ БАГАТОШАРОВОГО ПЕРСПЕТРОНУ В СИСТЕМАХ ЗАХИСТУ КОМП'ЮТЕРНОЇ ІНФОРМАЦІЇ

Стаття присвячена розробці методики оптимізації структури багатошарового перспетрону призначеного для використання в комп'ютерних системах захисту інформації з метою моніторингу та управління їх параметрами. Доведена доцільність використання двохшарового перспетрону. Отримано розрахункові вирази для визначення кількості нейронів у схованому шарі.

Ключові слова: захист інформації, двошаровий персептрон, нейронна мережа.

Постановка проблеми у загальному вигляді та її зв'язок із важливими практичними завданями. Одним із найбільш важливих напрямків розвитку систем захисту комп'ютерної інформації є вдосконалення математичного забезпечення підсистем моніторингу та управління. Оскільки використання явних моделей в багатьох випадках вже не приносить позитивних результатів все більшу популярність завойовують моделі які базуються на сучасних математичних теоріях, в тому числі і на теорії штучних нейронних мереж (НМ).

Використання НМ як правило полягає у застосуванні багатошарового перспетрону (БШП), призначеного для вирішення задачі розпізнавання образів. Хоча вказана нейромережева модель однозначно довела свою перспективність, але результати досліджень [3, 5] вказують на необхідність підвищення ефективності її застосування в першу чергу за рахунок оптимізації структури, що і визначає загальну проблематику даної статі.

Аналіз останніх досліджень та постановка проблеми. В загальному випадку БШП, структурна модель якого показана на рис.1, представляє собою НМ, яка складається із декількох послідовно з'єднаних між собою шарів штучних нейронів. Будемо розглядати розглядається класичний БШП у якого кожен нейрон схованого шару приймає всі вихідні сигнали попереднього шару, а його вихідний сигнал надсилається всім нейронам наступного шару.

Основними параметрами, що визначають структуру БШП являються: кількість вхідних нейронів, кількість схованих шарів нейронів, кількість нейронів у кожному схованому шарі та кількість вихідних нейронів. Базуючись на результатах [3, 4] в першому наближені можна вважати, що кількість вхідних та вихідних нейронів відповідає кількості контролюємих параметрів та кількості розпізнаваємих образів відповідно.

Дані величини задаються апріорно і не підлягають оптимізації. Тому оптимізувати структуру БШП можливо лише за рахунок кількості схованих шарів нейронів та кількості схованих нейронів. В [2, 4] наведено відповідні розрахункові вирази. Однак практичний досвід та результати [3] вказують на те, що на сьогодні теоретичні аспекти задачі визначення кількості схованих шарів та нейронів у схованому шарі вирішені далеко не повністю. Навіть в підходах до такого визначення існують деякі протиріччя. Так в [4]

використовується принцип – мінімізація кількості синаптичних зв'язків необхідних для навчання мережі на заданій множині прикладів.

Указаний принцип суперечить [3] де стверджується, що редукція розмірів мережі не призводить до зростання її узагальнюючих можливостей. При цьому в більшості робіт присвячених застосуванню БШП для розв'язку задач захисту інформації методика розрахунку структури не обґрунтована.

Таким чином *метою* даної роботи є оптимізація кількостей схованих шарів та схованих нейронів БШП, призначеного для вирішення актуальних задач захисту інформації.

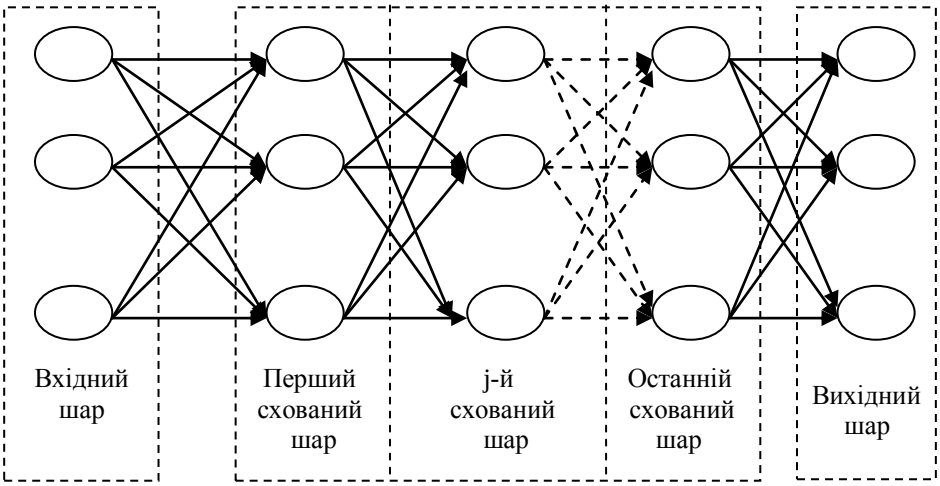

Рис. 1. Структурна модель БШП

В основу такого визначення покладемо критерій мінімізації відносної помилки БШП при дотриманні обмежень, характерних актуальним задачам захисту комп'ютерної інформації

$$\begin{cases} G_{MLP} \to min, \\ O \le O_d. \end{cases}, \tag{1}$$

де G_{MLP} – відносна помилка БШП, O_d –обмеження.

У даному випадку під терміном відносної помилки БШП будемо розуміти відношення помилки узагальнення (ε) до кількості синаптичних зв'язків (L_w)

$$G_{MLP} = \frac{\varepsilon}{L_w}. \tag{2}$$

Перепишемо (1) з урахуванням (2)

$$\begin{cases} \dfrac{\varepsilon}{L_w} \to min, \\ O \le O_d. \end{cases}. \tag{3}$$

Для визначення точки мінімуму, яка відповідає оптимальній кількості синаптичних зв'язків (L_w^{opt}) слід розв'язати наступне рівняння

$$\frac{\partial L_w}{\partial \varepsilon} = 0. \tag{4}$$

Відповідно [1, 4] помилка узагальнення БШП розраховується як сума помилок апроксимації (ε_a) та опису моделі (ε_o)

$$\varepsilon = \varepsilon_a + \varepsilon_o. \tag{5}$$

По аналогії з біологічним прототипом помилка апроксимації (навчання) співвідноситься з запам'ятовуванням БШП навчальних даних, а помилка опису моделі співвідноситься з узагальненням (стисненням) цих даних. Зазначимо, що як запам'ятовування так і стиснення навчальних даних відбувається за рахунок зміни вагових

коефіцієнтів синаптичних зв'язків. Вважають [1, 2, 4], що помилка апроксимації БШП (ε_a) пропорційна відношенню кількості синаптичних зв'язків до кількості компонент вхідного вектора (N_X)

$$\varepsilon_a \sim \frac{N_X}{L_w}. \tag{6}$$

Відповідно, помилка опису моделі БШП пропорційна відношенню кількості синаптичних зв'язків до кількості навчальних прикладів (P)

$$\varepsilon_o \sim \frac{L_w}{P}. \tag{7}$$

Узагальнений вираз для оцінки загальної помилки отримаємо підставивши (6, 7) в (5)

$$\varepsilon \sim \left(\frac{N_X}{L_w} + \frac{L_w}{P} \right). \tag{8}$$

Після тривіальних перетворень отримаємо точку максимуму

$$L_w^{opt} \sim \sqrt{P \times N_X}\ . \tag{9}$$

Отриманий вираз (9) для розрахунку оптимальної кількості синаптичних зв'язків дозволяє перейти до визначення оптимальної кількості схованих нейронів. Зазначимо, що співвідношення між кількістю синаптичних зв'язків та кількістю схованих нейронів БШП задається виразом

$$L_w = N_X \times N_1 + \sum_{s=1}^{S-1} \left(N_s \times N_{s+1} \right) + N_S \times N_Y, \tag{10}$$

де N_X – кількість вхідних нейронів, N_1 – кількість нейронів в першому схованому шарі, N_s – кількість нейронів в s–ому схованому шарі, N_Y – кількість нейронів у вихідному шарі, S – кількість схованих шарів нейронів.

Врахуємо теорему Хехта-Нільсена [1] в якій доведено – для представлення довільної функції достатньо двохшарової НМ прямого розповсюдження сигналу з повними зв'язками, що складається з n вхідних нейронів, $(2n+1)$ схованих нейронів та m вихідних нейронів. Це дозволяє спростити структуру моделі БШП до двохшарового перспетрону (ДШП). Вказане спрощення хоча і суперечить [2] в контексті зменшення обчислювальних можливостей, однак відповідає таким вимогам до НМ в задачах ЗІ, як максимальна простота та надійність.

Адаптований до ДШП вираз (10) виглядає так

$$L_w = \left(N_X + N_Y \right) \times N_1. \tag{11}$$

Для багатьох постановок задач ЗІ вихід НМ повинен представляти ймовірність (впевненість) виникнення очікуваної події, наприклад реалізації мережевої атаки на КС. В цьому випадку НМ повинна мати один вихідний елемент ($N_Y = 1$), що зумовлює зміну (11) наступним виразом

$$L_w = \left(N_X + 1 \right) \times N_1. \tag{12}$$

Прирівняємо (9) до (11). Отримаємо

$$\sqrt{P \times N_X}\ \sim \left(N_X + N_Y \right) \times N_1^{opt}, \tag{13}$$

$$N_1^{opt} \sim \frac{\sqrt{P \times N_X}}{N_X + N_Y}, \tag{14}$$

де N_1^{opt} – оптимальна кількість схованих нейронів в ДШП з довільною кількістю вихідних нейронів.

Для ДШП з одним вихідним зв'язком (14) можна спростити так

$$Ns_1^{opt} \sim \frac{\sqrt{P \times N_X}}{N_X + 1}, \tag{15}$$

де Ns_1^{opt} – оптимальна кількість схованих нейронів в ДШП з одним вихідним нейроном.

Зазначимо, що (14), (15) представляють собою пропорцію, а значить не дозволяютьбезпосередньо розрахувати оптимальну кількість схованих нейронів в ДШП, призначеному для вирішення задач ЗІ. Для переходу до точних розрахунків введемо в (14) коефіцієнт пропорційності

$$N_1^{opt} = k \times \frac{\sqrt{P \times N_X}}{N_X + N_Y}, \tag{16}$$

де k – деякий коефіцієнт пропорційності.

Проведемо оцінку означеного коефіцієнту. В загальному випадку [1, 4] , мінімально допустима кількість схованих нейронів визначається теоремою Хехта-Нільсена, а максимально допустима кількість обмежується кількістю навчальних прикладів. Тобто

$$N_1^{min} \geq 2N_X + 1, \tag{17}$$

$$N_2^{max} \leq P, \tag{18}$$

де N_1^{min}, N_2^{max} – мінімальна та максимальна кількість схованих нейронів.

Порівнявши (16) до (17) та (16) до (18) отримаємо

$$\begin{cases} k \times \dfrac{\sqrt{P \times N_X}}{N_X + N_Y} \geq 2N_X + 1 \\ k \times \dfrac{\sqrt{P \times N_X}}{N_X + N_Y} \leq P \end{cases}. \tag{19}$$

Як наслідок:

$$\begin{cases} k \geq \dfrac{(2N_X + 1) \times (N_X + N_Y)}{\sqrt{P \times N_X}} \\ k \leq \dfrac{P \times (N_X + N_Y)}{\sqrt{P \times N_X}} \end{cases}. \tag{20}$$

У теорії НМ [1-4] загальноприйнято, що кількість навчальних прикладів повинна перевищувати кількість вхідних параметрів як мінімум в 10 разів. Тобто

$$N_X \times P \geq 10N_X^2. \tag{21}$$

Проведемо уточнення меж діапазону величин коефіцієнту пропорційності з урахуванням специфіки актуальних задач ЗІ. Відповідно проведеному [3] оцінюванню в НМ, призначених для вирішення вказаних задач кількість вихідних параметрів не перевищує кількості вхідних параметрів, а кількість навчальних прикладів повинна перевищувати кількість розрізняємих класів (вихідних параметрів) як мінімум в 10 разів.

Підставивши (21) в (20) та провівши тривіальні спрощення отримаємо

$$\begin{cases} k \geq \dfrac{(2N_X + 1) \cdot (N_X + N_Y)}{10N_X} \\ k \leq \dfrac{10N_X^2 + P \times N_Y}{10N_X} \end{cases}. \tag{22}$$

Тому не порушуючи нерівності (22) можна вважати

$$N_Y \approx N_X, \tag{23}$$

$$N_Y \times P \approx 10N_X^2. \tag{24}$$

Підставивши (23, 24) в (22) та отримаємо:

$$\begin{cases} k \geq \dfrac{(2N_X+1)\cdot(N_X+N_X)}{10N_X} \\ k \leq \dfrac{10N_X^2+10N_X^2}{10N_X} \end{cases}, \tag{25}$$

$$\begin{cases} k \geq 0{,}4N_X + 0{,}2 \\ k \leq 2N_X \end{cases}. \tag{26}$$

Підстановка (26) в (16) дозволяє оцінити діапазон оптимальної кількості схованих нейронів ДШП так

$$N_1^{opt} \geq (0{,}4N_X + 0{,}2) \times \frac{\sqrt{P \times N_X}}{N_X + N_Y}, \tag{27}$$

$$N_1^{opt} \leq \frac{2\sqrt{P \times N_X}}{N_Y}. \tag{28}$$

Врахувавши в (27, 28) те, що кількість схованих нейронів має бути цілим числом отримаємо остаточні розрахункові вирази

$$N_1^{opt}(min) = Round\left((0{,}4N_X + 0{,}2) \times \frac{\sqrt{P \times N_X}}{N_X + N_Y} \right), \tag{29}$$

$$N_1^{opt}(max) = Round\left(\frac{2\sqrt{P \times N_X}}{N_Y} \right), \tag{30}$$

де $N_1^{opt}(max)$, $N_1^{opt}(min)$ – максимальна та мінімальна межа діапазону оптимальної кількості схованих нейронів, *Round*(*X*) – операція визначення найближчого цілого числа від аргументу *X*.

Висновки

Доведено, що для розв'язання задач захисту комп'ютерної інформації доцільно використовувати БШП з одним шаром схованих нейронів, тобто ДШП. Одержано розрахункові вирази, які дозволяють оптимізувати кількість схованих нейронів ДШП відповідно умов поставленої задачі захисту інформації, що дозволяє підвищити достовірність моніторингу параметрів захисту комп'ютерної інформації.

Перспективи подальших досліджень у даному напрямку полягають у вдосконаленні методики застосування БШП для розв'язання актуальних задач захисту інформації.

ЛІТЕРАТУРА

1. Бодянский Е.В. Исскуственные нейронные сети: архитектура, обучение, применение / Е.В. Бодянский, О.Г. Руденко. – Х. : ТЕЛЕТЕХ, 2004. – 369 с.

2. Каллан Р. Основные концепции нейронных сетей / Каллан Р. ; пер. с англ. А. Г. Сивака. – М. : Вильямс, 2003. – 288 с.

3. Терейковський І. Нейронні мережі в засобах захисту комп'ютерної інформації / І. Терейковський. – К. : ПоліграфКонсалтинг. – 2007. – 209 с.

4. Хайкин С. Нейронные сети: полный курс, 2-е изд., испр. / Хайкин С. ; пер. с англ. Н. Н. Куссуль – М. : Вильямс, 2006. – 1104 с.

5. Хорошко В. О. Основи інформаційної безпеки / В. О. Хорошко, В. С. Чередниченко. – К. : ДУІКТ, 2008. – 186 с.

Надійшла: 22.07.2012 р.

Рецензент: д.т.н., професор Щербак Л.М.