

УДК 681.3.06

І.А. Терейковський

Київський національний університет будівництва і архітектури, Київ

ОПТИМІЗАЦІЯ СТРУКТУРИ ДВОШАРОВОГО ПЕРСЕПТРОНУ ПРИЗНАЧЕНОГО ДЛЯ РОЗПІЗНАВАННЯ АНОМАЛЬНИХ ВЕЛИЧИН ЕКСПЛУАТАЦІЙНИХ ПАРАМЕТРІВ КОМП'ЮТЕРНОЇ МЕРЕЖІ

Розроблено методику визначення оптимальної кількості синаптичних зв'язків та схованих нейронів двошарового персептрону, призначеного для розпізнавання аномальних величин експлуатаційних параметрів комп'ютерної мережі.

Ключові слова: *Комп'ютерна мережа, надійність, нейронна мережа, двошаровий персептрон*

Постановка проблеми

Складність логічної та фізичної організації сучасних комп'ютерних мереж призводить до об'єктивних труднощів при вирішенні задач, пов'язаних з їх управлінням та захистом. В багатьох випадках розв'язання цих задач пов'язане з оперативним виявленням аномалій в процесі функціонування мережі [3]. Означені аномалії можуть бути як наслідком мережевих атак, так і наслідком збоїв в роботі апаратного чи програмного забезпечення або виникненням нештатних умов експлуатації. Досить часто при розробці систем виявлення аномалій використовується така методика:

- обирається підконтрольний об'єкт, технічний стан (стан захищеності) якого визначається величинами одного або декількох параметрів, які контролюються в експлуатації;

- створюється шаблон нормальної поведінки підконтрольного об'єкта який характеризується діапазоном значень контрольованих експлуатаційних параметрів;

- створюється механізм порівняння поточних значень експлуатаційних параметрів з шаблоном нормальної поведінки.

Створення шаблону нормальної поведінки та механізму порівняння ускладнюється багатоваріантністю та неоднозначністю співставлення. Адже для реальних мережевих об'єктів, як правило, складно чітко визначити всі можливі комбінації співвідношень поточних значень параметрів, що відповідають аномалії або нормальному стану. Перспективним шляхом виходу із цієї ситуації є використання нейромережевої моделі, задачею якої буде класифікація стану підконтрольного об'єкту (аномальний/нормальний) відповідно величин відхилень поточних значень параметрів від шаблону нормальної поведінки [4]. Вказане твердження визначає основну проблему даної статті –

підвищення ефективності розпізнавання аномальних величин експлуатаційних параметрів комп'ютерних мереж за допомогою нейронних мереж. Проблема безпосередньо пов'язана з таким важливим науково-практичним завданням, як забезпечення надійності функціонування комп'ютерних мереж.

Аналіз останніх досліджень і публікацій, на які спирається автор

Висновки [1; 2; 4; 5] вказують на те, що ефективність застосування нейромережевої моделі безпосередньо залежить від того наскільки її тип та параметри оптимізовані відповідно умов поставленої задачі. Також зазначається, що в більшості випадків процес оптимізації є багатокритеріальним. Однак, можливо застосовувати ряд універсальних критеріїв, серед яких найбільш важливими є: обчислювальна потужність нейронної мережі, достовірність розпізнавання, тип та термін навчання, потенційна можливість нейронної мережі навчатись на зашумлених прикладах, обсяг та корельованість навчальних прикладів, вид вихідних даних, швидкість розрахунку вихідного сигналу. При цьому відповідно [4] для розв'язання задачі розпізнавання аномальних величин експлуатаційних параметрів оптимальним типом нейромережевої моделі є багатошаровий персептрон.

В загальному випадку багатошаровий персептрон, структура якого показана на рис. 1, являє собою нейронну мережу, яка складається із декількох послідовно з'єднаних між собою шарів штучних нейронів [1; 5]. Зовнішня інформація надходить у вхідний шар, основними завданнями якого є прийом та розповсюдження вхідної інформації по іншим шарам нейронної мережі. Далі знаходиться один або декілька схованих шарів нейронів, в яких відбувається основна обробка інформації, результати якої відображаються у

вихідному шарі. Для класичного багатшарового перцептронув зв'язки між нейронами сусідніх шарів створюють повнов'язну структуру. Зв'язки між нейронами одного шару відсутні. Інформація розповсюджується тільки у напрямку "вхід→вихід". Кількість вхідних нейронів дорівнює кількості компонент вхідного вектору, а кількість вихідних нейронів визначається вимогами до вихідного сигналу.

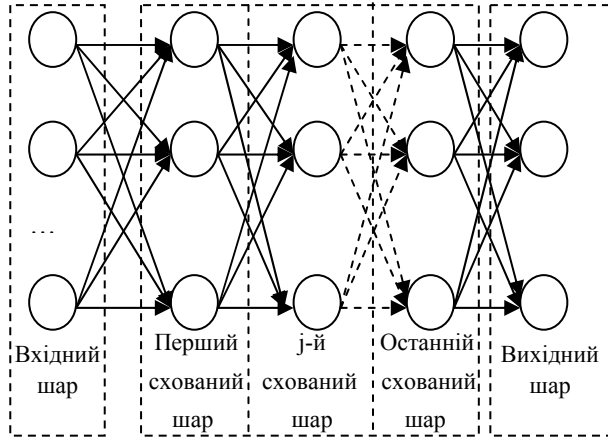


Рис. 1. Структура багатшарового перцептронув

Для вхідних нейронів найчастіше використовується лінійна функція активації. Для схованих нейронів як правило використовують сигмоїдальну функцію активації або гіперболічний тангенс

$$F(NET) = \frac{1}{1 + e^{-a \times NET}}, \quad (1)$$

$$F(NET) = \frac{e^{a \times NET} - e^{-a \times NET}}{e^{a \times NET} + e^{-a \times NET}}, \quad (2)$$

де NET – сумарний зважений вхідний сигнал; $F(NET)$ – вихідний сигнал нейрону.

В більшості випадків вихідні нейрони виконують тільки розрахунок власних вхідних сигналів, тому функція активації для них не використовується. Розрахунок параметрів j -го нейрону в l -му схованому шарі здійснюється так

$$NET_j^l = \sum_{i=1}^{K_j^l} w_{ij}^l x_{ij}^l, \quad (3)$$

$$OUT_j^l = F(NET_j^l - \theta_j^l), \quad (4)$$

$$x_{ij}^{l+1} = OUT_j^l, \quad (5)$$

де i – номер входу, j – номер нейрону в шарі; l – номер схованого шару; K_j^l – кількість вхідних зв'язків j -го нейрону в l -му шарі; w_{ij}^l – ваговий коефіцієнт i -го входу j -го нейрону в l -му шарі; θ_j^l – пороговий рівень активації j -го нейрону в l -му шарі; x_{ij}^l – i -й вхідний сигнал нейрону в l -му шарі;

F – функція активації; OUT_j^l – вихідний сигнал; NET_j^l – сумарний вхідний сигнал j -го нейрону в l -му схованому шарі.

Для першого схованого шару кількість вхідних зв'язків нейрону дорівнює кількості нейронів у вхідному шарі. Для інших схованих шарів нейронів K_j^l дорівнює кількості нейронів у попередньому схованому шарі. Навчання багатшарового перцептронув виконується методом "з вчителем" та полягає у визначенні вагових коефіцієнтів зв'язків схованих нейронів.

Відповідно [1, 2, 5] одним з основних напрямків вдосконалення моделі багатшарового перцептронув є оптимізація його структури, яка за своєю суттю є оптимізацією кількості синаптичних зв'язків. Кількість синаптичних зв'язків безпосередньо залежить від кількості нейронів. При цьому кількість вхідних та вихідних нейронів визначається умовами поставленої задачі і оптимізації не підлягає. Таким чином, оптимізація структури класичної моделі багатшарового перцептронув є оптимізацією кількості схованих нейронів. В [2; 5] наведено вирази для оцінки оптимальної кількості синаптичних зв'язків та кількості схованих нейронів в двошаровому перцептронув з сигмоїдальними функціями активації

$$\frac{N_0 P}{1 + \log_2 P} \leq L_w \leq N_1 \left(\frac{P}{N_1} + 1 \right) (N_1 + N_0 + 1) + N_1, \quad (6)$$

$$\frac{P}{10} - N_1 - N_0 \leq L_w \leq \frac{P}{2} - N_1 - N_0, \quad (7)$$

$$L_w < P \times \varepsilon_{max}, \quad (8)$$

де L_w – кількість синаптичних зв'язків; ε_{max} – максимальна допустима помилка узагальнення; N_0 – кількість вхідних нейронів; N_1 – кількість схованих нейронів.

Також наведено формулу для визначення максимальної кількості образів P , яку може запам'ятати двошаровий перцептронув з пороговою функцією активації

$$\frac{L_w}{N_0} < P < \frac{L_w}{N_0} \log_2 \left(\frac{L_w}{N_0} \right), \quad (9)$$

Зазначимо, що вирази (6-9) наведені в [2; 5] без належного теоретичного обґрунтування, методика їх визначення описана недостатньо, а практичний досвід свідчить про низьку точність. Також слід зазначити, що оптимальна кількість синаптичних зв'язків для перцептронув з декількома схованими шарами теоретично не визначена, хоча і вважається дещо меншою, ніж для двошарового. Перспективним шляхом виправлення вказаних

недоліків може стати створення відповідної методики оптимізації.

Формулювання мети статті

Розробити методику розрахунку оптимальної кількості синаптичних зв'язків та схованих нейронів для двохшарового перцептронів, призначеного для розпізнавання аномальних величин експлуатаційних параметрів комп'ютерних мереж.

Виклад основного матеріалу дослідження

Одним із найбільш важливих критеріїв ефективності функціонування багатшарового перцептронів є помилка узагальнення [1; 2; 5]. За допомогою цього критерію можна оцінити наскільки точно нейронна мережа здатна класифікувати дані, які не були представлені в навчальній вибірці. Виходячи з цих міркувань, в якості критерію оптимізації кількості синаптичних зв'язків оберемо помилку узагальнення двохшарового перцептронів. Запишемо відповідну математичну модель

$$\varepsilon(L_w) \rightarrow \min, L_w \in \{1, \infty, 1\}, \quad (10)$$

де ε – помилка узагальнення; L_w – кількість синаптичних зв'язків.

В [1] доведено, що в загальному випадку помилка узагальнення, складається із помилки апроксимації (ε_a) та помилки опису моделі (ε_o)

$$\varepsilon = \varepsilon_a + \varepsilon_o, \quad (11)$$

Також показано, що помилку апроксимації багатшарового перцептронів можливо оцінити так

$$\varepsilon_a \sim \frac{N_1}{L_w}, \quad (12)$$

де N_1 – кількість компонент вхідного вектора (розмірність вхідного вектору).

В першому наближенні помилку опису моделі багатшарового перцептронів можливо оцінити так

$$\varepsilon_o \sim \frac{L_w}{P}, \quad (13)$$

де P – кількість навчальних прикладів.

Зазначимо, що в (12;13) знак " \sim " означає пропорційність. Перепишемо пропорції (12, 13) так:

$$\varepsilon_a = k_a \times \frac{N_1}{L_w}, \quad (14)$$

$$\varepsilon_o = k_o \times \frac{L_w}{P}, \quad (15)$$

де k_a, k_o – фіксовані коефіцієнти, що належать масиву натуральних чисел.

У першому наближенні можна прийняти, що

$$k_a = k_o = k, \quad (16)$$

де k – деяке натуральне число.

Підстановка (14), (15), (16) в (11) дозволяє отримати наступний вираз для оцінки помилки узагальнення

$$\varepsilon = \left(\frac{k \times N_1}{L_w} + \frac{k \times L_w}{P} \right). \quad (17)$$

Розглянемо фізичну суть залежності помилки узагальнення від кількості синаптичних зв'язків. З одного боку, відповідно (12) збільшення кількості синаптичних зв'язків, а значить, і схованих нейронів призводить до більш точної апроксимації навчальних даних. Однак, з позицій (13) збільшення кількості синаптичних зв'язків призводить до збільшення помилки опису моделі. Тобто, модель може бути занадто точно підігнана під конкретні навчальні дані і втратити узагальнюючі властивості. Таким чином, зміна кількості синаптичних зв'язків призводить до двох взаємно протилежних тенденцій зміни помилки узагальнення. Наявність цих тенденцій свідчить про потенційну можливість розв'язку (10).

Для знаходження точки мінімуму функції (10) з урахуванням (17) продиференціюємо функцію $\varepsilon(L_w)$

$$\frac{d\varepsilon}{dL_w} = \frac{d\left(\frac{k \times N_1}{L_w} + \frac{k \times L_w}{P}\right)}{dL_w}. \quad (18)$$

Проведемо спрощення (18)

$$\frac{d\varepsilon}{dL_w} = \frac{d(k \times N_1 \times L_w^{-1})}{dL_w} + \frac{d(k \times P^{-1} \times L_w)}{dL_w}. \quad (19)$$

Остаточний диференціал описується виразом

$$\frac{d\varepsilon}{dL_w} = \frac{-k \times N_1}{L_w^2} + k \times P^{-1}. \quad (20)$$

Перейдемо до знаходження критичних точок функції $\varepsilon(L_w)$. Для цього, враховуючи (20), слід розв'язати рівняння

$$\frac{d\varepsilon}{dL_w} = 0. \quad (21)$$

Після перетворень (22) – (24) отримаємо вираз (25), для розрахунку критичної точки функції $\varepsilon(L_w)$.

$$\frac{-k \times N_1}{L_w^2} + k \times P^{-1} = 0 \quad (22)$$

$$\frac{k}{P} = \frac{k \times N_1}{L_w^2} \quad (23)$$

$$L_w^2 = N_1 \times P \quad (24)$$

$$L_w^0 = \sqrt{N_1 \times P}, \quad (25)$$

де L_w^0 – критична точка.

Перевірка (26), дозволяє стверджувати, що точка L_w^0 є точкою мінімуму функції $\varepsilon(L_w)$.

$$\begin{cases} \text{якщо } L_w < L_w^0 \text{ то } \frac{d\varepsilon}{dL_w} > 0, \\ \text{якщо } L_w > L_w^0 \text{ то } \frac{d\varepsilon}{dL_w} < 0. \end{cases} \quad (26)$$

Враховуючи (14) та (25) запишемо рівняння для приблизної оцінки оптимальної кількості синаптичних зв'язків багат шарового перцептрон, що відповідає мінімуму ε_0

$$L_w^{opt} = \sqrt{N_1 \times P} \quad (27)$$

Як видно, із показаної на рис. 2, типової структури двошарового перцептрон, призначеного для розпізнавання аномальних величин експлуатаційних параметрів комп'ютерної мережі, кількість його синаптичних зв'язків (L) розраховується так

$$L_w = N_1 \times L + N_0 \times L, \quad (28)$$

де L – кількість нейронів в схованому шарі; N_1 – кількість вхідних нейронів; N_0 – кількість вихідних нейронів.

Значимо, що особливістю типової структури є наявність тільки одного вихідного нейрону, вихід якого сигналізує про наявність або відсутність аномалій.

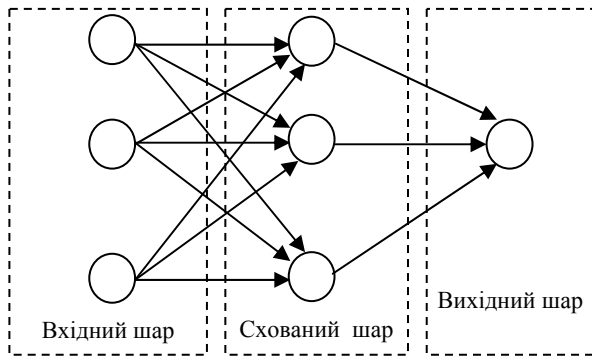


Рис. 2 Структура двошарового перцептрон

Відповідно (28), з врахуванням того, що $N_0=1$, кількість нейронів в схованому шарі двошарового перцептрон визначається так

$$L = \frac{L_w}{N_1 + 1}. \quad (29)$$

Для розрахунку оптимальної кількості схованих нейронів слід в (29) підставити (27)

$$L^{opt} = \frac{\sqrt{N_1 \times P}}{N_1 + 1}, \quad (30)$$

де L^{opt} – оптимальна кількість схованих нейронів в двошаровому перцептроні.

При достатньо великих значеннях N_1 можна вважати, що

$$N_1 + 1 \approx N_1 \quad (31)$$

Підстановка (31) в (30) дозволяє записати остаточний вираз для розрахунку оптимальної кількості схованих нейронів в двошаровому

перцептроні

$$L^{opt} \approx \sqrt{\frac{P}{N_1}} \quad (32)$$

Значимо, що по відношенню до (6), (7) вираз (32) дозволяє більш точно визначити діапазон оптимальної кількості схованих нейронів у двошаровому перцептроні. Проведені чисельні експерименти підтвердили ефективність виявлення аномальної поведінки комп'ютерних мереж за допомогою двошарового перцептрон, структура якого була розроблена за допомогою запропонованої методики.

Висновки

Запропоновано підхід та розроблено методику оптимізації структури двошарового перцептрон, призначеного для розпізнавання аномальних величин експлуатаційних параметрів комп'ютерної мережі. Результатом методики є розрахункові вирази, які дозволяють на основі обсягів навчальної вибірки та вхідних параметрів, визначити оптимальну кількість синаптичних зв'язків та оптимальну кількість схованих нейронів.

Основні перспективи подальших розвідок у даному напрямку полягають у розробці методів оптимізації структури синаптичних зв'язків двошарового перцептрон при визначеній оптимальній кількості схованих нейронів. Ще одним важливим напрямком досліджень має бути розробка подібної методики для оптимізації структури багат шарового перцептрон.

Список літератури

1. Ежов А. А. *Нейрокомп'ютинг и его применения в экономике и бизнесе* / А. А. Ежов, С. А. Шумский. – М. : МИФИ, 1998. – 224 с.
2. Каллан Р. *Основные концепции нейронных сетей* / Каллан Р. ; пер. с англ. А. Г. Сивака. – М. : Вильямс, 2003. – 288 с.
3. Менаске Д. *Производительность Web-служб. Анализ, оценка и планирование* / Менаске Д., Виргилио А. ; пер. с англ. – СПб. : ДиаСофтЮп, 2003. – 480 с.
4. Терейковський І. *Нейронні мережі в засобах захисту комп'ютерної інформації* / І. Терейковський. – К. : ПоліграфКонсалтинг. – 2007. – 209 с.
5. Хайкин С. *Нейронные сети: полный курс, 2-е изд., испр.* / Хайкин С. ; пер. с англ. Н. Н. Куссуль – М. : Вильямс, 2006. – 1104 с.

Стаття надійшла до редколегії 15.02.2011

Рецензент: д-р техн. наук, проф. С.В. Цюцюра, Київський національний університет будівництва і архітектури, Київ.