УДК 004.056.5
Погорелов В.В., аспірант
Національний технічний університет України "КПІ ім. Ігоря Сікорького"

ПРОБЛЕМАТИКА ВИКОРИСТАННЯ НЕЙРОМЕРЕЖЕВИХ СИСТЕМ РОЗПІЗНАВАННЯ КІБЕРАТАК.

Погорелов В.В. Проблематика використання нейромережевих систем розпізнавання кібератак. Розглянуто проблематику побудови нейромережевих систем розпізнавання сигнатур кібератак. Проаналізовано ефективність, адаптивність, ресурсоємність та оперативність роботи і навчання різних класів нейромереж, визначені типові проблеми та алгоритми їх мінімалізації. Показано особливості використання у даній області штучних нейромереж третього покоління та спільної роботи декількох нейромереж.
Ключові слова: системи розпізнавання кібератак, штучна нейронна мережа, нейромережеві мережі прямого поширення, машина Больцмана, глибинна мережа переконань, згорткова нейромережа.

Рассмотрена проблематика построения нейросетевых систем распознавания сигнатур кибератак. Проанализирована эффективность, адаптивность, ресурсоемкость и оперативность работы и обучения разных классов нейросетей, определены типичные проблемы и алгоритмы их минимизации. Показаны особенности использования в данной области искусственных нейронных сетей третьего поколение и совместной работы нескольких нейронных сетей.
Ключевые слова: системы распознавания кибератак, искусственная нейронная сеть, нейросетевые сети прямого распространение, машина Больцмана, глубокая сеть доверия, сверточные нейросети.

Pogorelov V.V. Problems of neural network usage for cyber-attack detection systems. Problems of construction of neural network systems for recognizing cyber-attacks signatures were considered. Efficiency, adaptation, resource intensity and operability of work and training of different ANNs' classes were analyzed, typical problems and solving algorithms were found. Special features of third generation ANNs application and cooperative work of several ANNs were shown.
Keywords: cyber-attack detection systems, artificial neural network, feed forward neural networks, Boltzman machine, deep belief network, convolutional neural networks.

Постановка наукової проблеми. Кібератака — це сукупність протиправних дій, що порушують вільний доступ до інформаційної системи, цілісність її роботи та конфіденційність її даних. При побудові систем розпізнавання кібератак (СРК) на мережеві ресурси інформаційних систем (МРІС) активно використовуються нейромережеві методи розпізнавання (НММР). Статистичний аналіз застосування нейромережевих моделей (НММ) вказує на ключові фактори, що характеризують ефективність роботи даних систем [1-4]:

- статистика неспрацьовувань та помилкових спрацьовувань СРК;
- час розробки та формування навчальної вибірки НММ;
- час та стабільність навчання НММ;
- адаптація НММ до особливостей МРІС,

Робота НММР полягає визначенні та кодуванні множини вхідних параметри сигнатури кібератаки, побудові множини навчальних прикладів для навчання НММ, а також у розрахунку параметрів функціонування НММ у відповідності до класу МРІС. У схемі НММ має бути передбачена можливість формування навчальної вибірки за допомогою експертної оцінки (рис. 1).

Першим етапом розробки НММР і НММ на їх основі є аналіз ключових параметрів функціонування МРІС, зокрема їх ефективності, оперативності та ресурсоємності. Поняття оперативності та ресурсоємності є інтуїтивно зрозумілими і відповідають часу виконання функцій системою та апаратному ресурсу [2-4], який при цьому використовується, в той час як термін ефективність слід розглянути більш детально. Функції ефективності відображає пристосованість НММ до конкретного класу задач і у даному дослідженні ми розглянемо інтегральну функцію ефективності процесу $E_\Sigma = f(E_{ANN}, E_{TS})$, де $E_{ANN} = f(e_O, e_P, e_R)$ – ефективність роботи штучної нейронної мережі (ШНМ), що є функцією від e_O (визначення оптимального НММ), e_P (пошук параметрів НММ), e_P (ресурсоємність використання ШНМ), а $E_{TS} = f(e_{TE}, e_{TS})$ – ефективність створення навчальної вибірки, що є функцією від e_{TE} (пошук параметрів навчальних прикладів) і e_{TS} (формування навчальної вибірки). При розрахунку ефективності роботи СРК проводиться аналіз множини ефективних видів НММ разом з оцінкою кожної моделі, визначається час очікування вихідного сигналу сигнатури кібератаки (СК) і параметри обчислювальних ресурсів сервера та системи експертних знань.

Рис. 1. Діаграма системи нейромережевого розпізнавання кібератаки на мережеві ресурси.

Важливим етапом є визначення ефективності навчання НММ, що залежить від сумарного часу побудови навчальної вибірки та процесу навчання $t_\Sigma = a(t_{TS}, t_{TP})$. Слід зауважити, що:

$$t_\Sigma \leq t_{max}(n_i), n_i \in N_a, \quad (1)$$

де $t_{max}(n_i)$ – максимальний допустимий сумарний час побудови навчальної вибірки та навчання НММ, n_i – окремий вид НММ серед допустимої множини N_a. Розрахунок функції ефективності кожного окремого НММ відбувається за наступним критерієм:

$$E_i = \sum_{k=1}^{K} \alpha_k R_k(n_i); \alpha_k = 0..1; n_i \in N_a, i = 1..I, \quad (2)$$

де I – кількість видів НММ, R_k – критерій ефективності, α_k – ваговий коефіцієнт критерія ефективності, K – кількість критеріїв ефективності.

Аналіз досліджень. Аналіз публікацій за даною темою надає можливість визначити перспективи розробки єдиної методології побудови алгоритмів СРК. У роботах [1-4] розглянуто сучасну методику адаптації типових НММ до мережевих ресурсів інформаційних систем та їх оптимізації. Певна частина наведених публікацій і монографій присвячена основам побудови ШНМ: мережам прямого поширення [5] та машинам Больцмана [8], а також НММ, що можуть бути побудовані на основі даних архітектур [5-9] . Показано, що для розпізнавання сигнатур кібератак високу точність показують ресурсоємні нейромережі третього покоління, зокрема згорткові нейромережі [11] та мережі глибинних переконань [10]. Також, у роботі [12] було розглянуто методику розробки СРК на основі кількох згорткових ШНМ, що працюють разом.

Виклад основного матеріалу й обґрунтування результатів дослідження. Систематичний аналіз показує, що побудова єдиної моделі застосування ШНМ у СРК дозволить звести задачу розробки та адаптації НММ до математичної задачі пошуку екстремумів цільових функцій. Необхідність вирішення даної задачі зумовлює потребу у побудові методологічної бази використання нейромерж третього покоління при захисті інформаційних ресурсів від кіберзагроз, що є ***метою*** даного дослідження.

1. Нейромережеві методи розпізнавання кібератак, що базуються на мережах прямого поширення

Мережі прямого поширення (МПП) представляють собою базову структуру ШНМ, що складається з шару вхідних нейронів I_n, одного або декількох шарів прихованих нейронів H_n^k та шару вихідних нейронів O_n (рис. 2); причому кожен нейрон пов'язаний зі всіма нейронами сусідніх шарів і не пов'язаний з жодним іншим нейроном [5]. Типовим методом навчання МПП є метод зворотного поширення помилки, що відноситься до методів навчання зі вчителем.

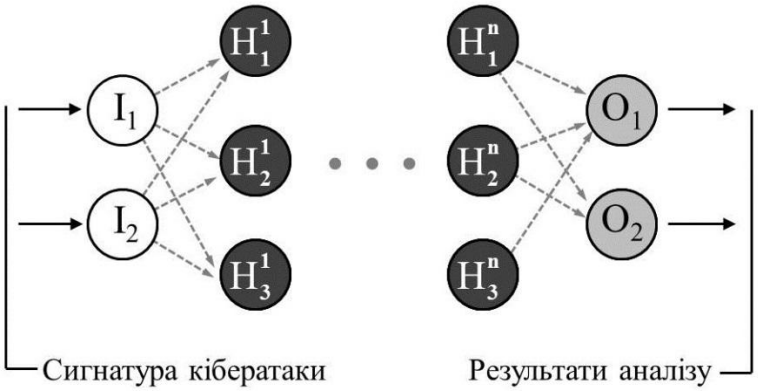

Рис. 2. Базова схема штучної нейронної мережі прямого поширення.

На основі МПП найбільш зручно описати принцип роботи ШНМ. Інформаційний сигнал (у даному випадку вхідний інтернет-трафік, що потенційно утримує у собі сигнатуру кібератаки) нормалізується функцією активації та попадає на шар вхідних нейронів. Значення прихованих та вихідних нейронів визначаються як суми добутків вихідних параметрів нейронів, що пов'язанні з даним нейроном на ваговий коефіцієнт w.

$$\begin{cases} H_n = \sum_m I_m \cdot w_{m-n} \\ H_k = \sum_n H_n \cdot w_{n-k}. \\ O_z = \sum_y H_y \cdot w_{y-z} \end{cases} \quad (3)$$

Повний набір даних яким оперує ШНМ називається тренувальним сетом, загальна кількість використаних тренувальних сетів — кількістю ітерацій, лічильник проходження повного тренувального сету — значенням епохи. Натренованість мережі визначається через розходження між вірною відповіддю та отриманим результатом. Збіжність системи вказує та те, що з кожною новою епохою натренованість ШНМ зростає.

Схема МПП при наявності достатньої кількості прихованих нейронів здатна до навчання розпізнавання сигнатур кібератак, але розглядається, як малоефективний вид НММ. Тим не менш, МПП може бути використана разом з іншими мережами чи взята за основу при розробці більш складної ШНМ. Одним з таких варіантів НММ є рекурентні нейронні мережі [6], що організовані подібно до мереж прямого поширення, але зміщенням у часі, що дозволяє прихованим нейронам отримувати інформацію як від сусіднього шару, так і від себе (рис. 3), тобто використовують данні, що формуються при попередньому проході.

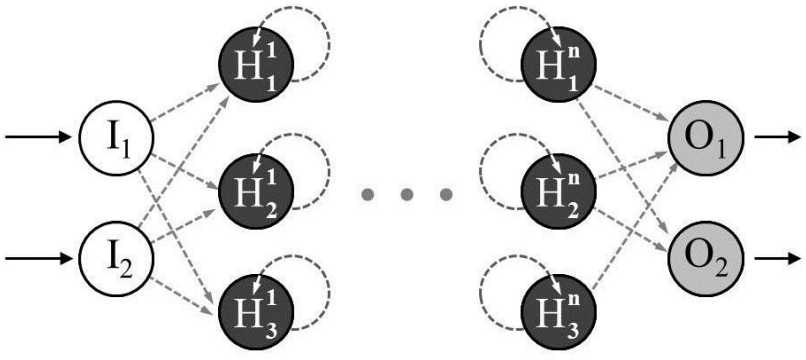

Рис. 3. Базова схема рекурентної нейронної мережі.

Таким чином, система найбільш ефективно працює з данними, що надаються послідовно і залежать від часу, тому знайшла своє галузь використання при розробці ШНМ. На сьогоднішній день рекурентні нейронні мережі відносяться до НММ, що ефективно працюють з малим об'ємом трафіку у системах з обмеженим апаратним ресурсом. Основною проблемою даного класу НММ є проблема вибухового градієнта, що призводить до втрати інформації на рівні вагових коефіцієнтів. Тому для рекурентних нейронних мереж перевага надається використанню стохастичного методу градієнтного спуску, при якому поправки на вагові коефіцієнти вводяться безпосередньо після обчислення вихідних данних мережі (для кожного зразка сету). При застосування пакетного та мініпакетного методів градієнтного спуску, що мають перевагу у швидкості обробки інформації, залишається ризик недосягнення збіжності системи.

Сучасним видом рекурентних нейромереж, що знайшли своє застосування у СРК є мережі з відлунням стану [7]. Вони виділяються тим, що ваги прихованих нейронів є незмінними й призначаються за допомогою функції випадкових значень. Шар вхідних нейронів у такій мережі слугує для ініціалізації системи, а вихідний шар працює в якості спостерігача за порядком активації нейронів; причому під час навчання зв'язок змінюється тільки між спостерігачем і прихованими шарами. Таким чином, функція похибки є квадратичною відносно параметрів і диференціюється до лінійної системи. Варіантом організації мереж з відлунням стану є мережі нестійких станів (рідкі скінченні автомати). Мережі нестійких станів відносяться до імпульсних ШНМ і замість сигмоідальної кривої використовують порогові функції, а отже кожен нейрон також є блоком пам'яті. Коли стан нейрона оновлюється, його значення складається з самим собою. Після перевищення порогу нейрон посилає імпульс нейронам сусіднього шару. Мережі нестійких станів показали високу ефективність при роботі у режимі реального часу і активно використовуються як при аналізі інтернет-трафіку, так і відеоінформації з камер об'єктів, що знаходяться під охороною.

Актуальним на сьогоднішній день способом використання МПП є автокодувальник [3-5, 8]. Принцип роботи автокодувальника полягає у кодуванні вхідної інформації шляхом її стиснення, таким чином, половина прихованих шарів кодує вхідну інформацію а друга половина декодую, причому середній шар, що є найменшим, утримує код (рис. 4). Як і більшість МПП автокодувальник навчають методом зворотного поширення помилки, подаючи вхідні дані і задаючи помилку як різницю між входом і виходом. Автокодувальник не використовують безпосередньо для розпізнавання сигнатур кібератак, але його можна застосувати як допоміжну систему з метою стиснення вхідних данних та зменшення шумів. У випадку СРК, це означає, що система з функціональним блоком автокодувальника здатна ефективно виділяти сигнатури кібератак за умов, коли схема кібератаки частково відрізняється від схеми з навчальної вибірки

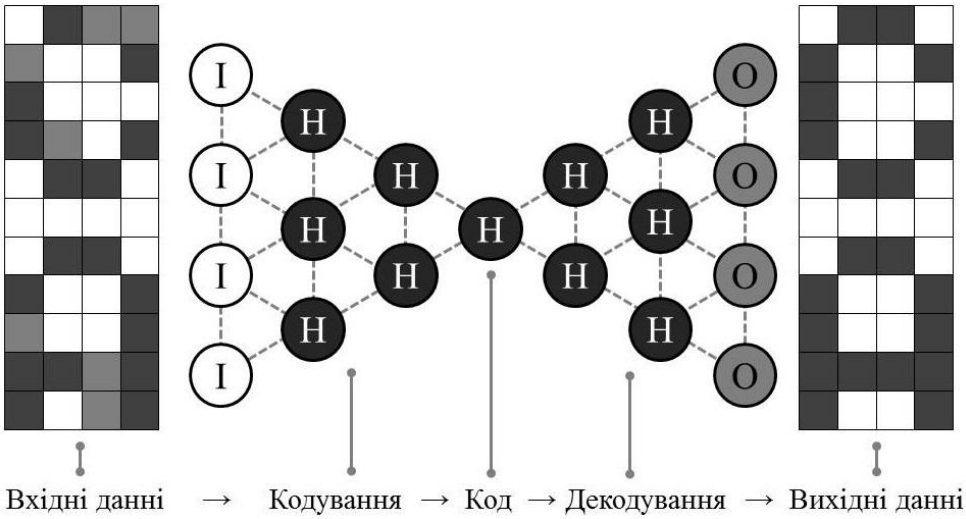

Рис. 4. Архітектура нейронної мережі прямого поширення типу «автокодувальник».

У сучасних СРК існує кілька схем застосування НММ автокодувальник, серед яких найбільш популярні: розріджений автокодувальник, варіаційний автокодувальник та завадостійкий автокодувальник. Архітектура розрідженого авто кодувальника [5, 8] протилежна архітектурі

показаній на рис. 4., кодування збільшує об'єм інформації і середній прихований шар є найбільшим. При навчанні розрідженого автокодувальника за системою зворотного поширення помилки можна НММ розподіляє вагові коефіцієнти таким чином, що вихідний сигнал повністю відповідає вхідному, тому у якості додаткової умови додається штраф за кількість активованих нейронів в прихованому шарі. Це в якісь мірі нагадує біологічну нейронну мережу, в якій більшість нейронів під час виконання задачі знаходяться не в збудженому стані. Мережі цього типу можуть застосовуватися для вилучення особливостей сигнатури кібератаки і застосовуються у якості додаткового вузла разом з основною системою. Варіаційні автокодувальники [8], у свою чергу, здатні більшою мірою аналізувати ступінь впливу однієї події на іншу, що актуально при аналізі трафіку у режимі реального часу. Даний тип автокодувальників ближче до машин Больцмана, хоча і спирається на Байєсова математику імовірнісних суджень. Такий підхід надає можливість виключати вплив одних нейронів на інші, показав свою ефективність при розробці СРК на основі НММ. Завадостійкий автокодувальник [8] найбільш ефективний при роботі з сигнатурами кібератак, що певною мірою відрізняються від схем навчальної вибірки. Система сприймає відмінності як шум і орієнтується на основні риси притаманні сигнатурі.

2. Нейромережеві моделі, що базуються на ланцюгах Маркова

Ланцюгом Маркова називають випадковий процес, що задовольняє властивості Маркова і приймає скінченну кількість станів [8]. У ланцюгах Маркова задаються ймовірності переходу з поточного стану в сусідні; причому подальший стан залежить від поточного і не залежить від минулих станів. Ланцюги Маркова не є ШНМ, але їх розглядають як теоретичну основу машин Больцмана. Машина Больцмана є видом стохастичної рекурентної нейронної мережі. Даний тип НММ не використовується при розпізнаванні сигнатур кібератак, але з каскаду обмежених машин Больцмана можна побудувати глибинну мережа переконань, як основу СРК. Характерною особливістю машини Больцмана є наявність тільки двох типів нейронів: відкритих V та прихованних H (рис. 5). При цьому відкриті нейрони можна розглядатися як вхідні, що стають вихідними, коли всі нейрони у мережі оновлюють свої стани. Процес навчання даного класу НММ відбувається шляхом присвоєння вагових коефіцієнтів випадковим чином та навчання методом зворотного поширення або за допомогою марковського ланцюга. Основним недоліком машин Больцмана є схильність до «стабілізації» стану мережі в локальному мінімумі, з метою подолання зазначеної проблеми була застосована концепція «теплового шуму» та алгоритм імітації відпалу. Якщо ввести параметр рівня теплового шуму t, то ймовірність активності окремого нейрона k визначається на основі ймовірнісної функції Больцмана:

$$P_k = \frac{1}{1+e^{-E_k/t}}, \quad (4)$$

де енергія E_k визначається як сума ваг зв'язків нейрона k зі всіма активними на даний момент нейронами. Отже активація регулюється значенням загальної температури, при зниженні якої зменшується і енергія нейронів, а зменшення енергії викликає стабілізацію нейронів. Таким чином, якщо температура задана вірно, система досягає рівноваги.

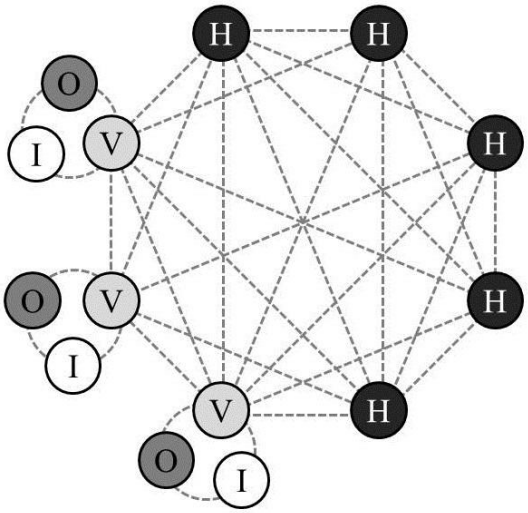

Рис. 5. Базова схема нейронної мережі типу «машина Больцмана».

У більш актуальній для побудови СРК обмеженій машині Больцмана [9] існують зв'язки лише між групами нейронів, що значно спрощує модель. НММ на основі обмеженої машини Больцмана навчаються аналогічно до мереж прямого поширення, але замість передачі даних від першого до останнього шару, в даному випадку данні у кінці ітерації повертаються до першого шару, що одночасно є вхідним і вихідним; отже у схемі застосовується пряме-зворотнє поширення.

Головною проблемою розпізнавання сигнатури кібератаки є відсутність повної інформації про образ, що має бути розпізнаним через внесення зловмисниками змін у стратегію атаки. У даному випадку певну ефективність можуть показати НММ у основу яких покладено архітектуру мереж Хопфілда та мереж Хеммінга, що є близькими за реалізацією до машин Больцмана. Мережа Хопфілда базується на матриці зв'язків *W* і векторі порогових значень активації нейронів *T*. Образом мережі Хопфілда є вихідний вектор *O*, що приймає значення -1 або 1. Образ, що подається на вхід мережі залишається до тих пір, доки стан мережі не зміниться. Кожне з'єднання мережі Хопфілда характеризується вагою w_{ij}, причому:

$$\begin{cases} w_{ij} = w_{ji} \\ w_{ii} = 0 \end{cases}, \tag{5}$$

що вказує на симетричність матриці зв'язків.

Кожен нейрон мережі слугує вхідним до навчання, прихованим під час навчання і вихідним після навчання. Після навчання декільком сигнатурам кібератак НММ буде завжди сходитися до однієї з них, як до стаціонарного стану. Система стабілізується тільки частково через те, що загальна енергія E_k знижується під час навчання. Кожен нейрон має поріг активації, що співвідноситься з даною енергією, і якщо сума вхідних даних перевищує поріг, нейрон може переходити в один з двох станів; при цьому вузли мережі можуть оновлюватися як паралельно, так і послідовно. Коли кожен з нейронів оновився і їх стани більше не змінюються, мережа приходить в стаціонарний стан. Таким чином, мережа Хопфілда є базовим засобом розпізнавання сигнатур кібератак, що характеризується низькою ресурсоємністю. Для інформаційним систем з низьким апаратним ресурсом також використовують мережі Хеммінга. Даний тип НММ складається з двох шарів, перший і другий шари складаються з m нейронів, нейрони першого шару мають по n синапсів, які з'єднані з входами мережі, а нейрони другого шару пов'язані між собою. Робота мережі полягає в знаходженні числа незбіжних компонент двох бінарних векторів між сигнатурою кібератаки, що подається на вхід мережі і кожним з образів, що система запам'ятала до процесу навчання.

3. Побудова систем розпізнавання кібератак на основі нейромереж третього покоління

Більшість сучасних нейромережевих СРК базуються на концепції глибинного навчання. Навчання глибинних ШНМ складається з трьох етапів: навчання кожного шару автоасоциативної

мережі за методом «навчання без вчителя», етапу ініціалізації нейронів прихованих шарів мережі прямого поширення, та етапу навчання НММ зі вчителем. Під час першого етапу вагові коефіцієнти нейронів шару стають вхідними данними для нейронів наступного шару, що призводить для узагальнення інформації про сигнатуру кібератаки. Третій етап займає менший час ніж для ШНМ другого покоління, завдяки ініціліазації нейронів. СРК на основі ШНМ третього покоління показують кращі результати при роботі з принципово новими стратегіями кібератак, але при цьому характеризуються високою ресурсоємністю по відношенню до завантаження апаратного комплексу та значним терміном підготовки до навчання. На сьогоднішній день у СРК найбільш найбільш активно використовуються НММ, що базуються на глибинній мережі переконань (ГМП) та згорткові ШНМ.

ГМП [10] представляють собою композицію декількох обмежених машин Больцмана або варіаційні автокодувальників, що пов'язані таким чином, щоб одна мережа навчалася кодувати іншу. Цей метод полягає в прийнятті оптимального на конкретний момент рішення, з метою отримання відповідного результату, що особливо важливо при роботі у режимі реального часу. Структура ГМП мережі складається з шарів латентних змінних і з'єднань між шарами. При тренування на навчальному наборі образів шари виступають в ролі детекторів ознак на входах. ГМП зручно розглядами як набір обмежених машин Больцмана, де прихований шар підмережі n слугує видимим шаром для підмережі $n+1$. Навчання ГМП складається з тренування машини Больцмана на вхідній матриці, перетворення матриці за допомогою машини Больцмана для отримання нових даних (для всіх шарів мережі) та тонкого налаштування параметрів глибинної архітектури. Спільний розподіл даної НММ для прихованих шарів h^n з $P\left(h^k \vee h^{k+1}\right)$ —умовним розподілом $P\left(h^k \vee h^{k+1}\right)$, та спільним розподілом машини Больцмана верхнього рівня $P(h^{n-1},h^n)$ визначається як:

$$P(x,h^1,h^2,\dots,h^n) = \left(\prod_{k=0}^{n-2} P\left(h^k \vee h^{k+1}\right)\right) P(h^{n-1},h^n), \quad (6)$$

Згорткові ШНМ [11] характеризуються тим, що обробляють поля вхідного образу за допомогою декількох шарів. Важливо зауважити, що для збільшення точності (через збільшення ресурсоємності) даної НММ пропонується накладати вибірки кожного шару з перекривання областей, що також зменшує вплив паралельних перенесень образу. Для нейронів вихідного шару згорткової ШНМ використовується одна матриця вагових коефіцієнтів, що називається ядром згортки (рис. 6). Шар отриманий в результаті операції згортки демонструє наявність ознаки в шарі обробки, за рахунок чого формується карта ознак. Під час етапу субдискретізації відбувається зменшення розмірності сформованих карт ознак і формується підвибірка. Таким чином, подальші обчислення прискорюються, а ШНМ стає інваріантною до масштабу вхідного образу (сигнатури кібератаки). Далі сигнал знов проходить шари згортки, в яких повторюються операції згортки і субдискретізації, причому на кожному наступному шарі карти ознак зменшуються, а кількість каналів — збільшується (рис. 6). СРК на основі згоркових ШНМ здатні розпізнавання складні ієрархій ознак, є ефективними і в більшості реалізацій менш ресурсоємними ніж ГМП.

Оберненим типом ШНМ по відношенню до згорткових ШНМ можна вважати розгорткові нейронні мережі, що генерують образ, а не розпізнають його. Замість шарів об'єднання в данному випадку застосовуються операції інтерполяції та екстраполяції. На основі концепції згорткових та розгорткових ШНМ також були побудовані та набули певну нішу застосування глибокі згорткові зворотні глибинні мережі, що базуються на варіаційних автокодувальниках у яких згорткові і розгорткові мережі працюють як функціональні блоки, що кодують та декодують сигнал. Такі мережі здатні розрізнити кілька видів кібератак, що застосовуються одночасно. і виділити їх з основного трафіку. Тим не менш даний тип ШНМ більш активно використовується при роботі з зображеннями.

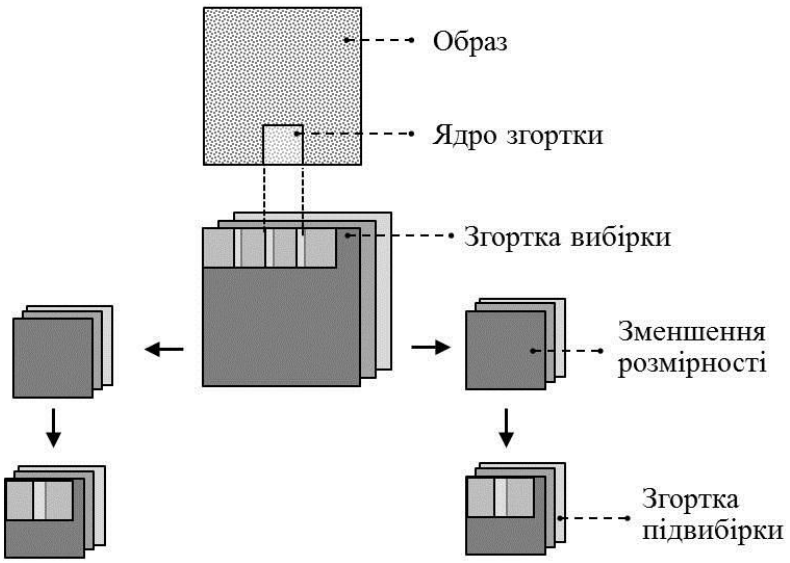

Рис. 6. Типова архітектура згорткової нейронної мережі.

При наявності достатнього апаратного ресурсу при розробці СРК має сенс використовувати комбінацію ШНМ третього покоління [12]. Базовою концепцією у даному випадку можуть слугувати генеративні мережі змагання (Generative adversarial networks), що включають у себе дві мережі, які працюють разом. На практиці зазвичай застосовують дві згорткові ШНМ, одна з яких генерує дані (генератор), а інша їх аналізує (дискримінатор). Дискримінатор отримує на вхід навчальні дані або дані згенеровані першою мережею і далі точність визначення дискримінатором образу слугує для оцінки помилок генератора. Дана схема ефективно працює у режимі навчання та використовується на етапі розробки архітектури нейромережевих СРК.

Висновки та перспективи подальшого дослідження. Штучні нейронні мережі є ефективним засобом розпізнавання кібератак на інформаційні системи. Для побудови методологічної бази розробки нейромережевих систем розпізнавання сигнатур кібератак необхідно визначити параметри цільових функцій таких комплексів та сформувати єдиний підхід до аналізу різних класів штучних нейромереж. Практика показує, що нейромережі третього покоління можуть бути розглянуті як сукупність базових нейромереж і, таким чином, піддаються систематичному аналізу на основі функцій ефективності, оперативносі та ресурсоємності.

1. Y. Mo, T. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopol, "Cyber-physical security of a smart grid infrastructure," Proceedings of the IEEE, vol. 100, no. 1, pp. 195-209, January 2012.
2. Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, and W. Zhao, "On false data-injection attacks against power system state estimation: Modeling and countermeasures," IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 3, pp. 717 - 729, March 2014.
3. Y. Liu, L. Yan, J. Ren, and D. Su, "Research on efficient detection methods for false data injection in smart grid," in International Conference on Wireless Communication and Sensor Network (WCSN), Wuhan, China, December 2014, pp. 188 - 192.
4. Z. Hu, Y. Wang, X. Tian, X. Yang, D. Meng, and R. Fan, "False data injection attacks identification for smart grids," in Third International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), Beirut, Lebanon, April - May 2015, pp. 139 - 143.
5. Annema, Jouke. Feed-forward neural networks: vector decomposition analysis, modelling and analog implementation. Place of publication not identified: Springer-Verlag New York, 2013. Print.
6. Yi, Zhang, and K. K. Tan. "Discrete Recurrent Neural Networks." Network Theory and Applications Convergence Analysis of Recurrent Neural Networks (2004): 195-217.
7. Čerňanský, Michal, and Peter Tiňo. "Predictive Modeling with Echo State Networks." Artificial Neural Networks - ICANN 2008 Lecture Notes in Computer Science (n.d.): 778-87.
8. Rowe, Daniel B. Multivariate Bayesian statistics: models for source separation and signal unmixing. Boca Raton, FL: Chapman & Hall/CRC, 2003
9. Safarikas, Al. Stock picking using an artificial neural network. N.p.: n.p., 1994. Print.
10. Y. Yan, X. Yin, S. Li, M. Yang, and H. Hao, "Learning document semantic representation with hybrid deep belief network," Computational Intelligence and Neuroscience, vol. 2015, pp. 1 - 9, 2015.
11. Sercu, T., & Goel, V. (2016). Advances in Very Deep Convolutional Neural Networks for LVCSR. Interspeech 2016. doi:10.21437/interspeech.2016-1033
12. Shinohara, Y. (2016). Adversarial Multi-Task Learning of Deep Neural Networks for Robust Speech Recognition. Interspeech 2016.