УДК 004.32.26:004.056.523
DOI 10.47049/2226-1893-2025-1-212-227

ІДЕНТИФІКАЦІЯ КОРИСТУВАЧІВ ЗА КЛАВІАТУРНИМ ПОЧЕРКОМ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ

Ю.В. Даус
к.геогр.наук, доцент кафедри «Технічна кібернетика
й інформаційні технології ім. проф. Р.В. Меркта»
ORCID: 0000-0001-9737-4663

Одеський національний морський університет, Одеса, Україна

С.В. Самойлов
к.ю.наук, начальник 3-го управління (інформаційних технологій та програмування)
ORSID: 0009-0004-0057-1126

Департамент кіберполіції Національної поліції України

М.Є. Даус
к.геогр.наук, доцент кафедри «Безпека життєдіяльності, екології та хімії»
ORCID:*0000-0001-5298-795X*

Д.Г. Ларін
к.т.наук,доцент кафедри «Технічна кібернетика
й інформаційні технології ім. проф. Р.В. Меркта»
ORSID: 0009-0006-4882-0683

Одеський національний морський університет, Одеса, Україна

Анотація. *В наш час цифрових технологій одною з провідних задач в кібербезпеці є надійна ідентифікація користувачів. В останні роки стрімко набувають популярності біометричні методи, які дають змогу чітко ідентифікувати користувача. Одним з найбільш поширених кіберзлочинів є викрадання паролів, та вхід с цими даними до різноманітних сервісів: банківського, страхового, учбового, податкового. Там кіберзлочинці, отримуючи права та доступ до інформації користувачів, завдають громадянам фінансових, репутаційних та моральних втрат. Ще більше втрат можуть отримати приватні та державні установи внаслідок витоку персональних даних, приватної інформації про майновий та фінансовий стан. Щороку потужність та кількість кібератак з використанням викрадених паролів тільки зростає. Тому необхідно розробити комплекс методів та заходів для протидії таким кібератакам. Одним з біометричних методів ідентифікації користувача є поведінковий біометричний метод з використанням клавіатурного почерку, який є індивідуальною особливістю кожної людини. Для отримання характеристик клавіатурного почерку не потрібно ніяких додаткових датчиків та сенсорів – достатньо звичайної клавіатури та деякого програмного забезпечення.*

У статті розглянуті основні характеристики клавіатурного почерку, застосована нейронна мережа для розпізнавання авторського та неавторського введення пароля. Загалом вдалося побудувати доволі просту та ефективну нейронну мережу для ідентифікації користувачів по клавіатурному почерку. Використання даної методики може значно підсилити достовірність ідентифікації користувачів і кібербезпеку інформаційних телекомунікаційних систем.

Ключові слова: *кібербезпека; вразливості; кіберзлочинці, ідентифікація користувачів, клавіатурний почерк, нейромережі.*

UDC 004.32.26:004.056.523
DOI 10.47049/2226-1893-2025-1-212-227

USERS IDENTIFICATION BY KEYBOARD HANDWRITING USING NEURAL NETWORKS

Y. Daus
Ph.D., docent of the Department «TechnicalCybernetics and Information Technologies named after prof. R.V. Merkt»
ORCID: 0000-0001-9737-4663

Odesa National Maritime University. Odesa, Ukraine

S. Samoilov
Ph.D., Head of the 3rd department (information technologies and programming)
ORSID: 0009-0004-0057-1126

Cyberpolice Department of the National Police of Ukraine

M. Daus
Ph.D., docent of the Department «Safety of Life, Ecology and Chemistry»
ORCID:*0000-0001-5298-795X*

D. Larin
Ph.D., docent of the Department «Technical Cybernetics and Information Technologies named after prof. R.V. Merkt»
*Scopusid:*6602177580

Odesa National Maritime University. Odesa, Ukraine

Abstract. *In our time of digital technologies, one of the leading tasks in cybersecurity is reliable user identification. In recent years, biometric methods that allow for clear user identification have been rapidly gaining popularity. One of the most common cybercrimes is password theft and logging in with these data to various services: banking, insurance, education, tax. There, cybercriminals, gaining rights and access to user information, cause citizens financial, reputational and moral losses. Private and public institutions can suffer even greater losses due to the leakage of personal data, private information about property and financial status. Every year, the power and*

number of cyberattacks using stolen passwords only increases. Therefore, it is necessary to develop a set of methods and measures to counter such cyberattacks. One of the biometric methods for user identification is the behavioral biometric method using keyboard handwriting, which is an individual feature of each person. To obtain the characteristics of keyboard handwriting, no additional sensors are required - a regular keyboard and some software are enough. The article considers the main characteristics of keyboard handwriting, a neural network is used to recognize the author's and non-author's password entry. In general, it was possible to build a fairly simple and effective neural network for identifying users by keyboard handwriting. The use of this method can significantly enhance the reliability of user identification and cybersecurity of information telecommunication systems.

Keywords: *cybersecurity; vulnerabilities; cybercriminals, user identification, keyboard handwriting, neural networks.*

Вступ. У наш час, коли все більше і більше пристроїв та сервісів потребують ідентифікації користувачів, на перший план виходить необхідність ідентифікувати особу або користувача не тільки фізично, а й віддалено. Віддалена ідентифікація (без фізичної присутності особи) робить привабливими та безпечнішими банківські сервіси, покупки в інтернет-магазинах, дистанційне навчання в університетах та дистанційну роботу. Завдяки цьому, роботодавець та продавець можуть значно розширити ареал своїх робітників та споживачів. Тому ідентифікація користувачів стає нагальною задачею для системних адміністраторів, підрозділів кібербезпеки та захисту інформації. Шкода, яку можуть завдати зловмисники, може стати непід'ємним тягарем фінансових та репутаційних втрат для підприємства чи організації, та навіть спричинити закриття цих організацій.

Загалом ідентифікацію користувачів можна поділити на кілька основних видів. Основні переваги та недоліки таких методів приведені в таблиці 1. Одним із самих надійних засобів є біометрична ідентифікація особи. Кожна людина від народження має свої унікальні особливості, які не змінюються протягом всього життя людини. До таких особливостей потрібно віднести унікальність відбитків пальців, практично неможливо знайти ще одну таку людину з такими же відбитками (крім близнюків). Недоліком є необхідність контактно зчитати відбиток пальця пристроєм.

Такими ж унікальними є райдужна оболонка та сітківка ока. Особливістю ідентифікації даного метода є безконтактний спосіб зчитування інформації, що дає можливість дистанційно проводити ідентифікації. Але для цього потрібно камери з великою роздільною здатністю.

Ідентифікація по біометрії обличчя в наш час стала дуже популярна. Практична вірогідність ідентифікації досягає 99 %. Цю технологію зараз взяли за основу банківська сфера, прикордонна служба (біометричні паспорти). В той же час зменшується популярність ідентифікації особи по біометрії руки та кисті, оскільки ці параметри можуть змінюватися протягом життя, а особливо після отримання фізичних травм.

Ідентифікація по термограмі не має особливого розповсюдження, оскільки для цього потрібно проводити термографічне сканування обличчя. Але в зв'язку з розповсюдженням вірусу COVID-19 в аеропортах установили таке обладнання для виявлення хворих людей. В подальшому, при виготовленні біометричних паспортів можливо також робити термограму обличчя, яка є унікальною для кожної людини, та в майбутньому проводити ідентифікацію особи за термограмою.

Таблиця 1

Види ідентифікації користувачів в інформаційних системах

Тип ідентифікації	Перевірка	Переваги	Недоліки
Парольний	Користувач вводить пароль, пін-код	Швидкий доступ, простота використання	Невідомо, хто вводить пароль
Електронні ключі	Використовується електронно-цифровий підпис (ключ + пароль)	Більша надійність за простий пароль. Прирівнюється до звичайного підпису особи	У випадку втрати ключа зловмисник отримує повний доступ та може підписуватися за користувача
Біометрична перевірка	Використовуються біометричні характеристики людини (відбитки пальців, сітківка ока, геометрія обличчя, геометрія руки, термограма тіла)	Дуже надійна ідентифікація	Потрібні апаратні засоби зчитування біологічних характеристик, деякі з них доволі дорогі
Динамічна перевірка	Клавіатурний почерк, рукописний текст, тембр голосу	Дозволяє отримувати додаткові механізми перевірки	Для зчитування деяких характеристик необхідне допоміжне обладнання (сканер, мікрофон)

Ми пропонуємо звернути увагу ще на один біометричний метод ідентифікації користувача – поведінковий. До поведінкової біометрії відносять впізнавання користувача по клавіатурному почерку по голосу, особливостям ходьби [1].

Для ідентифікації по голосу також необхідні датчики та обладнання для передачі такої інформації. При ідентифікації по клавіатурному почерку не потрібні додаткові датчики, а тільки програмне забезпечення, що можливо встановити практично на любому комп'ютері.

Динамічна перевірка менш надійніша за біометричну, але дозволяє використовувати наявні апаратні ресурси. Більш ефективними ці методи стають в

поєднанні з іншими методами. Наприклад, поєднання клавіатурного почерку з парольною ідентифікацією. З одного боку для аналізу клавіатурного почерку ми скорочуємо об'єм для аналізу, з іншого боку ми також перевіряємо правильність вводу паролів.

Вибір характеристик клавіатурного почерку та моделі розпізнавання. Напевно першою датою ідентифікацію особи по клавіатурному почерку можна вважати 1850 рік, коли в широке застосування ввійшов телеграф. Тоді телеграфісти спілкувалися та передавали сповіщення за допомогою азбуки Морзе. І вже тоді стало зрозуміло, що оператори могли впізнавати один одного по динаміці та манері передавання сповіщень. Метод розпізнавання радиста по манері, швидкості та ритму передавання інформації широко застосовувався під час другої світової війни. Це дозволяло встановити, що інформації передавала людина, якій довіряли, а не ворожий агент чи дезінформатор.

Виходячи з особливостей роботи з клавіатурою, можна виділити основні характеристики клавіатурного почерку:

1. Швидкість введення символів.
2. Динаміка введення – характеризується часом утримання клавіші та часом між відпусканням клавіші та натиснення на наступну клавішу.
3. Частота виникнення помилок при введенні – натискання клавіш Delete та Backspace.
4. Сила натискання кнопок.

Ідентифікація користувача за клавіатурним почерком можлива як за набором довільного тексту, так і за набором ключової фрази, а в нашому випадку це введення пароля. В першому випадку ми маємо доволі суттєвий набір дослідницьких даних, а в іншому жорстко обмежені довжиною пароля.

У ході дослідження встановлено [2], що часові інтервали між натисканнями клавіш точніше характеризують клавіатурний почерк користувача, ніж час утримання клавіш. Таким чином, на думку автора [2], для ідентифікації користувача по клавіатурному почерку слід використовувати:

1. Швидкість набору.
2. Час утримання клавіші та число перекриттів між клавішами.
3. Інтервали між натисканнями клавіш.
4. Ступінь аритмічності при наборі символів.

Значний вплив на ритмічність набору мають технічні навички у людини – здатність друкувати одним, два або всіма пальцями. Завжди відчувається, коли людина послідовно вводить два або три символи різними пальцями, так звані диграфи та триграфи [3]. Диграфи більш потужно себе проявляють, коли людина набирає текст однією рукою, а триграфи − коли людина набирає текст або пароль двома руками.

Після отримання характеристик клавіатурного почерку, необхідно вибрати модель розпізнавання. На основі диграфів та триграфів була розроблена модель [4] з використанням квадратичного індексу нечіткості. Такий підхід дозволив дослідникові отримати правильну ідентифікацію в 82,3 % процентів випадків.

Іншим підходом є застосування розпізнавання методом описання даних опорними векторами (SVDD) і однокласове навчальне векторне квантування (LVQ) [5]. Кращі результати були отримані при застосуванні метода LVQ.

Ще одним із сучасних методів є побудова нейронної мережі для розпізнавання клавіатурного почерку. Вчені Махерхварі та Вікрама Пуді [6] побудували таку нейрону мережу, використовуючи три прихованих шара по 100-400-100 нейронів відповідно.

Для практичного застосування необхідно розробити просту в реалізації модель і швидку в обчисленнях. Найбільш ефективним, на думку авторів, є побудова нейронної мережі з мінімально можливою кількістю прихованих шарів.

Опираючись на роботу [6], для розпізнавання користувачів по клавіатурному почерку було обрано повнозв'язану нейронну мережу прямого поширення, приклад якої представлено на рис. 1.

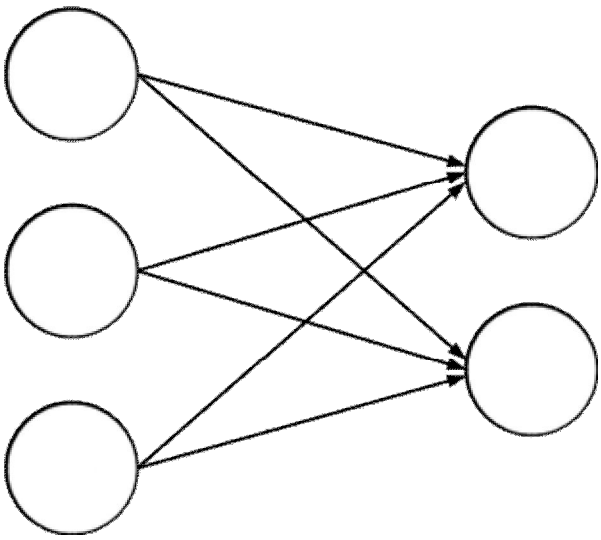

Рис. 1. Повнозв'язана нейронна мережа прямого поширення

Основною невідомою змінною нейронної мережі є її функція активації. Найбільш вдалою функцією, на нашу думку, є функція сигмоїду, яка є неперервно диференційована на всьому відрізку та представлена в формулі 1. Нелінійність функції дозволяє гнучко передавати та відтворювати сигнали в неронній мережі.

$$\sigma(x) = \frac{1}{1+e^{-x}} = \frac{e^{x}}{1+e^{x}} \tag{1}$$

Для обчислення вагових коефіцієнтів використаємо метод зворотного розповсюдження похибки. Цей метод є модифікацією метода градієнтного спуска, та дозволяє в деяких випадках обходити локальні мінімуми.

Отримання експериментальних даних. Нейронна мережа, розроблена на основі багатошарового перцептрону, має бути налаштована таким чином, що отримання вхідного сигналу на початковий шар генерувало певний вихід даних, який відповідав би нашим сподіванням. Для цього потрібно, щоб відповідні вагові коефіцієнти в нейронах відображали відповідність очікуваним зв'язкам в мережі. Таким чином, нам необхідно вирішити задачу на визначення цих вагових коефіцієнтів – провести навчання нашої нейронної мережі. Типовий алгоритм навчання здійснює пошук у множині всіх можливих ваг, щоб отримати такий набір ваг, який найкраще відповідає заданим прикладам. Для керованого начання нейронній мережі необхідно спочатку отримати експериментальні дані для калібрування такої мережі та перевірки на незалежному матеріалі.

Для отримання таких данних нами був розроблений додаток KeyPress4 на мові програмування C#. Цей додаток фіксував технічні характеристики при введені пароля: час утримання клавіші, час між відпусканням клавіші та натисненням наступної клавіші, загальний час введення пароля, інтегральні характеристики (загальний час введення пароля, час натиснення диграфів та триграфів). Додаток може працювати з багатьма користувачами індивідуально, збираючи та записуючи дані в файл. Оскільки багато інформаційних систем, на сьогоднійшний момент ставлять до паролів особливі вимоги:

- одна літера обов'язково повинна бути великою;
- пароль повинен містити число та спеціальний символ (спецсимвол);
- довжина пароля від 8 до 12 символів.

Такі вимоги, на нашу думку, підвищують надійність пароля, а для нашого випадку це допомагає збільшити точність ідентифікації клавіатурного почерку, оскільки переходи від маленьких букв до великих, а також при введенні спецсимволу яскраво фокусують нашу увагу на особливості клавіатурного почерку людини.

На рис. 2 представлені отримані експериментальні дані по авторському введенню пароля, де Tm – час утримання клавіші в мілісекундах; TS – порядковий номер введення парольного символу.

Загалом нами було отримано 145 авторських введень пароля. При отриманні даних проводилось не більше ніж два введення пароля на день, що відповідає стандартній поведенці користувача в інформаційній системі, де паролі вводяться зранку, та після обідньої перерви.

Для отримання неавторського введення пароля були залучені добровольці, яким для входу в систему був виданий логін та пароль. Кожному добровольцю дозволялось зробити не більше трьох спроб введення пароля, щоб врахувати обмеження системи на невірно введений пароль. Загалом було отримано 48 відбитків неавторського введення пароля. З врахуванням авторського введення пароля ми отримали 193 відбитки введення пароля.

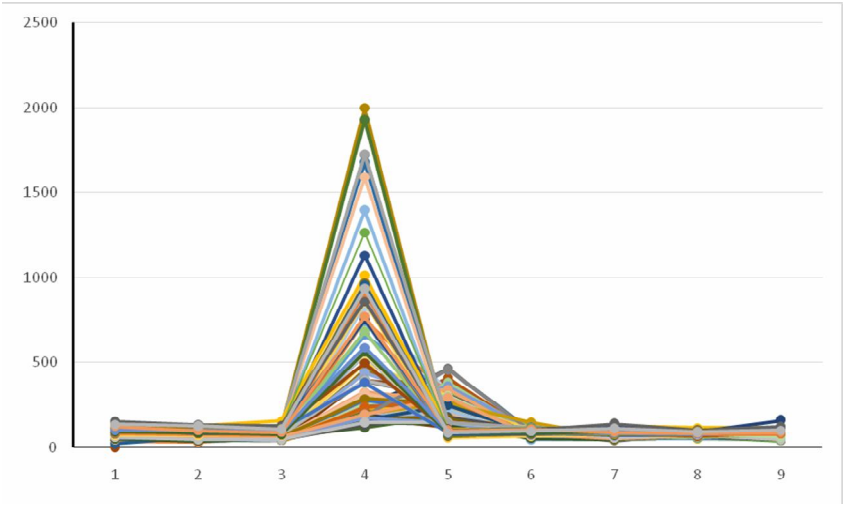

Рис. 2. Отримані експериментальні дані авторського введення пароля

Загалом при порівнянні авторського та неавторського введення пароля різниця в технічних характеристиках помітна зразу, а ще більше вона проявляється при аналізі інтегральних характеристик. На рис. 3 представлено графік загальної тривалості введення пароля.

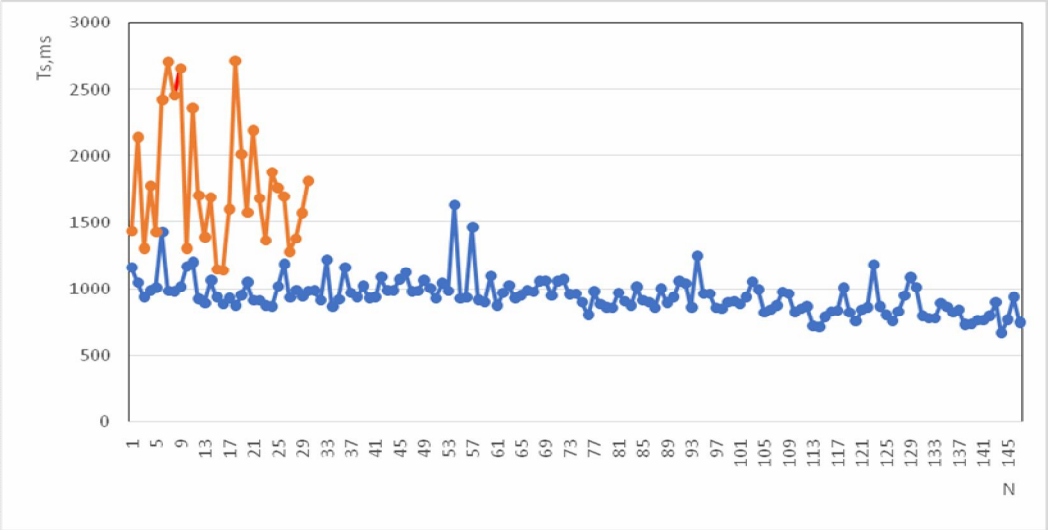

Рис. 3. Тривалість введення пароля:
авторський (нижня лінія) та неавторський (верхня лінія)

Звичайно людина – це не машина і не робот, людина може перебувати в різноманітному фізичному та психологічному стані, що відображається на загальному часі введення пароля.

Аналізуючи графік, ми можемо сказати, що загалом при неавторському введенні пароля користувач використовував значно більше часу, оскільки текст для введення був незнайомий та не містив змістовного навантаження. А при відсутності змістовного навантаження в тексті, буфер пам'яті людини значно скорочується, що веде до значного сповільнення швидкості набору.

Навчання нейронної мережі та задання начальних параметрів. Для побудови нейронної мережі, що адекватно описує та ідентифікує користувачів за клавіатурним почерком, необхідно вирішити задачу знаходження вагових коефіцієнтів $\mathbf{W_i}$ при кожному нейроні. У багатошаровому перцептроні між вхідним та вихідним шарами знаходяться приховані шари. У цих шарах знаходиться функція підсумовування, яка додає всі зважені входи та зміщення. У кінці штучного нейрона сума попередньо зважених входів та зміщення проходить через функцію активації. Потім штучний нейрон передає оброблену інформацію на вихід. Таким чином, можна вважати сучасну модель нейронної мережі – моделлю з багатошаровим перцептроном. Принцип роботи штучного нейрона приведений на рис. 4 [7; 8].

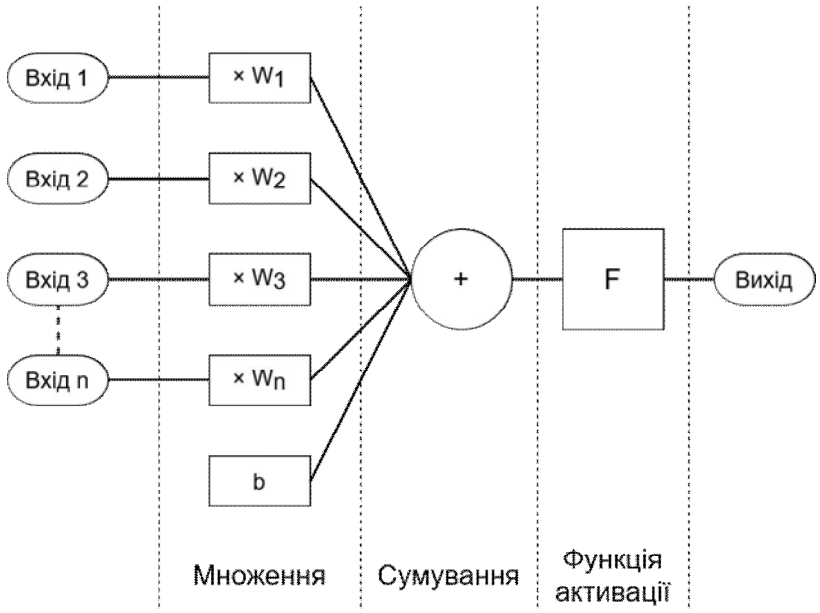

Рис. 4. Принцип роботи штучного нейрона

Для знаходження значень вагових коефіцієнтів використаємо метод зворотнього розповсюдження похибки, а в якості функції активації виберемо функцію сигмоїду [9; 10]. Для навчання нашої нейронної мережі ми розробили власний програмний комплекс KeyNet. Цей комплекс дозволяє створити нейронну мережу будь якої кофігурації і можемо гнучко задати: кількість прихованих шарів, кіль-

кість нейронів кожногш шару, кількість епох навчання, крок змінення вагового коефіцієнта, кількість нейронів на вхідному та вихідному шарі. Всі ці налаштування зчитуються з файлу. Також попередньо проводиться нормализація отриманих даних.

Загалом було отримано 193 випадки введення пароля. Ми розділили отримані дані на навчальні – 145 випадків та перевірочні – 48 випадків. У перевірочних даних 24 випадки авторського введення пароля і 24 випадки – неавторського введення пароля. Дані були відібрані випадковим чином. Для ідентифікації авторського введення пароля ми використовували додаткове поле з значенням 1, для неавтоського введення пароля – значення 0.

Після нормалізації даних, перед початком обчислень, необхідно ініціювати початкові значення вагових коефіцієнтів. Застосували два варіанта: перший – задавались однакові значення для всіх вагових коефіцієнтів, а другий варіант − ініціалізували випадковими значеннями в діапазоні від 0 до 1. Найбільш ефективним виявився метод ініціалізації випадковими значеннями. При ініціалізації однаковими значеннями при деяких значеннях кроку змінення вагового коефіцієнта спостерігався навіть колапс нейронної системи. Можливо це також пов’язано з тим, що сигмоїдна функція активації може мати дуже маленькі значення, чим фактично виключає деякі нейрони з процесу навчання, або навпаки вплив деяких нейронів нівелює вклад інших нейронів.

Важливим являється також визначення кроку змінення та кількості епох навчання. В нашому випадку ми приймали кількість епох від 10 000 до 100 000, а крок змінення вагових коефіцієнтів 0,01. Така кількість епох дозволяла з достатньою швидкістю провести обчислення вагових коефіцієнтів нейронів. Ми проводили чисельні експерименти по зміненню значення кроку вагових коефіцієнтів. При встановленні дуже великого кроку (> 0,1), а також при встановленні дуже малого кроку змінення (< 0,001) призводить до колапсу нашої нейронної мережі. Евристичним шляхом ми встановили, що найбільш ефективними являються значення від 0,008 до 0,03. Такий крок забезпечує на нашій вибірці оптимальне навчання нейронів. Великі кроки, в нашому випадку, доводять до пропуску мінімуму та виходять на плато похибки, в такому випадку встановлюється однакове значення виходу.

В роботі [6] автори побудували свою модель, використовуючи загалом 5 шарів: зовнішній, вихідний, та три внутрішніх шари. Внутрішні приховані шари були по 100-400-100 нейронів відповідно. З функцій активації використовували LeakyReLU, що являється модифікацією звичайною ReLU з деякими обмеженнями. Ці обмеження дозволяють обійти проблему виключення з навчання деяких нейронів, а також проблему з навчанням при великих градієнтах. Звичайно така побудова має свої переваги та недоліки. З переваг – це можливість налаштувати нашу нейрону мережу в дуже широкому діапазоні. З недоліків – необхідні великі обчислювальні потужності та великий часовий проміжок для навчання нашої нейронної мережі.

Наша задача, створити доволі просту та надійну мережу, що може давати сигнал про неавторське введення пароля. Шляхом підбору та експериментування з використанням багатьох варіантів найбільш простою та ефективною виявилась схема з двома прихованими шарами по 7 та 6 нейронів відповідно, а також з одним вихідним шаром та одним вхідним. Вихідний шар складається всього з одного нейрону і він видає результат розпізнання користувача по клавіатурному почерку, а вхідний шар складається з кількості символів пароля. В нашому випадку ми використовували пароль, що складається з 9 символів. Застосування такої схеми нейронної мережі дозволяє використати всього 102 вагових коефіцієнта замість 8100 в моделі, представленій в роботі [6].

Ще одним важливим параметром в налаштуванні нашої нейронної мережі є кількість епох навчання. Для знаходження найменшого значення кількості епох, що дозволяє якісно навчити нашу нейронну мережу розпізнавати авторське та неавторське введення пароля, також були проведені чисельні експерименти із змінної кількості епох. Загалом значення кількості епох змінювали від 100 до 1 000 000. Для оцінки якості навчання нейронної мережі використовували середньоквадратичну похибку (StdErr) розпізнавання клавіатурного почерку. На рис. 5 представлена залежність величини середньоквадратичної похибки від кількості епох.

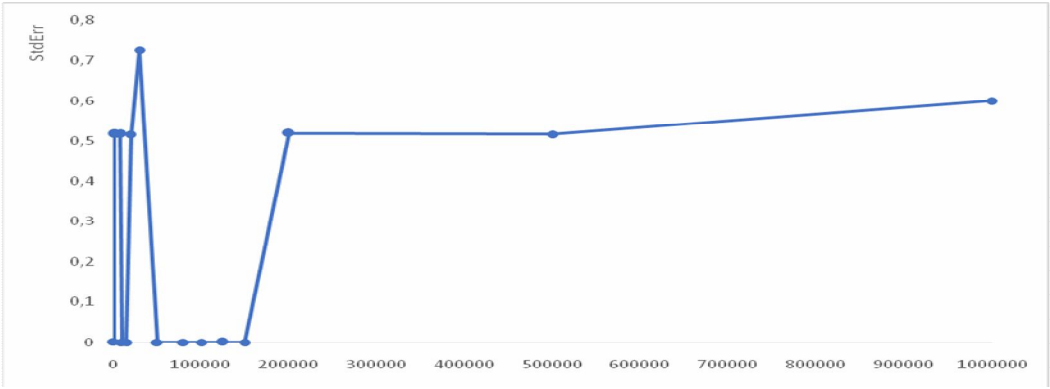

*Рис. 5. Залежність середньоквадратичної похибки StdErr *10^{-5} від кількості епох*

Як видно з рисунка, ландшафт зміни похибки досить цікавий. Наразі, для цього набору даних, існують три локальних мінімуми на значення кількості епох 500, 1000 та 1500. Потім величина StdErr значно збільшується. В діапазоні значень від 50 000 до 150 000 значення StdErr виходять на мінімальні значення і практично не змінюються. Після 150 000 значення похибки спочатку різко зростає, а потім зростання набуває форму плавного підйому. При використанні методу градієнтного спуску, напевно значення першого локального мінімуму привело до зупинки навчання нашої нейронної мережі та не дало б нам можливості вийти на інший мінімум цільової функції.

Таким чином, в даному випадку використання методу зворотного розповсюдження похибки має свої переваги перед методом градієнтного спуску.

Рекомендуємо використовувати значення кількості епох в районі 100 000 [11; 12]. Це дає якісне навчання нашої нейронної мережі.

Ми провели аналіз значень функції сигмоїду при виході на плато похибки. Ці значення приймали дуже маленькі значення в області близької до 0 – значення були менші 10^{-6}- 10^{-10}. Це приводило до дуже маленьких значень вагових коефіцієнтів, і як наслідок виключення деяких нейронів з процесу навчання.

Ефективним методом подолання такого ефекту може бути штучне обмеження верхнього та нижнього значень сигмоїду, тим самим ми отримуємо деякий модифікований сигмоїд схожий на LeakyReLU [6], в якому є штучні обмеження.

Аналіз отриманих результатів. В нашій нейронній мережі для авторського введення пароля цільову функцію ми встановили 1, а якщо це неавторське введення паролю – то цільова функція встановлюється 0. Приймаємо також гіпотезу, що при значенні 0,5 та вище цільової функції, – це було авторське введення, а при значеннях < 0,5 – неавторське введення. Загалом прийняття такої гіпотези − це умовність і дослідники можуть встановити для себе будь які значення, що відповідають предмету дослідження та сенсовому навантаженню на ці значення.

На рисунку 6 представлені фактичні дані (авторське та неавторське введення пароля) **F** та обчислені за допомогою нейронної мережі **S**. Внаслідок дуже малих відхилень ці два графіка практично співпадають і при такому масштабі дуже важко розгледіти відмінності. Загальна середньоквадратична похибка становить $2{,}98*10^{-5}$. Досить часто науковці використовують такий термін як точність (accuracy) обчислення. По суті це вираження в процентах наскільки точно прогноз приближається до фактичного значення: 100 % прогноз співпадає повністю, 0 % – прогноз повністю не співпадає.

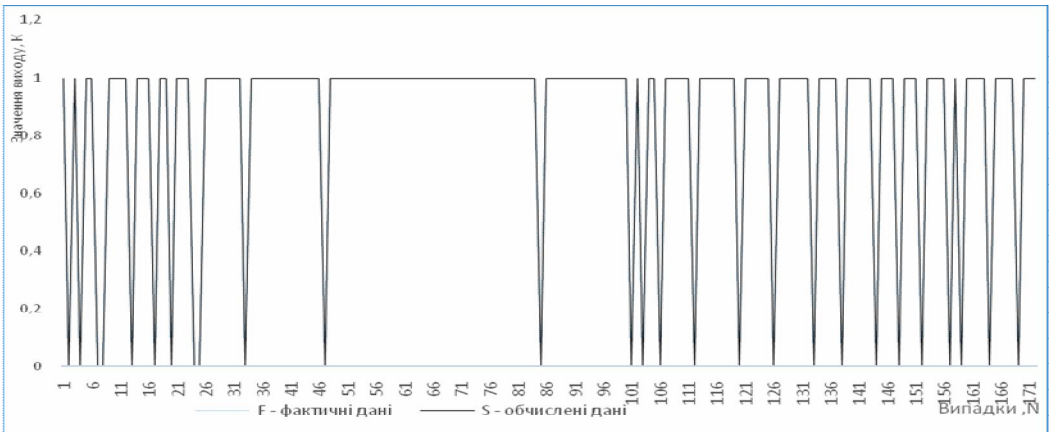

Рис. 6. Фактичні дані (F) та обчислені дані (S)

Цікавим було б порівняння отриманих нами результатів з результатами, що були отримані іншими вченими з використанням різноманітних методів.

У табл. 2 наводиться точність ідентифікації клавіатурного почерку при використанні різних методів [13].

Таблиця 2

Точність ідентифікації по клавіатурному почерку

Назва методу	Точність, %
Метод Леггета, Умпресса і Вільяма	89,5
Метод Джойса й Гопала	78,0
Метод Расторгуєва	90,0
Метод Махерхварі та Вікрама Пуді	82,22-93,59
Метод Нура Ханура	90-92

Найбільш близьким до нашого метода є метод Махерхварі та Вікрама Пуді. Вони також використовували нейронну мережу для ідентифікації користувачів за клавіатурним почерком. В якості вхідних даних використовували дані введення 51 користувача одного і того ж пароля. Загалом 200 випадків введення пароля на одного користувача. Для навчання нейронної мережі брали дані одного користувача і добавляли як неавторське введення по 5 випадків з інших 50 користувачів. Таким чином, загалом було задіяно 450 введень пароля, 10 % – 45 випадків було задіяно в перевірці і не брали участь у навчанні.

Напевно для порівняння наших результатів з роботою Махерхварі та Вікрама Пуді найкраще зробити таблицю, що наглядно показує різницю. Також, для порівняння, ми відтворили модель з кількістю прихованих шарів як в моделі Махерхварі та Вікрама Пуді.

Таблиця 3

Порівняння нашого метода з методом Махерхварі та Вікрама Пуді

Назва показника	Наш метод	Метод Махерхварі та Вікрама Пуді
Кількість прихованих шарів	2	3
Кількість нейронів в прихованих шарах	7-6	100-200-100
Вагових коефіцієнтів	102	81000
Швидкість обчислення моделі, секунди	0,0036	0,84
Точність обчислення (accuracy)	98 %	85,22 – 93,59%
Кількість для навчання нейронної мережі	145	405
Кількість для перевірки на незалежному матеріалі	48	45

Як видно з таблиці, в нашій моделі перевага в меншій кількості прихованих шарів і загальній кількості нейронів. Точність обчислення в нашій моделі також краща – 98 % проти 93,59 %. Моделі Махерхварі та Вікрама Пуді використовують більше експериментальних даних 405 проти 145 в нашій моделі, але кількість випадків перевірки на незалежному матеріалі майже однакова.

Велика кількість нейронів в моделі Махерхварі та Вікрама Пуді забезпечує дуже глибоке навчання, але використання 205 випадків введення зловмисниками пароля викликає запитання. Відомо, що загальна більшість інформаційних систем налаштована таким чином, що після трьох випадків неправильного введення пароля, користувач блокується. Відповідно, 5 випадків неправильного введення пароля інформаційна система не допустить.

Вважаємо, що великою перевагою нашої моделі, є швидкість обчислення моделі: 0,0036 с. проти 0,84 с. Така швидкість дозволяє використовувати дану модель в реальному режимі часу.

Висновки. В наш час нейроні мережі все частіше застосовуються в самих різноманітних областях, починаючи від побуту і закінчуючи військовими технологіями та безпекою. Особливо бурхливо розвивається застосування нейронних мереж та штучного інтелекту в кібербезпеці: це і firewall останнього покоління з штучним інтелектом і аналізатори подій і детекторів шкідливого програмного коду.

В даній роботі було розглянуто можливість застосування нейронних мереж для аналізу та розпізнавання авторського клавіатурного почерку при введенні пароля користувачем. Отримані результати показують, що дана нейронна мережа досить точно відрізняє авторське та неавторське введення пароля.

Побудована нейронна мережа має невелику кількість прихованих шарів, та дозволяє проводити швидке навчання, та миттєво видавати результат розпізнавання. Точність розпізнавання клавіатурного почерку складає 98 %

Дана нейронна мережа може бути застосована в комплексних системах захисту інформаційно-телекомунікаційних систем, спільного використання з firewall останнього покоління з штучним інтелектом. Застосування ідентифікації по клавіатурному почерку може значно знизити загрозу несанкціонованої автентифікації та зменшити площину кіберзагроз.

СПИСОК ЛІТЕРАТУРИ

1. Горошко М.П., Миклуш С.І., Хомюк. П.Г. Біометрія – Львів: Вид-во «Камула», 2004. – 236 с.
2. Іванов А.І. Біометрична ідентифікація особи за динамікою підсвідомих рухів. – Пенза: Вид-во Пенз. держ. ун-ту, 2000. – 188 с.
3. Чалая Л.Є. Модель ідентифікації користувачів по клавіатурному почерку. «Штучний інтелект», № 4. 2004 , С. 811-817.
4. Bryan W.L. & Harter N. (1897). Studies in the physiology and psychology of the telegraphic language. PsychologicalReview, 4(1), P. 27-53. – https://doi.org/10.1037/h0073806

5. Shaffer L.H. Reading and Typing –https://www.researchgate.net/publication/233266615_Reading_and_Typing.
6. Saket Mahesh wary, SoumyajitGanguly, Vikram Pudi. DeepSecure: A Fastand Simple Neural Network based approach for User Authentication and Identification via Key stroke Dynamics. https://www. research gate. net/ publication/322952671_Deep_Secure_A_Fast_and_Simple_Neural_Network_based_approach_for_User_Authentication_and_Identification_via_Keystroke_ Dynamics.
7. Брюхоміцький Ю.А. Метод навчання нейромережевих біометричних систем з урахуванням копіювання областей / Ю.А. Брюхоміцький, М.М. Казарин. – Перспективні інформаційні технології та інтелектуальні системи (Електронний журнал). – 2003. – № 3 (15). – С. 17-23.
8. Літвінчук І.С, Корчомний Р.О., Борисов І.В., Коршун Н.В. Розробка рекомендацій щодо мінімізації ризиків злому облікових даних на основі аналізу найпоширеніх методів злому. Кібербезпека:освіта, наука, техніка, № 4-12, 2021, С. 163-171.
9. Back propagation and stochastic gradientdes cent method / Amari S. // Neurocomputing – 1993. – № 5. – P. 185-96.
10. Stochastic gradient learning in neural networks / Bottou L. // Proceeding sof Neuro-Nımes –1991. – № 91 – 12 p.
11. Nielsen M.A. NeuralNetworksandDeepLearning. – DeterminationPress, 2015.
12. The Delta Rule [Electronicresource] /Russell I. – University of Hartford, 2012. – Mode of access:https://web.archive.org/web/20160304032228/http:/uhavax. hartford.edu/compsci/neural-networks-delta-rule.html. – Date of access: 12.06.2024.
13. Данилюк І.І., Карпінець В.В., Приймак А.В., Яремчук Ю.Є. Костюченко О.І. Метод ідентифікації користувача за клавіатурним почерком на основі нейромереж. стор. Реєстрація, зберігання і обробка даних, 2018, Т.20 № 2, С. 68-76.

REFERENCES

1. Goroshko M.P., Myklush S.I., Khomyuk P.G. Biometrics – Lviv: Publishing house «Kamula», 2004. – 236 p.
2. Ivanov A.I. Biometric identification of a person based on the dynamics of subconscious movements. – Penza: Publishing house Penza State University, 2000. – 188 p.
3. Chalaya L.E. User identification model based on keyboard handwriting. «Artificial Intelligence», No. 4. 2004, P. 811-817.
4. Bryan W.L., & Harter N. (1897). Studies in the physiology and psychology of the telegraphic language. Psychological Review, 4(1), 27-53. – https://doi.org/10.1037/h0073806

5. L.H. Shaffer Reading and Typing –https://www.researchgate.net/publication/233266615_Reading_and_Typing
6. Saket Mahesh wary, Soumyajit Ganguly, Vikram Pudi. DeepSecure: A FastandSimpleNeuralNetworkbasedapproachfor User Authentication and Identification via Key stroke Dynamics. https://www.researchgate.net/ publication/322952671_Deep_Secure_A_Fast_and_Simple_Neural_Network_based_approach_for_User_Authentication_and_Identification_via_Keystroke_Dynamics.
7. Bryuhomitsky, Yu.A. Method of training neural network biometric systems taking in to account copying of regions/Ya.A. Bryuhomitsky, M.M. Kazarin. – Promising in for mation technologies and intelligent systems (Electronic journal). – 2003. – No. 3 (15). – P. 17-23.
8. Litvinchuk I.S, Korchomny R.O., Borisov I.V., Korshun N.V. Development of recommendations for minimizing the risks of hacking credentials based on the analysis of the most common hacking methods. Cybersecurity: education, science, technology, No. 4-12, 2021, P. 163-171.
9. Back propagation and stochastic gradientdes cent method / Amari S. // Neurocomputing – 1993. – № 5. – p. 185-96.
10. Stochastic gradient learning in neural networks / Bottou L. // Proceeding sof Neuro-Nımes –1991. – № 91 – 12 p.
11. Nielsen M. A. Neural Networks and Deep Learning. – Determination Press, 2015.
12. TheDeltaRule [Electronicresource] / Russell I. – University of Hartford, 2012 – Mode of access: https://web.archive.org/web/20160304032228/http:/uha-vax. hartford.edu/compsci/neural-networks-delta-rule.html. – Dateofaccess: 12.06.2024.
13. Danylyuk I.I., Karpinets V.V., Priymak A.V., Yaremchuk Yu.E. Kostyuchenko O.I. Method of user identification by key board hand writing based on neural networks. pp. Registration, storageanddataprocessing, 2018, Vol. 20 No. 2, P. 68-76.

Стаття надійшла до редакції 12.12.2024

Посилання на статтю: Даус Ю.В., Самойлов С.В., Даус М.Є., Ларін Д.Г. Ідентифікація користувачів за клавіатурним почерком з використанням нейронних мереж // *Вісник Одеського національного морського університету: Зб. наук. праць*, 2025. № 1 (75). С. 212-227. DOI 10.47049/2226-1893-2025-1-212-227.

Article received 12.12.2024

Reference a journal artic: Daus Y., Samoilov S., Daus M., Larin D. Users identification by keyboard handwriting using neural networks // *Herald of the Odesa national maritime university: Coll. scient. works*, 2025. № 1 (75). P. 212-227. DOI 10.47049/2226-1893-2025-1-212-227.