Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)
Кафедра комп'ютерних наук
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр
(назва освітнього ступеня)

на тему: Система динамічної аутентифікації користувача на основі аналізу його роботи на клавіатурі комп'ютера

Виконав: студент IV курсу, групи СН-42
спеціальності 122 Комп’ютерні науки
(шифр і назва спеціальності)

_____ Погребняк Д.В.
(підпис) (прізвище та ініціали)

Керівник _____ Литвиненко Я.В.
(підпис) (прізвище та ініціали)

Нормоконтроль _____ Шимчук Г.В.
(підпис) (прізвище та ініціали)

Завідувач кафедри _____ Боднарчук І.О.
(підпис) (прізвище та ініціали)

Рецензент _____ _____
(підпис) (прізвище та ініціали)

Тернопіль - 2026

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних наук
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Боднарчук І.О.
(підпис) (прізвище та ініціали)
«_08_» ___06_____ 2026 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Бакалавр
(назва освітнього ступеня)

за спеціальністю 122 Комп'ютерні науки
(шифр і назва спеціальності)

Студенту Погребняку Данилу Віталійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Система динамічної аутентифікації користувача на основі аналізу його роботи на клавіатурі комп'ютера

Керівник роботи Литвиненко Ярослав Володимирович, д.т.н., проф.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «14» 06 2026 року № 4/9-239

2. Термін подання студентом завершеної роботи 21.06.2026 р.

3. Вихідні дані до роботи наукові літературні джерела

4. Зміст роботи (перелік питань, які потрібно розробити)
Вступ
1. Огляд предметної області.
2. Теоретична частина
3. Проектування та реалізація системи. Аналіз результатів
4. Безпека життєдіяльності, основи охорони праці
Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)
1. Титулка. 2.Мета, задачі дослідження. 3. Огляд предметної області.
4. Загальна архітектура системи безперервної автентифікації.
5. Етапи клавіатурної автентифікації. 6. Діаграми частотності.
7. Методи розпізнавання. 8, Використовувані технології.
9. Діаграми клас.ів та діяльності. 10. Діаграма використання.
11. Інтерфейс програми. 12. Результати досліджень за методами.
13. Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності, основи охорони праці			

7. Дата видачі завдання 26.01.2026 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	26.01.26	*Виконано*
2.	Підбір джерел про динамічну аутентифікацію		
	користувача на основі клавіатурного почерку	27.01 – 13.02.26	*Виконано*
3.	Опрацювання літературних джерел	14.02 – 18.04.26	*Виконано*
4.	Виконання дослідження щодо розробки програмного		
	забезпечення	19.04 – 22.05.26	*Виконано*
6.	Оформлення розділу «Огляд предметної області»	24.05 – 25.05.26	*Виконано*
7.	Оформлення розділу «Теоретична частина»	26.05 – 27.05.26	*Виконано*
8.	Оформлення розділу «Проектування та реалізація		
	системи. Аналіз результатів»	28.05 – 30.05.26	*Виконано*
9.	Виконання завдання до підрозділу «Безпека		
	життєдіяльності, основи хорони праці»	31.05 – 03.06.26	*Виконано*
10.	Оформлення кваліфікаційної роботи	20.05 – 03.06.26	*Виконано*
11.	Нормоконтроль	01.06 – 05.06.26	*Виконано*
12.	Перевірка на плагіат	04.06 – 10.06.26	*Виконано*
13.	Попередній захист кваліфікаційної роботи	01.06 – 18.06.26	*Виконано*
14.	Захист кваліфікаційної роботи	27.06.26	

Студент _____ Погребняк Д.В..
(підпис) (прізвище та ініціали)

Керівник роботи _____ Литвиненко Я.В.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Система динамічної аутентифікації користувача на основі аналізу його роботи на клавіатурі комп'ютера // Погребняк Данило Віталійович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп’ютерно-інформаційних систем та програмної інженерії, кафедра комп’ютерних наук, група СН–42 // Тернопіль, 2026 // С. – 54, рис. – 26, табл. – 3, слайдів – 13, бібліогр. – 37.

Ключові слова: клавіатурний почерк, інформаційна безпека, автентифікація, біометрія, динамічні характеристики клавіатурного почерку

Кваліфікаційна робота присвячена дослідженню параметрів динамічної аутентифікації користувача за його клавіатурним почерком.

У процесі дослідження проводилися роботи з вивчення методів автентифікації користувачів за клавіатурним почерком. У ході роботи було розглянуто існуючі підходи до вирішення задачі аутентифікації, а також запропоновано рішення щодо оптимізації існуючих алгоритмів.

Областю застосування є автентифікація користувачів за динамічними характеристиками клавіатурного почерку, визначення психофізіологічного стану користувача, прихований моніторинг користувачів корпоративної мережі з метою визначення підміни оператора.

В результаті дослідження було розроблено програмний застосунок для аутентифікації користувача на основі динамічних характеристик клавіатурного почерку. Було вивчено, а також протестовано з точки зору точності аутентифікації алгоритми розпізнавання користувача на основі динамічних характеристик клавіатурного почерку.

Об'єктом дослідження є система, що розробляється, для аутентифікації користувача на основі аналізу роботи на клавіатурі.

ANNOTATION

Dynamic User Authentication System Based on Keyboard Typing Behavior Analysis // Pohrebniak Danylo // Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Computer Science // Ternopil, 2026 // P. - 54, Fig. - 26, Table - 3, Slide - 13, References - 37.

Keywords: keyboard handwriting, information security, authentication, biometrics, dynamic characteristics of keyboard handwriting

Thesis deals with the study of the parameters of dynamic user authentication based on his keyboard handwriting.

In the course of the research, work was carried out to study the methods of user authentication based on keyboard handwriting. During the work, existing approaches to solving the authentication problem were considered, and solutions were proposed to optimize existing algorithms.

The scope of application is user authentication based on dynamic characteristics of keyboard handwriting, determination of the user's psychophysiological state, covert monitoring of corporate network users in order to determine operator substitution.

As a result of the research, a software application was developed for user authentication based on dynamic characteristics of keyboard handwriting. User recognition algorithms based on dynamic characteristics of keyboard handwriting were studied and tested in terms of authentication accuracy.

The object of the study is a user authentication system being developed based on an analysis of their keyboard activity.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ СКОРОЧЕНЬ І ТЕРМІНІВ

FAR (англ. False Accept Rate, коефіцієнт помилкових прийняття) – ймовірність того, що система помилково прийме неавторизовану особу.

FRR (англ. False Reject Rate, коефіцієнт помилкових відхилень) – ймовірність того, що система помилково відхилить уповноважену особу.

DET (англ. Detection Error Tradeoff, компроміс помилок виявлення) – графік коефіцієнтів помилок для бінарних систем класифікації, що відображає коефіцієнт помилкових відхилень порівняно з коефіцієнтом помилкових прийнять.

KM dataset (Killourhy & Maxion) – набір даних про динаміку клавіатурних натискань.

ROC (англ. Receiver Operating Characteristic, робоча характеристика приймача) – крива похибок, графік, що дозволяє оцінити якість бінарної класифікації, відображає співвідношення між часткою об'єктів від загальної кількості носіїв ознаки, правильно класифікованих до загальної кількості об'єктів, що не несуть ознаки, помилково класифікованих, як такі, що мають ознаку.

SVM (англ. Support vector machine, метод опорних векторів) – метод аналізу даних для класифікації та регресійного аналізу.

TAR (англ. True accept rate, Коефіцієнт істинних прийнять) – ймовірність того, що система правильно прийме уповноважену особу.

TRR (англ. True Reject rate, Коефіцієнт істинних відхилень) - ймовірність того, що система правильно відхилить уповноважену особу

Евклідова відстань – найкоротша відстань між двома точками в просторі, що визначається як довжина прямої лінії, що їх з'єднує.

Манхеттенська відстань (метрика міських кварталів) – сума абсолютних різниць координат двох точок, що визначає найкоротший шлях між ними на прямокутній сітці (як вулиці міста), де неможливий рух по діагоналі.

ПЗ – програмне забезпечення.

ЗМІСТ

ВСТУП

Актуальність теми. Ми проживаємо у час інформаційних технологій та інтернет сервісів, саме тому безпека і захист персональних даних є одним із найбільш актуальних та важливих завдань. За даними статистики, у 2025 році у світі було понад 5,2 мільярда користувачів інтернету, які використовували різні види автентифікації для доступу до своїх облікових записів, додатків та сервісів. Однак, більшість із них володіють певними недоліки, в т.ч. незручність використання, можливість крадіжки чи підробки, висока вартість обладнання та обслуговування.

У зв'язку з цим, за останні роки значна увага приділяється розробці та застосуванню так званих поведінкових методів аутентифікації, які ґрунтуються на аналізі індивідуальних особливостей поведінки користувача під час роботи з комп'ютером або мобільним пристроєм. Одним із таких методів є динамічна автентифікація користувача на основі аналізу його роботи на клавіатурі комп'ютера.

Даний метод полягає в тому, що при введенні тексту на клавіатурі комп'ютера користувач демонструє певний стиль набору тексту, який характеризується різними параметрами, такими як швидкість набору, тривалість утримання клавіш, інтервали натискання клавіш і т.д. Ці параметри формують унікальний профіль користувача, який може бути застосований з метою його ідентифікації та автентифікації..

Цей метод має ряд переваг у порівнянні з іншими видами аутентифікації, такі як зручність використання, низька вартість, висока точність та неможливість підробки. Однак, він також має ряд недоліків та загроз, пов'язаних з безпекою та захистом персональних даних користувача. Наприклад, кіберзлочинці можуть використовувати різні способи отримання доступу до профілю користувача або заміни його поведінкових характеристик. Такі способи можуть включати фішинг, кілоггер, імітацію стилю набору тексту і т.д. Це може призвести до витоку або компрометації даних користувача або порушення його

конфіденційності.

Тому, для підвищення безпеки та захисту персональних даних необхідно поєднувати поведінкову аутентифікацію та інші способи, такі як шифрування даних, оновлення ПЗ, резервне копіювання даних, використання надійних паролів та двофакторної аутентифікації.

Мета роботи – розвиток завдань теорії розпізнавання користувачів на основі динамічних характеристик клавіатурного почерку та створення на цій основі ПЗ системи аутентифікації користувача.

Для досягнення мети виділено ряд завдань:

– провести огляд існуючих методів аутентифікації користувача та порівняти їх із методом динамічної аутентифікації на основі роботи на клавіатурі;

– вивчити основні параметри та характеристики роботи користувача на клавіатурі комп'ютера та способи їх вимірювання та аналізу;

– підібрати набір даних для експерименту;

– розробити ПЗ обробки даних роботи користувача на клавіатурі комп'ютера;

– провести експериментальне тестування методу динамічної аутентифікації користувача на основі його роботи на клавіатурі комп'ютера;

– оцінити точність та надійність методу динамічної аутентифікації користувача на основі роботи на клавіатурі комп'ютера;

– виявити можливі загрози та ризики для безпеки та захисту персональних даних користувача при використанні даного методу аутентифікації;

– запропонувати рекомендації щодо покращення методу динамічної аутентифікації користувача на основі роботи на клавіатурі комп'ютера.

Практичне значення одержаних результатів Розробка може бути успішно застосована для визначення психофізіологічного стану користувача, прихованого моніторингу користувачів корпоративної мережі з метою визначення підміни оператора.

РОЗДІЛ 1. ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Питання клавіатурної автентифікації та ідентифікації

Питання клавіатурної аутентифікації та ідентифікації відносяться до галузі біометрії поведінки, яка використовує манеру та ритм набору тексту на клавіатурі для визначення особи користувача. Аутентифікація клавіатури може бути заснована на різних параметрах, таких як швидкість набору, тривалість натискання клавіш, інтервали між клавішами і помилки.

Це показано на рисунку 1.1.

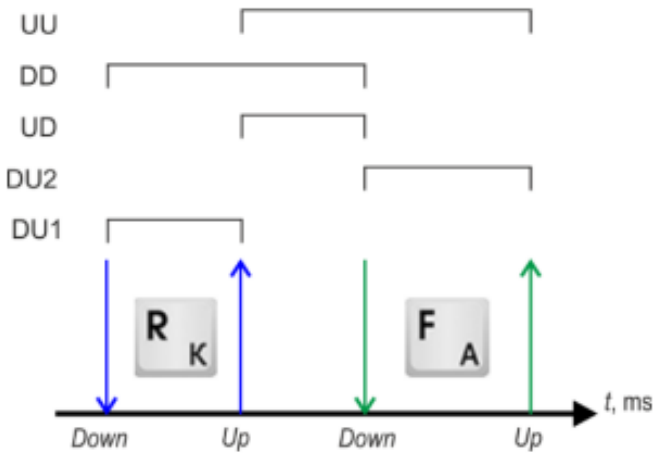

Рисунок 1.1 – Показники клавіатурного почерку

Де, DU - час утримання клавіші, UD - пауза між натисканнями, UU або DD інтервал між натисканням або відпусканням однієї клавіші та натисканням або відпусканням наступної клавіші відповідно.

Один із можливих способів захисту системи від несанкціонованого доступу – використовувати двоетапний процес верифікації:

– первинна ідентифікація особи;

– динамічна автентифікація особи.

Однак кожна людина має індивідуальний ритм набору тексту. З огляду на цю особливість біометрична система розпізнавання особи може використовувати

клавіатурний почерк. Для наочності на рисунку 1.2 представлена швидкість набору тексту 8 користувачів з датасета KM.

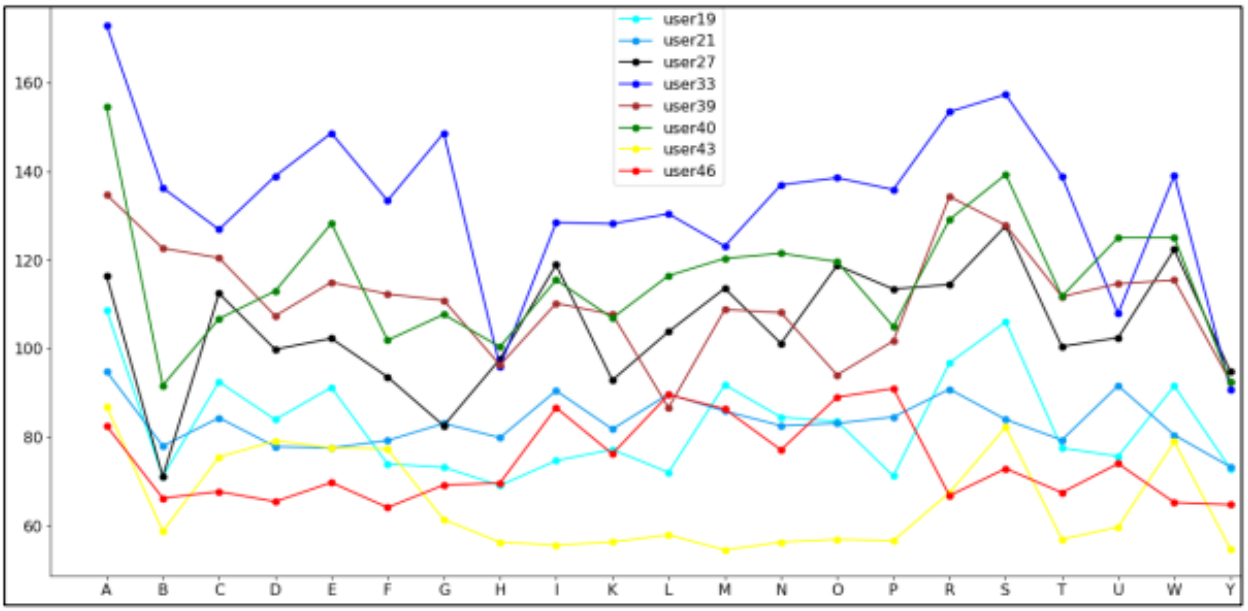

Рисунок 1.2 – Клавіатурні шаблони користувачів

Візуальний аналіз демонструє певні розбіжності між часом натискання літер на клавіатурі. Цей розкид демонструє унікальність ритму натискання клавіш кожного користувача. З технічної точки зору, чим більше клавіш натискає користувач, тим більш точно алгоритм може зрозуміти і відтворити клавіатурний шаблон користувача. Унікальність клавіатурного шаблону підвищує точність системи розпізнавання.

Клавіатурна ідентифікація - це спосіб визначення користувача серед багатьох інших потенційних користувачів на основі того, як він друкує на клавіатурі. Водночас беруться до уваги різні фактори, котрі чинять вплив на стиль набору, наприклад швидкість друкування, ритм натискання клавіш, тривалість утримання клавіші та інтервали між натисканнями. Будь-яка особа має свій неповторний клавіатурний почерк, який відрізняє його від інших і може бути використаний як біометрична ознака для його ідентифікації.

Для реалізації клавіатурної аутентифікації та ідентифікації можуть використовуватись різні техніки, від статистичних методів до підходів штучного інтелекту, зокрема нейронних мереж. Перевагою клавіатурної автентифікації є

те, що вона не потребує спеціального обладнання, такого як сканери відбитків пальців або обличчя, а може працювати з будь-якою стандартною клавіатурою. Однак клавіатурна автентифікація також має свої недоліки, такі як вплив фізичного чи емоційного стану користувача, зміна стилю набору в часі та можливість підробки чи імітації.

Клавіатурна автентифікація та ідентифікація можуть застосовуватися для різних цілей, таких як підвищення безпеки входу в систему, контроль доступу до конфіденційної інформації, моніторинг поведінки користувачів у мережі або виявлення вторгнень. Клавіатурна ідентифікація також може бути властиво застосована для психологічного аналізу особистості за її способом друкування.

1.2 Методи аутентифікації

Існує множина методів автентифікації користувача. Методи можна умовно розділити на чотири групи, що показано на рисунку 1.3:

– на основі знання унікальної інформації;
– на основі володіння унікальним предметом;
– на основі біометрії;
– інші ознаки.

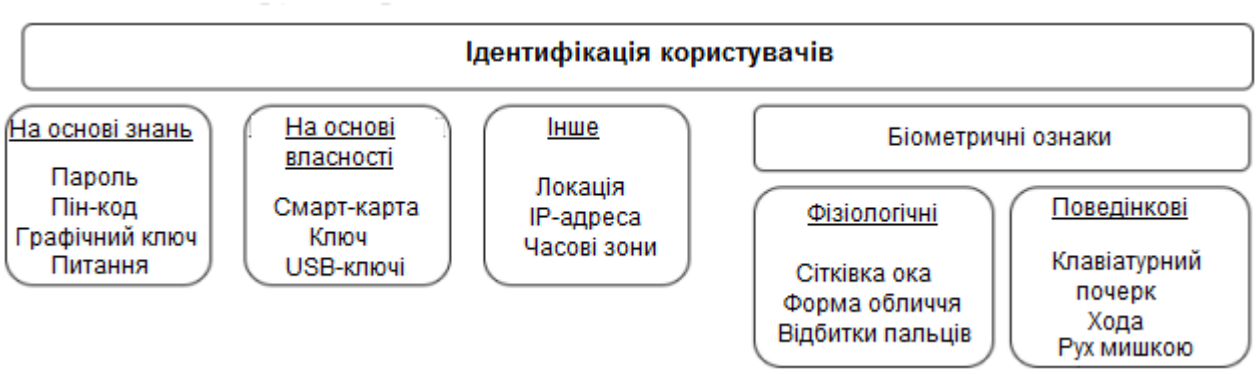

Рисунок 1.3 – Методи аутентифікації

Аутентифікація з урахуванням знань особистої інформації (імені, пароля, секретного питання). Ці способи прості у використанні та недорогі. Однак вони

забезпечують низький рівень безпеки [1].

Аутентифікація на основі володіння особистими об'єктами користувача, такими як смарт-картки та ключі. Це найменш популярний метод електронної аутентифікації, тому що особисті предмети можуть бути вкрадені або скопійовані [1].

Інші ознаки аутентифікації засновані на місцезнаходження, часовій зоні, IP-адресі та ін [1]

Біометричні ознаки аутентифікації поділяються на фізіологічні та поведінкові. Фізіологічні характеристики включають форму обличчя, райдужну сітківку ока, відбитки пальців тощо. Поведінкові - голос, хода, підпис, рух миші, рукописний та клавіатурний почерк. Аутентифікація на основі фізіологічних ознак є точною, але технічні пристрої розпізнавання досить дорогі [1]. Клавіатурний почерк відноситься до біометрії поведінки.

Методи клавіатурного розпізнавання – це способи ідентифікації та аутентифікації користувача за його індивідуальним стилем набору тексту на клавіатурі комп'ютера. Ці методи належать до поведінкової біометрії і можуть використовуватися для статичної або динамічної (безперервної) автентифікації користувача.

Частота використання методів клавіатурного розпізнавання залежить від різних факторів, таких як тип тексту (структурований або вільний), мова тексту (англійська або інша), мета аутентифікації (первинна або вторинна), алгоритм класифікації (заснований на метричних відстанях, статичних методах або машинному навчанні) і т.д.

Згідно з [3] існує множина різних методів клавіатурного розпізнавання, котрі можливо рподілити на три визначальні групи щодо розпізнавання образів:

– оцінка метричних відстаней;
– статистичні методи;
– методи машинного навчання.

Відносна частота використання різних методів клавіатурного розпізнавання представлена на рисунку 1.4 у порядку спадання [3].

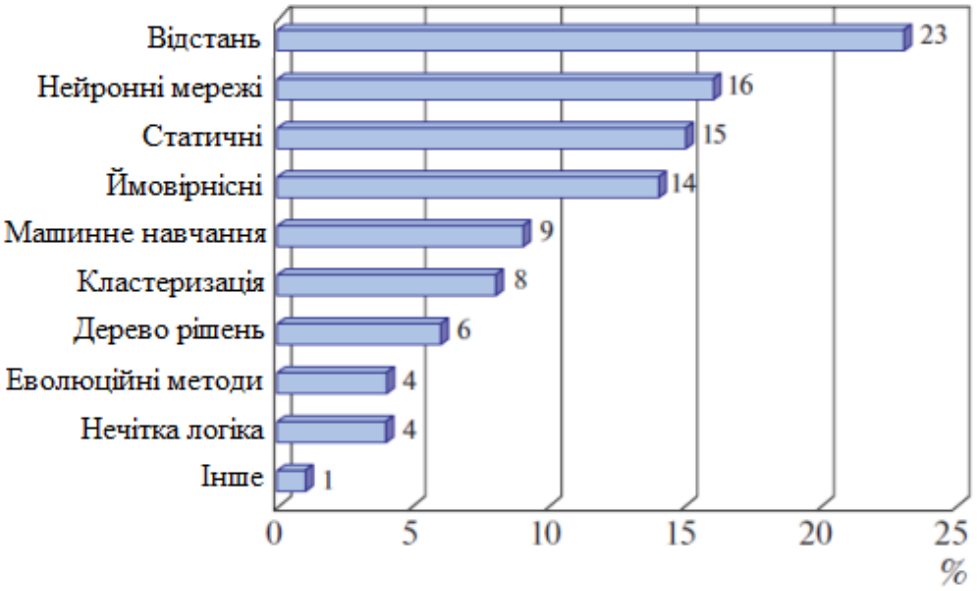

Рисунок 1.4 – Відносна частота використання методів клавіатурного розпізнавання

Як видно з ис.1.4, методами клавіатурного розпізнавання, що найбільш часто використовуються, є ті, які засновані на оцінці метричних відстаней між поточним і еталонним профілями користувача. Ці методи прості у реалізації і вимагають складних обчислень. Однак, вони також мають свої недоліки, такі як низька точність, висока чутливість до змін у поведінці користувача та необхідність вибору порогового значення для ухвалення рішення.

Статичні методи клавіатурного розпізнавання використовують різні статичні моделі для опису розподілу параметрів клавіатурного розпізнавання та обчислення ймовірності приналежності поточного профілю до еталонного. Ці методи більш точні та стійкі до шумів та варіацій даних, але вони також складніші в реалізації та вимагають більшого обсягу даних для навчання моделі.

Методи машинного навчання клавіатурного розпізнавання застосовують різні алгоритми класифікації, такі як нейронні мережі, SVM, дерево рішення і т.д., для навчання моделі класифікатора на основі наявних даних та прогнозування належності поточного профілю до одного із заздалегідь визначених класів (користувачів). Ці методи можуть досягати високої точності та адаптивності до нових даних, але вони також вимагають великого обсягу даних для навчання, а також вибору оптимальних параметрів для кожного

алгоритму.

В цілому, можна сказати, що частота використання методів клавіатурного розпізнавання визначається різними факторами і залежить від конкретного завдання та умов автентифікації користувача.

1.3 Режими аутентифікації

Найбільш обґрунтований для системи розпізнавання та комфортний для користувача спосіб аутентифікації особи – це постійний та прихований моніторинг динаміки його роботи.

Динамічні характеристики клавіатурного почерку складніші для розпізнавання, ніж фізіологічні. Однак цей факт компенсується більш трудомістким процесом підміни користувача, що сприятливо позначається на рівні захищеності системи. Крім того, динамічні характеристики можуть відображати не тільки особу користувача, але і його емоційний стан, що може бути корисним для аналізу його поведінки та мотивації.

Існує два види аутентифікації: статична та динамічна.

При статичній аутентифікації користувачеві системи надається певний текст фіксованої довжини, який має ввести для підтвердження своєї особи. Цей текст може бути паролем, PIN-кодом або іншою комбінацією символів. Перевагою цього є простота реалізації та перевірки. Недоліком є можливість крадіжки чи забування тексту, і навіть необхідність постійного запам'ятовування нових текстів за її зміни.

Динамічна автентифікація є складнішим процесом моніторингу натискань клавіш користувачем. За певної заданої умови це може бути часте використання службових символів, що не властиво користувачеві, або занадто повільний друк, система може обмежити доступ до облікового запису і попросить повторно пройти процес ідентифікації. Перевагою цього способу є можливість безперервної автентифікації користувача протягом усієї сесії роботи з системою, а також відсутність необхідності запам'ятовувати спеціальні тексти. Недоліком є складність реалізації та налаштування параметрів розпізнавання.

Обидва способи можуть доповнювати один одного залежно від поставленої організацією завдання. Наприклад, статична автентифікація може бути як перший рівень захисту. Динамічна автентифікація виступатиме як друга. Таким чином, можна підвищити надійність та безпеку системи розпізнавання особи.

РОЗДІЛ 2. ТЕОРЕТИЧНА ЧАСТИНА

2.1 Життєвий цикл автентифікації

Життєвий цикл клавіатурного розпізнавання – це послідовність етапів, які необхідно виконати для ідентифікації чи аутентифікації користувача, як наведено на рисунку 2.1.

Рисунок 2.1 – Життєвий цикл автентифікації

Перший етап життєвого циклу розпізнавання є збиранням даних про натискання клавіш користувача - це стартовий і важливий етап життєвого циклу клавіатурного розпізнавання. Цей етап полягає в тому, що на комп'ютері користувача встановлюється спеціальна програма або пристрій, який фіксує та зберігає натискання клавіш, які користувач робить під час введення різних текстів. Ці тексти можуть бути структурованими чи вільними, різними мовами тощо.

Існує два основних типи збору даних про натискання клавіш користувача:

– апаратний;
– програмний.

Апаратний спосіб даних про натискання клавіш користувача здійснюється

за допомогою спеціальних пристроїв, що під’єднуються між клавіатурою та комп'ютером або вбудовуються у саму клавіатуру. Ці пристрої записують всі дані, які передаються від клавіатури до комп'ютера, у внутрішню пам'ять або зовнішній носій. Прикладом такого пристрою може бути так званий keyloggerr, який показано на рисунку 2.2.

Рисунок 2.2 – Апаратний keyloggerr

Програмний збір даних про натискання клавіш користувача виконується із застосуванням спеціальних програм, котрі інсталюються на комп'ютер користувача та перехоплюють усі дані, що надходять від клавіатури до ОС або програм. Ці програми можуть бути різними за рівнем доступу та складністю реалізації. Наприклад, існують програми, які працюють у режимі користувача і використовують API-функції ОС для отримання даних про натискання клавіш [6]. Також існують програми, які працюють у ядрі ОС та мають прямий доступ до драйверів клавіатури [6]. Яскравою ілюстрацією такого ПЗ може бути WhatPulse, інтерфейс котрого відображено на рисунку 2.3.

a

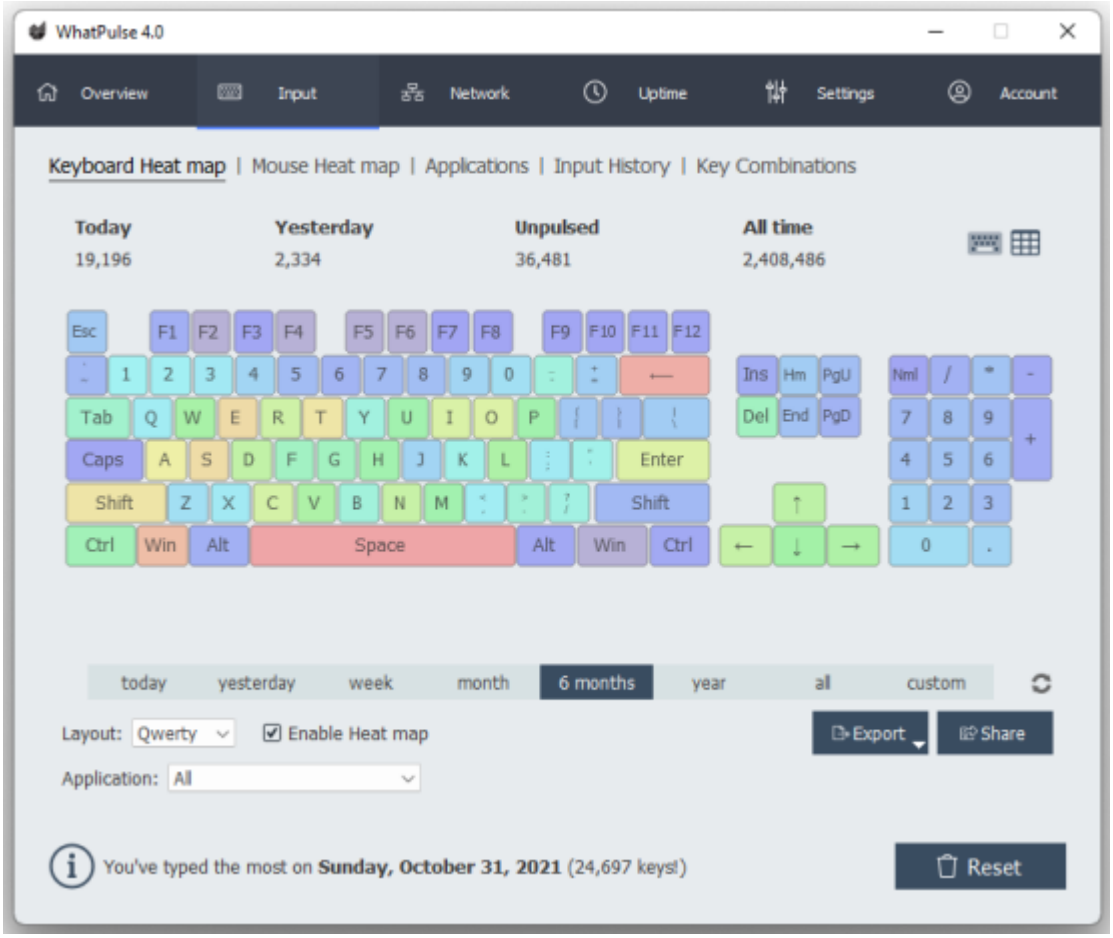

Рисунок 2.3 – Інтерфейс програмного keyloggerr WhatPulse

Збір даних про натискання клавіш користувача має свої переваги та недоліки. З одного боку, він дозволяє отримати великий обсяг інформації про стиль набору тексту користувача, який може бути використаний для його ідентифікації та автентифікації. З іншого боку, він може порушувати приватність та безпеку користувача, якщо ці дані потраплять до рук зловмисників або будуть використані без згоди користувача.

Видобуток ознак клавіатурного почерку - це другий етап життєвого циклу клавіатурного розпізнавання. Цей етап полягає в тому, що із зібраних даних про натискання клавіш користувача виділяються ті характеристики, які найкраще відображають його індивідуальний стиль набору тексту та дозволяють відрізнити його від інших користувачів. Ці властивості називаються ознаками клавіатурного почерку.

Існує багато різних видів ознак клавіатурного почерку, які можуть бути вилучені з даних про натискання клавіш користувача. Наприклад, існують:

– статистичні ознаки, які описують розподіл і варіабельність різних параметрів натискання клавіш, таких як швидкість набору, тривалість утримання клавіші, інтервали між натисканнями клавіш і т.д.;

– символьні ознаки, які описують частоту і послідовність використання різних символів на клавіатурі, такі як букви, цифри, розділові знаки і т.д.

Етап розпізнавання – це третій етап життєвого циклу клавіатурного розпізнавання. Цей етап полягає в тому, що одержані ознаки клавіатурного почерку користувача порівнюються з еталонними ознаками або профілями інших користувачів за допомогою алгоритму класифікації, який визначає, якого класу або категорії належить користувач. Цей етап може бути втілений при допомозі різноманітних алгоритмів та методів, таких як:

– методи, які засновані на оцінці близькості, що ґрунтуються на оцінці метричних відстаней між поточним та еталонним профілями користувача;

– SVM, який знаходить оптимальну гіперплощину, яка поділяє одержані ознаки клавіатурного почерку користувачів;

– нейронні мережі, які навчаються на ознаках клавіатурного почерку користувача і видають вектор ймовірностей приналежності користувача до різних класів.

Останні дослідження, присвячені розпізнаванню користувачів за клавіатурним почерком, дозволили узагальнити дані про ефективність безперервної аутентифікації. Узагальнені дані наведені в таблиці 2.1. Дані отримані та адаптовані з оглядових статей [13, 19] та адаптовані з оглядових статей [11,12, 14, 16, 18-25].

Етап прийняття рішення – четвертий та заключний етап життєвого циклу клавіатурного розпізнавання Цей етап полягає в тому, що на основі результату етапу розпізнавання, тобто вектора ймовірностей приналежності користувача, приймається остаточне рішення про те, чи користувач є справжнім або підробленим, емоційним або спокійним і т.д.

Цей етап може бути втілений із застосуванням порогового рішення (threshold decision), яке порівнює ймовірність приналежності користувача до

певного класу із заданим порогом і приймає рішення залежно від того, чи більша чи менша ймовірність порога.

Таблиця 2.1 – Дослідження динамічної ідентифікації

Рік	**Посилання, автор**	**Параметр КП**	**Метод**	**Ефективність**
2005	[17] Gunetti	FT	Відстань (R та A)	FAR 0.005%, FRR 5%
2010	[23] Shimshon		Кластеризація	FAR 3,47%, FRR 0%
2011	[24] Messerman		Статистичні, відстань	FAR 2.02%, FRR 1.84%
2011	[28] Solami		Кластеризація	Точність 100%
2013	[19] Alsultan	діграф	Змішана (Fusion)	FAR 21%, FRR 17%
2014	[26] Ahmed	діграф	Нейронні мережі	FAR 0.015% FRR 4.82%
2015	[30] Antal	DT, FT	Статистичні SVM Нейронні мережі Дерево рішень	93.04% Точність
2014	[31] Locklear		Статистичні	EER 4,55 - 13,37%
2015	[32] Kang	DT, FT	Кластеризація, Відстань	3.8% EER
2015	[33] Matsubara	діграф, DT	Відстань	99% Точність
2016	[15] Morales	діграф, n- граф	k-NN найближчий сусід, Відстань	90% Точність
2017	[23] Alsultan	діграф, DT	SVM	0.169 FAR, 0.423 FRR
2017	[20] Mondal Bours	діграф, DT	Відстань	182 keystrokes

Продовження таблиці 2.1

2017	[27] Goodkind	Contextual features	Наївний Байєс	82.2% Точність
2017	[22] Ali		k-NN метод	EER 3,7%
2021	[25] Chang	DT, FT	CNN-GRU	Точність 99% EER 0,0690

2.2 Прихований моніторинг та динамічне розпізнавання

Прихований моніторинг за клавіатурним начерком — це процес аналізу часових параметрів натискання клавіш користувачем для його динамічного розпізнавання (автентифікації та ідентифікації). Динамічне розпізнавання використовує різні методи для порівняння поточного почерку користувача з його еталоном для видачі рішення про аутентифікацію. Архітектура динамічної (безперервної) аутентифікації (рисунок 2.4) включає три підсистеми:

– реєстрація;
– аутентифікація;
– адаптація.

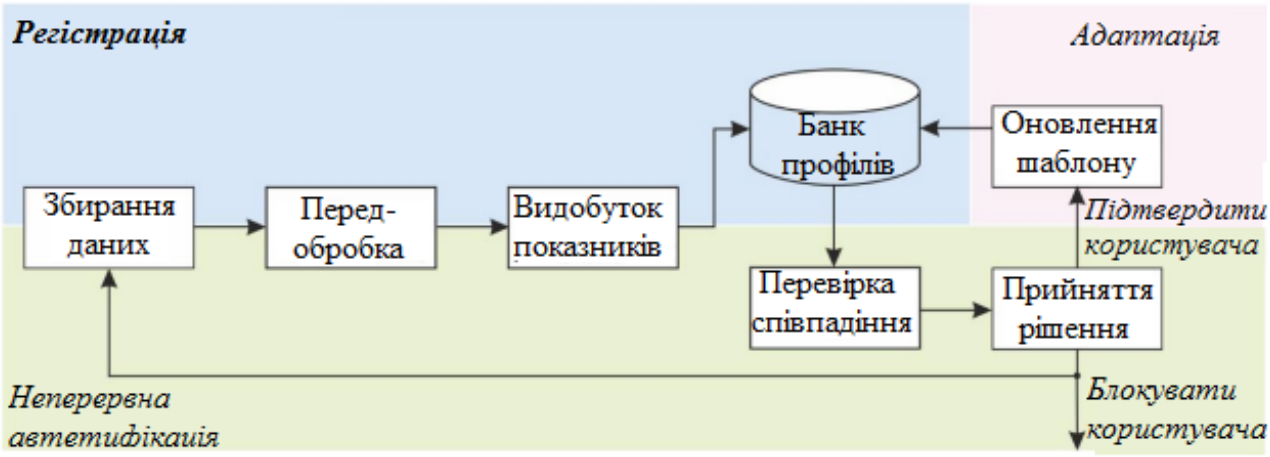

Рисунок 2.4 – Архітектура системи безперервної автентифікації

Під час процесу реєстрації відбувається збір даних про клавіатурні натискання, далі проводиться передобробка та вилучення показників. Усі ці дії

ведуть до поповнення банку профілів (шаблонів). Далі під час процесу безперервної автентифікації відбувається перевірка збігу шаблонів та прийняття рішення про допуск/відмову у допуску користувачеві в систему. Оновлення шаблону користувача відбувається за умови підтвердження його особи з подальшим оновленням даних у банку профілів [1].

Для динамічного розпізнавання користувача необхідні такі дані:

– часові характеристики набору тексту (швидкість введення, час утримання клавіші, інтервали між натисканнями тощо);

– еталонні зразки клавіатурного почерку кожного користувача системи, записані раніше;

– алгоритми розпізнавання, які порівнюють введений текст та його динаміку з еталонними зразками та видають рішення про автентифікацію користувача.

Для проведення досліджень з динамічної аутентифікації по клавіатурному почерку необхідні спеціальні набори даних, що містять інформацію про швидкість та ритм набору тексту різними користувачами. Такі набори даних можна отримати двома способами: зібрати їх локально або завантажити готовий датасет.

2.3 Оцінки ефективності аутентифікації

Для оцінки ефективності системи аутентифікації за клавіатурним начерком використовуються різні показники частоти помилок.

Один з таких показників - False Rejection Rate (FRR), який означає оцінку помилкового відхилення або помилку I роду. FRR визначає процент випадків, коли законний користувач помилково відхиляється:

$$FRR = \frac{FR}{TA + FA + TR + FR} \tag{2.1}$$

Другий показник - це False Acceptance Rate (FAR), який означає помилку хибного ухвалення або помилку II роду. FAR визначає відсоток випадків ухвалення нелегальних користувачів:

$$FAR = \frac{FA}{TA + FA + TR + FR} \tag{2.2}$$

У формулах (2.1) і (2.2) прийнято такі позначення:

– True Accept (TA) - правильний допуск до системи законного користувача;

– True Reject (TR) - вірна відмова у доступі незаконному користувачеві;

– False Accept (FA) - хибний допуск незаконного користувача;

– False Reject (FR) - помилкова відмова у доступі законному користувачеві.

Сума перерахованих вище показників становить загальну кількість спроб.

Гіпотетично помилки FRR і FAR варіюються залежно від рівня чутливості алгоритму (порогового значення) і мають протилежний характер: коли помилка зменшується, інша збільшується.

Більш високі значення FAR, зазвичай, є кращими у системах, де безпека не має першорядної важливості, тоді як більш високі значення FRR є переважними у застосунках з високим ступенем захисту.

Ще один показник - це Equal Error Rate (EER), який є значенням помилки, коли FAR та FRR набувають рівних значень і не залежить від рівня чутливості. ERR використовується визначення загальної точності системи розпізнавання.

Перелічені показники ефективності (рисунок 2.5) вимагають додаткового аналізу під час використання завдань аутентифікації та ідентифікації користувачів. Прийняття рішення не може базуватися лише на показниках FAR та FRR. Бажано мати узагальнений просторовий показник, доповнений пороговим значенням (чутливістю) та граничними значеннями показників.

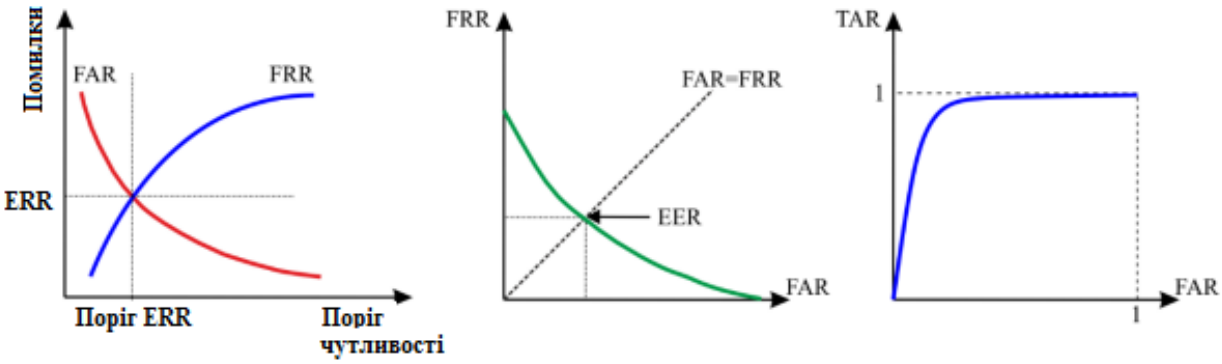

Рисунок 2.5 – Показники ефективності клавіатурної автентифікації

2.4 Етапи клавіатурної автентифікації

Етапи клавіатурної аутентифікації - це процес автентифікації користувача за його способом набору тексту на клавіатурі. Це один із видів біометричної аутентифікації, яка заснована на вимірі унікальних фізичних чи поведінкових характеристик людини.

Згідно з деякими джерелами [2, 3], основні етапи клавіатурної аутентифікації такі:

– збір інформації про клавіатурні натискання користувача під час введення тексту;

– видобування ознак клавіатурного почерку, таких як тривалість натискання і відпускання клавіш, інтервали між натисканнями і т.д.;

– порівняння ознак із заздалегідь збереженим шаблоном або еталоном для цього користувача;

– ухвалення рішення про легітимність користувача на основі обраного алгоритму та порогового значення.

Деякі системи також можуть використовувати додаткові фактори, такі як частота використання літер у текстах [3], географічне розташування [2] або контекст введення, для підвищення точності та надійності розпізнавання.

2.4.1 Безперервний збір даних про клавіатурні натискання

Для досягнення найкращих результатів при зборі даних необхідно

проводити його у безперервному режимі. Безперервний збір даних про клавіатурні натискання - це процес відстеження динаміки набору тексту користувача в режимі реального часу для цілей автентифікації. Це означає, що система не тільки перевіряє користувача при вході в систему, але й продовжує контролювати його поведінку на клавіатурі упродовж усієї сесії роботи. Таким чином, система може виявити та заблокувати несанкціонований доступ до комп'ютера або програми, якщо виявить аномалії у клавіатурному почерку користувача.

Для безперервного збору даних про клавіатурні натискання необхідно використовувати спеціальне ПЗ, яке може перехоплювати та аналізувати сигнали від клавіатури. Таке ПЗ може бути інстальоване на ПК користувача чи на віддаленому сервері. У будь-якому випадку, ПЗ має бути здатним витягувати ознаки клавіатурного почерку з будь-якого введеного тексту, незалежно від його змісту та мови.

Безперервний збір даних про клавіатурні натискання має ряд переваг перед статичною автентифікацією за паролем або одноразовим кодом. По-перше, він підвищує рівень безпеки, так як утруднює підбір чи крадіжку пароля. По-друге, він підвищує зручність використання, тому що не вимагає від користувача запам'ятовувати або вводити додаткові дані для автентифікації. По-третє, він підвищує ефективність роботи, оскільки не переривает процес введення тексту і відволікає користувача від основного завдання [2].

2.4.2 Актуалізація динамічних наборів даних

Це процес оновлення та корекції еталонних шаблонів користувачів на основі їхньої поточної поведінки на клавіатурі. Це необхідно для підвищення точності та надійності системи аутентифікації за клавіатурним почерком, яка відстежує динаміку натискання клавіш у режимі реального часу [3].

Актуалізація динамічних наборів дозволяє враховувати зміни у клавіатурному почерку, викликані різними чинниками, як-от втома, стрес, емоції, здоров'я тощо. [4]. Для актуалізації динамічних наборів даних необхідно використовувати спеціальні алгоритми, які можуть визначати ступінь подібності

між поточним та еталонним зразком клавіатурного почерку та вносити необхідні корективи до шаблону [3, 4].

2.5 Дані для експерименту

Даними є набори даних про клавіатурні натискання. Більшість досліджень у науковому середовищі використовує готові набори даних.

У цьому дослідженні буде використано дані, зібрані для статті [10].

Набір даних містить CSV -файли з характеристиками часу (час утримання та паузи) кожного натискання клавіш у наборі довільного тексту. Двадцять піддослідних виконали зіставні завдання транскрипцію і вільний твір; реалізовано два класифікатори динаміки натискань клавіш; Кожен класифікатор оцінювався з використанням зразків як вільного складу, і зразків транскрипції [10]. Натискання клавіш було згруповано для кожного користувача незалежно від своїх сеансів. Таким чином, для кожного набору даних, користувача, завдання, коду віртуальної клавіші та функції було створено профіль, що складається з набору мчасових значень. Імена файлів вказані з використанням наступної схеми: DATASET-TASK-USER-FEATURE-VK та організовані в папки відповідно до їх набору даних та завдання. Оскільки кількість файлів перевищує сто тисяч, вони упаковуються в файл DISTRIBUTIONS.zip. Для ілюстрації угоди про ім'я додано п'ять файлів, які також входять до комплекту постачання. Наприклад, KM-transcribed-USERs019-FT-VK32.csv містить часові спостереження за часом польоту (FT) клавіші пробілу (VK32, код віртуальної клавіші 32) при натисканні користувачем s019 у наборі даних KM, коли він виконує завдання транскрипції.

У процесі збору даних було отримано дані про клавіатурні натискання по одній сесії для кожного з двадцяти користувачів. Оскільки в дистрибутиві не передбачено поділу на сесії для одного користувача, ми припустили, що кожен користувач має тільки одну сесію. Зважаючи на малу кількість даних, було прийнято рішення змоделювати сесії отриманих користувачів.

Для моделювання сесій було використано десктопну програму. При цьому

слід враховувати, що отримані при моделюванні вибірки з однієї генеральної сукупності з близьким математичним сподіванням та середнім квадратичним відхиленням.

Моделювання сесій реальних користувачів було з розрахунку з 10 сесій на користувача, кожна сесія містить у собі 1000 символів. Діаграма розкиду значень згенерованих сесій користувача user19 зображено на рисунку 2.6.

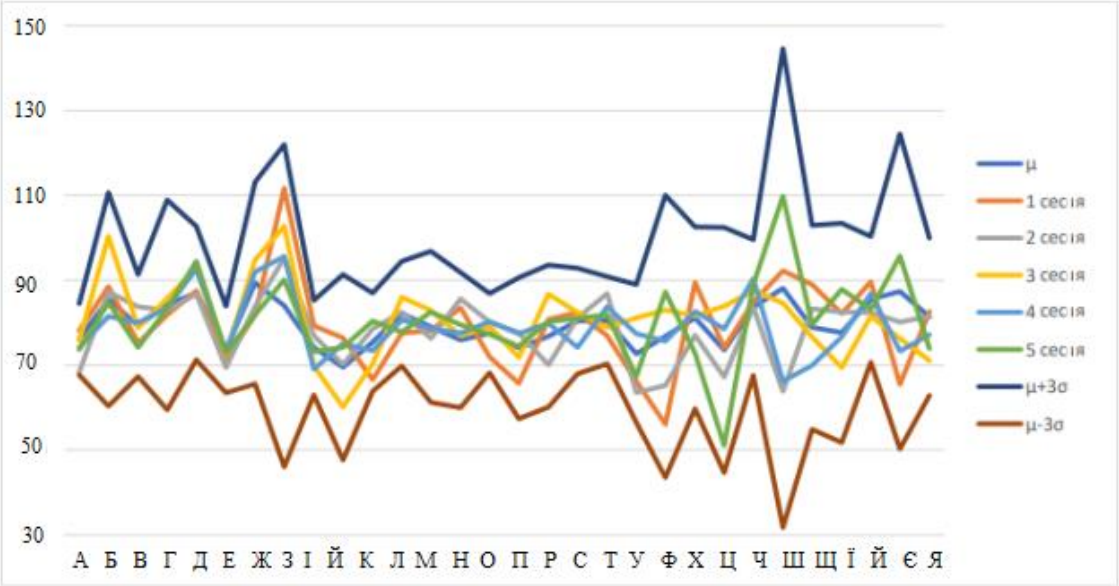

Рисунок 2.6 – Згенеровані сесії

Зібрані та змодельовані дані зберігаються у форматі txt та зображені на рисунку 2.7.

Рисунок 2.7 – Дані формату txt

2.6 Формування часового показника

Формування часового показника клавіатурного почерку — це процес вилучення та аналізу характеристик, пов'язаних із тривалістю та інтервалами натискання та відпускання клавіш під час введення тексту. Ці характеристики відображають індивідуальний стиль та ритм набору користувача, який може бути використаний для його ідентифікації або автентифікації [3]. Формування часового показника клавіатурного почерку потребує спеціального ПЗ, яке може перехоплювати сигнали від клавіатури та вимірювати час між ними. Потім ці дані піддаються статистичного аналізу виділення ознак, які можуть відрізняти одного користувача від іншого [5]. Формування часового показника клавіатурного почерку має низку переваг перед іншими біометричними методами, такими як відбитки пальців, розпізнавання обличчя чи голосу. По-перше, він не вимагає додаткового обладнання, окрім звичайної клавіатури. По-друге, він не порушує приватність користувача, оскільки не збирає особисті дані. По-третє, він є ненав'язливим і непомітним користувача, оскільки вимагає від нього спеціальних дій чи поз [5].

У роботі для підвищення якості часових характеристик, була використана частотність букв англійського алфавіту. Використання частотності дозволяє оптимізувати алгоритм через нормування значення окремої літери в шаблоні користувача.

Частотності англійського алфавіту зображено на рисунках 2.8 та 2.9.

Як видно з рис. 2.9, деякі літери використовуються користувачами досить часто, деякі дуже рідко - менше 2%. У ході дослідження було прийнято рішення не використовувати літери чия частотність нижче 1%, цими літерами є J, Q, V, Z.

Рисунок 2.8 – Діаграма частотності букв англійського алфавіту

Рисунок 2.9 – Діаграма частотності букв англійського алфавіту, відсортована за частотою використання

2.7 Створення шаблонів користувачів

Створення шаблонів клавіатурного почерку користувачів – це процес формування еталонних зразків, які відбивають індивідуальні особливості набору тексту клавіатурі. Ці зразки можуть бути застосовані з метою ідентифікації чи аутентифікації користувачів за клавіатурним почерком.. Створення шаблонів клавіатурного почерку користувачів вимагає збору та обробки даних про часові характеристики натискання та відпускання клавіш, а також частотність

використання літер у текстах [5].

Створення шаблонів клавіатури користувачів має низку переваг щодо інших методів автентифікації, таких як паролі, відбитки пальців або розпізнавання обличчя. По-перше, воно посилює захист паролів, оскільки додає додатковий фактор автентифікації. По-друге, воно спрощує процес аутентифікації, оскільки не вимагає від користувача запам'ятовування складних комбінацій. По-третє, воно підвищує зручність використання системи, оскільки не вимагає від користувача введення пароля щоразу за доступом до ресурсів [3].

2.8 Алгоритми та методи розпізнавання

Одні й самі методи застосовується для динамічної і статичної аутентифікації. Поділ методів розпізнавання на класи досить умовний, але можна виділити методи машинного навчання, статистичні методи та засновані на оцінці метричних відстаней.

У цій роботі використовуються такі методи.

1. Метод на основі Евклідової відстані. Ця відстань обчислюється за теоремою Піфагора:

$$\rho(x,y) = \sqrt{\sum_{i}^{n}(x_i - y_i)^2} \tag{2.3}$$

2. Метод на основі міських кварталів (Манхеттенська відстань). Це відстані між двома точками у К-вимірному векторному просторі, яка дорівнює сумі модулів різниць їх координат.

$$\rho(x,y) = \sum_{i}^{n}|x_i - y_i| \tag{2.4}$$

3. SVM – це метод машинного навчання, який використовується для

задач класифікації та регресії. Основна ідея методу - знайти таку гіперплощину у просторі ознак, яка максимально поділяє об'єкти різних класів (рис. 2.10).

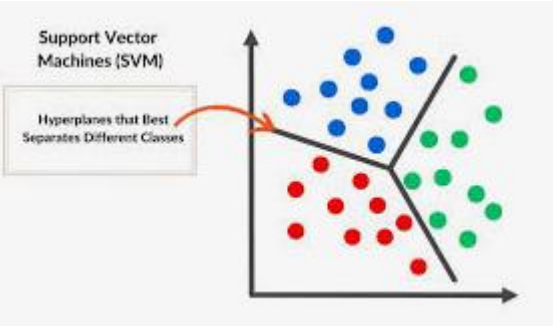

Рисунок 2.10 – Графічне подання SVM

І тому метод спирається деякі об'єкти, звані опорними векторами, які лежать найближче до межі поділу. Для клавіатурного розпізнавання SVM може використовуватися для класифікації користувачів за натисканням клавіш на клавіатурі. Для цього необхідно зібрати навчальну вибірку, що складається з набору ознак, що описує кожне натискання користувачем. Потім на основі цієї вибірки можна навчити SVM-класифікатор, який класифікуватиме користувачів за їх клавіатурним почерком.

РОЗДІЛ 3. ПРОЕКТУВАННЯ ТА РЕАЛІЗАЦІЯ СИСТЕМИ

3.1 Функціональні можливості

Функціональними можливостями системи є здатність перевіряти справжність користувача за його індивідуальним стилем набору тексту .

Для цього система аналізує різні параметри клавіатурного почерку, за допомогою яких формує вектор біометричних характеристик, які порівнюються з еталонними зразками користувачів і приймають рішення в залежності від обраного алгоритму розпізнавання.

3.2 Аналіз та вибір інструментів

Для розробки веб-програми були використані наступні технології та інструменти.

Мови програмування: для клієнтської частини веб-програми були використані HTML і CSS, а для серверної частини Python. HTML визначає структуру та зміст сторінки, CSS задає стилі та оформлення елементів. Python - це високорівнева мова програмування, яка дозволяє писати чистий та ефективний код.

Фреймворки та бібліотеки: для спрощення та прискорення розробки веб-програми були використані різні фреймворки та бібліотеки. Наприклад, я використовую Django для створення повноцінного веб-сервера, scikit-learn для втілення різноманітних алгоритмів машинного навчання, json для опрацювання даних і т.д. Django - це багатофункціональний фреймворк для веб-розробки Python, який надає всі необхідні інструменти для створення веб-зстосунку.

Середовища та інструменти розробки: для написання та налагодження коду було використано інтегроване середовище розробки PyCharm, яке має безліч корисних функцій та розширень для роботи з Python. Для контролю версій коду була використана система Git, яка дозволяє відслідковувати зміни, синхронізувати код із віддаленим репозиторієм на GitHub та співпрацювати з

іншими розробниками.

3.3 Архітектура програми

Діаграма класів програми зображена на рисунку 3.1.

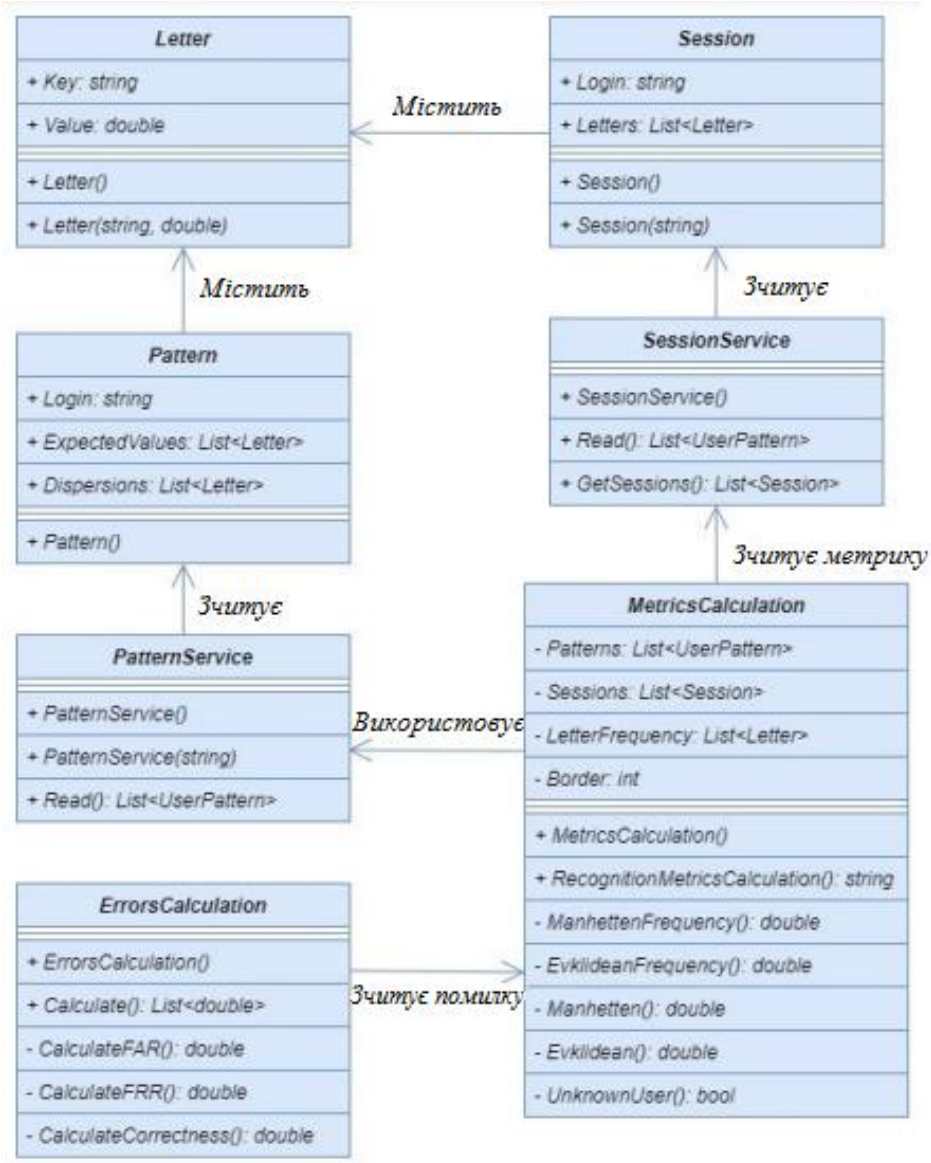

Рисунок 3.1 – Діаграма класів

Метою класу PattemService є вилучення шаблонів сесій зі сховища даних. Об'єкти класу Pattern є записами шаблонів сесій.

Клас SessionService відповідає за отримання сесій зі сховища даних. Об'єкти класу Session містять дані про сесії.

Клас Letter описує літеру, яку натиснув користувач, та час утримання клавіші.

Клас MetricsCalculation вирішує завдання ідентифікації шляхом обчислення різних методів.

Клас ErrorsCalculation обчислює помилки методів.

На рисунку 3.2 представлена діаграма використання веб- застосунку.

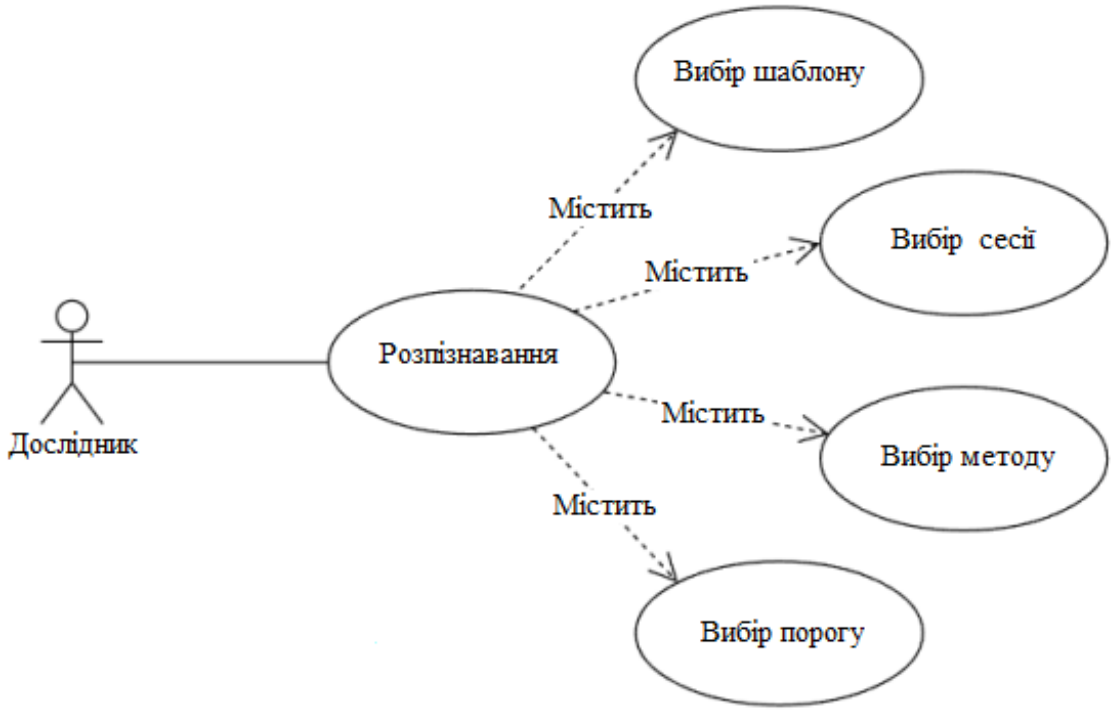

Рисунок 3.2 – Діаграма використання

Користувачем веб-застосунку є дослідник з клавіатурного розпізнавання, він може використовувати прецедент розпізнавання, який включає вибір шаблону, сесії, методу, а також граничне значення розпізнавання.

На рисунку 3.3 представлена діаграма діяльності процесу розпізнавання. Дослідник вибирає файл з шаблонами користувачів і файл, що містить сесії, далі він (дослідник) задає параметри розпізнавання такі як номер сесії для розпізнавання, значення порогу і вибирає один із запропонованих методів розпізнавання (Евклідова відстань, Манхеттенська відстань, SVM, Евклідова чи Манхеттенська відстані з впливом частотності букв). Після всіх перерахованих

вище дій користувач системи отримує результат ідентифікації обраної сесії.

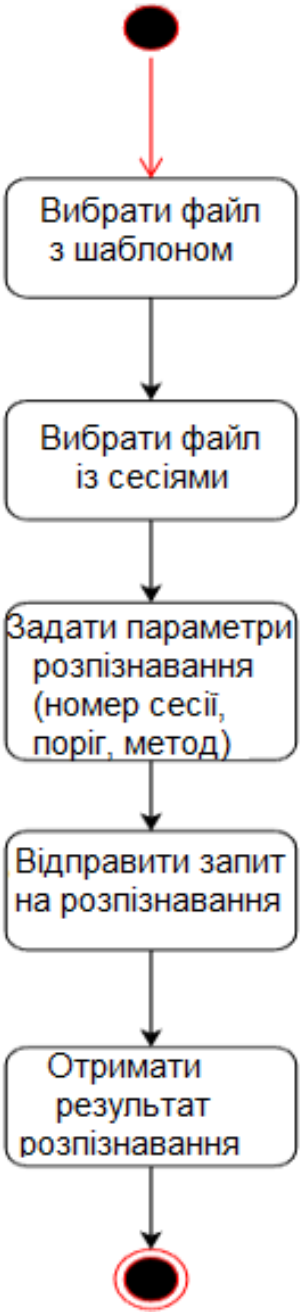

Рисунок 3.3 – Діаграма діяльності

3.4 Інтерфейс програми

Застосунок представлено у вигляді двох веб-сторінок (рисунок 3.4).

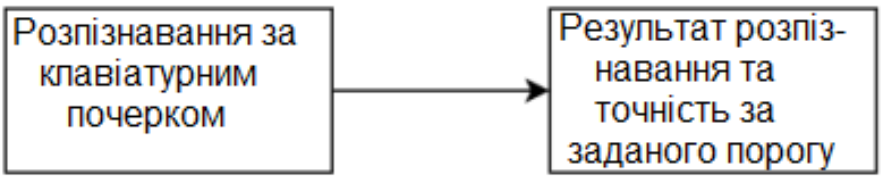

Рисунок 3.4 – Структура програми

На рисунку 3.5 представлено веб-форму для заповнення параметрів розпізнавання.

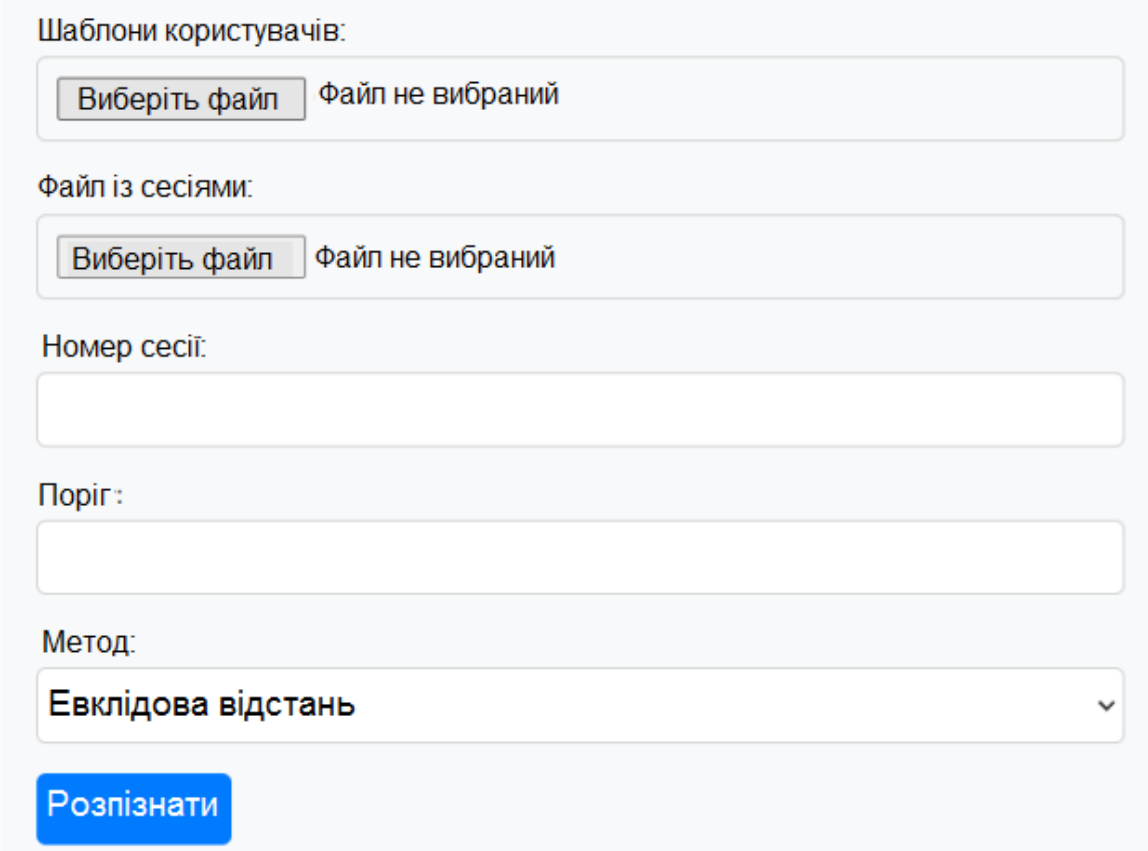

Рисунок 3.5 – Веб-форма для розпізнавання

Веб-форма у своєму складі містить:

1. Два поля введення типу file для вибору файлів шаблонів користувачів та файл із сесіями користувачів.
2. Поле з номером сесії для розпізнавання
3. Поле для встановлення порога розпізнавання
4. Поле списку з методами розпізнавання користувача

Для початку процесу розпізнавання необхідно заповнити всі поля форми перед натисканням кнопки «Розпізнати» (рисунок 3.6).

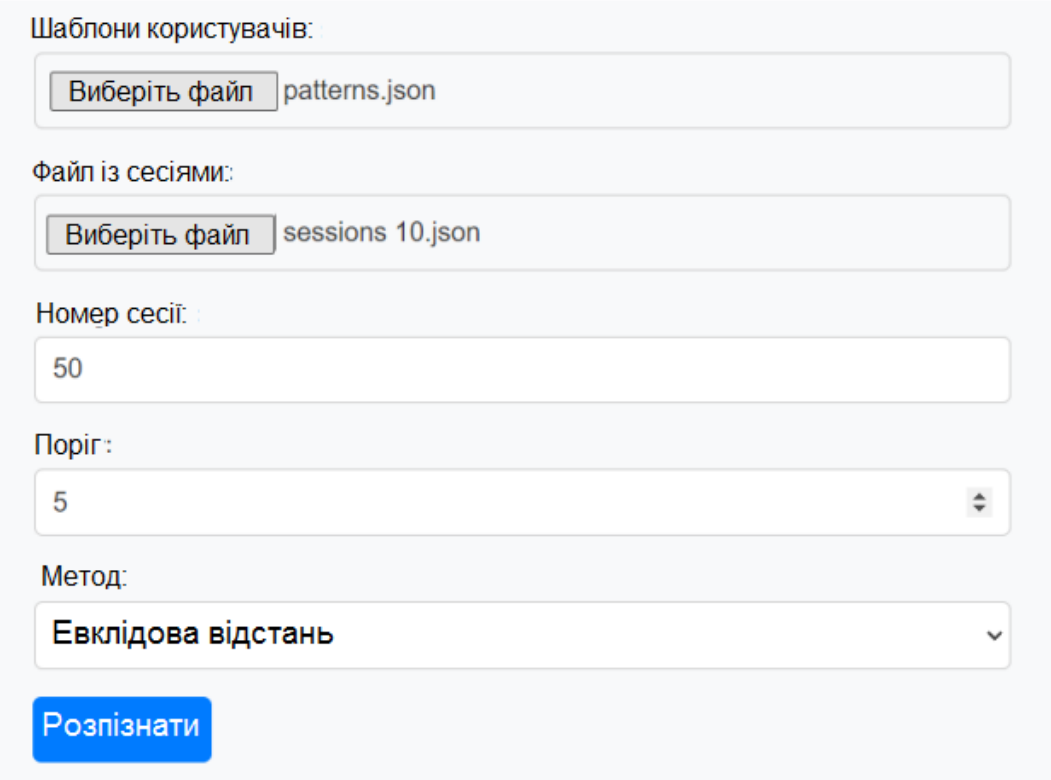

Рисунок 3.6 – Заповнення полів для початку процесу розпізнавання

Результат розпізнавання сесії номер 50, а також точність методів наведено на рисунку 3.7.

Системою ідентифікований користувач User27

Точність алгоритму

Метод	TAR	FAR	FRR	TRR
Евклідова відстань	97	25	44	34
Манхетенська відстань	93	29	48	30
Евклідова відстань + частотність	41	81	26	52
Манхетенська відстань + частотність	44	78	23	55
SVM (метод опорних векторів)	97	25	44	34

Рисунок 3.7 – Результат та точність методів розпізнавання

3.5 Результати досліджень

Цикл безперервної аутентифікації користувача складається з двох основних етапів: реєстрації даних та аутентифікації особи. На першому етапі відбувається постійний збір та аналіз даних про клавіатурні натискання користувача, на основі яких витягуються характеристики його клавіатурної динаміки.

Шаблони користувачів є динамічними, тобто адаптуються до змін у клавіатурному почерку користувача, викликаних різними факторами, такими як психоемоційний стан, втома тощо. Шаблони формуються на основі випадкових натискань на клавіатуру під час роботи з будь-якими програмами ОС. Для проведення дослідження було обрано датасет про клавіатурні натискання англійською мовою КМ [9].

Наступним етапом система обчислює середній час утримання клавіші в поточному сеансі кожної літери і оновлює відповідний шаблон у базі даних. Це дозволяє підвищити точність ідентифікації.

На рисунку 3.8 подано візуальні відображення шаблонів користувачів банку даних КМ. По горизонтальній осі вказані літери англійської абетки, по вертикалі - час утримання клавіші в мілісекундах. З рис. 3.8 видно, що шаблони користувачів мають відмінності.

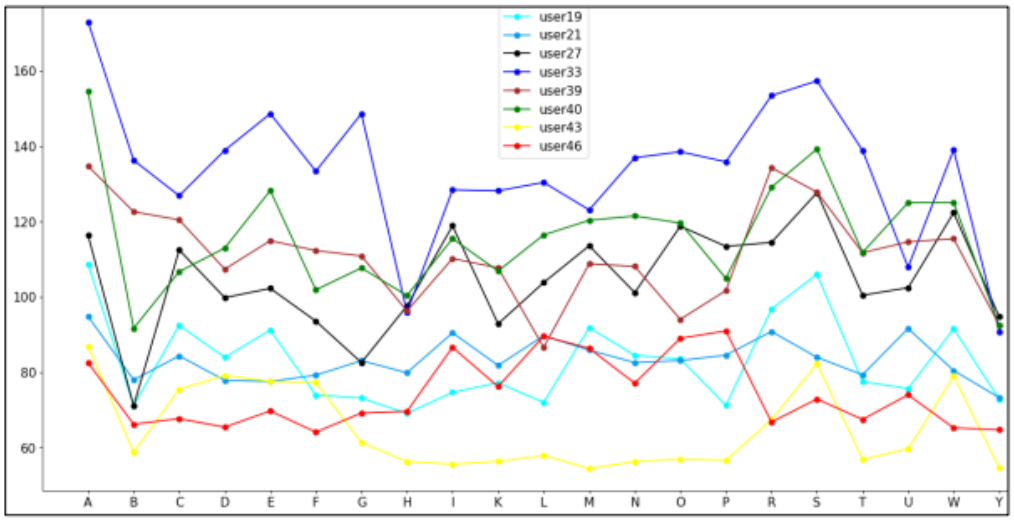

Рисунок 3.8 – Клавіатурні шаблони датасету КМ

З рис. 3.8 випливає, що кожен шаблон унікальний, це відбувається через відмінності швидкості і ритму набору тексту різними користувачами.

Рисунок 3.9 показує щільність розподілу часу утримання клавіші.

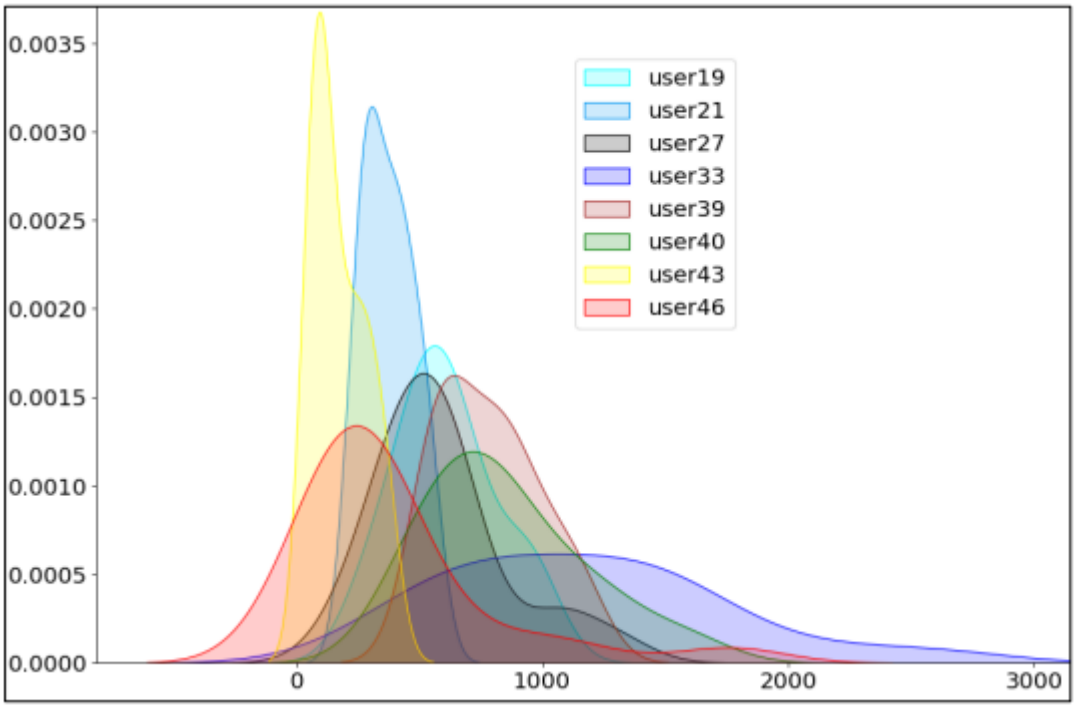

Рисунок 3.9 – Щільності розподілу часу утримання клавіші користувачів

Аналіз графіків на рис. 3.9 показує відмінності у швидкості та якості набору тексту у користувачів. User43 і user46 набирають текст найшвидше (жовта і червона лінії), але у User46 більш рівномірний час утриманні клавіш для різних літер, що свідчить про кращі навички роботи з клавіатурою. User33 набирає повільніше за всіх (синя лінія) і має великий розкид часу утримання, що вказує на слабкі навички набору тексту. Статичний аналіз також може виділити гендерні відмінності та психоемоційну нестабільність користувача.

Для підтвердження легітимності користувача системи порівнюються його поточний та збережений у базі даних шаблони. Можливі два варіанти:

– шаблони збігаються;

– шаблони не збігаються.

У цьому дослідженні для аналізу збігу шаблонів використовувалися SVM та Евклідова s Манхеттенська метрики відстаней.

У будь-якому методі важливо вибрати поріг ухвалення рішення в залежності від поставлених завдань. Низький поріг означає малу різницю між шаблонами та складний доступ до системи. Високий поріг означає велику різницю між шаблонами та простий доступ до системи. На рисунку 3.10 показуються візуальні інструменти помилок першого і другого для Манхеттенської метрики.

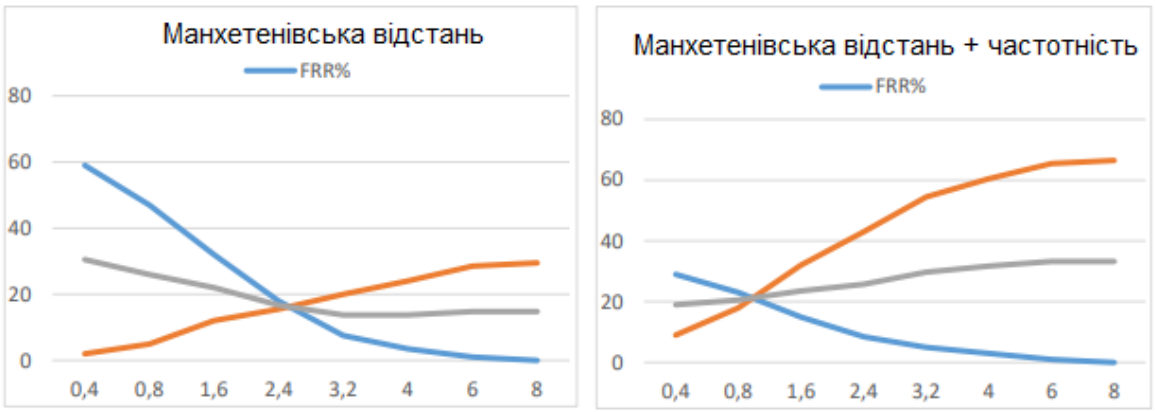

Рисунок 3.10 – Оцінка ефективності розпізнавання користувачів

Як видно з рис. 3.10, досягнуто зменшення порогового значення приблизно в 3 рази з 2,4 до 0,8 мс при незначному збільшенні помилки.

У таблиці 3.1 зведено результати дослідження за методами розпізнавання.

Таблиця 3.1 - Результати дослідження за методами

	Відстань				**Метод**
	Евклідова		**Манхеттенська**		**SVM**
	частотність		**частотність**		
	-	+	-	+	
ERR, (0-100)%	12,25	21	16,75	20,5	14,75
Поріг (FAR = FRR), мс	2,5	0,9	2,6	0,8	2,3
Accuracy, (0-100)%	86,3	80,32	82,08	89,47	88,3
Precision, (0-1)	0,64	0,47	0,58	0,57	0,64
Recall, (0-1)	0,73	0,16	0,69	0,2	0,73

Порівняльний аналіз показав, що показники, отримані на основі Манхеттенської та Евклідової метрик ідентичні. SVM показав результат ERR 14,75% при пороговому значенні 2,3 мс.

Для розпізнавання легітимних користувачів та шпигунів використовується матриця відповідності помилок, представлена в таблиці 3.2. У ній відображено чотири можливі результати розпізнавання:

– вірний допуск (TA);
– вірна відмова (TR);
– хибний допуск (FA);
– хибна відмова (FR).

Таблиця 3.2 - Матриця відповідності помилок

		Користувач при розпізнанні	
		законний	незаконний
Фактичний користувач	законний	TA	FR
	незаконний	FA	TR

На основі цієї матриці можна обчислити такі показники якості розпізнавання:

– Accuracy - частка вірних рішень щодо всіх користувачів

$$Accuracy = \frac{TA + TR}{TA + TR + FA + FR}. \quad (3.1)$$

– Precision – частка легітимних користувачів серед усіх допущених

$$Precision = \frac{TA}{TA + FA}. \quad (3.2)$$

– Recall – частка допущених користувачів серед усіх легітимних

$$Recall = \frac{TA}{TA + FR}. \quad (3.3)$$

Усі три показники відображають точність аутентифікації «свого» користувача, але з різними акцентами. Accuracy показує загальну точність вірних допусків та відмов. Precision показує, як часто система пропускає шпигунів. Recall показує, як часто система відмовляє легітимним користувачам.

У цьому дослідженні для порівняння шаблонів користувача були використані метричні відстані (Евклідова та Манхеттенська) та SVM. Для оцінки візуалізації якості класифікації розпізнавання були побудовані ROC та DET на рисунках 3.11 та 3.12 відповідно.

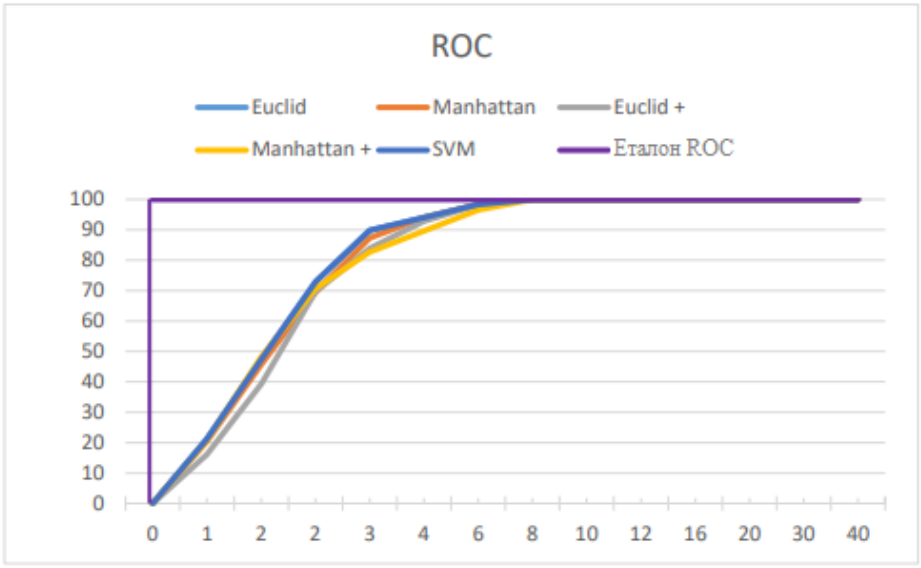

Рисунок 3.11 – Графік ROC

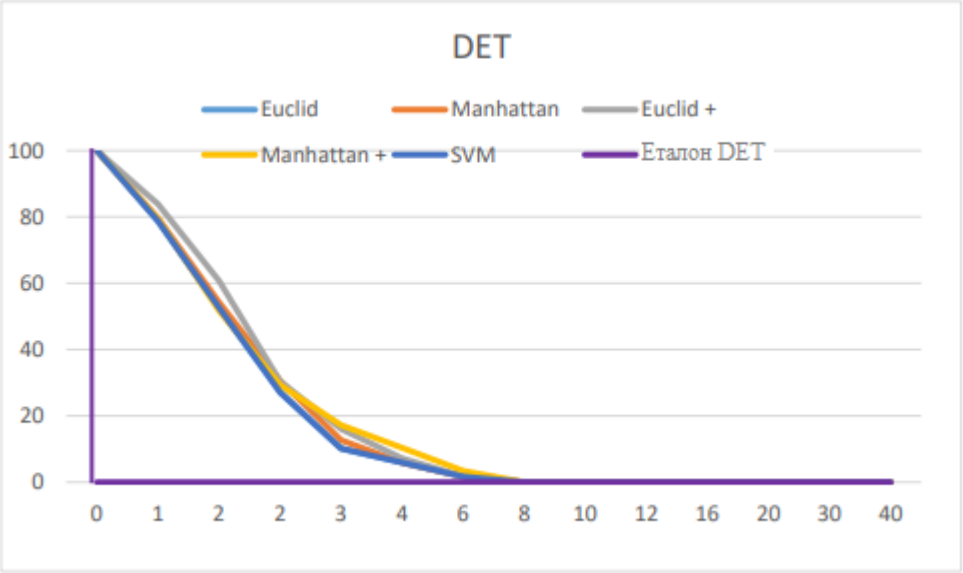

Рисунок 3.12 – Графік DET

На графіку ROC видно, що метод опорних векторів має більшу площу по кривій (AUC), ніж інші алгоритми, що означає, що він якісніше поділяє користувачів на противагу іншим алгоритмам. На графіку DET видно, що манхеттенська відстань з поправкою на частотність має меншу частоту помилок виявлення та пропуску при середніх порогах класифікації, ніж інші, що означає, що він точніший і надійніший.

Крім того, для кожного методу аутентифікації було створено таблицю Excel з результатами розрахунків програми. На основі цих таблиць було побудовано графіки залежності помилок FAR, FRR, ERR.

Результати розрахунків показали, що найменше значення ERR має Евклідова відстань, яка дорівнює 12,25%. Наступним зростанням ERR – це SVM зі значенням 14,75%. Потім йдуть Манхеттенська відстань, Манхеттенська відстань та Евклідова відстань із поправкою на частотність букв англійського алфавіту 16,75%, 20,5% та 21% відповідно.

У ході дослідження з'ясувалося, що облік частотності букв англійської мови значно зменшує час допуску шляхом підвищення чутливості втричі, на шкоду точності розпізнавання в районі 5%. Допустимі значення помилок першого та другого роду, а також значення ERR необхідно підбирати в залежності від конкретної прикладної задачі.

РОЗДІЛ 4. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

4.1 Долікарська допомога при ураженні електричним струмом

Удар електричним струмом є поширеною травмою і часто може закінчитися летальним випадком. Ураження електричним струмом відбувається під час контакту тіла людини з джерелом електричної енергії під напругою. Це реакція організму людини на проходження електричного струму через тіло. Вона проявляється по різному, від легких потрясінь до небезпечних здоров'ю травм, які можуть вплинути на тканини в організмі.

Шкода завдана електричним струмом залежать від кількох факторів: наскільки висока була напруга, яка область тіла була вражена і від виду струму. Фізичні наслідки для людини можуть коливатися від опіків частин тіла до серйозних вражень внутрішніх органів [36].

Часто люди стикаються з підвищеним ризиком ураження високою напругою. Низька напруги не наносить серйозних травм для люди, а з іншої сторони висока напруги, яка має більше 500 В, може призвести до серйозного пошкодження тканин. У дітей або підлітків при враженні електричним струмом в діапазоні від 110 до 200 В можуть вникнути значні травми. Зазвичай це трапляється підчас порушення техніки безпеки при роботі з електричними приладами, які є в побуті. Це можуть бити електричні шнури, подовжувачі, несправні розетки та багато чого іншого.

Виділяють чотири основні фактори від яких залежить вплив електричного струму на організм: величина струму, що протікає через тіло; органи, через які проходить струм; час, протягом якого струм вражає тіло; частота струму.

Фізичні наслідки на пряму залежать від величини струму, яка вражає тіло людини. Струм менший за 1 мА не завдає жодної шкоди людина фізичного ефекту. При 1 мА можуть відчуватися слабкі поколювання, а при 5 мА – легкий поштовх. Однак при струмі ведичною від 6 до 25 мА людина може відчувати больовий шок і деяку втрату контролю над м'язами.

При ураженні електричним струмом від 50 до 150 мА може спричинити у людини сильний біль, м’язове скорочення і навіть призвести до зупинку дихання. У певних випадках можливий летальний результат для людини. Струм від 1000 до 4300 мА призводить до ймовірної смерть, тому що дана напруга спричиняє пошкодження нервових зав’язків, м’язових скорочення та порушення ритму серця. До 10 000 мА електричний струм спричиняє важкі опіки шкіри людини та зупинку серця. Висока ймовірність смерті.

Найпоширеніші ознаки та симптоми враження електричним струмом включають:

- втрата свідомості;
- ускладнення або зупинка дихання;
- опіки, які виникають там, де струм входить і виходить з тіла;
- зупинка серця;
- слабкий і непостійний пульс або його припинення.

Перш за все, що потрібно зробити при першій допомозі, це відключити джерело живлення. Вимкнути електропостачання, від’єднати електроприлад від джерела електричного струму або вимкнути блок запобіжників, якщо він знаходиться неподалік. Не потрібно намагатися підходити близько до жертви, якщо не переконані, що це безпечно і живлення вимкнено.

Потрібно бути обережним у вологих місцях, тому що вода є електричним провідником і рятівник може стати теж жертвою. Якщо людина не впевнена щодо вологості поверхні, потрібно відключити основне електропостачання будинку. У випадку коли це неможливо, використати підручний предмет, який не є провідником і відокремити людину від джерела струму. Це може бути дерев’яна або пластмасова річ.

Після того як постраждалого відокремили від джерела електрики, потрібно викликати швидку і надати першу медичну допомогу. Далі потрібно визначити стан жертви. Перевірити, чи людина у свідомості і дихає. У складних випадках у жертви може бути слабкий пульс або його відсутність. Можливо, що дихання зупинилося. Якщо людина втратила свідомість і перестала дихати, потрібно почати серцево-легеневу реанімацію. Руки розташовуємо в центрі грудної, одна

на іншу. Сильно і швидко натиснути 30 раз приблизно до третини діаметра грудної клітки. Після кожного натискання на грудну клітку робиться два рятувальні вдихи. Потрібно відкинути голову потерпілого назад і підняти підборіддя. Затиснути ніс і створити повне ущільнення. Далі подути потерпілому в рот і подивитися, чи підніметься грудна клітка. Потрібно продовжувати робити натискань на грудну клітку та вдих, поки не прибуде медична допомога або людина не почне сама дихати. Якщо потерпілий живий, перемістити його у зручне йому положення подальше від небезпеки. Можна запобігти шоку, поклавши людину рівно на землю, з головою трохи нижче тіла [36].

Якщо людина при свідомості, нормально дихає і на тілі є опіки, потрібно накрити їх звичайною харчовою плівкою або іншою неклейкою пов’язкою, але без мазі чи лосьйону. Якщо кровотеча у потерпілого, може знадобитися компресія та джгут.

При роботі з соціальною мережею користувач повинен знати і вміти як правильно поводитися з ПК, тому що людина перебуває у непосередньому контакті з джерелом напруги. Удар електричним струмом є потенційно смертельною травмою. Негайна медична допомога важлива, щоб запобігти серйозним травмам і смерті. Для запобігання уражень електричним струмом при роботі за ПК слід встановити додаткові захисні пристрої, що забезпечують недоступність токопровідних частин для дотику. Для зменшення небезпеки використовувати розділовий трансформатор для розв'язки з основною мережею.

Удар електричним струмом є потенційно смертельною травмою. Негайна медична допомога важлива, щоб запобігти серйозним травмам і смерті.

4.2 Вимоги ергономіки до організації робочого місця оператора ПК

Робоче місце – це ділянка простору, яка облаштована необхідним обладнанням, відповідно до трудової діяльності, для виконання поставлених завдань.

Правильно побудоване робоче місце повинне забезпечувати:

– найкраще розміщення обладнання і предметів праці;

– не допускати дискомфорту;

– підвищувати продуктивність праці;

– зменшувати втому працівника.

Розмір робочого місця повинен бути таким, щоб людина не виконувала лишніх рухів і не відчувала дискомфорту під час виконання роботи. Також для працівника важливо мати змогу змінити робочу позу, наприклад, положення тулуба, рук або ніг. Потрібно мінімізувати або звести до нуля всі незручності положення тіла [37].

Різні дослідження заявляють, що при правильному проектуванні робочого місця продуктивність людини може зрости від 15-25%. Такі фактори як рівень освітлення, вологість повітря, температура, шум, вібрація, токсичність, мають значний вплив на умови життєдіяльності і працездатності людини.

Антропометричні вимоги визначають відповідність робочого місця до фізіологічних параметрів тіла людини як зріст і розміри тіла. Індикатором цього є правильна робоча поза, відсутність дискомфорту, оптимальні зони досягнення, раціональні рухи. Психофізіологічні та фізіологічні вимоги формулюють відповідність обладнання і робочого місця можливостям співробітника щодо розуміння, обробки даних, пошук і реалізації рішень.

Організація робочого місця передбачає наступні пункти:

– раціональне положення робочого місця у приміщенні;

– вибір робочих меблів відповідно до фізіологічних характеристик працівника;

– правильне компонування і розміщення обладнання на робочих місцях;

– урахування особливостей та характеру професійної діяльності.

До загальних принципів організації робочого місця відносять:

– робоче місце повинне містити тільки ті предмети, які беруть учать у робочому процесі, але не заважати йому;

– предмети, які часто використовуються у роботі, розміщуються ближче, ніж ті речі, якими користуються рідше;

– предмети, які беруться лівою рукою, повинні розміщуватися зліва, а предмети, які використовуються правою рукою — справа;

– якщо при роботі з предметом працівник використовує дві руки, то він розміщується з урахуванням зручності захоплювання його двома руками;

– робоче місце не повинно бути засмічене;

– необхідна оглядовість повинна бути забезпечена при правильній організації робочого місця.

Робоча поза – це найбільш тривале положення тіла працівника протягом робочого дня. При зручній робочій позі забезпечується стійкість положення тулуба, ніг, рук, голови і витрачається мінімальний запас енергії та максимальну продуктивність [37].

Сидячи і стоячи – дві найпопулярніших пози у робочому процесі. При проектуванні робочого місця потрібно враховувати, що з фізичним навантаженням бажана поза стоячи, а при малих зусиллях – сидячи. При роботі стоячи, людина стомлюється більше ніж сидячи. У відсотковому еквіваленті це на 10% більше енергії. При додатковому навантаженні підвищується артеріальний і венозний тиск крові, розширення вен, пошкоджуються ступені та викривляється хребет. У свою чергу при сидячій роботі нижня частина тіла розслаблена, а основне навантаження спрямоване на м'язи шиї, спини, таза, стегон. При неправильній сидячій позі розвивається застій крові у ногах, а якщо пальці виконують багато роботи можливе запалення суглобів.

Організація робочого місця при використанні ПК повинна відповідати усім ергономічним вимогам. Ключові ергономічні вимоги до проектування робочого місця оператора ПК зображені на рисунку 4.1.

При роботі з персональним комп’ютером потрібно:

– зменшувати кількість статичних напружень;

– розподіляти кількість і час статичних напружень;

– змінювати робочі пози під професійної діяльності.

Саме вибір правильної робочої пози визначається від впливу багатьох факторів. Одні з найважливіших це – кількість зусиль яка прикладається, велична робочої зони, відношення висоти робочої поверхні і ростом працівника.

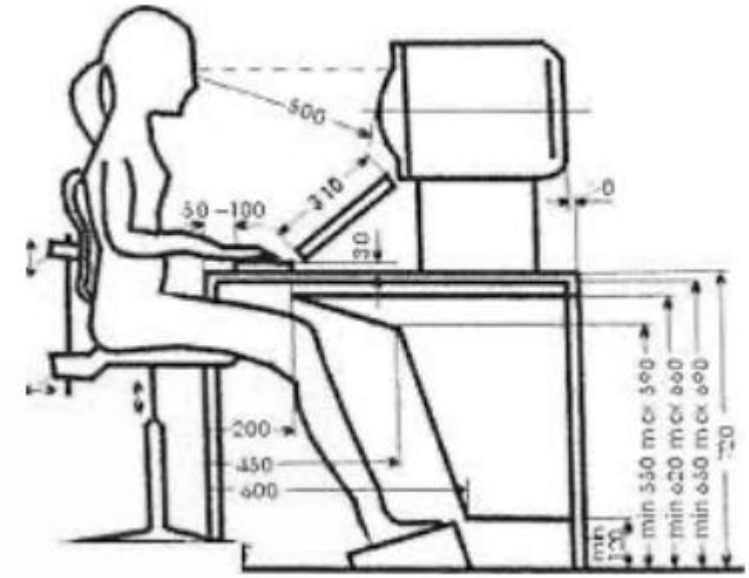

Рисунок 4.1 – Робочий стіл і розміщення користувача ПК

При використанні соціальної мережі користувач повинен дотримуватися вище зазначених правил, які будуть сприяти комфортній і продуктивній роботі. Важливо регулярно робити короткі перерви. Часта зміна заняття – кращий спосіб уникнути можливих неприємностей.

ВИСНОВКИ

У рамках виконання роботи було вивчено спосіб вирішення проблеми підтвердження особи користувача на основі аналізу його роботи на клавіатурі.

Було виявлено, що автентифікація користувача може здійснюється на основі безперервного контролю його клавіатурних натискань у будь-якому програмному середовищі.

На основі проведених досліджень були сформульовані такі висновки:

– необхідно коригувати зразки клавіатурного почерку при використанні динамічної аутентифікації користувача за почерком клавіатури;

– поліпшення якості даних, що збираються системою, а також виділення стійких ознак клавіатурного почерку суттєво впливають на наступні обчислення і, зрештою, на здатність системи правильно підтверджувати законність користувача;

– впровадження в систему частотності букв значно впливає на чутливість до аутентифікації;

– використаний у дослідженні SVM показав середній результат поступившись лише розпізнаванню на основі Евклідової відстані.

ПЕРЕЛІК ДЖЕРЕЛ

1. Коваленко С. В., Красножон О. В. Методи первинної обробки даних у системах ідентифікації користувачів на основі клавіатурного почерка. Технічні науки та технології. 2025. № 2(40). С. 294-302.

2. Аутентифікація і авторизація: що це і в чому [Електронний ресурс] – Режим доступу: https://qagroup.com.ua/publications/autentyfikatciia-i-avtoryzatciia/ (Дата звернення 16.05.2026).

3. Rashik Shadman Ahmed Anu Wahab Michael Manno Matthew Lukaszewski Daqing Hou Faraz Hussain (2024). Keystroke Dynamics: Concepts, Techniques, and Applications https://arxiv.org/ pdf/2303.04605.

4. Ложников П.С., Сулавко А.Е., Бура Е.В. Аутентифікація користувачів комп'ютера по клавіатурноve почерку і особливостях особи. Питання кібербезпеки. 2017. №4. С. 24-34..

5. Guide to Biometrics. Ruud M. Bolle, onathan H. Connell, Sharath Pankanti. 2013.

6. What is a keylogger. [Електронний ресурс] – Режим доступу: https://www.csoonline.com/article/3326304/keyloggers-explained-howattackers-record-computer-inputs.html (дата звернення: 16.05.2026).

7. Lytvynenko, I., Lupenko, S., Nazarevych, O., Shymchuk, G., & Hotovych, V. (2021). Mathematical model of gas consumption process in the form of cyclic random process. 2021 IEEE 16th International Conference on Computer Sciences and Information.

8. Методичні вказівки до виконання кваліфікаційної роботи ор Бакалавр для студентів спеціальності 122 – Комп'ютерні науки, всіх форм навчання / укладачі: Готович В.А., Дуда О.М. Никитюк В.В. Тернопіль: Тернопільський національний технічний університет імені Івана Пулюя, 2024. 43 с.

9. Keystroke Dynamics - Benchmark Data Set. [Електронний ресурс] – Режим доступу: https://www.cs.cmu.edu/~keystroke/ (дата звернення: 18.05.2026).

10. González N., Calot E.P., Ierache J.S., Hasperué W. On the shape of timings distributions in free-text keystroke dynamics profiles // Heliyon. – 2021. – Vol. 7, no. 11. – e08413. – DOI: 10.1016/j.heliyon. 2021.e08413.

11. Kim J., Kim H., Kang P .Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature 89 extraction and novelty detection // Applied Soft Computing .– 2018 .– Vol .62 .– P .1077–1087 .– DOI: 10 .1016 / j.asoc .2017 .09 .045.

12. Kochegurova E.A., Martynova Y.A.Aspects of continuous user identification based on free texts and hidden monitoring // Program Comput Softw.– 2020.– Vol.–46 (1).– P.–12-24.– DOI:10.– 1134/S036176882001003X.

13. Zaidi A.Z., Chong C.Y., Jin Z., Parthiban R., Sadiq A.S.Touch-based continuous mobile device authentication: State-of-the-art, challenges and opportunities // J Network Comput Appl.– 2021.– Vol.–191.– P.– 103162.– DOI:10.–1016/j.jnca.– 2021.–103162.

14. Teh P.S., Teoh A.B.J., Yue S.A survey of keystroke dynamics biometrics // Sci World J.–2013;–2013:1-24;–DOI:10;– 1155/2013/408280

15. Morales A., Fierrez J., Tolosana R.Ortega-Garcia J.Galbally J.GomezBarrero M.Anjos A.Marcel S.KBOC: Keystroke biometrics OnGoing competition // 2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS).–2016;–DOI:10;– 1109/BTAS;–2016;–7791180.

16. Pisani P.H.Lorena A.C.A systematic review on keystroke dynamics // J Braz Comput Soc;–2013;–19(4):573-587.

17. Gunetti D., Picardi C.;Keystroke analysis of free text // ACM Trans Inf Syst Secur;–2005;–8(3):312-347;–DOI:10;–1145/1085126;1085129.

18. Kochegurova E.Luneva E.Gorokhova E.On continuous user authentication via hidden free-text based monitoring // Adv Intell Sys Comput;–2019;–875:66-75;– DOI:10;1007/978-3-030-01821-4_8.

19. Alsultan A.Warwick K.Non-conventional keystroke dynamics for user authentication // Pattern Recogn Lett;–2017;89:53-59;doi:10;1016/j.patrec;2017;02;010.

20. Mondal S.Bours P.A study on continuous authentication using a combination of keystroke and mouse biometrics // Neurocomputing;– 2017;(230):1-22;DOI:10;1016/j.neucom;2016;11;031.

21. Zhong Y.Deng Y.A survey on keystroke dynamics biometrics: approaches, advances, and evaluations // Recent Advances in User Authentication Using Keystroke Dynamics Biometrics;2015;(2):1- 22.DOI:10;15579/gcsr.vol2.ch1.

22. Ali M.L., Monaco J.V., Tappert C.C. et al. Keystroke Biometric Systems for User Authentication // J Sign Process Syst. – 2017. – P. 175– 190. – DOI: 10.1007/s11265-016-1114-9.

23. Alsultan A., Warwick K., Wei H. Non-conventional keystroke dynamics for user authentication // Pattern Recogn Lett. – 2017. – Vol. 89. – P. 53-59. – DOI: 10.1016/j.patrec.2017.02.010.

24. Messerman T., Mustafić T., Camtepe S.A., Albayrak S.Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics // 2011 International Joint Conference on Biometrics (JCB).– 2011.– P.–1–8.– DOI:10.–1109/IJCB.–2011.– 6117552.

25. Chang HC, Li J, Wu C, Stamp M. Machine Learning and Deep Learning for Fixed-Text Keystroke Dynamics. arXiv:2107.07409v1 [cs. LG]; 2021. doi.org/10.48550/arXiv.2107.07409.

26. Ahmed AA, Traore I. Biometric recognition based on free-text keystroke dynamics/ Cybern. IEEE Trans 2014; 44(4): 458–472. DOI:10.1109/TCYB.2013.2257745.

27. Goodkind A, Brizan DG, Rosenberg A. Utilizing overt and latent linguistic structure to improve keystroke-based authentication. Image and Vision Computing 2017; 58: 230-238. DOI:10.1016/j.imavis.2016.06.003.

28. Al Solami E, Boyd C, Clark A, Ahmed I. User-representative feature selection for keystroke dynamics. 5th International Conference on Network and System Security (NSS'11) 2011: 229–233. DOI:10.1109/ICNSS.2011.6060005.

29. Eberz S, Rasmussen KB, Lenders V, Martinovic I. Evaluating behavioral biometrics for continuous authentication: challenges and metrics. 2017 ACM on Asia

Conference on Computer and Communications Security (ASIA CCS ’17). 2017: 386-399. DOI:10.1145/3052973.3053032.

30. Antal M, Szabó LZ, Laszlo I. Keystroke dynamics on Android platform. Procedia Technology 2015; 19: 820-826. DOI:10.1016/j.protcy.2015.02.118.

31. Locklear H, Govindarajan S, Sitova Z, etc. Continuous authentication with cognition-centric text production and revision features. IEEE/IAPR international joint conference on biometrics (IJCB 2014); 2014. DOI:10.1109/BTAS.2014.6996227.

32. Kang P, Cho S. Keystroke dynamics-based user authentication using long and free text strings from various input devices. Inf Sci 2015; 308: 72-93. DOI:10.1016/j.ins.2014.08.070.

33. Matsubara Y, Samura T, Nishimura H. Keyboard Dependency of Personal Identification Performance by Keystroke Dynamics in Free Text Typing. Journal of Information Security 2015; 6: 229-240. DOI: 10.4236/jis.2015.63023.

34. Lupenko, S. A., Lytvynenko, I. V., Sverstiuk, A., Shelestovskyi, B., & Horkunenko, A. (2021). Software for Statistical Processing and Modeling of a Set of Synchronously Registered Cardio Signals of Different Physical Nature. CMIS, 194-205.

35. Bodnarchuk, I., Skorenkyy, Y., Kramar, T., Duda, O., & Nykytyuk, V. (2022). Use of Analytical Hierarchy Process in Scenarios Design for a Digital Museum with XR components. ITTAP, 414–425

36. Заікіна Д., Глива В. Основи охорони праці та безпека життєдіяльності. 2019. URL: https://doi.org/10.31435/rsglobal/001 (дата звернення: 29.05.2026).

37. Безпека в надзвичайних ситуаціях. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання / укл.: Стручок В. С. Тернопіль: ФОП Паляниця В. А., 2022. 156 с.