Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет *Інформаційних радіотехнологій і технічного захисту інформації*

Кафедра *Комп'ютерної інженерії та систем технічного захисту інформації*

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти *другий (магістерський)*

Дослідження ефективності ідентифікації особи за клавіатурним почерком з урахуванням сили тиску на клавіші

Виконав:
студент *2* курсу, групи *СТЗІАм-22-1*
Зеленський Іван Борисович

Спеціальність *125 «Кібербезпека»*

Тип програми *освітньо-професійна*

Освітня програма *«Системи технічного захисту інформації, автоматизація її обробки»*

Керівник *доц. Горелов Д.Ю.*

Допускається до захисту
Зав. кафедри _____ *проф. Антіпов І.Є.*
(підпис)

2024 р.

Харківський національний університет радіоелектроніки

Факультет *Інформаційних радіотехнологій і технічного захисту інформації*
Кафедра *Комп'ютерної інженерії та систем технічного захисту інформації*
Рівень вищої освіти *другий (магістерський)*
Спеціальність *125 «Кібербезпека»*
Тип програми *освітньо-професійна*
Освітня програма *«Системи технічного захисту інформації, автоматизація її обробки»*

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«___» _____ 20 _____ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові *Зеленському Івану Борисовичу*
(прізвище, ім'я, по батькові)

1. Тема роботи *Дослідження ефективності ідентифікації особи за клавіатурним почерком з урахуванням сили тиску на клавіші*

затверджена наказом по університету від « 03 » 11 2023 р. № 1281 Ст

2. Термін подання студентом роботи до екзаменаційної комісії *10 січня 2024 р.*

3. Вихідні дані до роботи

Тип біометричної СКУД: динамічний аналіз клавіатурного почерку користувачів комп'ютеризованих інформаційних систем

Дослідні інформативні ознаки клавіатурного почерку:

1) монографи та диграфи (часові характеристики поодиноких та подвійних подій клавіатури);

2) динаміка зміну тиску на клавіші в процесі вводу парольної фрази.

4. Перелік питань, що потрібно опрацювати в роботі

Дослідити вплив на точність ідентифікації за клавіатурним почерком динаміки зміни тиску на клавіші в процесі вводу парольних фраз.

Для досягнення поставленої мети необхідно розв'язати наступні задачі:

1) провести пошук відкритих датасетів клавіатурного почерку, в яких присутні дані про динаміку зміни тиску на клавіші в процесі набору текстів, та обрати один з них для подальших досліджень; 2) на основі обраного датасету провести експериментальні дослідження впливу на точність ідентифікації за клавіатурним почерком ознак динаміки зміни тиску сили на клавіші та їх комбінацій з часовими параметрами клавіатурного почерку.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів)

1. Мета та задачі кваліфікаційної роботи. А4. Ел.ф.

2. Датасети параметрів клавіатурного почерку, що знаходяться у вільному доступі. А4. Ел.ф.

3. Queen Mary University Keystroke benchmark dataset. А4. Ел.ф.

4. Результати мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за часовими параметрами та динамікою зміни тиску на клавіші. А4. Ел.ф.

5. Результати двокласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за динамікою зміни тиску на клавіші. А4. Ел.ф.

6. Результати класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за комбінацією часових параметрів та динаміки зміни тиску на клавіші. А4. Ел.ф.

7. Висновки

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	*Огляд біометричних технологій контролю доступу*	*01.09.23 – 20.09.23*	
2	*Аналіз сучасних методів, алгоритмів та систем ідентифікації користувачів за клавіатурним почерком*	*21.09.23 – 31.10.23*	
3	*Проведення експериментальних досліджень*	*01.11.23 – 31.12.23*	
4	*Перевірка роботи на антиплагіат*	*03.01.24 – 05.01.24*	
5	*Представлення кваліфікаційної роботи на кафедрі*	*10.01.2024*	

Дата видачі завдання *02 вересня 2023 р.*

Студент _____
(підпис)

Керівник роботи _____ _____
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 81 с., 19 рис., 48 табл., 27 джерел, 1 додаток.

БІОМЕТРИЧНІ СИСТЕМИ, ІДЕНТИФІКАЦІЯ, КЛАСИФІКАЦІЯ ДАНИХ, DATA MINING, КЛАВІАТУРНИЙ ПОЧЕРК,

Об’єкт дослідження – біометричні системи ідентифікації особи.

Предмет дослідження – аутентифікація за клавіатурним почерком.

Метою цієї роботи є підвищення інформаційної безпеки комп’ютерних систем на основі аналізу клавіатурного почерку.

За допомогою бази даних «Queen Mary University Keystroke benchmark dataset» та програмного забезпечення Orange проведено дослідження впливу на точність ідентифікації за клавіатурним почерком часових параметрів, динаміки зміни тиску на клавіші та їх комбінацій. Експериментально підтверджено, що і для задач мультикласової класифікації і для задач двокласової класифікації на основі ознак тиску на клавіші можна отримати точність ідентифікації не менше 99 %.

ABSTRACT

Master thesis: 81 p., 4 tables, 19 fig., 27 sources, 1 annex.

BIOMETRIC SYSTEMS, IDENTIFICATION, DATA CLASSIFICATION, DATA MINING, USER BEHAVIOR ANALYTICS.

The objects of the research are the biometric systems of personal identification.

The subject of the research is the authentication by the personal keystroke dynamics usage pattern.

The objective of the work is explore the possibility of using keystroke dynamics in computer networks' user identification tasks.

Using the "Queen Mary University Keystroke benchmark dataset" database and the Orange software, a study of the influence of time parameters, dynamics of changes in key pressure and their combinations on the accuracy of identification by keyboard handwriting was carried out. It has been experimentally confirmed that both for multi-class classification problems and for two-class classification problems based on key pressure signs, it is possible to obtain an identification accuracy 99 %.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ ТА ТЕРМІНІВ

EER (Equal Error Rate) – коефіцієнт рівної імовірності помилок 1 і 2-го роду;

FAR (False Acceptance Rate) – помилка другого роду – випадок надання системою доступу неавторизованому користувачеві;

FRR (False Rejection Rate) – помилка першого роду – доступ заборонений користувачеві, зареєстрованому в системі;

ББД – біометрична база даних;

БСА – біометрична система аутентифікації;

БСКД – біометрична система контролю доступу.

ВСТУП

Традиційні засоби аутентифікації зазвичай ґрунтуються на паролях або на перевірці індивідуальних особливостей суб'єкта (біометричних ознак). Перші дуже схильні до «людського фактору», а біометричні системи захисту також не позбавлені недоліків. Щоб об'єднати переваги згаданих технологій, можна використовувати таємні біометричні ознаки, які можуть бути засновані тільки на динамічних біометричних ознаках, наприклад, індивідуальному клавіатурному почерку суб'єкта в процесі набору парольної фрази. Недолік такого методу полягає у порівняно низькій надійності прийнятих рішень. Підвищити надійність розпізнавання суб'єктів за клавіатурним почерком можна за допомогою використання додаткових ознак, що реєструються спеціальними датчиками та характеризують динаміку набору тексту на клавіатурі.

Основні (базові) ознаки – часові характеристики введення символів на клавіатурі – застосовуються у всіх системах розпізнавання. Однак додаткові ознаки, що ґрунтуються на врахуванні даних про динаміку зміни тиску на клавіші клавіатури в процесі набору тексту суб'єктом, мало досліджені.

Метою роботи є підвищення інформаційної безпеки комп’ютерних мереж на основі аналізу клавіатурного почерку.

Для досягнення поставленої мети необхідно розв’язати наступні задачі:

1) провести аналітичний огляд сучасних рішень і алгоритмів аутентифікації користувачів за клавіатурним почерком;

2) провести пошук відкритих датасетів клавіатурного почерку, в яких присутні дані про динаміку зміни тиску на клавіші в процесі набору текстів, та обрати один з них для подальших досліджень;

3) на основі обраного датасету провести експериментальні дослідження впливу на точність ідентифікації за клавіатурним почерком ознак динаміки зміни тиску сили на клавіші та їх комбінацій з часовими параметрами клавіатурного почерку.

1. БІОМЕТРИЧНА ІДЕНТИФІКАЦІЯ ОСОБИСТОСТІ

1.1. Біометричні характеристики особистості

У контексті інформаційної безпеки біометрія – це наука про ідентифікацію та аутентифікацію (підтвердження справжності) особи за її анатомічними (статичними) або поведінковими (динамічними) відмінними характеристиками. Анатомічні (статичні) біометричні характеристики особа набуває вже на момент народження. Більшість із них залишаються незмінними протягом усього життя людини і можуть успішно використовуватися для ідентифікації / аутентифікації її особистості. Статичність анатомічних біометричних показників проявляється в тому, що вони не залежить від часу при їх спостереженні і вимірі. Проте у багатьох конкретних системах під час реєстрації особи, вимірювання проводять кілька разів, щоб нейтралізувати можливі відносні зміщення аутентифікатора та вимірника.

Поведінкові (динамічні) біометричні характеристики особи засновані на індивідуальних особливостях добре завчених підсвідомих рухів (голосова артикуляція, рукопис, маніпуляція на клавіатурі тощо). Поведінкові характеристики є функціями часу, що обов'язково враховується при їх вимірі, обробці та зберіганні. Динамічний характер поведінкових характеристик зумовлює спеціальну процедуру отримання. По-перше, вимір має проводитися протягом певного проміжку часу. По-друге, при реєстрації особи вимірювання необхідно провести кілька разів для отримання усередненого зразка кожної динамічної характеристики.

У класі статичної біометрії в даний час практичне застосування отримали біометричні системи ідентифікації / аутентифікації особи, які використовують як анатомічні параметри особи:

- папілярний рисунок пальців рук;
- рисунок кровоносних судин очного дна (сітківка);
- рисунок райдужної оболонки ока (райдужка);

- геометрія кисті руки;
- рисунок вен кисті руки;
- геометрія обличчя;
- рисунок лицьових артерій та вен.

Основними перевагами статичної біометрії є:

- зручність зняття біометричних характеристик;
- компактність машинних репрезентацій біометричних характеристик (біометричних еталонів);
- малі витрати зусиль користувачів та обслуговуючого персоналу під час проходження процедури реєстрації та ідентифікації;
- можливість організації біометричної ідентифікації великих потоків;
- незалежність результатів ідентифікації від психофізичного стану особи.

Недоліки статичної біометрії:

- відкритість біометричних ознак, що створює потенційну можливість їхньої фальсифікації (хоча й дуже трудомістку);
- висока вартість, зумовлена використанням високих технологій та дорогого обладнання;
- негативне ставлення частини населення до реєстрації біометричних характеристик, пов'язане з анонімністю та конфіденційністю особи.

У класі динамічної біометрії в даний час практичне застосування отримали біометричні системи ідентифікації / аутентифікації, що використовують як образ особи динамічні параметри добре завченого процесу відтворення певного короткого тексту (парольної фрази), що реалізується за допомогою:

- голосу;
- рукопису;
- клавіатурного набору.

Першою перевагою систем динамічної біометрії є їхня відносно низька вартість. Це зумовлено тим, що вони не вимагають застосування дорогого

обладнання і можуть бути реалізовані, або виключно програмними засобами (клавіатурні БСА), або з використанням стандартних недорогих засобів мультимедіа (графічного планшета, мікрофонна та звукової карти). У підсумку вартість динамічних БСА визначається вартістю розробки додаткового програмного забезпечення.

Ще однією важливою перевагою динамічної біометрії є можливість збереження образу особи (еталона біометричних параметрів особи) у таємниці та можливість швидкої зміни цього образу шляхом зміни парольної фрази. Статичні біометричні характеристики особи (наприклад, відбиток пальця, геометрію кисті руки) не можна зберегти в таємниці або змінити, вони дані особи один раз і назавжди.

Недоліком динамічних БСА є те, що на їхню роботу впливає психофізичний стан особи (переляк, стрес, психотропні препарати та ін.). Водночас зазначений недолік динамічних БСА, як такий, проявляється по відношенню до статичних БСА. Існують сфери застосування динамічної біометрії, де виявлення відхилення психофізичного стану особи від норми може бути ефективно використано. Наприклад, необтяжливий допуск до роботи людей певних професій, що характеризуються високою ціною помилки (пілоти, диспетчери повітряного та залізничного транспорту, оператори атомних станцій, чергові офіцери стратегічних систем озброєнь та військової техніки тощо.). Також встановлено, що відхилення психофізичного стану особи від норми виникають при скоєнні особою різних неправомірних дій. Ця обставина також може ефективно використовуватися, наприклад, для виявлення інсайдерів (легальних користувачів інформаційних систем, які здійснюють правопорушення), в детекторах брехні тощо.

Біометричні характеристики особи не тільки поділяються на статичні і динамічні, а й мають низку інших властивостей, які також важливі з погляду їх застосування для ідентифікації. П'ять таких властивостей описані в [1]:

1. Загальність: кожна людина має біометричні характеристики.

2. Унікальність: немає двох людей, які мають однакові біометричні ха-

рактеристики.

3. Постійність: біометричні характеристики мають бути стабільними у часі.

4. Вимірюваність: біометричні характеристики повинні бути вимірювані будь-яким фізичним пристроєм, що зчитує.

5. Прийнятність: сукупність користувачів та суспільство загалом не повинні (сильно) заперечувати проти вимірювання/збору біометричних характеристик.

Комбінація всіх цих властивостей визначає ступінь ефективності тієї чи іншої біометрії та, як наслідок, ефективність відповідних БСА.

Важливо розуміти, що не існує біометричних параметрів, які абсолютно задовольняють будь-якій із зазначених вимог, так само як і тих, які поєднували б у собі всі ці властивості одночасно (особливо якщо враховувати п'яту властивість). Це означає, що будь-яка обираєма для конкретного застосування біометрія – це результат багатьох компромісів.

1.2. Біометричні системи аутентифікації

Будь-які системи аутентифікації особи, у тому числі і біометричні, містять два етапи використання:

- реєстрації образу особи;
- аутентифікації особи.

На першому етапі здійснюється реєстрація у БСА осіб усіх легальних користувачів. Процедура реєстрації починається з первинних вимірів, що відповідають типу БСА біометричних характеристик особи користувача. Вимірювання здійснюються, як правило, кілька разів з подальшим усередненням одержуваних результатів. Як правило, результати первинних вимірів є надмірно великим обсягом біометричних характеристик особи, які важко безпосередньо використовувати для формування біометричного еталона особи. Тому наступним кроком є вилучення з них обмеженої кількості найбільш значу-

щих властивостей та формування на їх основі компактної машинної репрезентації біометричного образу особи – біометричного еталона. Біометричні зразки зареєстрованих у БСА осіб користувачів заносяться до спеціальної біометричної бази даних (ББД) системи.

На етапі біометричної аутентифікації спочатку, як і під час реєстрації, здійснюється вимірювання біометричних характеристик особи користувача, що претендує на доступ, та формування машинної репрезентації його біометричного образу. Далі отримана машинна репрезентація біометричного образу зіставляється з біометричними еталонами, що зберігаються у ББД системи.

На етапі біометричної аутентифікації застосовуються два основні методи:

- Верифікація – процедура аутентифікації особи, яка пред'явила БСА свої біометричні параметри та деякий додатковий, не біометричний ідентифікатор (логін, ПІН, пароль тощо). За цим ідентифікатором, який виступає як адреса, БСА витягує зі своєї ББД відповідний даній особі біометричний еталон і порівнює його з машинною репрезентацією пред'явленого біометричного образу. Збіг свідчить у тому, що особа є тією, яку з себе видає.
- Ідентифікація – процедура аутентифікації особи користувача виключно на основі пред'явлених біометричних параметрів. У цьому випадку БСА почергово порівнює машинну репрезентацію пред'явленого біометричного образу з біометричними зразками всіх зареєстрованих у БСА користувачів. Результатом порівняння є відповідь на запитання – чи зареєстрований даний користувач у БСА чи ні.

З позиції теорії розпізнавання образів верифікація відповідає задачі класифікації вхідних образів на два класи: «свій» та «чужий» (порівняння здійснюється за принципом 1:1). Ідентифікація відповідає задачі класифікації вхідних образів на $(m + 1)$ класів, де m – число зареєстрованих в БСА користувачів («своїх»), плюс один клас, що включає всіх інших, не зареєстрованих у БСА користувачів («чужі») (порівняння здійснюється за принципом $1: m$).

Біометрична ідентифікація на відміну від верифікації може розглядати-

ся як «чиста» біометрична автентифікація, але її набагато складніше реалізувати та застосовувати: кожен пред'явлений біометричний зразок має бути зіставлений (у гіршому випадку) з кожним обліковим записом ББД. Ефективність та швидкодію такої системи забезпечити набагато важче. Разом з тим, для деяких завдань застосування «чистої» біометричної аутентифікації є єдиним можливим. Приклад – пошук біометричного образу доки не встановленої особи у базі даних злочинців («негативна аутентифікація»).

Структура біометричної системи аутентифікації. Будь-яку БСА можна представити як систему розпізнавання образів. Типова структура БСА показана на рис. 1.1.

БСА включає наступні основні функціональні блоки:

- Біометричні зчитувачі (сенсори) здійснюють вимірювання біометричних даних, що пред'являються особою відповідної модальності.
- Екстрактор властивостей витягує значні біометричні параметри з біометричних даних, що надходять, і формує на їх основі машинну репрезентацію біометричного образу особи (біометричний еталон особи).
- Пристрій зіставлення (метчер) здійснює зіставлення машинної репрезентації пред'явленого біометричного образу особи з біометричними стандартами.
- Біометрична база даних (ББД) зберігає біометричні зразки всіх зареєстрованих у системі осіб.

БСА функціонує у двох режимах: реєстрації та ідентифікації / аутентифікації особи.

Режим реєстрації використовується на попередньому етапі роботи БСА створення біометричних еталонів всіх легальних користувачів системи. У цьому режимі представлені потенційним користувачем біометричні параметри нормуються, потім екстрактором властивостей з них вилучається значуща біометрична інформація, яка представляється у вигляді компактного (за можливістю) біометричного препарату – так званої машинної репрезентації (від лат. repraesetatio: re і praesetare – представляти), що виконує роль біометрич-

ного еталона особи . За допомогою деякого додаткового ідентифікаційного параметра (номери, логіну, ПІН, паролю тощо) отриманий біометричний еталон зв'язується в єдиний масив з іншими еталонами, що зберігаються в ББД.

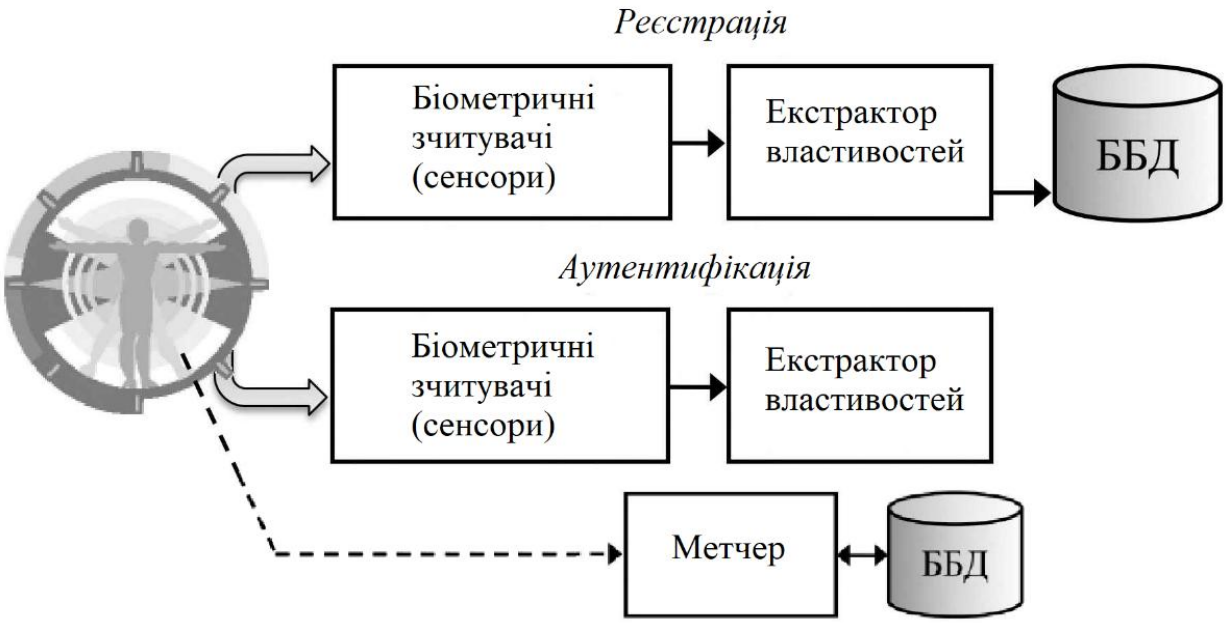

Рисунок 1.1 – Типова структура БСА

Режим ідентифікації / аутентифікації є основним (робочим) режимом БСА. У цьому режимі здійснюється ідентифікація / аутентифікація особи користувача, який пред’явив свої біометричні дані. Залежно від програми вирішується одне з двох завдань – біометрична верифікація або біометрична ідентифікація.

До БСА пред'являються такі основні вимоги:

- висока точність;
- малий час ухвалення рішення (швидкодія);
- низька вартість;
- можливість обробки виняткових випадків.

Точність роботи біометричної системи не може бути фізично виміряна, її можна лише приблизно підрахувати статистичними методами. У математичній статистиці це відповідає завданню перевірки за певним статистичним критерієм P^x двох статистичних гіпотез H_0 і H_1 про належність представленої вибірки X одному з двох розподілів (у разі БСА це «свій» та «чужий») Рс

і Рч. У цьому можливі чотири ситуації, показані у табл. 1.1.

Очевидно, що помилки першого та другого роду є взаємно симетричними. Якщо гіпотези H_0 і H_1 поміняти місцями, також поміняються місцями помилки першого і другого роду. Для визначеності практично прийнято вважати, що нульова гіпотеза H_0 відповідає стану «за умовчанням» (тобто природному, найочікуванішому стану речей), наприклад, що аутентифікуєма людина – є легальним, зареєстрованим суб'єктом.

Таблиця 1.1 – Перевірка статистичних гіпотез

		Вірна гіпотеза	
		H_0	H_1
Результат застосування критерію P^x	H_0	H_0 правильно прийнята	H_0 неправильно прийнята (помилка другого роду)
	H_1	H_0 неправильно відкинута (помилка першого роду)	H_0 правильно відкинута

У більшості БСА використовують саме таке тлумачення гіпотез H_0 і H_1. Проте існують додатки, у яких можлива зворотна ситуація. Наприклад, у розпорядженні правоохоронних органів є база даних, що містить біометричні дані людей, які раніше вчинили злочини. Позитивний результат біометричної ідентифікації людини за такою базою даних свідчить про те, що вона є злочинцем. Вочевидь, що з такого додатку стан «за умовчанням» відповідає негативній ідентифікації людини, коли його біометричні дані у базі даних не виявляються (гіпотеза H_1). Для таких додатків БСА застосовують зворотне тлумачення гіпотез H_0 і H_1.

При традиційному тлумаченні гіпотез H_0 і H_1. (коли станом «за замовчуванням» відповідає гіпотеза H_0) точність роботи БСА прийнято характеризувати відсотком помилок при проходженні процедури аутентифікації. При цьому розрізняють помилки трьох видів:

1. FRR (False Reject Rate) – помилка першого роду. Суть помилки –

ймовірність помилкової відмови авторизованому користувачеві (помилкова відмова «своєму»);

2. FAR (False Accept Rate) – помилка другого роду. Суть помилки – ймовірність пропуску незареєстрованого користувача (помилковий пропуск «чужого»);

3. EER (Equal Error Rates) – норма помилок першого та другого роду. Суть – рівна ймовірність помилок першого та другого роду.

Помилку першого роду також називають хибною тривогою, хибним спрацьовуванням, хибною відмовою доступу або хибнопозитивним спрацьовуванням.

Помилку другого роду також називають пропуском події, хибним доступом або хибнонегативним спрацьовуванням.

У реальних БСА, що використовують різні біометричні ознаки особистості, ці помилки дуже різні і лежать у широкому діапазоні від 10^{-1} до 10^{-9} .

Час прийняття рішення є важливим параметром біометричної системи, особливо у випадках, коли необхідно оперативно аутентифікувати великі потоки людей (прохідні великих підприємств та організацій). З цією характеристикою тісно пов'язане поняття масштабованості системи, тобто здатність БСА підтримувати ББД користувачів розмірності, що змінюється. Наприклад, для великої організації характерна висока динаміка зміни кадрового складу. Очевидно, що в таких випадках БСА повинна мати здатність оперативно, без тривалих простоїв коригувати вміст своєї ББД.

Вартість системи включає вартість всіх компонентів БСА, вартість обробки однієї операції, витрати на технічне обслуговування системи, витрати на навчання користувачів і персоналу. Також має бути заздалегідь розрахована вартість обробки виняткових випадків.

Обробка виняткових випадків передбачає спеціальний режим роботи БСА, коли аутентифікаційне рішення приймається вручну, без використання біометричних аутентифікаторів, на підставі інших осіб, що засвідчують дані (наприклад, пропуску або паспорта).

Розрізняють такі види виняткових випадків:

1) суб'єкт не може скористатися БСА (неможливість використання);

2) суб'єкт може потрапити в ту частину популяції, для якої біометричні параметри, що використовуються в БСА, не можуть бути зареєстровані;

3) суб'єкт може мати «поганий біометричний день» (переважно це характерно для БСА, що використовують динамічні біометричні параметри).

Важливим чинником, що впливає ефективну роботу БСА, є число виняткових випадків, яке складно визначити заздалегідь. Так чи інакше, при розробці системи має бути враховано прийнятну величину виняткових випадків та передбачено процедуру їх обробки.

Безпека. Біометричні системи аутентифікації особистості мають справу з високочутливими біометричними даними людей.

Помилкові рішення БСА можуть скомпрометувати особу і ціна таких помилок може бути дуже високою. Ця обставина призводить до необхідності забезпечення високої безпеки самої БСА.

У будь-якій інформаційній системі, у тому числі і БСА, атак найчастіше зазнають найслабші, вразливі місця. Ця закономірність формулюється як принцип найлегшого доступу. Найбільш відоме вразливе місце БСА – це фальсифікація вхідних біометричних даних. Тобто завжди існує деяка, ненульова ймовірність обману БСА за допомогою муляжу (у статичних системах) чи наслідування (у динамічних системах). Муляжі та наслідування імітують біометричні дані їх власника і пред'являються системі за його відсутності. Цей вид атаки на БСА призводить до необхідності використовувати біометричні ознаки, які важко фальсифікувати.

Фальсифікація біометричних даних – це найпростіший шлях атаки на БСА, оскільки він є зовнішнім по відношенню до самої БСА і не пов'язаний з необхідністю ретельного вивчення БСА пошуку вразливостей та шляхів проникнення всередину біометричної системи. На жаль, у БСА може бути досить багато вразливостей і всередині самої системи, які також можуть бути використані зловмисниками для атак, хоча і з меншою ймовірністю, ніж фальси-

фікація біометричних даних.

Безпека БСА забезпечується шляхом усунення вразливостей у точках можливих атак. Це називається захистом «цінних активів» програми. Біометрична аутентифікація повинна бути частиною комплексної системи безпеки додатку, яка включає в себе в тому числі засоби захисту самої біометричної системи. Там, де проблема забезпечення безпеки не є першочерговою, біометрія може застосовуватися просто для більшої зручності використання програми.

Конфіденційність. Анонімність і конфіденційність відносяться до одних із найбільш значимих цінностей у будь-якому вільному суспільстві. Багато захисників свободи особи вважають, що технології біометричної ідентифікації дегуманізують суспільство. Вони побоюються, що біометрія може використовуватися тоталітарними режимами як інструмент контролю над суспільством, тому що вона дозволяє пов'язувати між собою окремі (псевдо) особи, використовуючи їх характерні біометричні характеристики, і в такий спосіб порушує право анонімності. Тому всі технології забезпечення безпеки інформації, у тому числі й біометричні, повинні гарантувати дотримання конфіденційності особи.

Поєднання вимог. Проектування «хорошої» БСА пов'язано з необхідністю задоволення дуже багатьох вимог, які у значній частини несумісні один з одним (рис. 1.2.).

Так, природне прагнення підвищити точність аутентифікації в БСА призводить, як правило, до зниження швидкодії та підвищення вартості системи. Крім того, точність визначається насамперед типом використовуваних біометричних ознак. Однак БСА, що використовують надійні біометричні ознаки, наприклад сітківку ока, дуже дорогі, викликають негативне ставлення та дискомфорт у користувачів.

Прагнення підвищити швидкодію БСА призводить до її подорожчання та додаткових проблем із забезпечення масштабованості.

Підвищення безпеки БСА з метою забезпечення конфіденційності та-

кож призводить до підвищення вартості, а можливо, до зниження швидкодії та/або зручності використання.

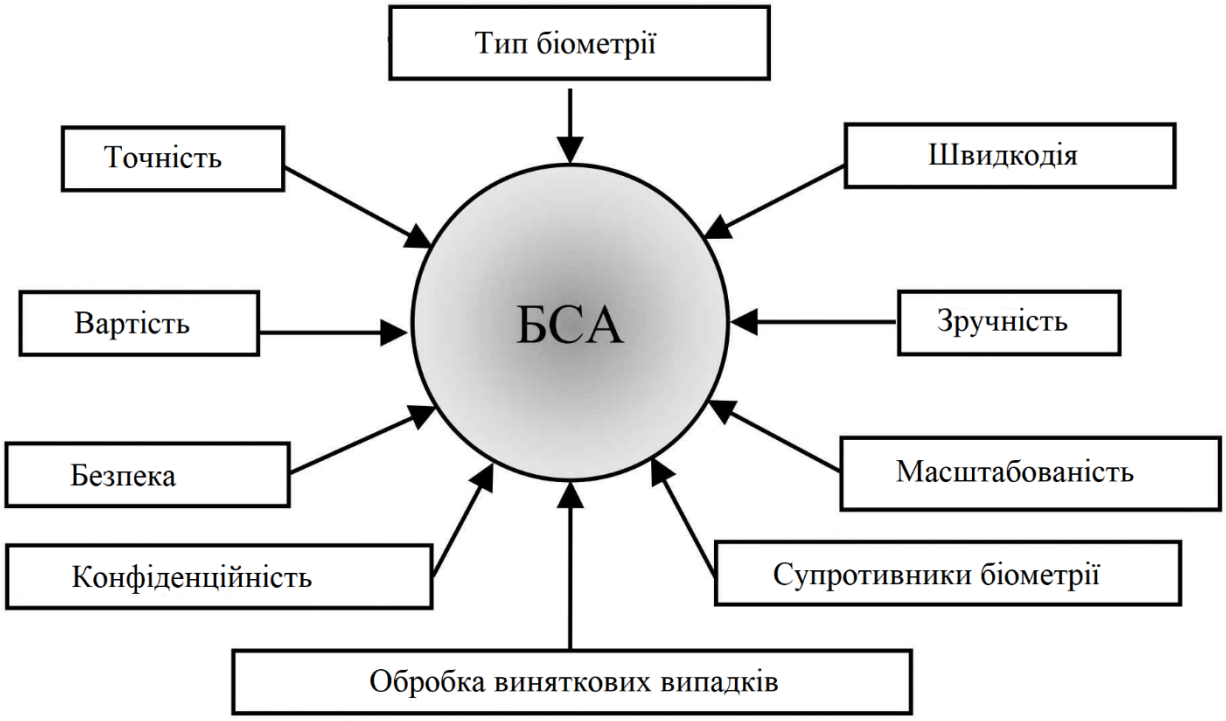

Рисунок 1.2 – Поєднання вимог при створенні якісної БСА

Прагнення зробити БСА зручнішою для користувачів обов'язково позначиться на її точності.

Спроби зниження вартості неминуче призведуть до погіршення більшості інших показників. Таким чином, створення «хорошої» БСА перетворюється на складне багатокритеріальне завдання, оптимальне, а по суті, квазіоптимальне рішення якої може бути знайдено тільки з орієнтацією на конкретний додаток та з розрахунку на широке залучення знань, досвіду, таланту та інтуїції розробників.

1.3. Режими роботи біометричної системи аутентифікації

Режим біометричної верифікації. Особливість БСА, що працює в режимі верифікації, полягає в тому, що аутентифікація здійснюється шляхом зіставлення пред'явлених суб'єктом біометричних характеристик лише з одним зареєстрованим записом ББД. Для реалізації такої процедури суб'єкт крім своїх біометричних характеристик пред'являє БСА деякий додатковий не біометричний ідентифікатор своєї особи, за яким з ББД зчитується біометричний еталон, що відповідає йому.

Зіставлення біометричних характеристик суб'єкта з еталоном може бути реалізовано двома способами: з використанням централізованої чи розподіленої ББД.

Структура БСА, що реалізує режим біометричної верифікації, з урахуванням централізованої ББД наведено на рис. 1.3.

Система складається з біометричних зчитувачів (сенсорів), пристрою виділення біометричних ознак (екстрактора властивостей), пристрою зіставлення (метчера) та централізованої ББД, що містить біометричні зразки всіх зареєстрованих у системі суб'єктів.

У робочому режимі суб'єкт пред'являє свої біометричні характеристики, які зчитуються за допомогою біометричних сенсорів, екстрактор властивостей формує їх машинну репрезентацію у форматі біометричних еталонів, що зберігаються в ББД. Крім біометричних характеристик суб'єкт пред'являє також деякий додатковий ідентифікатор (логін, пароль, ПІН тощо), що дозволяє системі знайти ББД відповідний даному суб'єкту біометричний еталон. Метчер здійснює зіставлення машинної репрезентації біометричних характеристик суб'єкта із вилученим із ББД біометричним еталоном. Результатом зіставлення є у відповідь питання – чи є даний суб'єкт тією особистістю, яку з себе видає.

Централізовані ББД переважно застосовують у БСА, призначених для контролю логічного доступу.

Структура БСА, що реалізує режим біометричної верифікації, з ураху-

ванням розподіленої ББД, наведена на рис. 1.4.

Біометрична база даних такої БСА є сукупністю розподілених серед легальних суб'єктів персональних знімних носіїв (смарт-карт, токенів тощо), що містять біометричні характеристики своїх господарів.

Суб'єкт пред'являє системі свої біометричні властивості і знімний носій, у якому записаний його індивідуальний біометричний стандарт. Для ініціювання транзакції зазвичай використовується ще додатковий ідентифікатор – ПІН. Метчер БСА зіставляє машинну репрезентацію пред'явленій біометрії з біометричним зразком суб'єкта. Результатом зіставлення є відповідь на питання – чи є даний суб'єкт тією особою, яку з себе видає. При цьому обмін інформацією між БСА та знімним носієм, що містить еталон, здійснюється за безпечним протоколом.

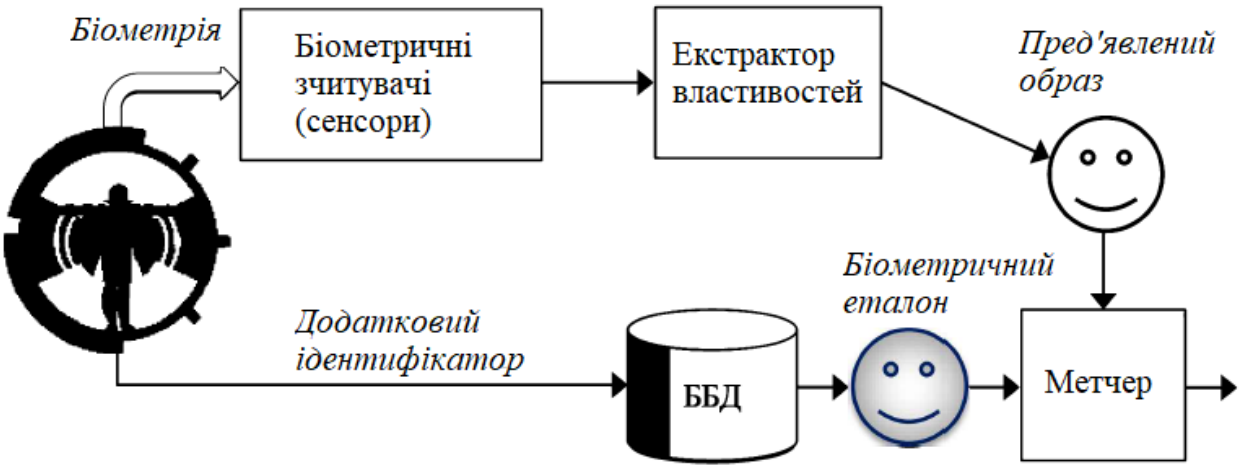

Рисунок 1.3 – Структура БСА, що реалізує режим біометричної верифікації, на основі централізованої ББД

Розподілені ББД застосовуються переважно в БСА, призначених для контролю фізичного доступу до приміщень та на території. Насправді у багатьох БСА використовуються ББД обох типів. Розподілена ББД застосовується для щоденної офлайн-верифікації, а централізована ББД – для онлайн-верифікації чи перевипуску знімних носіїв, у разі втрати, без повторного зняття біометричних параметрів.

Режим біометричної ідентифікації. Особливість БСА, що працює у

цьому режимі, полягає в тому, що аутентифікація особи здійснюється виключно виходячи з пред'явлених суб'єктом біометричних характеристик. На рис. 1.5 наведено структуру БСА, що реалізує режим біометричної ідентифікації.

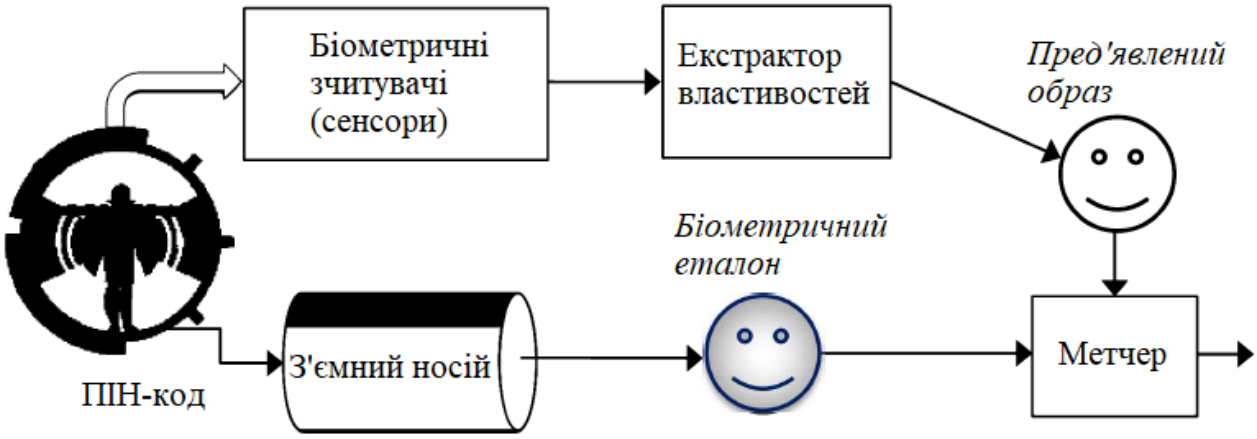

Рисунок 1.4 – Структура БСА, що реалізує режим біометричної верифікації, на основі розподіленої ББД

У робочому режимі суб'єкт пред'являє свої біометричні характеристики, які зчитуються за допомогою біометричних сенсорів, екстрактор властивостей формує їх машинну репрезентацію у форматі біометричних еталонів, які зберігаються в ББД. Метчер починає процедуру послідовного зіставлення пред'явленої машинної репрезентації з кожним біометричним стандартом ББД. Результатом цієї процедури є список ідентифікаторів, які мають найбільший ступінь подібності до пред'явленої репрезентації. Можлива також негативна (нульова) відповідь, що свідчить про відсутність у ББД ідентифікаторів, які мають достатню схожість з пред'явленим ідентифікатором.

Режим біометричної ідентифікації може бути реалізований у двох варіантах:

1. Позитивна ідентифікація. Система визначає – чи зареєстрована дана особа в ББД. При цьому можуть бути допущені помилки хибної відмови доступу та хибного доступу.

2. Негативна ідентифікація. Система перевіряє факт відсутності образу особистості у ББД. При цьому можуть бути допущені помилки хибне ви-

знання та хибне заперечення.

Режим позитивної ідентифікації відповідає традиційному використанню БСА, коли природним, найбільш очікуваним результатом є позитивна відповідь, що свідчить про те, що образ аутентифікованої особи є в ББД (стан «за умовчанням»). Такий режим ідентифікації використовується більшістю БСА для контролю фізичного та логічного допуску суб'єктів.

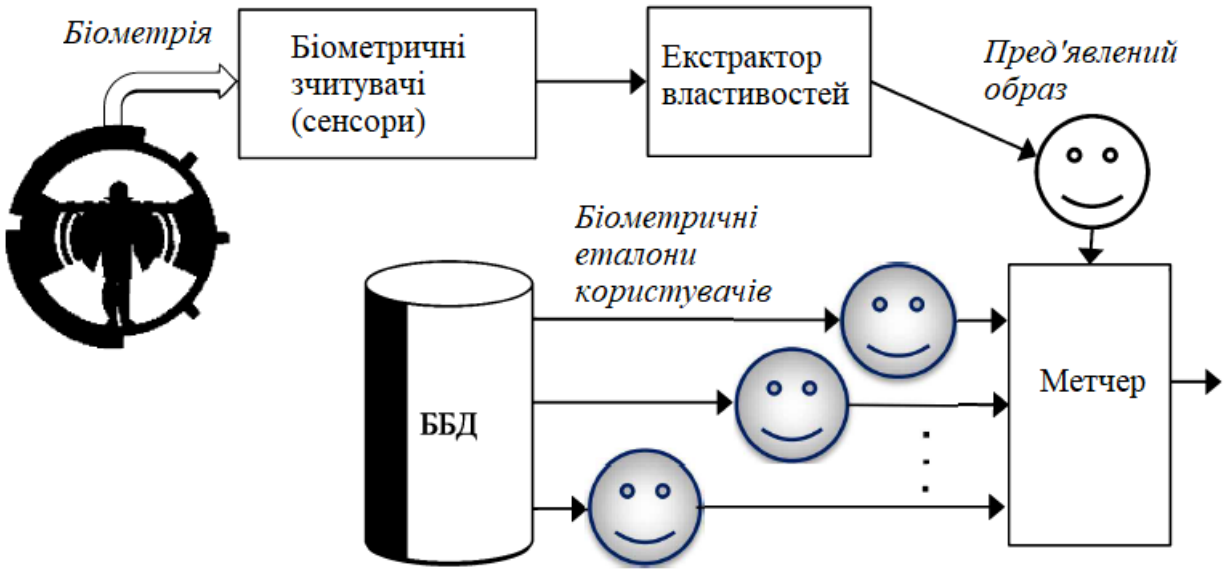

Рисунок 1.5 – Структура БСА, що реалізує режим біометричної ідентифікації

Режим негативної ідентифікації відповідає нетрадиційному використанню БСА, коли природним, найбільш очікуваним результатом є негативна відповідь, що свідчить про те, що образ аутентифікованої особи відсутній в ББД (стан «за умовчанням»). Такий режим ідентифікації використовується для пошуку особи у негативних ББД.

У режимі позитивної ідентифікації бажаним результатом роботи БСА є наявність у ББД єдиного образу особи, що відповідає пред'явленій системі біометрії.

У режимі негативної ідентифікації бажаним результатом роботи БСА є або повна відсутність у ББД образу аутентифікованої особи, або наявність невеликої кількості образів, близьких до пред'явленої системи біометрії, для зручності їх подальшого ручного аналізу.

1.4. Організація біометричного: контролю доступу

Контроль доступу. Процес взаємодії сторін при біометричній аутентифікації повинен чимось регулюватися, тобто проводитися за певними правилами. Ці правила регламентуються протоколами. Протокол – це певна послідовність кроків двох чи більше сторін, які збираються вирішити якесь завдання. Порядок кроків дуже важливий, тому протокол регулює поведінку всіх сторін. Усі сторони погоджуються з протоколом або принаймні розуміють його.

Аутентифікаційний протокол має бути:

1) встановлений заздалегідь. Протокол повністю визначається і розробляється до його застосування. Послідовність проходження протоколу та правила, що регулюють роботу, мають бути визначені. Також мають бути визначені критерії, за якими визначатиметься збіг аутентифікаційних засвідчених даних;

2) взаємно узгоджений. Усі сторони, які беруть участь, повинні бути згодні з протоколом і дотримуватися встановленого порядку;

3) недвозначним. Жодна зі сторін не може порушувати послідовність кроків через їхнє нерозуміння;

4) детальним. Для будь-якої ситуації має бути визначений порядок дій. Наприклад, протоколом має бути передбачена обробка виняткових випадків.

Розраховані на багато користувачів автоматизовані інформаційні системи та комунікації використовуються часто як засоби отримання доступу до послуг, привілеїв і різних додатків. Оператори таких систем зазвичай незнайомі з користувачами, і рішення про надання чи заборону доступу має більшою мірою визначатися без втручання оператора. Користувач не може довіряти операторам та іншим користувачам системи через анонімність реєстрації та віддаленості. Тому необхідні протоколи, за якими дві сторони, які не довіряють одна одній, змогли б успішно взаємодіяти. Ці протоколи, по суті, і регулюватимуть поведінку сторін. Аутентифікація буде проводитися згідно з протоколом між користувачем та системою, в результаті користувач зможе

авторизуватися та отримати доступ до програми.

Будь-яка система контролю доступу має забезпечувати:

1) аутентифікацію. Ця операція гарантує справжність комунікації. У процесі з'єднання система повинна гарантувати, що два суб'єкти є справжніми, тобто кожен суб'єкт є саме тим, ким він заявляє себе. Крім того, система повинна гарантувати, що в комунікацію не втручається третій суб'єкт, який видає себе за легітимну сторону;

2) неможливість відмовитися від авторства. Тобто, коли повідомлення надіслано, одержувач може перевірити, що воно було справді надіслано заявленим автором. Так само, коли повідомлення отримано, відправник може перевірити, що повідомлення справді було отримано заявленим одержувачем.

3) конфіденційність. Процес аутентифікації повинен проводитися конфіденційно, тобто потай від інших користувачів системи.

Схема організації системи контролю доступу наведена на рис. 1.6.

Через користувальницький інтерфейс B за допомогою пристрою введення, наприклад, зчитувального пристрою для смарт-карт і сканера для відбитків пальців, збираються посвідчувальні дані, що надаються суб'єктом A; користувальницький інтерфейс B може включати в себе і пристрій виведення, для того щоб надавати інформацію про якість отриманих біометричних зразків. Система контролю доступу C пропонує кілька сервісів, включаючи аутентифікацію, яка проходить згідно з протоколом. У разі успішного виконання цього протоколу об'єкту надається доступ до програми E за допомогою механічного або логічного перемикача D.

Аутентифікаційні протоколи. Будь-який аутентифікаційний протокол, у якому використовуються різні методи (і різні біометричні ідентифікатори), може бути визначений та виконаний на основі представлених посвідчувальних даних. Аутентифікаційний протокол – це автоматизований процес прийняття рішення: чи справді посвідчувальні дані об'єкта, є достатніми для підтвердження його особи, щоб дозволити йому доступ на основі цих посвідчувальних даних або інших знаків. Розглянемо приклади реалізації доступу.

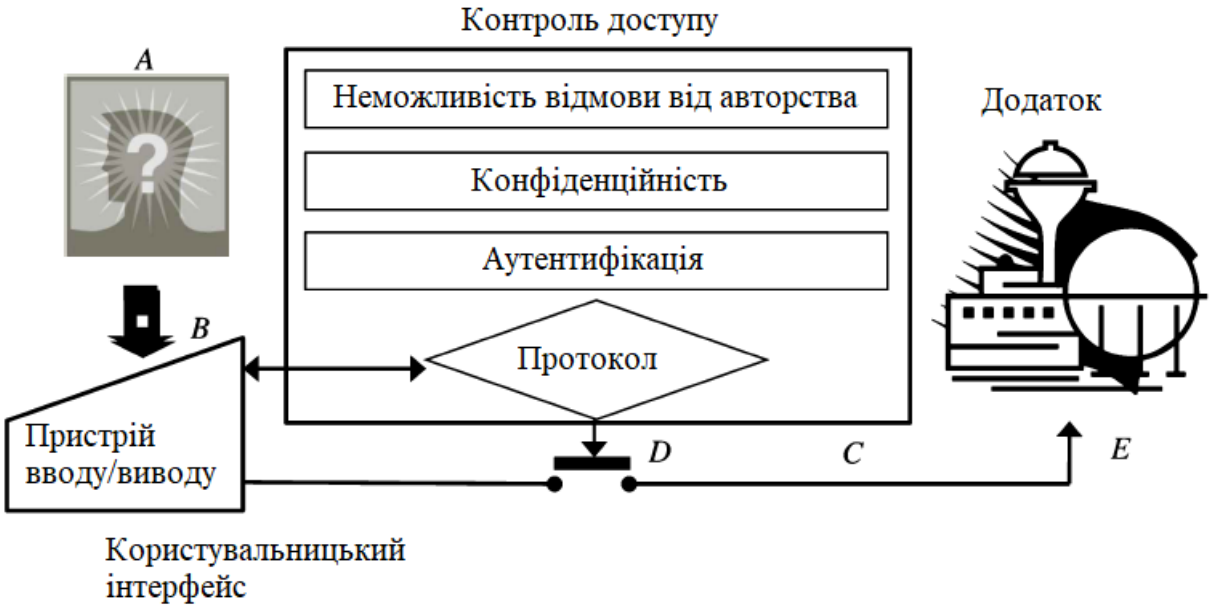

Рисунок 1.6 – Схема організації системи контролю доступу

1. P (property) – те, що ми маємо. Фізичний ключ законному чи незаконному його власнику дасть доступ до певного місця. Бейдж дозволяє його власнику (законному чи незаконному) користуватися послугами будь-якої установи чи заходу. Кредитна картка дає можливість власнику (законному чи незаконному) здійснювати грошові операції.

2. K (knowledge) – те, що ми знаємо. Якщо є схованка, то знання K – це шифр замку, який авторизує людину, щоб вона могла потрапити в схованку; ці знання повністю підтверджують право переміщення об'єктів із схованки. Тут аутентифікаційний протокол простий: користувач демонструє знання секретного коду.

3. P, K – те, що ми маємо, і те, що ми знаємо. Для проведення банківських операцій через банкомат протокол вимагає наявності пластикової картки (P) та ПІН (K). Будь-хто, хто має такі посвідчувальні дані, буде авторизований і зможе здійснювати банківські операції.

4. B (biometrics) – унікальні характеристики користувача. Це суто ідентифікаційний протокол, який вимагає лише представлення біометричних параметрів і не передбачає інших форм взаємодії з інтерфейсом користувача.

Без біометричних параметрів, тобто без інваріантних у часі властивос-

тей особи немає можливості вирішити проблему автоматичної ідентифікації за допомогою сенсорних пристроїв. Тільки біометричні ідентифікатори справді відрізняють одну людину від іншої.

Звідси випливає, що будь-який аутентифікаційний протокол безпеки включає аутентифікаційні посвідчувальні дані, які можна назвати знаками:

1) знаки власності P;

2) знаки знання K;

3) біометричні знаки B.

Тоді реєстрація буде комунікацією між користувачем та системою контролю доступу з метою обміну знаками. Система постачає користувача знаками власності P і знань K, користувач надає системі свої біометричні знаки B (зразки). Цей процес визначається реєстраційною політикою.

Ряд ознак $T = \{x_1 \dots x_n \,/\, x \in (P, K, B)\}$ – це лише частина аутентифікаційного протоколу. Крім них необхідний набір правил A_p, який визначить цей протокол, що використовує T як встановлену послідовність кроків або правил поведінки. Приклад аутентифікаційного протоколу:

$$A_p(T) = A_p(P, K, B). \tag{1.1}$$

З'єднання в аутентифікаційному протоколі декількох аутентифікаційних методів, особливо біометричних B, збільшує достовірність аутентифікації і зменшує ймовірність відмови від авторства або обману.

Наприклад, аутентифікаційний протокол

$A_p(T) = A_p(\{P, K, B_1, B_2\}) = A_p$ показує, які методи як і використовують у протоколі A_p – наборі правил, які у $T = \{P, K, B_1, B_2\}$. Зазначений протокол A_p можна описати приблизно так: хтось, у кого є кредитна карта P з підписом та фотографією, хто може відтворити підпис B_1 , який збігається з підписом на кредитній карті, хто схожий на зображенні на фотографії і, крім того, має знання K – ПІН, може використовувати цю кредитну картку.

Очевидно, що на додаток до ряду ознак $T = \{P, K, B\}$ та правил роботи з ними $A_p(T)$, в аутентифікаційному протоколі має бути визначено, що озна-

чає, якщо дві будь-які ознаки P, K або B вважаються такими, що збігаються.

Зіставлення біометричних зразків. Виконання аутентифікаційного протоколу, безумовно, потребує можливості зіставлення. Посвідчувальні дані P і K можуть бути перевірені шляхом детального порівняння. Біометричні зразки можна порівняти тільки за допомогою технік розпізнавання, так як дві машинні репрезентації, отримані з двох зразків біометричних параметрів β, ніколи не будуть однаковими через присутність шуму і спотворень в процесі їх отримання.

Біометричний шаблон – це машинна репрезентація біометричного зразка B в термінах властивостей (наприклад, відбитки пальців, відстань між очима, довжина пальців). Шаблони часто мають свою спеціальну назву (наприклад, голосовий відбиток, код райдужної оболонки, код пальця, лицьовий образ тощо).

На рис. 1.7 показаний стандартний біометричний метчер, який виконує простий протокол визначення ідентичності двох реальних біометричних параметрів β_1 і β_2, і вирішує питання, чи вони належать одному об'єкту.

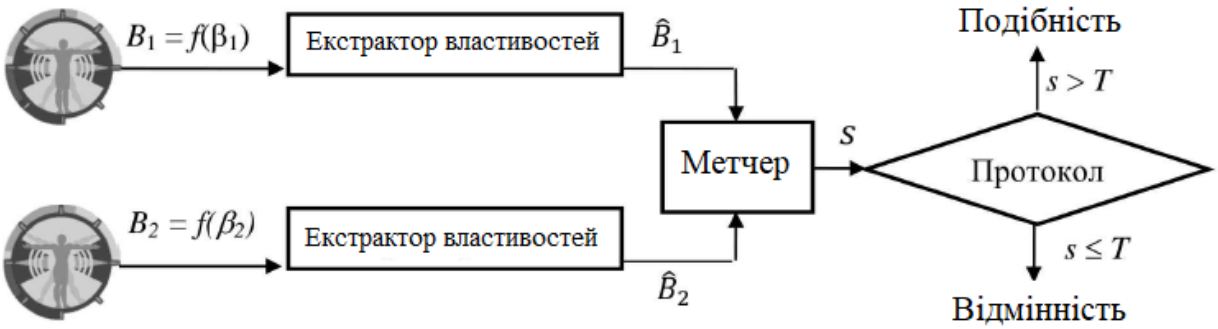

Рисунок 1.7 – Зіставлення біометричних зразків у метчері

Біометричний метчер обчислює величину s, яка виражає ступінь подібності $s(\hat{B}_1, \hat{B}_2)$ між шаблонами, отриманими з біометричних зразків $B_1 = f(\beta_1)$ та $B_2 = f(\beta_2)$.

$$s = s\left(\hat{B}_1, \hat{B}_2\right) = s(B_1, B_2) = s[f(\beta_1), f(\beta_2)]. \quad (1.2)$$

Величина s використовується далі для прийняття рішення, заснованого

на порівнянні величини s з пороговою величиною T:

якщо $s >$ T, значить β_1 і β_2 співпадають;

якщо $s <$ T, означає β_1 і β_2 не співпадають.

З описаного біометричного метчера випливають три визначальних аспекти розробки пристроїв контролю доступу:

1) одержання біометричних зразків або сигналів $B_2 = f(\beta_2)$);

2) визначення функції подібності між двома шаблонами $s(\hat{B}_1, \hat{B}_2)$,

3) встановлення граничної величини прийняття рішення чи подібності.

Далі розглянемо різновиди БСА, у яких використовуються аутентифікаційні протоколи.

Біометрична ідентифікація. Один з можливих варіантів використання біометричних технологій – це автоматизована позитивна ідентифікація суб'єкта d за допомогою порівняння його з m суб'єктами d_i, ББД M $(i = 1, 2, \ldots, m)$. Аутентифікаційний протокол у разі вимагає подання системі лише біометричного параметра β (рис. 1.8).

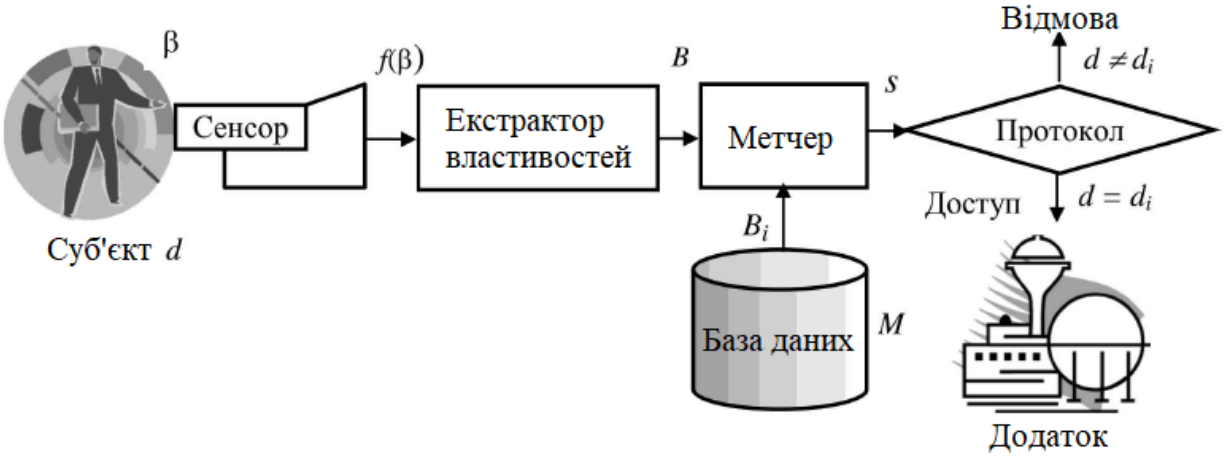

Рисунок 1.8 – Автоматизована позитивна ідентифікація

Суб'єкт d представляє системі реальний біометричний параметр β. Біометричний сенсор знімає з нього зразок $f(\beta)$. З цього зразка за допомогою екстрактора властивостей формується біометричний шаблон B Після цього біометричний метчер визначає подібність $s_i = s(B, B_i)$ між шаблоном B і шаблоном B_i згідно з записами в ББД.

Якщо $s_i > T$, то $d = d_i$ а якщо $s_i \leq T$, то $d \neq d_i$. Цьому критерію може задовольняти безліч $d_i \in M$, і тоді система зможе видати список кандидатів $C = \{d_a, d_b, \ldots\}$ для скорочення списку кандидатів до одного необхідний другий метчер або деякий додатковий біометричний параметр.

Відбір (негативна біометрична ідентифікація). Ця процедура проводиться з метою показати, що ця людина не входить до списку «цікавих» людей (наприклад, список злочинців).

Відбір – аутентифікаційний протокол, який розпоряджає порівняння всіх знаків чи посвідчувальних даних суб'єкта з базою даних, не звертаючись до особи людини (чи довіряючи інформації, що він себе дає). Аутентифікаційний протокол визначає, що суб'єкт може бути аутентифікований, якщо набір його ознак не збігається з наборами ознак, що є у базі даних. Найчастіше ця процедура застосовується для пошуку злочинців. Як ознаки відбору можуть бути і біометричні параметри, проте при поточному рівні розвитку біометрії важко очікувати точного результату, при використанні лише біометричних параметрів. Чисто біометричне зіставлення призведе до появи множини помилок, які буде складно обробити, особливо якщо список дуже великий. З іншого боку, біометричні ідентифікатори можуть використовуватися поряд зі списками, отриманими в результаті параметричних пошуків, які включають ім'я, дату народження і т. д. Такі списки можуть бути досить маленькими, щоб забезпечити прийняття рішення на основі біометричного зіставлення.

Біометрична верифікація. Посвідчувальні дані для біометричної верифікації крім біометричних ознак включають знаки власності та/або знання (ID). Тому, крім подання біометричного ідентифікатора, аутентифікаційний протокол включає уявлення знаків власності та/або знань. Ці додаткові посвідчувальні дані дозволяють однозначно визначити зареєстровану. в базі даних М особистість та пов'язану з нею машинну репрезентацію (шаблон B_i)

Типова схема біометричної верифікації показана на рис. 1.9. Суб'єкт d пред'являє ідентифікаційний номер ID_i і біометричний параметр β для отри-

мання зразка $f(\beta)$. Екстрактор властивостей обчислює біометричний шаблон B, пов'язаний з ID_i і витягує його з бази даних M. Тоді біометричний метчер підраховує значення $s_i = s(B, B_i)$. Якщо це значення (ступінь подібності між B і B_i досить велике, тобто $s > T$, то вважається, що $d = d_i$ і об'єкт допускається до додатку, якщо $s \leq T$, то вважається, що то $d \neq d_i$ і суб'єкт отримує відмову у доступі.

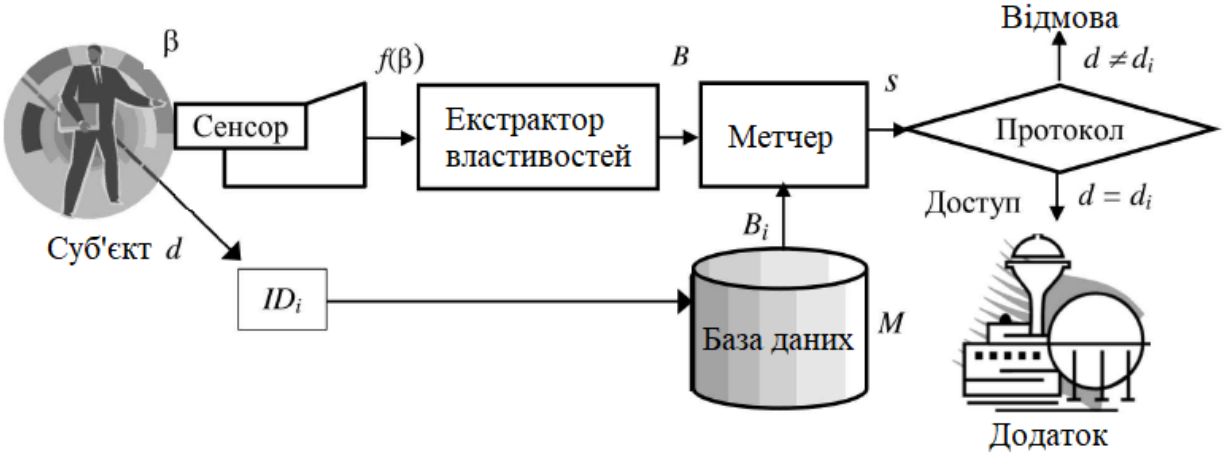

Рисунок 1.9 – Схема біометричної верифікації

Контроль цілісності особи. При позитивному результаті зіставлення біометричних зразків можна бути впевненим (до певної міри), що особа, яка верифікується, і є та сама людина, яка надала дані параметри. Однак через деякий час вже не буде жодних гарантій того, що додатком користується та сама людина. Тому біометрична система може періодично перевіряти ще раз посвідчувальні дані під час транзакцій. Це особливо прийнятно для біометричних параметрів, які можуть бути отримані пасивно без участі людини (наприклад, обличчя і голос), так як повторна перевірка в цьому випадку не вимагає будь-яких дій з боку користувача.

Інші системи можуть аналізувати «постійність особи», щоб контролювати незмінність посвідчувальних даних. Наприклад, місце аутентифікації можна забезпечити відеокамерою. У цьому випадку є можливість переконатися в тому, що аутентифікована людина не залишала місце аутентифікації і що жодна інша людина не замінила її під час проведення транзакцій. Постій-

ність особи можна також перевіряти у процесі візуального спостереження за переміщенням людини у просторі (такі технології вже використовуються). Тобто, поки система спостерігає за людиною, можна бути впевненим, що це той самий суб'єкт, і будь-яка виконана аутентифікація людини поширюється на весь період її відстеження.

Гібридні методи аутентифікації. Здатність одноразово порівнювати параметричні ідентифікатори, такі як власність і знання, і біометричні ідентифікатори, є однією з найважливіших якостей біометричної аутентифікації. Для гібридної аутентифікації використовується одна або кілька ознак $T = \{P, K, B\}$. Для персональної аутентифікації кожну ознаку, яку надає користувач, потрібно порівняти з ознакою, збереженою під час реєстрації. Щоб ухвалити рішення про схожість даних ознак, необхідно інтегрувати результати зіставлення метчерів різної модальності, які верифікують ознаки. При цьому при використанні ознак власності чи знань зіставлення зводиться до перевірки точного збігу.

При обговоренні гібридних методів аутентифікації в контексті біометрії можливі два способи поєднання ідентифікаторів:

1. Об'єднання біометричних параметрів. Запитані посвідчувальні дані T можуть включати різні біометричні параметри, наприклад $\{B_1, B_2\}$, де B_1 – папілярний візерунок пальця, а B_2 – зображення обличчя. Можливість об'єднання кількох біометричних параметрів підвищення ступеня захищеності додатків є нині об'єктом підвищеної уваги розробників БСА.

2. Об'єднання посвідчувальних даних. Найкращим варіантом, безумовно, було б об'єднання двох або більше аутентифікаційних методів. Як було зазначено вище, співвіднесення власності P або знання K з біометричними параметрами зводить завдання біометричної ідентифікації до завдання біометричної верифікації, коли зіставлення здійснюється за принципом 1:1 замість $1{:}\,m$.

Інтеграція різних біометричних параметрів – це проблема, що відноситься до галузі розпізнавання патернів, у той час як поєднання різних не бі-

ометричних аутентифікаційних методів відноситься до традиційних питань інформаційної безпеки додатків. Інтеграція традиційних аутентифікаційних методів із біометричними стала активно використовуватися в контексті інформаційної безпеки додатків.

Таким чином, використання будь-якого з перерахованих гібридних методів P, K або B означає, що повинна існувати можливість зіставлення за допомогою верифікації не біометричних параметрів і порівняння біометричних параметрів. При цьому знаки власності та знання при машинній обробці вимагають точного збігу, а біометричне зіставлення може бути до певної міри приблизним.

2 БІОМЕТРИЧНИЙ КОНТРОЛЬ ДОСТУПУ ЗА ДИНАМІЧНИМИ ХАРАКТЕРИСТИКАМИ

2.1. Особливості динамічних характеристик особистості

Динамічні біометричні характеристики особи засновані на індивідуальних особливостях добре завчених підсвідомих рухів, таких як хода, жестикуляція, голос, рукопис та ін. Виникає питання – чим пояснюється індивідуальність та стійкість динамічних характеристик людини? Правдоподібним поясненням цього є те, що рухи людини реалізуються в багатовимірному просторі її м'язових рухових можливостей, що породжує величезну надмірність варіантів досягнення мети.

У першому наближенні розмірність рухової задачі можна оцінити за кількістю задіяних м'язів. Так, під час письма виявляються задіяними м'язи більшості пальців та частина м'язів передпліччя. При цьому загальна кількість задіяних м'язів може бути більше 50. Навіть зважаючи на те, що найбільш істотну роль грають приблизно 10 м'язів, то виходить, що при письмі людина вирішує в реальному часі 10-вимірне завдання управління.

При роботі з клавіатурою додатково включаються ще приблизно 20 м'язів плеча і плечового пояса по кожній руці, що може задіяти в русі до 140 м'язів. Виходячи з припущення, що найбільший вплив мають лише 20 % від загальної кількості м'язів, отримуємо приблизно 28-мірне завдання управління.

При відтворенні мови беруть участь 44 м'язи грудей, 9 м'язів живота та черевної порожнини, 28 м'язів обличчя та щелеп, 12 м'язів язика, 9 м'язів глотки, 6 м'язів м'якого піднебіння та м'язи гортані. Тобто в сукупності може бути задіяно близько 120 м'язів. Виходячи з припущення, що найбільший вплив мають лише 20 % від загальної кількості м'язів, отримуємо приблизно 24-мірне завдання управління.

З точки зору розв'язання задачі управління, м'язи слід розглядати як

нелінійний по управлінню привод. Це збільшує мірність завдання управління руховими реакціями людини.

У результаті виходить, що людина при реалізації своїх рухових функцій повинна безперервно вирішувати найскладніші багатовимірні завдання управління. Мозок людини неспроможний робити це у процесі різноманітної рухової активності. Природа вирішила цю проблему через навчання. З самого народження протягом багатьох років людина шляхом проб, помилок і тривалих тренувань навчається найбільш оптимально для неї самої виконувати спочатку найпростіші, а потім все більш складні рухи та запам'ятовувати їх. Поступово в людини виробляються свої індивідуальні способи реалізації більшості рухових функцій (програм сукупного управління м'язами): хода, почерк, голос і т.д., якими він потім успішно користується протягом всього життя.

Фактично в людини здійснюється перенесення завдання управління на автоматичний підсвідомий рівень, який пов'язаний з появою індивідуальних особливостей (індивідуального почерку) складних рухів.

Слід зазначити, що підсвідомі рухи стабільні, якщо в них не втручається вищий свідомий рівень. При втручанні свідомості високого рівня відбувається переривання підпрограми управління до моменту прийняття рішення. Ця ситуація добре всім знайома. Якщо при відтворенні свого факсиміле людину хоч на мить щось відволікло, вона не може «правильно» закінчити свій підпис.

Ще одним наслідком багатовимірності завдання управління рухами є унікальність динаміки рухів кожної особи, коли подібних результатів можна досягти зовсім різними сукупностями управляємих впливів. З одного боку, подібна надмірність є гарантією працездатності системи загалом за поганої роботи частини м'язів (наприклад, при пораненнях). З іншого боку, через велику надмірність кожен індивід знаходить свої, оптимальні йому вирішення завдань управління. Одних і тих самих результатів різні люди можуть домагатися, працюючи різними групами м'язів.

Отже, унікальність динамічних портретів наших рухів не є наслідком того, що ми маємо різні голови, руки, ноги, горлянки. Це наслідок закладеної природою надмірності та наслідок випадковості процедури ітераційного статистичного навчання до виконання рухів. Кожен індивід змушений самостійно вирішувати завдання керування своїм тілом у період тривалого навчання. Голос, хода, почерк виробляються все життя і стоять їх власнику дуже великих зусиль. Саме в цьому причина унікальності та стабільності динаміки швидких підсвідомих рухів.

Загальні принципи побудови динамічних БСКД. Біометричні системи, побудовані на аналізі індивідуальних особливостей динаміки рухів (голос, рукописний та клавіатурний почерки), мають багато спільного. Це дозволяє для опису біометричних систем цього класу використовувати одну узагальнену схему (рис. 2.1)

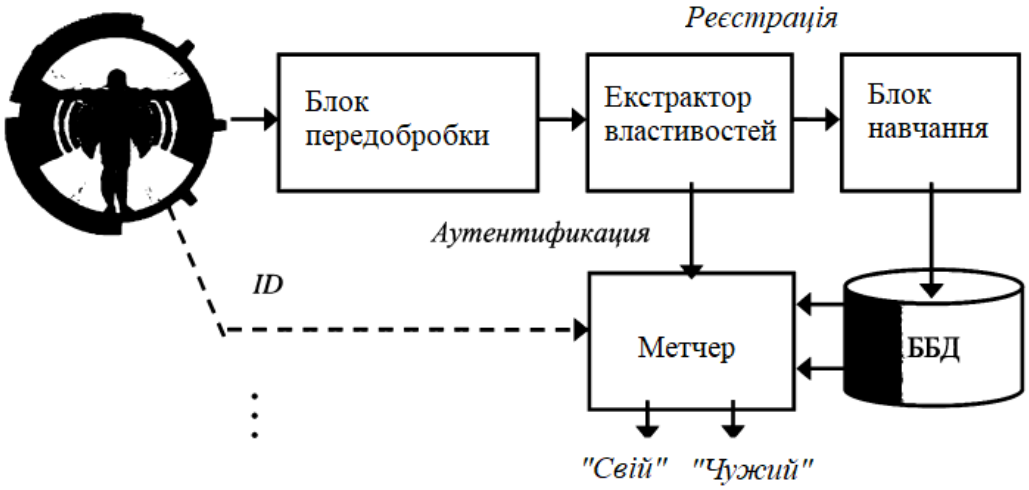

Рисунок 2.1 – Узагальнена структурна схема динамічних БСКД

Динамічні показники на відміну від статичних є функціями часу, що накладає ряд важливих особливостей їх отримання та використання.

Вхідні образи динамічної БСКД будь-якої модальності є сигналами, що відповідають відтворенню голосом, рукописом або набором на клавіатурі певного тексту.

Система має два режими роботи: реєстрації та аутентифікації. В обох режимах вхідні образи надходять у блок попередньої обробки, де вони перет-

ворюються на електричні сигнали, при необхідності оцифровуються і нормуються. Потім екстрактор властивостей з отриманих сигналів витягує значущу для ідентифікації особистості біометричну інформацію - контрольовані біометричні параметри, які подаються у вигляді компактних біометричних препаратів - репрезентацій машинних біометричних образів відповідної модальності. Формат машинних репрезентацій відповідає виду біометричного зразка особи.

В режимі реєстрації особа кілька разів пред'являє свій динамічний біометричний образ. Після його попередньої обробки образу, екстрактор властивостей витягує з нього контрольовані біометричні параметри особи, які надходять до блоку навчання і там усереднюються. Результатом усереднення є машинна репрезентація біометричного образу особи, що виступає далі як біометричний зразок цієї особи. Отриманий зразок заноситься до ББД, де зв'язується у єдиний масив з зразками інших.

Аутентифікація особи залежно від програми реалізується у двох можливих режимах: верифікації та ідентифікації.

В режимі верифікації особа пред'являє системі свій біометричний образ і додатковий ідентифікаційний параметр ID (логін, пароль, ПІН тощо.). Пред'явлений біометричний образ після своєї передобробки надходить у екстрактор властивостей, який витягує з нього контрольовані біометричні параметри особи, представляє їх у вигляді машинної репрезентації та направляє на вхід метчера. Одночасно, за пред'явленим особою ID, що виконує роль адреси, метчер витягує з ББД відповідний біометричний еталон і порівнює його з машинною репрезентацією пред'явленого біометричного образу. Якщо ступінь близькості між машинною репрезентацією пред'явленого біометричного образу і зразком виявляється досить високою (перевищує заданий поріг), метчер виносить рішення «Свій». В іншому випадку, виноситься рішення «Чужий» та відмова у доступі.

В режимі ідентифікації особа пред'являє системі лише свій біометричний образ, який після передобробки надходить до екстрактора властивостей.

Останній витягає з образу контрольовані біометричні параметри особи, представляє їх у вигляді машинної репрезентації та спрямовує на вхід метчера. У розпорядженні метчера в цьому випадку немає інформації у тому, з яким зразком має здійснюватися зіставлення пред'явленого образу. Тому метчер здійснює процедуру послідовного зіставлення машинної репрезентації пред'явленого образу з усіма стандартами, що є в ББД. Результатом цієї процедури є список еталонів, які мають найбільший ступінь схожості з пред'явленим образом. Можлива також негативна (нульова) відповідь, що свідчить про відсутність у ББД еталонів, які мають достатню схожість з пред'явленим образом.

З позиції теорії розпізнавання образів верифікація відповідає задачі класифікації вхідних образів на два класи: «свій»/«чужий». Наявність ID дозволяє здійснювати зіставлення образів за принципом 1:1. Ідентифікація відповідає задачі класифікації вхідних образів на (m+1) класів, де m - число зареєстрованих в БСА користувачів – «своїх», плюс один клас всіх інших не зареєстрованих в БСА користувачів – «чужий». У разі відсутності ID зіставлення здійснюється за принципом 1: m. Реалізація режиму ідентифікації потребує суттєвих додаткових ресурсів, тому застосовується у додатках, коли відсутня можливість використання ID особи (наприклад, ББД злочинців).

Вид використовуваного в метчері вирішального правила та вид біометричного зразка тісно пов'язані. Під час розробки БСКД найчастіше обраний спосіб подання біометричного зразка визначає вид вирішального правила. Але іноді розробники поступають навпаки, відштовхуються від виду вирішального правила, виходячи з якого потім конструюють відповідну машинну репрезентацію контрольованих параметрів та вид біометричного еталона.

У динамічних БСКД біометричний еталон особи створюється з урахуванням усереднення і встановлення границь варіації його біометричних властивостей. Для визначення границь варіації біометричних параметрів використовуються два способи. Вибір визначається кількістю зразків l кожного контрольованого параметра.

При малій кількості зразків l (1-й спосіб) біометричний еталон буду-

ється шляхом прямого вимірювання граничних значень варіації кожного контрольованого параметра v_j.

$$\left(\min_l v_j\,, \max_l v_j\right), \qquad j = 1, 2, \dots N. \tag{2.1}$$

Більшість динамічних параметрів особистості характеризуються розподілом ймовірностей, близьким до нормального. Тому за великої кількості зразків l (2-й спосіб) більш достовірним стає обчислення числових характеристик нормального розподілу контрольованих параметрів: математичних очікувань $m(v_j)$ та дисперсій $\sigma(v_j)$, на основі яких потім визначаються граничні значення варіації кожного параметра:

$$\min_l v_j = m(v_j) - \sigma(v_j), \tag{2.2}$$

$$\max_l v_j = m(v_j) + \sigma(v_j). \tag{2.3}$$

Вибір того чи іншого способу залежить від додатка і типу біометричних параметрів, що використовуються, та впливає на характеристики продуктивності БСКД в режимах реєстрації та аутентифікації.

2.2. Особливості клавіатурного почерку

За останні 30 років активно триваючих досліджень в області клавіатурного почерку численні автори в своїх роботах використовували різні методи, підходи і алгоритми для збору і обробки необхідних статистичних даних, їх подання, класифікації та оцінки ефективності для вивчення можливості побудови програмних додатків на базі великого парку пристроїв (персональних комп'ютерів, ноутбуків з механічними клавіатурами (МК), мобільних платформ (МП) з сенсорними дисплеями (СД), спеціалізованих пристроїв введення), що підвищують надійність аутентифікації.

При цьому найбільш поширеними параметрами для аналізу є наступні:

1) кількість помилок при наборі (частота натискання клавішу delete);

2) звукові сигнали, що відтворюються за допомогою клавіатури при наборі тексту користувачем;

3) час натискання (ЧНК) – це період часу, протягом якого клавіша перебуває у натиснутому стані;

4) час паузи між натисканнями (ЧПК) – це період між натисканнями клавіш;

5) наявність факту утримання однієї із клавіш;

6) відсутність факту утримання однієї із клавіш;

7) наявність факту утримання одночасно двох клавіш;

8) часто використовувані поєднання клавіш – притаманні користувачеві комбінації клавіш клавіатури, що прискорюють його роботу з ІКС;

9) наявність факту використання основної або додаткової частини клавіатури, клавіш Shift або CapsLook при введенні великих літер;

10) швидкість набору – кількість знаків, що набираються в хвилину;

11) загальний час набору парольної фрази;

12) кількість перекриттів між клавішами (накладення клавіш) – відбувається тоді, коли одна клавіша ще не відпущена, а інша вже натискається;

13) ступінь аритмічності під час набору – характеризує рівномірність швидкості набору користувачем символів парольної фрази;

14) сила тиску на клавіші;

15) flight time – період між відпусканням однієї клавіші і відпусканням іншої, що натиснено;

17) Up to Up – період між послідовним відпусканням однієї натиснутої клавіші і потім іншої натиснутої клавіші;

18) положення кистей рук відносно клавіатури;

19) кількість випадків використання додаткових клавіш за одиницю часу;

20) швидкість руху клавіш при натисканні їх користувачем – обчислюється як швидкість зміни ємності контактного датчика у часі;

21) вібрація кнопки при натисканні на неї;

23) частота використання функціональних клавіш і комбінацій;

На рис. 2.2 дві послідовні події клавіатури формують дограф, три події – три граф, n послідовних події – n -граф.

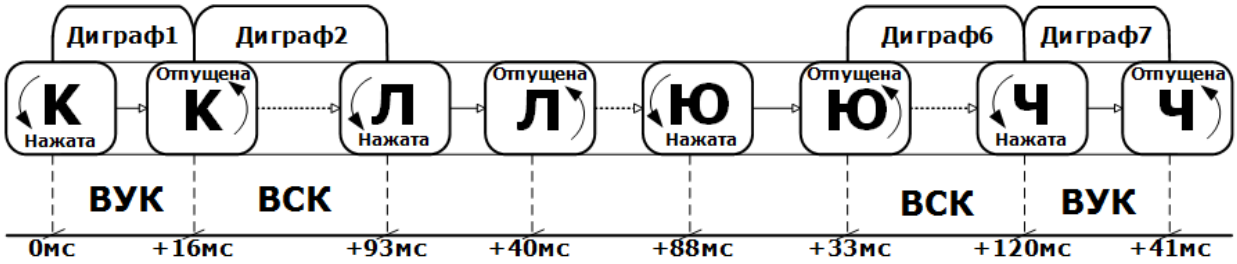

Рисунок 2.2 – Представлення часових подій клавіатури
за допомогою дографів

Більш ранні дослідні роботи містили, як правило, аналіз фіксованого тексту. Так, наприклад, в були представлені ідеї використання поведінкових біометричних особливостей як доповнення до традиційної моделі аутентифікації. Часові параметри натискань клавіш використовувалися для посилення безпеки при введенні пароля фіксованої довжини. Далі з'являлися роботи, в яких в якості вихідних даних використовувався довгий структурований текст певної довжини, а також вільний текст довільної довжини.

Основними показниками якості роботи системи біометричної аутентифікації особистості є помилки трьох видів, виражені в процентному співвідношенні: FRR, FAR та EER. Більш низький показник EER вказує на кращу ефективність системи аутентифікації. Інші критерії оцінки системи включають в себе технологічність, продуктивність і зручність використання.

Однією з найважливіших частин системи аутентифікації, заснованої на аналізі будь-якої поведінкової моделі, є базовий алгоритм обробки отриманих даних. Аналіз літератури в даній предметній області показує, що в більш ранніх роботах більшість методів класифікації представляли собою ймовірносно-статистичні підходи, однак, в даний час дослідники зосередилися на вивченні і апробації підходів по класифікації параметрів клавіатурного почерку, в основі яких лежать сучасні методи машинного навчання.

В рамках ймовірносно-статистичних підходів обробки отриманих да-

них найчастіше застосовуються:

– математичне сподівання та середнє квадратичне відхилення (Mean and STD);

– алгоритми, засновані на обчисленні оцінок подібності між об'єктами, наприклад, метод найближчих сусідів k-NN (k-nearest neighbors algorithm);

– класифікатори, засновані на визначенні геометричних відстаней – Евклідова відстань (Euclidean distance), відстань Махаланобіса (Mahalanobis distance), Манхеттенський відстань (Manhattan distance), відстань Хеммінга (Hamming distance);

– методи, засновані на величині міри ентропії (невпорядкованості) системи;

– алгоритми динамічної трансформації тимчасової шкали (dynamic time warping);

– приховані марковські моделі (hidden Markov model);

– байєсовські класифікатори (Bayes classifier);

– методи дисперсійного аналізу ANOVA (Analysis of variance).

Методи машинного навчання включають в себе:

– штучні нейронні мережі ANN (artificial networks);

– дерева прийняття рішень (decision trees);

– рішення на базі елементів нечіткої логіки (fuzzy logic);

– еволюційне моделювання (evolutionary computation);

– методи опорних векторів (support vector machines).

Архітектура штучних нейронних мереж може бути представлена у вигляді:

– багатошарового перцептрона MLP (multilayer perceptron);

– мережі радіально-базисних функцій RBFN (radial basis function network);

– навчального векторного квантування LVQ (learning vector quantization);

– самоорганізованих карт Кохонена SOM (self-organizing map) або

SOFM (self-organizing feature map).

В якості основи еволюційного моделювання застосовуються:

– генетичні алгоритми GA (genetic algorithms);

– метод рою частинок PSO (particle swarm optimization);

– мурашиний алгоритм ACO (ant colony optimization algorithms).

У табл. 2.1 представлені результати порівняння різних підходів аналізу характеристик клавіатурного почерку, що використовують найбільш поширені методи машинного навчання та ймовірнісно-статистичні моделі.

Таблиця 2.1 – Результати порівняння сучасних методів аналізу характеристик клавіатурного почерку

Рік публікації	Кількість учасників експерименту	Аналізовані параметри	Метод класифікації	Вид та довжина тексту, що вводиться	Пристрій	Результати, %
Методи, засновані на розрахунку середнього значення і дисперсії (Mean and STD)						
2005	205	ЧНК	Mean and STD	довгий	МК	FAR: 0.5 FRR: 5
2009	30	ЧНК, ЧПК	Mean and STD	текст	МП	EER: 13
2009	+1254	ЧНК, ЧПК	Mean and STD	короткий	МК	FAR: 16 FRR: 1
2012	51	ЧНК, ЧПК	Mean and STD	короткий	МК	EER: 8.4
2013	152	ЧНК, ЧПК, тиск	Mean and STD	цифровий	МП	FAR: 4.19 FRR: 4.59
Підходи, засновані на методі k найближчих сусідів (k-NN)						
2002	7	ЧНК, ЧПК,	k-NN	цифровий	СУВ	EER: 78-99
2008	10	тиск	k-NN	цифровий	ТС	EER: 1.00
2010	120	ЧНК, ЧПК	k-NN	короткий	МК	EER: 1.00
2010	100	ЧНК, ЧПК	k-NN	текст	МК	EER: 2.7
2010	30	ЧНК, ЧПК	k-NN	цифровий	МК	EER: 0.5
2013	40	ЧНК, ЧПК	k-NN	довгий	МК	EER: 6.1

Продовження таблиці 2.1

Рік публікації	Кількість учасників експерименту	Аналізовані параметри	Метод класифікації	Вид та довжина тексту, що вводиться	Пристрій	Результати, %
Методи, засновані на визначенні геометричних відстаней (Euclidean distance)						
2007	21	ЧНК, ЧПК	Euclidean distance	короткий	МК	EER: 3.8
2008	30	ЧНК, ЧПК тиск	Euclidean distance	цифровий	СУВ	FAR: 15 FRR: 0 EER: 10
2009	16	ЧНК, ЧПК	Euclidean distance	короткий	МК	EER: 4.28
2010	100	ЧНК, ЧПК	Euclidean distance	текст	МК	EER: 2.7
2010	189	ЧНК, ЧПК	Euclidean distance	довгий	МК	FAR: 0.01 FRR: 3
2011	51	ЧНК	Euclidean distance	довгий	МК	EER: 0.84
2011	20	ЧНК	Euclidean distance	довгий	МК	FAR: 2 FRR: 4
Методи, засновані на деревах прийняття рішень (Random forest decision tree, RFDT)						
2010	21	ЧНК, ЧПК,	RFDT	довгий	МК	FAR: 3.47 FRR: 0 EER: 1.73
2010	28	ЧНК, ЧПК,	RFDT	цифровий	МК	FAR: 0.03 FRR: 1.51 EER: 1
Методи, засновані на штучних нейронних мережах (ANN)						
2007	100	ЧНК, ЧПК,	ANN	короткий	МК	FAR: 1 FRR: 8
2010	25	ЧНК, ЧПК, тиск	ANN	цифровий	МК	FAR: 4.12 FRR: 5.55
Підходи, засновані на величині ентропії системи (Entropy)						
2005	31	ЧНК	Entropy	довгий	МК	FAR: 1.99 FRR: 2.42
2005	205	ЧНК	Entropy	довгий	МК	FAR: 0.5 FRR: 5
2009	21	ЧНК, ЧПК	Entropy	довгий	МК	FAR: 0.14 FRR: 1.59
2011	50	ЧНК	Entropy	довгий	МК	EER: 10
2011	186	ЧНК, ЧПК	Entropy	довгий	МК	FAR: 1.65 FRR: 2.75

Продовження таблиці 2.1

Рік публікації	Кількість учасників експерименту	Аналізовані параметри	Метод класифікації	Вид та довжина тексту, що вводиться	Пристрій	Результати, %
Підходи, засновані на інших статистичних методах, а також їх комбінаціях						
1990	26	ЧНК	Baysian, Minimum Distance	короткий	МК	FAR: 2.8 FRR: 8.1
2004	41	ЧНК, ЧПК	Gaussian mixture modeling	короткий	МК	FAR: 4.3 FRR: 4.8 EER: 4.4
2005	9	ЧНК, ЧПК, тиск,	ANOVA	цифровий	МК	EER: 2.4
2006	100	ЧНК, ЧПК, тиск,	Dynamic time warping	цифровий	МК	EER: 1.4
2006	20	ЧНК, ЧПК	Hidden Markov model	цифровий	МК	EER: 3.6
2006	20	ЧНК, ЧПК	Euclidian, Mahalanobis	цифровий	ТС	FAR: 0 FRR: 2.5
2009	25	ЧНК, ЧПК	Gauss, Parzen, K-NN, K-mein	короткий	МК	EER: 1.00
2009	100	ЧНК, ЧПК	Bayesian, Euclidean, Hamming	короткий	МК	EER: 6.96
2010	51	ЧНК, ЧПК	Manhattan distance	короткий	МК	EER: 7.16
2011	100	ЧНК, ЧПК	Gaussian PDF	короткий	МК	EER: 1.401
2011	55	ЧНК, ЧПК	Spearmans's foot rule distance	довгий	МК	FAR: 2.02 FRR: 1.84
2011	33	ЧНК, ЧПК	Naive Bayesian	довгий	МК	EER: 1.72
2013	152	ЧНК, ЧПК, тиск, датчики	k-mean	цифровий	МП	FAR: 4.19 FRR: 4.59
2013	10	ЧНК, ЧПК, датчики	Bayesian	цифровий	ТС	FAR: 0.02 FRR: 0.018
2014	30	ЧНК, ЧПК	SMD, SED	цифровий	МК	EER: 26

Продовження таблиці 2.1

Рік публікації	Кількість учасників експерименту	Аналізовані параметри	Метод класифікації	Вид та довжина тексту, що вводиться	Пристрій	Результати, %
Підходи, засновані на методі опорних векторів (SVM)						
2007	24	ЧНК, ЧПК	SVM	довгий	МК	FAR: 0.76 FRR: 0.81 EER: 1.57
2007	61	ЧНК, ЧПК	SVM	короткий	-	FAR: 14.5 FRR: 1.78
2007	5	тиск	SVM	цифровий	МК	FAR: 0.95 FRR: 5.6
2011	117	ЧНК, ЧПК	SVM	короткий	МК	EER: 11.8
2014	30	ЧНК, ЧПК, тиск	SVM	цифровий	СД	EER: 2.8
Підходи, засновані на інших менш поширених методах машинного навчання						
2005	43	ЧНК, ЧПК	Decision trees, Monte Carlo	короткий	МК	FAR: 0.88 FRR: 9.62
2005	53	ЧНК, ЧПК	Fuzzy ARTMAP	короткий	МК	FAR: 0.87 FRR: 4.4
2007	30	ЧНК, ЧПК	Sequence alignment algorithms	короткий	МК	FAR: 0.2 FRR: 0.2 EER: 0.4
2014	42	ЧНК, ЧПК, тиск, датчики	Naive, Bayesian	короткий	МП	EER: 12.9
2015	42	ЧНК, ЧПК, тиск, датчики	Two-class	короткий	МП	EER: 3

1. Аналізуючи наведені в табл. 2.1 дані, можна зробити наступні висновки.

1. Кінцева точність отриманих результатів визначається цілою низкою факторів, головними серед яких є: основний алгоритм класифікації отриманих даних, кількість учасників експерименту з різною величиною досвіду роботи з клавіатурою, спосіб і організація введення даних і апаратна платформа, на базі якої проводиться тестування системи аутентифікації.

2. Сукупність факторів (зашумлень), що впливають на точність іденти-

фікації:

1) емоційний стан;

2) залежність особливостей роботи користувача від характеру виконуваних задач та часу доби;

3) ступінь покриття символами, що використовуються для вводу парольної фрази, поля клавіатури;

4) можливість набору користувачем парольної фрази однією рукою (або одним пальцем) за його бажанням або через травму кінцівки;

5) технічні характеристики клавіатури: форма (пряма, ергономічна і т. д.), ступінь легкості натискання клавіші, розташування клавіш (QWERTY, AZERTY і т. д.);

6) залежність між стабільністю клавіатурного почерку та рівнем користувача;

7) нерегулярний характер роботи користувача за комп'ютером;

8) залежність точності вимірювання таймера, що захоплює час, протягом якого відбулася подія на клавіатурі, від встановленої на комп'ютері операційної системи та мови програмування.

3. Використання комбінацій ймовірносно-статистичних підходів аналізу особливостей клавіатурного почерку в рамках одного алгоритму дозволяє підвищити рівень точності системи в цілому.

4. Використання додаткових просторових параметрів клавіатурного почерку (тиск, координати пальців на сенсорних дисплеях мобільних пристроїв) дають значний приріст в показниках точності в порівнянні зі стандартними механічними клавіатурами за рахунок виділення додаткових інформативних ознак аналізованого почерку суб'єкта.

5. Суттєво підвищити якість систем аутентифікації користувачів ПК можна шляхом переходу до комплексних моделей, які враховують, наприклад, такі характеристики:

• дані, які дозволять однозначно ідентифікувати користувача (унікальний ідентифікатор, пароль, цифрові підписи обладнання тощо);

• інформаційний почерк користувача – клавіатурний почерк і динаміку системи «користувач-миша»;

• активність користувача в рамках операційної системи (середній відсоток використання центрального процесора, середній обсяг займаної пам'яті, тип найбільш часто відкриваються файлів і т.д.);

• мережева активність користувача (найбільш часто використовувані мережеві сервіси та додатки, тип активності користувача в мережі і т.д.);

• програмно-апаратні зміни в конфігурації ПК (установка нового програмного забезпечення, установка або заміна зовнішніх пристроїв і т.д.).

3 ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ ІДЕНТИФІКАЦІЇ ОСОБИСТОСТІ ЗА КЛАВІАТУРНИМ ПОЧЕРКОМ З УРАХУВАННЯМ СИЛИ ТИСКУ НА КЛАВІШІ

3.1. Queen Mary University Keystroke benchmark dataset

Датасет «Queen Mary University Keystroke benchmark dataset» [26, 27] було опубліковано у 2017 році. Датасет складається з двох датасетів.

Перший датасет містить параметри тиску користувачем на клавіші в процесі вводу паролю «.try4-mbs». Кількість користувачів, що брали участь в експерименті – 30 користувачів. Кількість вводів паролю – від 487 (користувач 24) до 907 (користувач 3) раз. Дані представляють файл Excel format з 11 стовпчиків. Поле «user» це ідентифікатор користувача. Останні 10 стовпчиків – параметри тиску на клавіші в процесі вводу паролю – *p_k1, p_k2, p_k3, p_k4, p_k5, p_k6, p_k7, p_k8, p_k9, p_k10* (десята клавіша – enter).

Другий датасет містить часові параметри вводу паролю «.try4-mbsz», який набирали ті ж 30 користувачів. Тут кількість вводів паролю однакова – 400 раз. Дані представляють файл Excel format з 32 стовпчиками. Поле «user» це ідентифікатор користувача. Останні Останні 28 стовпчик – часові параметри (в одиницях системного часу) вводу парольної фрази: *H_k1, DD_k1k2, UD_k1k2, H_k2, DD_k2k3, UD_k2k3, H_k3, DD_k3k4, UD_k3k4, H_k4, DD_k4k5, UD_k4k5, H_k5, DD_k5k6, UD_k5k6, H_k6, DD_k6k7, UD_k6k7, H_k7, DD_k7k8, UD_k7k8, H_k8, DD_k8k9, UD_k8k9, H_k9, DD_k9k10, UD_k9k10, H_k10.* Тут «*H.kA*» – час натискання клавіші A, «*DD_kAkB*» – час між натисканням клавіш A та B, «*UD_kAkB*» – час між відпусканням клавіші A та натисканням клавіші B.

В якості платформи розробки програмно-апаратного комплексу для реєстрації додаткових ознак клавіатурного почерку було використано контролер Arduino Uno R3, який побудований на чіпі ATmega328, що забезпечує перетворення аналогового сигналу в цифрову форму за допомогою вбудова-

ного АЦП, і може використовуватися для розробки інтерактивних систем, керованих різними датчиками та перемикачами. Зведені характеристики програмованого контролера Arduino Uno R3 представлені на рис. 3.1.

Рисунок 3.1 – Основні характеристики програмованого контролера Arduino Uno R3

Структурну схему програмно-апаратного комплексу для реєстрації додаткових ознак клавіатурного почерку від датчиків тиску представлено на рис. 3.2.

До Arduino Uno R3 було підключено 5 сенсорів тиску. Для визначення сили натискання на клавіші використаний датчик тиску Interlink 408 FSR, який являє собою силовимірювальний резистор, виконаний у вигляді тонкого плоского пасивного компоненту, опір якого пропорційний зусиллю, що діє на його поверхню. Без навантаження опір перевищує 1 МОм і варіюється від 100 кОм до кількох сотень Ом залежно від сили натискання на поверхню датчика. Для визначення кодів клавіш та моментів їх натискання використано модуль USB Host Shield, призначений для підключення пристроїв HID та емуляції їх роботи в операційній системі. Підсилювачем аналогового сигналу

служить модуль на основі мікросхеми LM358. До Arduino Uno R3 через USB Host Shield підключено клавіатуру Logitech K120.

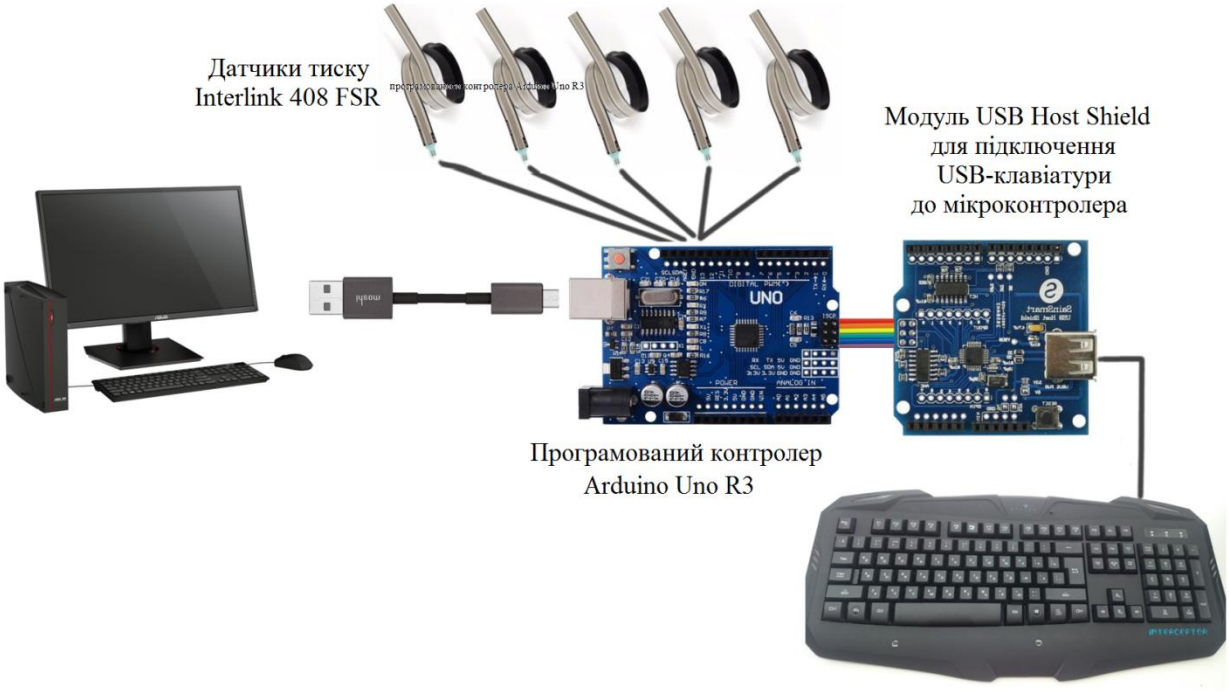

Рисунок 3.2 – Структурна схема програмно-апаратного комплексу визначення параметрів клавіатурного почерку (час та тиск)

Корпус клавіатури був розкритий і під ряди її клавіш встановлені датчики тиску (рис. 3.3). Датчики тиску також підключаються до Arduino Uno R3.

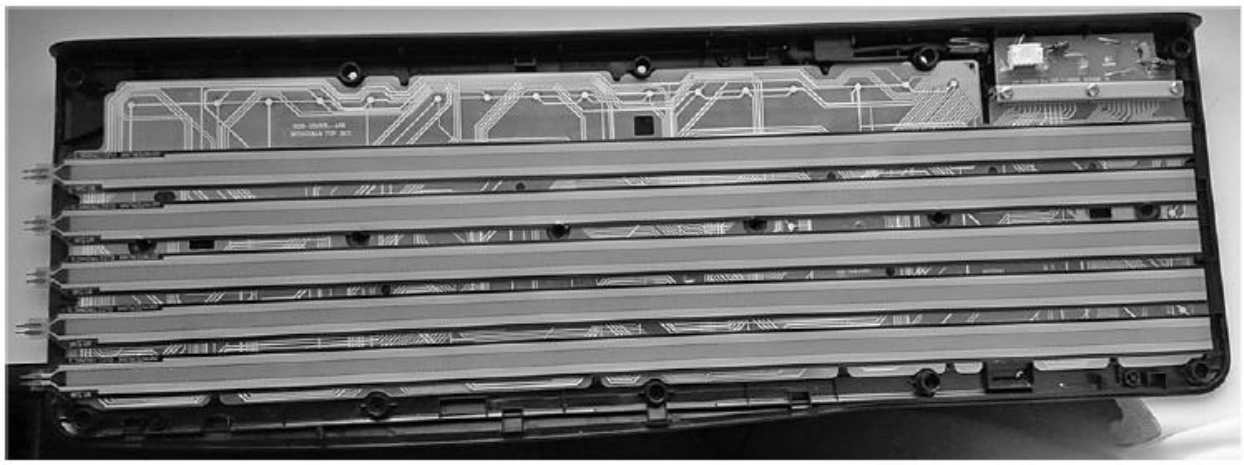

Рисунок 3.3 – Розташування датчиків тиску на клавіатурі

Частота опитування датчиків мікроконтролером Arduino Uno R3 становить 3000 Гц, але оскільки послідовно опитуються 5 каналів (датчиків), реальна частота дискретизації кожного з реєстрованих сигналів становить 600 Гц. Щоб оцінити максимальну інформативну частоту сигналів, які формуються під час набору тексту на клавіатурі, потрібно орієнтуватися на найвищу можливу швидкість друку користувача на клавіатурі. Використовуючи клавіатуру Дворака (варіант розкладки клавіатури, що забезпечує більш високу швидкість набору тексту в порівнянні з традиційною розкладкою QWERTY), в 2005 Барбара Блекберн (Barbara Blackburn) встановила світовий рекорд зі швидкості набору тексту англійською мовою. Вона друкувала із середньою швидкістю 150 слів за хвилину протягом 50 хвилин, часом її швидкість піднімалася до 170 слів за хвилину, а в короткий проміжок часу досягла швидкості 212 слів за хвилину. В англійській мові середня довжина слова дорівнює 5.2 літери, проте WPM – кількість слів за хвилину – нерідко дорівнює 5 символам (в інших країнах швидкість набору вимірюється також у CPM – символах за хвилину або у SPM – ударах за хвилину). Таким чином, рекорд Барбари Блекберн становив близько 750 символів за хвилину. За іншими оцінками, нормальною швидкістю набору для клавіатури з розкладкою QWERTY вважається 150 – 200 символів на хвилину, високою – 250 – 300 символів. Максимальній швидкості набору на клавіатурі, зазначеній у відкритих джерелах, відповідає частота 12.5 Гц, нормі – 2.5 Гц. Відповідно до теореми відліків частота дискретизації повинна бути вдвічі вищою за частоту сигналу. З цього виходить, що частоти дискретизації сигналів 25 Гц цілком достатньо для фіксації всіх частотних змін, що відбуваються в клавіатурному почерку людини.

Зразки клавіатурного почерку формувались розробленим програмним модулем, в результаті кожен зразок був перетворений на вектор значень ознак.

3.2. Схема експерименту та отримані результати

Експериментальні дослідження проводились в пакеті Orange Data Mining. Оскільки задача ідентифікації математично є задачею класифікації, тобто задачею розбиття множини об'єктів (тестові вектори біометричних ознак) на апріорно задані класи (імена користувачів), всередині кожного з яких вони вважаються схожими один на одного, та мають приблизно однакові властивості й ознаки (вектори біометричних характеристик одного користувача дуже схожі один на одного), то в якості алгоритму класифікації користувачів використовувався метод Random forests. Кількість дерев прийняття рішень дорівнювала 30 – адже саме стільки користувачів в дослідних датасетах.

Точність класифікації перевірялась за вбудованим у віджет «Test and Score» алгоритмом 10-fold cross-validation, у відповідності до якого дослідний набір даних розбивається на 10 однакових за розміром блоків. З 10 блоків один залишається для тестування моделі, а інші 9 блоків використовуються як тренувальний набір. Процес повторюється 10 разів, і кожен з блоків використовується один раз як тестовий набір. Наприкінці аналізу отримують 10 результатів, по одному на кожен блок, вони усереднюються і дають одну оцінку. Перевага такого способу в тому, що всі спостереження використовуються і для тренування, і для тестування моделі, і кожне спостереження використовується для тестування в точності один раз.

Результати мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за часовими ознаками клавіатурного почерку наведено на рис. 3.4 та рис. 3.5.

Як можна бачити з рис. 3.4, інтегральна помилка першого роду FRR становить:

$$FRR = 1 - Recall = 1 - 0.933 = 6.7\ \%,$$

інтегральна помилка другого роду FAR становить:

$$FAR = 1 - Specificity = 1 - 0.998 = 0.2\ \%.$$

Як можна бачити, рівень FAR відповідає високому рівню ідентифікації, а рівень FRR – не відповідає.

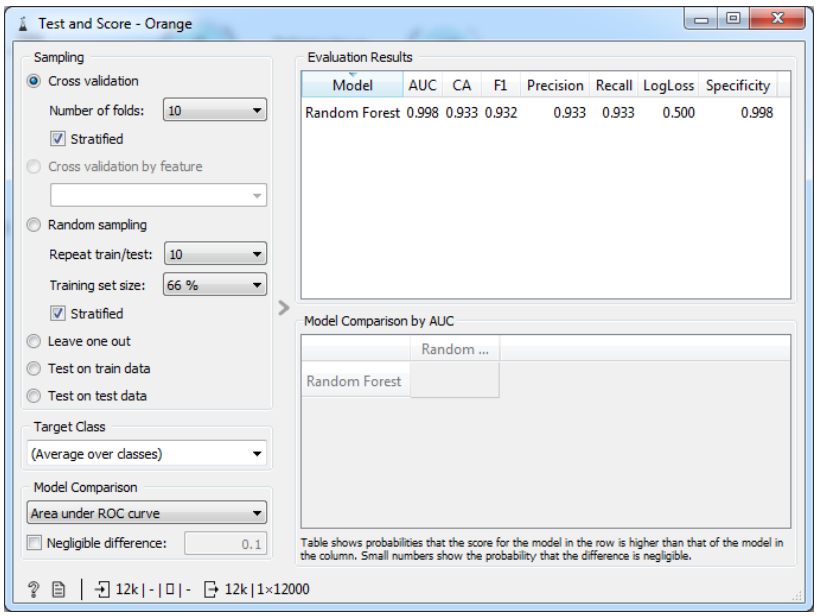

Рисунок 3.4 – Результати мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за часовими ознаками клавіатурного почерку

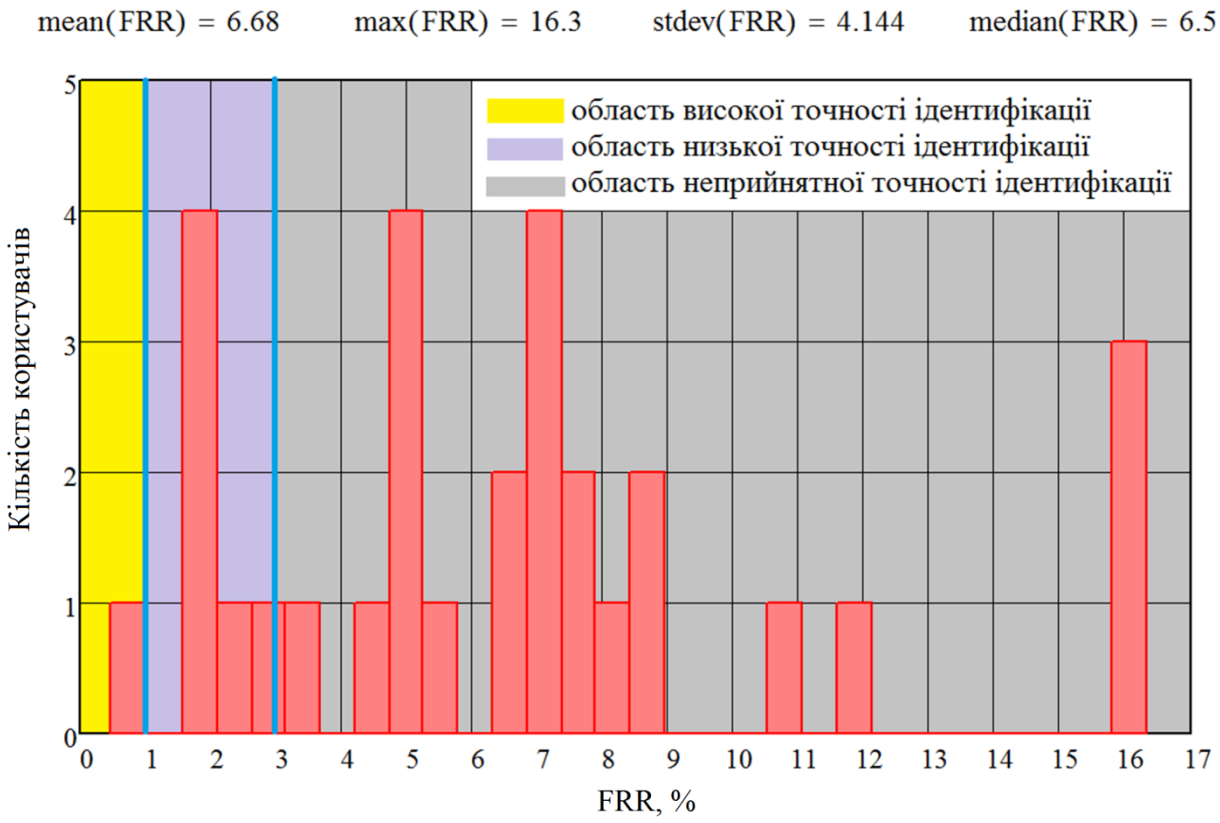

Рисунок 3.5 – Гістограма помилок першого роду мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за часовими ознаками клавіатурного почерку

Результати мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за ознаками тиску на клавіші наведено на рис. 3.6 та рис. 3.7 (кольорві позначення аналогічні рис. 3.5).

Як можна бачити з рис. 3.6, інтегральна помилка першого роду FRR становить:

$$FRR = 1 - Recall = 1 - 0.993 = 0.7\ \%,$$

інтегральна помилка другого роду FAR становить:

$$FAR = 1 - Specificity = 1 - 1 = 0\ \%.$$

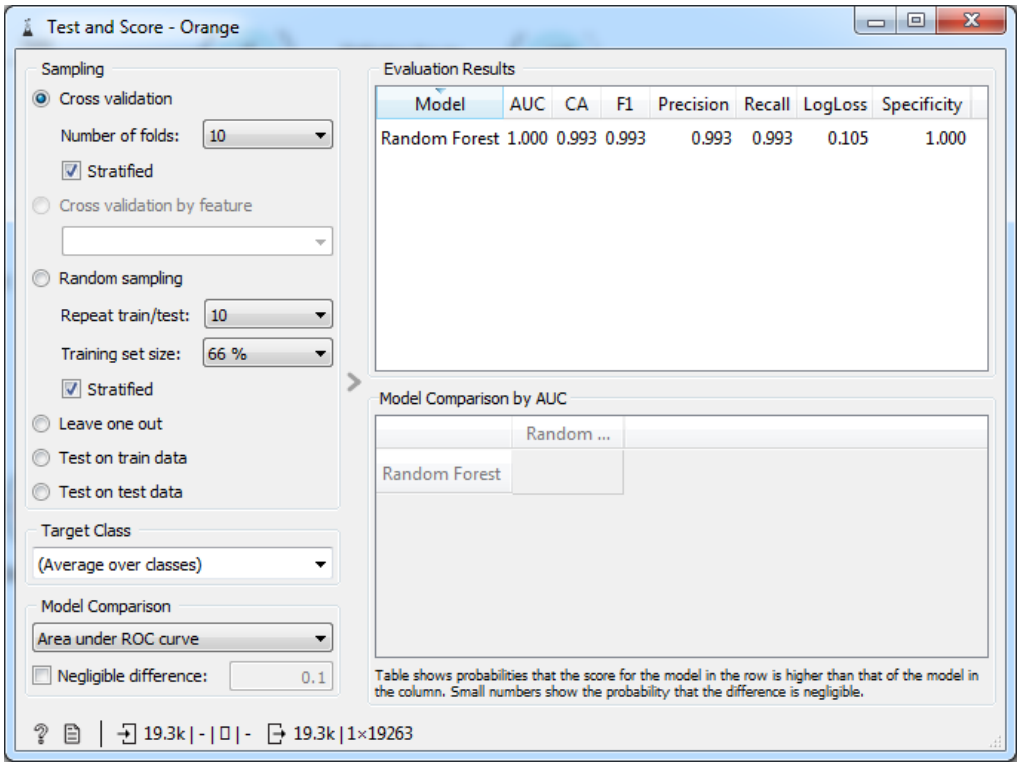

Рисунок 3.6 – Результати мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за ознаками тиску на клавіші

Як можна бачити, рівень помилки другого роду відповідає дуже високому рівню ідентифікації. Для 27 з 30 користувачів (90 %) значення FAR дорівнює нулю, тобто система абсолютно точно визначає зловмисника та блокує йому доступ до інформаційної системи. Для 3 користувачів рівень FAR склав 0.1 %,тобто для одного випадку з тисячі система надасть доступ зловмиснику. Графік гістограми значень помилки першого роду також ілюструє

високу точність ідентифікації – для 24 користувачів (80 %) значення FRR становить менше 1 %, що відповідає високій точності, ще для 5 користувачів значення FRR лежить в зоні низької (або прийнятної) точності, і лише для одного користувача значення FRR більше 3 %, що відповідає неприйнятній точності ідентифікації.

Отже, перший висновок: для задач мультикласової класифікації ознаки тиску на клавіші є набагато більш інформативними, ніж часові параметри натискань на клавіші. За умови великої навчальної вибірки точність ідентифікації зловмисника може становити 100 %, а точність ідентифікації зареєстрованих користувачів – менше 1 %.

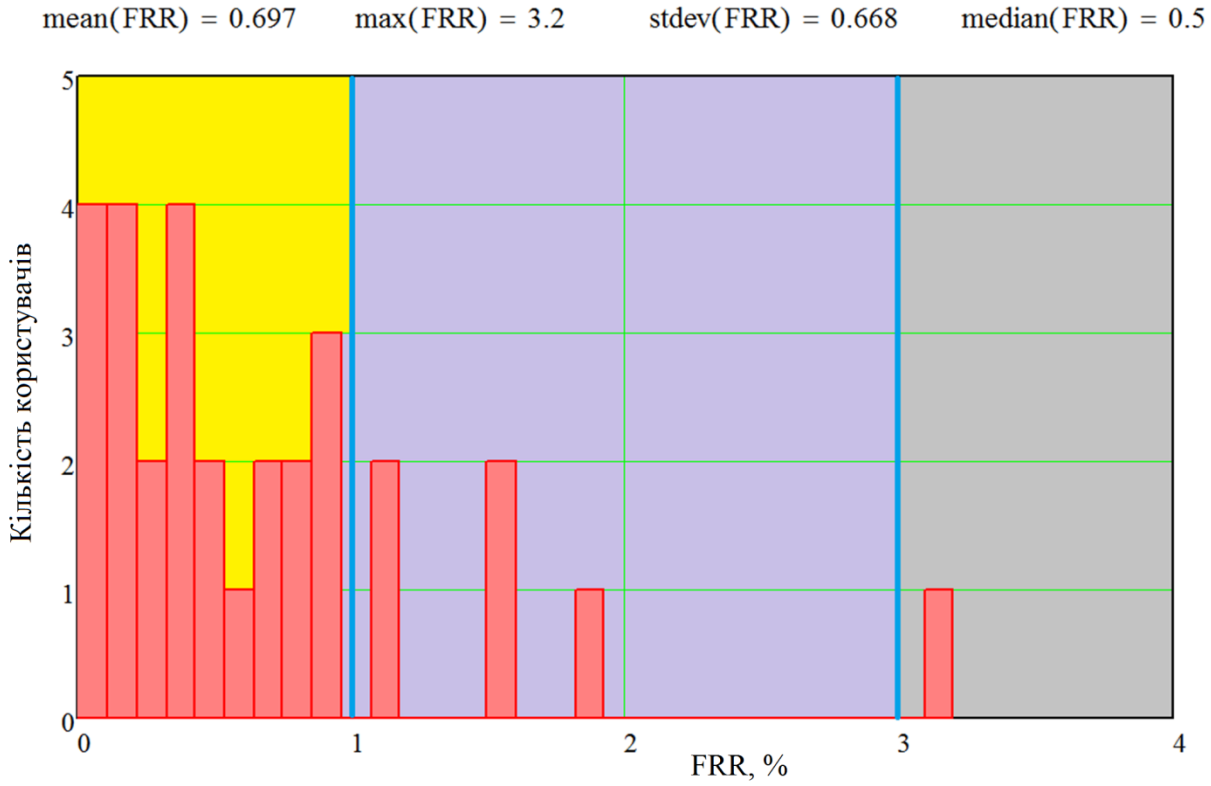

Рисунок 3.7 – Гістограма помилок першого роду мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за ознаками тиску на клавіші

Наступним кроком є дослідження точності двокласної класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за ознаками тиску на клавіші. Для експерименту було обрано 8 користувачів:

user26 (FRR=3.2 %, FAR=0.1 %), user29 (FRR=1.8 %, FAR=0.1 %), user12 (FRR=1.6 %, FAR=0.1 %), user23 (FRR=1.6 %, FAR=0 %), user22 (FRR=0 %, FAR=0 %), user28 (FRR=0 %, FAR=0 %), user2 (FRR=0.5 %, FAR=0 %), user16 (FRR=0.5 %, FAR=0 %). Результати класифікації наведено в табл. 3.1.

Таблиця 3.1 – Результати двійкової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за ознаками тиску на клавіші

FRR / FAR, %							
	user12	user16	user22	user23	user26	user28	user29
user2	0.0/0.0	0.0/0.2	0.0/0.0	0.3/0.0	0.4/0.0	0.4/0.0	0.3/0.0
user12		0.0/0.0	0.0/0.3	0.4/0.0	0.0/0.0	0.0/0.0	0.0/0.0
user16			0.2/0.2	0.2/0.0	0.3/0.0	0.0/0.0	0.0/0.0
user22				0.0/0.0	0.0/0.0	0.0/0.0	0.0/0.0
user23					0.4/0.0	0.0/0.4	0.2/0.0
user26						0.0/0.4	0.0/0.0
user28							0.4/0.0

Як можна бачити з табл. 3.1, для всіх можливих пар користувачів значення помилок першого та другого родів менше 0.5 відсотка, тобто відповідають високій точності ідентифікації.

Отже, другий висновок: у випадку двокласової класифікації на основі ознак тиску на клавіші можна будувати основну систему контролю та управління доступом.

Наступним кроком є дослідження точності ідентифікації на основі комбінованих ознак клавіатурного почерку: часових та тиску. Оскільки розміри датасетів відмінні і невідомо чи відповідають в датасетах однакові номери спроб вводу пароля конкретній події, то було прийнято рішення обмежити датасети на позначці 350 спроб вводу паролю та об’єднати їх. Результати класифікації наведено в табл. 3.2.

Таблиця 3.2 – Результати мультикласової класифікації користувачів датасету «Queen Mary University Keystroke benchmark dataset» за комбінаціями ознак клавіатурного почерку

Комбінація ознак	FRR, %	FAR, %	max(FRR)	3-й квартиль(FRR)
Pressure	0.0	0.0	1.4	0.3
Pressure+ HoldTime	0.0	0.0	2.5	0.6
Pressure+ DownDownTime	0.0	0.0	2.2	0.3
Pressure+ UpDownTime	0.0	0.0	2.2	0.3
Pressure+ HoldTime+ DownDownTime	0.0	0.0	2.2	0.3
Pressure+ HoldTime+ UpDownTime	0.0	0.0	2.2	0.3
Pressure+ DownDownTime+ UpDownTime	0.0	0.0	2.2	0.3
Pressure+ HoldTime+ DownDownTime+ UpDownTime	0.0	0.0	2.8	0.3

Як можна бачити з табл. 3.2, отримані результати дуже суперечливі. По-перше, значення помилки першого роду для зменшеного датасету ознак тиску є меншими, ніж значення FRR для повного датасету. Цей ефект дуже нагадує помилку перенавчання, коли складна модель показує добрі результати на навчальній вибірці, але не працює на тестовій. Було проведено декілька експериментів, в яких змінювалась кількість дерев в алгоритмі Random Forest, а також використані класифікатори на основі *k* найближчіх сусідів та нейронна мережа з одним прихованим шаром. Результати усюди однакові: рівень помили FRR у зменшеного датасету менший. Отже, можна стверджувати про актуальність задачі пошуку та фільтрації шумових даних, отриманих з датчиків тиску.

По-друге, комбінація ознак тиску з будь-якими часовими ознаками

клавіатурного почерку також призводить до погіршення точності ідентифікації. Отже, рекомендувати використовувати ознаки тиску на клавіші в мультимодальних біометричних системах неможна. Виключення становлять випадки, коли усі модальності використовуються окремо в залежності від сценарію роботи системи (наприклад, вхід в систему за паролем та одночасно моніторинг динаміки вводу паролю на клавіатурі, а далі прихований моніторинг за користувачем в процесі роботи на основі ознак тиску на клавіші).

Підводячи підсумки проведених експериментів, можна зробити наступні висновки.

По-перше, для задач мультикласової класифікації ознаки тиску на клавіші є набагато більш інформативними, ніж часові параметри натискань на клавіші. За умови великої навчальної вибірки точність ідентифікації зловмисника може становити 100 %, а точність ідентифікації зареєстрованих користувачів – менше 1 %.

По-друге, у випадку двокласової класифікації на основі ознак тиску на клавіші можна будувати основну систему контролю та управління доступом, бо значення помилок першого та другого родів менше 0.5 відсотка.

По-третє, ознаки тиску на клавіші є більш зачумленими, ніж часові ознаки клавіатурного почерку. Отже, актуальною є задача пошуку та фільтрації шумових даних, отриманих з датчиків тиску.

По-четверте, комбінація ознак тиску з будь-якими часовими ознаками клавіатурного почерку також призводить до погіршення точності ідентифікації. Отже, рекомендувати використовувати ознаки тиску на клавіші в мультимодальних біометричних системах неможна. Виключення становлять випадки, коли усі модальності використовуються окремо в залежності від сценарію роботи системи (наприклад, вхід в систему за паролем та одночасно моніторинг динаміки вводу паролю на клавіатурі, а далі прихований моніторинг за користувачем в процесі роботи на основі ознак тиску на клавіші).

По-п’яте, враховуючи високу точність ідентифікації, ознаки тиску на клавіші в процесі роботи за клавіатурою можуть бути використані в системах

виявлення потенційних внутрішніх порушників інформаційної безпеки. Як відомо, часто внутрішніми інсайдерами є звичайні співробітники, які вимушені з тієї чи іншої причини чинити протизаконні дії. Такі інсайдери часто характеризуються високим рівнем тривожності, схильні до інтрапунітивних реакцій (самообвинувачування, похмурість, злість, незадоволення, напруженість, сердитість), чим відрізняються від справжніх злочинців.

Для такої категорії порушників можливе проактивне (на ранніх етапах) виявлення потенційних, схильних до протиправних дій осіб шляхом контролю їхньої діяльності та оцінки психоемоційного стану.

Одним з технічних індикаторів, які можуть вказувати на наявність потенційної інсайдерської загрози, може виступати клавіатурний почерк у сукупності таких показників, як сила тиску на клавіші, динаміка введення, системні друкарські помилки та використання певних літер, символів та «гарячих» клавіш. Сукупність певних значень кожного з цих показників утворює досить індивідуальну картину, що відповідає конкретній людині у певному психоемоційному стані.

ВИСНОВКИ

1. Виконано огляд основних методів біометричної аутентифікації. Використання клавіатурного почерку має потенціал для застосування в комп'ютерних інформаційних системах як додаткова міра, що підвищує загальний рівень безпеки при аутентифікації. Відсутність необхідності впровадження додаткового устаткування, а також можливість тривалого збору статистики під час усього сеансу роботи користувача з клавіатурою (приховане спостереження) роблять цей підхід перспективним для практичної реалізації.

2. Проведено порівняльний аналіз сучасних систем та перспективних технологій аутентифікації користувачів комп'ютерних систем за клавіатурним почерком. Як зазначають більшість дослідників, точність аутентифікації користувачів за сенсорним почерком визначається силою тиску на екран та розміром плями від пальця. Тому актуальною є задача дослідження ідентифікаційного потенціалу клавіатурного почерку з урахуванням сили тиску на клавіші.

3. Для задач мультикласової класифікації ознаки тиску на клавіші є набагато більш інформативними, ніж часові параметри натискань на клавіші. За умови великої навчальної вибірки точність ідентифікації зловмисника може становити 100 %, а точність ідентифікації зареєстрованих користувачів – менше 1 %.

4. У випадку двокласової класифікації на основі ознак тиску на клавіші можна будувати основну систему контролю та управління доступом, бо значення помилок першого та другого родів менше 0.5 відсотка.

5. Ознаки тиску на клавіші є більш зачумленими, ніж часові ознаки клавіатурного почерку. Отже, актуальною є задача пошуку та фільтрації шумових даних, отриманих з датчиків тиску.

6. Комбінація ознак тиску з будь-якими часовими ознаками клавіатурного почерку також призводить до погіршення точності ідентифікації. Отже,

рекомендувати використовувати ознаки тиску на клавіші в мультимодальних біометричних системах неможна. Виключення становлять випадки, коли усі модальності використовуються окремо в залежності від сценарію роботи системи (наприклад, вхід в систему за паролем та одночасно моніторинг динаміки вводу паролю на клавіатурі, а далі прихований моніторинг за користувачем в процесі роботи на основі ознак тиску на клавіші).

7. Враховуючи високу точність ідентифікації, ознаки тиску на клавіші в процесі роботи за клавіатурою можуть бути використані в системах виявлення потенційних внутрішніх порушників інформаційної безпеки – чим вища точність, тим точніше можна визначити діапазони зміни дослідних ознак клавіатурного почерку для кожного користувача.

Як відомо, часто внутрішніми інсайдерами є звичайні співробітники, які вимушені з тієї чи іншої причини чинити протизаконні дії. Такі інсайдери часто характеризуються високим рівнем тривожності, схильні до інтрапунітивних реакцій (самообвинувачування, похмурість, злість, незадоволення, напруженість, сердитість), чим відрізняються від справжніх злочинців.

Для такої категорії порушників можливе проактивне (на ранніх етапах) виявлення потенційних, схильних до протиправних дій осіб шляхом контролю їхньої діяльності та оцінки психоемоційного стану.

Одним з технічних індикаторів, які можуть вказувати на наявність потенційної інсайдерської загрози, може виступати клавіатурний почерк у сукупності таких показників, як сила тиску на клавіші, динаміка введення, системні друкарські помилки та використання певних літер, символів та «гарячих» клавіш. Сукупність певних значень кожного з цих показників утворює досить індивідуальну картину, що відповідає конкретній людині у певному психоемоційному стані.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Biometric Systems: Technology, Design and Performance Evaluation. James L. Wayman, Anil K. Jain, Davide Maltoni. 2005.

2. Biometric Security. Jiankun Hu, David Chek Ling Ngo, Andrew Beng Jin Teoh. 2015.

3. Guide to Biometrics. Ruud M. Bolle, Jonathan H. Connell, Sharath Pankanti. 2013.

4. Handbook of Biometric Anti-Spoofing: Trusted Biometrics Under Spoofing Attacks. Sébastien Marcel, Mark S. Nixon, Stan Z. Li. 2014.

5. Handbook of Biometrics. Anil K. Jain, Patrick Flynn, Arun A. Ross. 2007.

6. Biometrics in Identity Management: Concepts to Applications. Shimon K. Modi. 2011.

7. The Biometric Computing: Recognition and Registration. Karm Veer Arya, Robin Singh Bhadoria. 2019.

8. Obaidat M.S., Sadoun B. Verification of computer users using keystroke dynamics. IEEE Trans. Syst., Man. Cybem. B, 1997, vol. 27, iss. 2, pp. 261-269.

9. Shepherd S.J. Continuous authentication by analysis of keyboard typing characteris¬tics. European Convention on Security and Detection, 1995, pp. 111-114.

10. Monrose F., Rubin A. Authentication via keystroke dynamics. Proc. 4th ACM CCCS, 1997, pp. 48-56.

11. Chandrasekar V., Kumar S. Biometric based keystroke dynamics authentication — a review. Asian J. Res. Soc. Sci. Human., 2016, vol. 6, no. 9, pp. 698-718.

12. D. El Menshawy, H.M.O. Mokhtar, and O. Hegazy. A keystroke dynamics based approach for continuous authentication. In 10th International Conference on Beyond Databases, Architectures, and Structures (BDAS), Vol. 424 of CCIS,

pp. 415–424, 2014.

13. D. Gunetti and C. Picardi. Keystroke analysis of free text. ACM Transactions on Information and System Security, Vol. 8(3), pp.312–347, 2005.

14. D. Umphress and G. Williams. Identity verification through keyboard characteristics. International Journal of Man-Machine Studies, Vol. 23(3), pp. 263–273, 1985.

15. F. Bergadano, D. Gunetti, and C. Picardi. User authentication through keystroke dynamics. ACM Transactions on Information and System Security, Vol. 5(4), pp.367–397, 2002.

16. J.A. Robinson, V.W. Liang, J.A.M. Chambers, and C.L. MacKenzie. Computer user verification using login string keystroke dynamics. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Vol. 28(2), pp.236–241, 1998.

17. K.S. Balagani, Vir V. Phoha, A.Ray, and S. Phoha. On the discriminability of keystroke feature vectors used in fixed text keystroke authentication. Pattern Recognition Letters, Vol. 32(7), pp.1070–1080, 2011.

18. Kyle O.Bailey, James S. Okolica, Gilbert Peterson, "User identification and authentication using multi-modal behavioral biometrics," Computers and Security journal, vol.43, pp.77-89, June 2014.

19. Md Liakat Ali, John Monaco, Charles Tappert, Meikang Qiu, "Keystroke Biometric Systems for User Authentication". Journal of Signal Processing Systems, vol. 86(2) pp.175–190, March 2017.

20. R. Gaines, W. Lisowski, S. Press, and N. Shapiro. Authentication by keystroke timing: some preliminary results. Technical Report Rand Rep. R-2560-NSF, RAND Corporation, p51, 1980.

21. R.A. Maxion and K.S. Killourhy. Keystroke biometrics with number-pad input. In IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 201–210, 2010.

22. R.Spillane. Keyboard apparatus for personal identification. IBM Technical Disclosure Bulltin, 17(11), 1975.

23. Roman V.Yampolskiy, Venu Govindaraju, “Behavioural biometrics: a survey and classifiation,” International Journal of Biometrics, vol.1, pp.81-113, Novemberv 2008.

24. S.Z.S. Idrus, E. Cherrier, C. Rosenberger, and P. Bours. Soft biometrics for keystroke dynamics. In 10th International Conference on Image Analysis and Recognition (ICIAR), Vol. 7950 of LNCS, pp. 11–18, 2013.

25. T. Sim and R. Janakiraman. Are digraphs good for free-text keystroke dynamics? In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–6, 2007.

26. C.C. Loy, C.P. Lim, and W.K. Lai. Pressure-based typing biometrics user authentication using the fuzzy ARTMAP neural network. In International Conference on Neural Information Processing (ICONIP), 2017.

27. Queen Mary University Keystroke benchmark dataset. URL: https://personal.ie.cuhk.edu.hk/~ccloy/downloads_keystroke100.html #:~:text=Keystroke100%20benchmark%20dataset%20is%20a,password%20%22try4%2Dmbs%22 (дата звернення: 20.05.2023).