МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет програмної інженерії та бізнесу

Кафедра інженерії програмного забезпечення

Пояснювальна записка до дипломного проекту

магістра
(освітній ступінь)

на тему «Експериментальне дослідження методів перетворення біометричних даних людини в задачах аутентифікації особистості»

ХАІ.603.667п1.121.156336.20О

Виконав: студент 6 курсу групи № 667п1
Напрям підготовки 121 Інженерія програмного забезпечення
(код та найменування)
Освітня програма Хмарні обчислення та Інтернет речей
Килимчук Б. О.
(прізвище й ініціали студента)
Керівник: Данова М. О.
(прізвище й ініціали)
Рецензент: Філімончук Т. В.
(прізвище й ініціали)

Харків - 2020

Міністерство світи і науки України
Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет програмної інженерії та бізнесу
(повне найменування)
Кафедра інженерії програмного забезпечення
(повне найменування)
Рівень вищої освіти другий (магістерський)
Спеціальність 121 – інженерія програмного забезпечення
(код та найменування)
Освітня програма хмарні обчислення та Інтернет речей
(найменування)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ І. Б. Туркін
(підпис) (ініціали та прізвище)
“_____” _____ 2020 року

З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЄКТ СТУДЕНТУ

Килимчуку Богдану Олеговичу
(прізвище, ім’я, по батькові)
1. Тема дипломного проекту *Експериметальне дослідження методів перетворення біометричних даних людини в задачах аутентифікації особистості*
керівник дипломного проекту *Данова Марія Олександрівна, к.т.н., доцент*
(прізвище, ім’я, по батькові, науковий ступінь, вчене звання)
затверджені наказом Університету № _____ від “____” _____ 2020 року
2. Термін подання студентом проекту _____
3. Вихідні дані до проекту дослідження існуючих методів перетворення біометричних даних людини в задачах аутентифікації особистості; програмне забезпечення для перетворення біометричних даних в код доступу с ціллю наступної аутентифікації

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити) огляд та аналіз предметної області з технологій перетворення біометричних даних; експериментальне дослідження методів перетворення персональних даних користувача в біометричних системах; обробка та аналіз результатів експериментального дослідження методів перетворення персональних даних користувача в біометричних системах.

5. Перелік графічного матеріалу _____
РПЗ – стор. 84, рисунків – 34 шт., таблиць – 1 шт., презентація – 23 слайди.

6. Консультанти розділів проекту

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1	Данова М.О., доцент каф. 603		
2	Данова М.О., доцент каф. 603		
3	Данова М.О., доцент каф. 603		

Нормоконтроль _____ _В. А. Постернакова_____ «__»_____20__р.
(підпис) (ініціали та прізвище)

7. Дата видачі завдання «___»_____20_____ р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту	Строк виконання етапів проекту	Примітка
1	Отримання і затвердження теми диплому	03.09.2019	
2	Аналіз предметної області	20.09.2019	
3	Постановка задачі	25.09.2019	
4	Проведення теоретичних досліджень	27.12.2019	
5	Розробка прототипу ПЗ	03.09.2020	
6	Підготовка пояснювальної записки	16.10.2920	
7	Оформлення пояснювальної записки до дипломного проекту	13.11.2020	
8	Передзахист дипломного проекту	20.11.2020	
9	Захист дипломного проекту	10.12.2020	

Студент _____ ____Килимчук Б.О.
(підпис) (прізвище та ініціали)

Керівник проекту _____ Данова М.О.
(підпис) (прізвище та ініціали)

РЕФЕРАТ

Пояснювальна записка до дипломного проекту містить 84 стор., 34 рис., 21 джерел.

Об'єкт дослідження – процес багатофакторної аутентифікація користувача.

Предмет дослідження – методи перетворення біометричних даних користувача для аутентифікації особистості.

Метою роботи є підвищення ефективності аутентифікації користувача шляхом знеособлення персональних даних користувача в біометричних системах за рахунок розроблення програмного забезпечення для перетворення біометричних даних в код доступу с ціллю наступної аутентифікації.

Для досягнення поставленої мети необхідно вирішити такі задачі: провести огляд та аналіз предметної області з технологій перетворення біометричних даних; провести експериментальне дослідження методів перетворення персональних даних користувача в біометричних системах; провести обробку та аналіз результатів експериментального дослідження методів перетворення персональних даних користувача в біометричних системах.

Наукова новизна. Удосконалено метод знеособлення персональних даних в біометричних системах, якій на відміну від існуючих використовує біометричний контейнер в знеособлених даних, що дозволяє обмежити доступ операторам до інформації, що ідентифікує людину, а приналежність знеособлених даних конкретного суб'єкта визначати за допомогою біометричної аутентифікації.

Практична значимість отриманих результатів. Результати дослідження дозволять підвищити надійність систем з біометричною аутентифікацією особистості.

БІОМЕТРИЧНА АУТОТЕНТИФІКАЦІЯ, ПЕРСОНАЛЬНІ ДАНІ, ШТУЧНІ НЕЙРОННІ МЕРЕЖИ, БІОМЕТРИЧНІ СИСТЕМИ, ЗАХИСТ ІНФОРМАЦІЇ

ABSTRACT

Explanatory note to the graduate work contains 84 pp., 34 fig., 21 sources.

The object of study - the process of multifactor user authentication.

The subject of research - methods of converting user biometric data for identity authentication.

The aim of the work is to increase the efficiency of user authentication by depersonalizing the user's personal data in biometric systems through developing software for converting biometric data into access code for subsequent authentication.

To achieve this goal it is necessary to solve the following tasks: to review and analyse the subject area of biometric data conversion technologies; to conduct an experimental study of methods for converting personal data of the user in biometric systems; to process and analyse the results of experimental research of methods of conversion of personal data of the user in biometric systems.

Scientific novelty. The method of depersonalization of personal data in biometric systems has been improved, which, unlike the existing ones, uses a biometric container in depersonalized data, which allows operators to restrict access to personally identifiable information and determine the identity of depersonalized data by a specific subject using biometric authentication.

The practical significance of the results obtained. The results of the study will increase the reliability of systems with biometric authentication.

BIOMETRIC AUTOTENTIFICATION, PERSONAL DATA, ARTIFICIAL NEURAL NETWORKS, BIOMETRIC SYSTEMS, INFORMATION PROTECTION

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

БЧХ – код Боуза-Чоудхурі-Хоквінгема;

ВБП – вектор біометричних параметрів;

ШНМ – штучна нейронна мережа;

НСД – несанкціонований доступ;

ПД – персональні дані.

ВСТУП

У даний час перспективним заходом захисту даних від несанкціонованого доступу є багатофакторна аутентифікація користувача, що включає в себе декілька факторів різного роду: дані, що відомі користувачеві, об’єкти, що маються у користувача та унікальні біометричні характеристики користувача. Засоби аутентифікації на основі паролів, що видаються авторизованим користувачам легко реалізуються та мають низьку вартість. Головним недоліком таких засобів при захисті від несанкціонованого доступу є можлива втрата конфіденційності секретів користувачів. Людині важко запам’ятати довгі паролі. Тому таку паролі зберігають в паперовому блокноті або в незахищеному файлі на комп’ютері. Короткі змістові паролі легко підбираються шахраями. Це робить аутентифікацію за допомогою пароля слабко захищеною.

Одним з видів безпечного заходу аутентифікації вважають персональні пристрої зберігання ключової інформації: пластикові картки, смарт-картки, токени. При наявності такого пристрою, користувачу немає необхідності запам’ятовувати довгі паролі. Щоб зламати захист шахраю необхідно викрасти пристрій користувача. Помітивши крадіжку, власник пристрою може відразу повідомити службу безпеки про факт викрадення. Вартість засобів аутентифікації з використанням носії ключової інформації вища в порівнянні з засобами аутентифікації з використанням звичайних паролів.

Засоби біометричної аутентифікації використовують фізіологічні й поведінкові особливості, котрими володіє користувач і нерозривно пов’язаний з ними. К засобам біометричної аутентифікації відносять геометрію обличчя, геометрію рук, рисунок відбитків пальців, рисунок сітківки ока, рукописний почерк, характеристики голосу. Біометричні характеристики неможливо забути, втратити, викрасти, передати іншій людині. За допомогою біометрії не тільки перевіряється достовірність користувача, а й підтверджується його особистість. Системи біометричної аутентифікації повинні гарантувати високу надійність,

щоб підтверджувати авторизованого користувача та відкидати шахрая із схожими біометричними параметрами, а також забезпечувати конфіденційність біометрії як персональних даних користувача.

Тому надійна ідентифікація користувача системою може бути гарантована, тільки якщо одним із факторів звірення є його унікальна біометрична характеристика. Ціни на біометричні пристрої стали більш-менш прийнятними для покупців. Таким чином біометричні технології досягли етапу практичного застосування а їх використання в засобах захисту інформації цілком доцільне.

Проте все таки залишаються деякі важливі питання використання біометричних даних в засобах захисту. Справа в тому, що використання унікальних біометричних даних людини потребує їх надійного захисту від втрати, яка може призвести до серйозних наслідків. Фізіологічні й поведінкові особливості людини дозволяють несанкціоновано встановити його особистість, організувати за ним приховане стеження. Через цю причину деякі користувачі відмовляються від зберігання своїх біометричних даних, наприклад, дані рисунка відбитку пальця в яких-небудь базах даних інформаційних систем, побоюючись загроз, пов’язаних з порушенням приватного життя й конфіденційності особистості.

Актуальною задачею є безпечне зберігання неточно (неоднозначно) відтворюваних біометричних даних. Саме ця проблема робить неможливим використання криптографічних хеш-функцій для безпечного зберігання біометричної інформації. З плином часу повилось два напрямки з рішення даної проблеми безпеки. Перший напрямок, створений й розвивається за кордоном, характеризується використанням нечітких екстракторів на основі кодів з виявленням і виправленням помилок. В Україні створено й активно розвивається напрямок нейромережевого перетворення біометричних даних людини в код доступу. Розроблені стандарті високонадійної нейромережевої біометричної аутентифікації.

Одним з ефективним засобів захисту персональних даних є знеособлення. Знеособлення дозволяє знизити ризики несанкціонованого використання і

заподіяння збитку в разі викрадення. Біометричні дані також відносяться до персональної інформації, яку необхідно захищати від несанкціонованого використання. Проте, традиційні засоби знеособлення надто неефективні для інформаційних систем, що обробляють біометричні дані. Тому актуальною задачею є розроблення шляхів знеособлення, враховуючих особливості біометричних даних й систем, що їх використовують.

Об'єкт дослідження – процес багатофакторної аутентифікація користувача.

Предмет дослідження – методи перетворення біометричних даних користувача для аутентифікації особистості.

Метою роботи є підвищення ефективності аутентифікації користувача шляхом знеособлення персональних даних користувача в біометричних системах за рахунок розроблення програмного забезпечення для перетворення біометричних даних в код доступу с ціллю наступної аутентифікації.

Для досягнення поставленої мети необхідно вирішити такі задачі:

– провести огляд та аналіз предметної області з технологій перетворення біометричних даних;

– провести експериментальне дослідження методів перетворення персональних даних користувача в біометричних системах;

– провести обробку та аналіз результатів експериментального дослідження методів перетворення персональних даних користувача в біометричних системах.

Методи досліджень. У роботі було використано методи штучного інтелекту, методи планування експерименту, статистичний аналіз результатів експерименту.

Наукова новизна. Удосконалено метод знеособлення персональних даних в біометричних системах, якій на відміну від існуючих використовує біометричний контейнер в знеособлених даних, що дозволяє обмежити доступ операторам до інформації, що ідентифікує людину, а приналежність

знеособлених даних конкретного суб'єкта визначати за допомогою біометричної аутентифікації.

Практична значимість отриманих результатів. Результати дослідження дозволять підвищити надійність систем з біометричною аутентифікацією особистості.

1 КРИТИЧНИЙ ОГЛЯД ТЕХНЛОГІЙ ПЕРЕТВОРЕННЯ БІОМЕТРИЧНИХ ДАНИХ

1.1 Характеристика класичних біометричних систем

Біометричні системи проектуються й розробляються для здійснення контролю доступу до приміщень під охороною на військових об’єктах. [1] Створюються системи автоматичної ідентифікації особистості, які використовуються в оперативно-пошуковій діяльності правоохоронних органів.[2] У світі отримують розповсюдження системи біометричної ідентифікації громадян при доступі до соціальних послуг, наданні медичних послуг, голосуванні, переписі населення, обліку тривалості робочого часу, придбанні товарів, тощо. [3, 4]

В якості нормативної бази при розробленні біометричних систем виступають як міжнародні, так і українські стандарти. [5, 6]

Галузь біометричних стандартів, орієнтованих на створення біометричних систем на основі біометричних образів, що зберігаються в таємниці, поки активно створюється лише в Європі.

Згідно з міжнародними стандартами найпростіша біометрична система – це автоматизована система, що здійснює:

− реєстрацію біометричного образа користувача за допомогою біометричного сканера;

− обрахування біометричних параметрів й формування біометричного шаблона;

− порівняння біометричних параметрів з параметрами, що містяться в біометричному шаблоні;

− прийняття рішення й видача результату ідентифікації та верифікації.

Біометрична система складається з підсистем реєстрації та верифікації (ідентифікації) користувачів. Підсистема реєстрації здійснює збирання

біометричних даних й інших відомостей про користувача (суб’єкта доступу), дії якого регламентуються згідно яких-небудь правил розмежування доступу. При цьому даних повинно бути достатньо для ідентифікації (присвоєння ідентифікатора) користувача й перевірки належності суб’єкту доступу ідентифікатора, що їм пред’являється.

Типова структура біометричної системи зображена на рисунку 1.1.

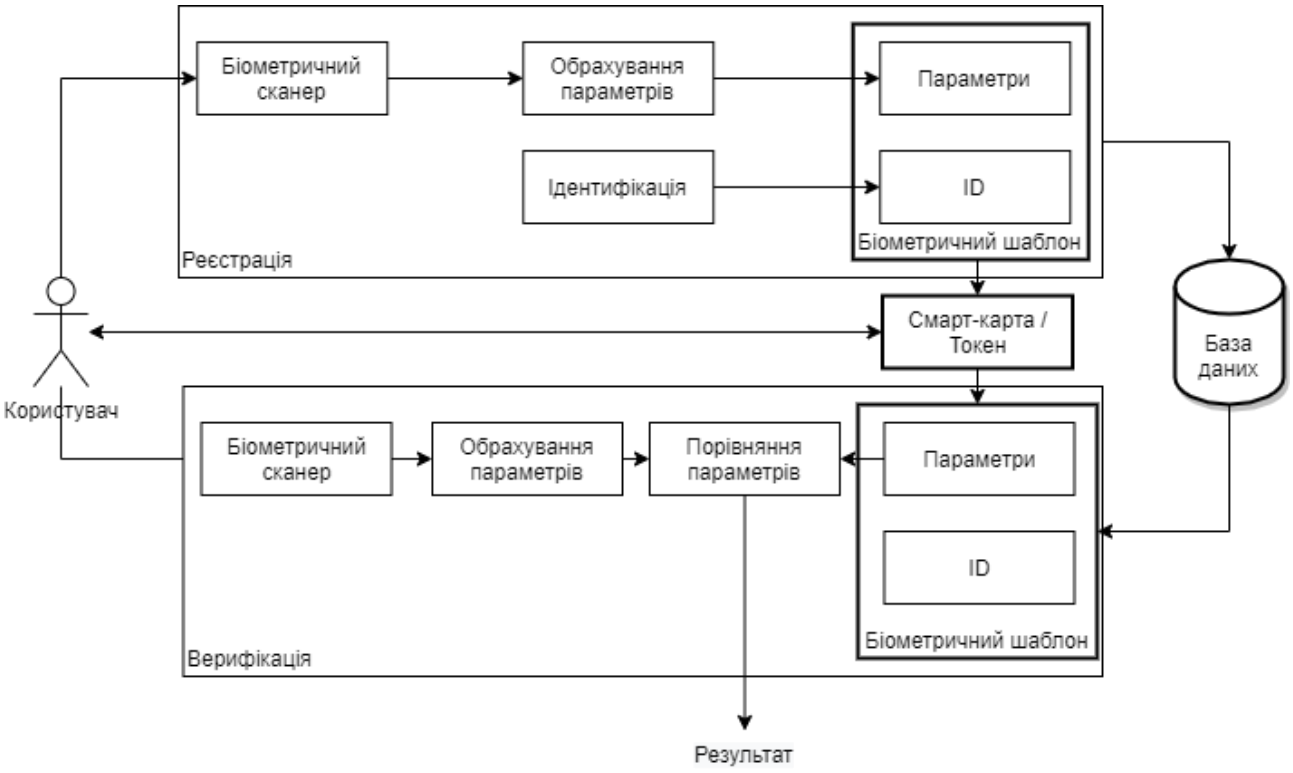

Рисунок 1.1 – Типова структура біометричної системи

У процесі реєстрації біометричних образ зчитується біометричним пристроєм введення, формується біометричний шаблон й зберігається в базі даних разом з яким-небудь ідентифікатором, який має користувач або присвоюється йому в процесі реєстрації. Ідентифікатор користувача та біометричний шаблон можуть бути розміщені на смарт-картці або токені. При верифікації користувач пред’являє системі ідентифікатор і біометричні дані. Потім обраховуються значимі параметри користувача, які порівнюються з

параметрами біометричного шаблона, отриманого із бази даних за пред'явленим ідентифікатором. Після чого приймається рішення й видається відповідь чи це та людина за яку вона себе видає.

Також біометрична система може відповісти на питання приналежності пред'явлених біометричних даних кому-небудь з користувачів системи, тобто ідентифікувати особистість.

При порівнянні біометричних параметрів обраховується деяка міра схожості S, яка порівнюється з пороговим значенням T. В тому випадку, якщо S більша або дорівнює T, приймається рішення про те, що пред'явлені біометричні дані й дані біометричного шаблону належать одній і тій самій людині. Якщо S менша за T, приймається рішення про те, що параметри належать різним людям.

Основними характеристиками біометричних систем є похибки першого та другого роду. Коли пред'явлені біометричні параметри й параметри біометричного шаблону дійсно належать одній людині, а результат їх порівняння менший за порогове значення, виникає похибка першого роду P_I. Коли пред'явлені біометричні параметри й параметри біометричного шаблона належать різним людям, а результат порівняння більший за порогове значення, виникає похибка другого роду P_{II}. В таблиці 1.1 наведені середні значення P_I і P_{II} для різних біометричних систем. [7]

Таблиця 1.1 – Характеристика відомих біометричних систем

Хар.	Лицо	Отпечаток пальца	Вены ладони	Вены пальца	Радужка	Подпись	Голос
P_I	$2{,}6 \cdot 10^{-2}$	10^{-2}	10^{-2}	10^{-2}	10^{-3}	10^{-2}	10^{-1}
P_{II}	$1{,}3 \cdot 10^{-2}$	10^{-5}	10^{-7}	10^{-5}	10^{-6}	10^{-6}	$3 \cdot 10^{-2}$

Біометричні системи, які відповідають виключно вимогам міжнародних стандартів прийнято називати класичними біометричними системами. Такі системи прості в реалізації, але в той же час мають наступні значимі недоліки:

− біометричний шаблон, що зберігається й обробляється в біометричній системі незахищений. Компрометація біометричного шаблону призводить до компрометації як таємного так і відкритого образу людини;

− однобітовий результат («так»/«ні») робить засоби аутентифікації уразливими до атак, спрямованих на підміну результату аутентифікації;;

− зберігання таємного криптографічного ключа разом з біометричним шаблоном також неприпустимо.

Перелічені вразливості класичних біометричних систем можуть бути усунуті за рахунок засобів захисту біометричної аутентифікації, які засновані на наступних положеннях. По-перше біометрію людини необхідно пов’язувати з деяким кодом доступу. По-друге біометричні шаблони не повинні використовуватися для прийняття рішення й видачі результату ідентифікації та верифікації.

1.1.1 Біометричні характеристики людини

Всі біометричні характеристики людини можна поділити на два класи: статичній й динамічні. Статичні біометричні характеристики є фізіологічними особливостями, які є незмінними з плином часу. До таких характеристик відносять:

- геометрія обличчя;
- рисунок відбитків пальців;
- рисунок вен кисті руки;
- сітківка ока;
- ДНК.

Фізіологічні особливості можуть бути втрачені внаслідок хвороби або при фізичному (хірургічному) впливі на його органи.

До динамічних біометричним характеристикам відносять:

- динаміка рукописного почерку;
- характеристики голосу;
- серцевий ритм;
- хода.

Не існує єдиної біометричної характеристики, що була б ідеальна у всіх випадках. Кожна біометрична характеристика має свої переваги та недоліки. Процес отримання біометричної характеристики має бути за можливості швидким й простим, без заподіяння яких-небудь незручностей для користувача.

На сьогоднішній день біометричні дані широко використовуються в інформаційних системах для органолептичної ідентифікації громадян. Але з розвитком засобів обчислювальної техніки біометричні дані обличчя стали успішно застосовуватися автоматичними засобами розпізнавання людини для встановлення й підтвердження її особистості. В загальному випадку можна виділити два класи засобів біометричної ідентифікації людини по обличчю. Перший клас засобів аналізую плоскі двомірні зображення обличчя людини (2D портрети). Двовимірний аналіз плоского зображення полягає у виділенні на ньому характерних точок й обрахування геометрії, а саме відстаней між центрами очей, між лінією очей та кінчиком носа, тощо.

Двомірний аналіз плоского зображення обличчя людини дозволяє отримати зовсім не багато біометричної інформації. Перехід до більш складного трьохвимірного аналізу геометрії обличчя людини дозволяє значно збільшити об’єм одержуваної біометричної інформації.

Особливості папілярного рисунка відбитку пальця широко використовується в дактилоскопії, яка на сьогодні є невід’ємною частиною криміналістики, оперативно-пошукової діяльності. Поява компактних сканерів зчитування папілярного рисунка відбитку пальця дозволяє успішно використовувати дані біометричних характеристик для підтвердження особистості людини в системі контроля і управління доступом. Слід відмітити,

що виникла гостра необхідність застосування відбитків пальців паспортно-візовими службами та міграційною політикою різних держав світу.

Відбиток сітківки ока є унікальною фізіологічною характеристикою людини. Сканування засновано на поглинанні інфрачервоного випромінювання меланіном, відповідального за пігментацію сітківки ока. Параметри вимірюються по отриманому зображенню.

Наразі існує два основні підходи до ідентифікації геометрії руки людини. Перший підхід базується на геометричних характеристиках кісті людини. Другий відноситься до сучасних підходів, оскільки окрім геометричних параметрів базується ще й на тому як розташовані кровоносні судини.

Рукописний почерк – це динамічна характеристика, яка пов’язана з поведінкою людини. Вимірюваними параметрами є залежності координат кінців пера від часу, одержувані, як правило, за допомогою графічного планшета.

Певну міру унікальності також мають біометричні характеристики голосу людини. Встановлення належності різних мовних фраз одній і тій же людині, ідентифікація особистості за голосом використовується в криміналістиці. Існує автоматичні засоби розпізнавання, здатні визначити параметри на основі характерних для звуків мови сигналів, що мають свою форму коливань тиску. Відомо, що частина звуків мови – голосні (є періодичними), а інша частина звуків є шиплячими (не періодичні).

Серцевий ритм й хода людини не мають високого ступню унікальності біометричних параметрів, тому не отримали широкого застосування.

Таким чином, можна виділити позитивні властивості біометричних характеристик людини. Біометрію неможливо втратити, забути, віддати іншій людині, як наприклад інші ідентифікатори: пароль, пін-код, смарт-карту, токен. Данні властивості особливо важливі для забезпечення безпеки таких систем, які потребують надійну аутентифікацію користувачів.

1.1.2 Таємні та відкриті біометричні образи

Серед біометричних характеристик людини можна виділити ті характеристики, котрі можуть залишатися в таємниці. Дані на рисунку 1.2 відображають сучасну практику використання відкритих біометричних образів при ідентифікації людини.

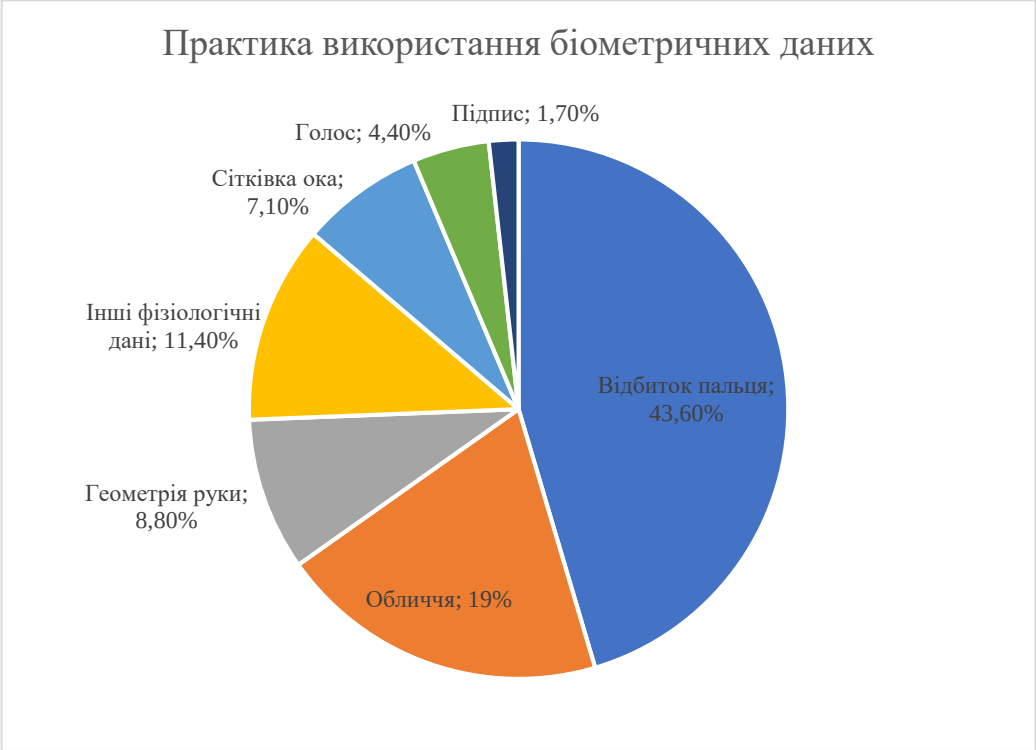

Рисунок 1.2 – Розподіл затребуваності біометричних технологій

Найбільш зручними біометричними технологіями з точки зору забезпечення таємниці образів є біометричні технології аналізів голосу та рукописний почерк людини. Забезпечити тайну таких біометричних образів простіше всього, достатньо голосом або своїм почерком відтворити пароль (парольну фразу). Для того, щоб забезпечити таємницю статичних образів, котрі не змінюються за волею людини, необхідно користуватися системою організаційно-технічних заходів, що забезпечують знеособлення людини. Як тільки сторонній спостерігач дізнається ім’я людини, знайти його і скомпрометувати дані його статичного

біометричного образа не важко. Можна зробити висновок, що таємниця біометричного образа на багато сильніша, ніж біометричні технології.

На сьогоднішній день спостерігається дві виражені тенденції в розвитку засобів захисту інформації:

− масове використання асиметрично криптографії (пари з відкритого та закритого ключа);

− посилення процедур ідентифікації й аутентифікації за допомогою біометричних технологій.

Згідно до експертних оцінок вітчизняні й зарубіжні досліджень безпечності використання засобів асиметричної криптографії може бути за рахунок зв'язування відкритих та закритих ключів з біометрією їх власника. При авторизації прав особистості на управління закритим криптографічним ключем необхідно забезпечити зберігання в таємниці закритого біометричного образа. Відкритий ключ може бути пов'язаний з відкритим (загальнодоступним) біометричним образом людини, наприклад з 3д-маскою обличчя людини. Закритий ключ обов'язково повинен бути пов'язаний з таємним (закритим) образом людини, наприклад з його рукописним паролем.

1.2 Засоби захищеної біометричної аутентифікації

В 2001 році закордонними вченими вперше були запропоновані засоби усунення загроз порушення конфіденційності біометричних шаблонів в системах біометричної аутентифікації та ідентифікації. Зокрема запропоновані засоби трансформації біометричного шаблону й його перетворення за допомогою поліноміальних незворотних функцій.

Поєднання біометричних функцій та криптографії призвів до утворення біометричних криптографічних систем. [8] В таких системах біометричні дані користувача перетворюються в його особистий криптографічний ключ, містяться механізми захисту біометричного шаблона. Даний напрямок отримав назву «біометричне шифрування» («biometric encryption»).

Оскільки біометричні дані неточно відтворювані, задача отримання їх на основі точного значення криптографічного ключа є доволі складною. На даний момент ця технологія розвивається за декількома напрямками і нараховує декілька видів біометричних криптографічних систем:

- система звільнення ключа (key release cryptosystem);
- система прив’язки ключа (key binding cryptosystem);
- система генерації ключа (key generation cryptosystem).

Розглянемо кожну систему більш детально, зокрема принципи й методи виконуваних перетворень.

В системі звільнення ключа зберігаються криптографічний ключ та біометричний шаблон. Звільнення ключа (надання доступу до ключа) відбувається у випадку успішної біометричної аутентифікації. В процесі аутентифікації потрібен доступ до біометричного шаблону, тому він зберігається відкрито, що є недоліком. Таким чином можлива модифікація ключової інформації в даних шаблонах. Перевагою системи із звільненням ключа є простота її реалізації.

В системі з прив’язкою ключа криптографічний ключ пов’язується з біометричним шаблоном шляхом заміщення його значимих даних даними ключа, що витягаються. Стійкість такої системи залежить від конфіденційності алгоритму пов’язування. Проте даний вид систем може бути застосований для захисту біометричних шаблонів.

В системі з генерацією ключа криптографічний ключ не зберігається, а витягується безпосередньо із біометричних даних користувача, що є незаперечною перевагою системи. Однак реалізація системи з генерацією ключа набагато важча розглянутих вище систем. Складність полягає в необхідності створення алгоритмів обчислювання якісних біометричних параметрів та генерації двійкового вектору фіксованої довжини, що дає точне значення криптографічного ключа легітимного користувача при пред’явленні системі одного й того самого біометричного образа «Свій» і рівноймовірний розподіл нулів і одиниць при пред’явленні випадкових значень біометричних даних.

Наразі відомі наступні підходи до перетворення біометричних даних в криптографічний ключ і застосування для різних біометричних технологій:

- нечіткий екстрактор («fuzzy extractors»);
- нечітке сховище («fuzzy valt»);
- біометрично-нейромережеві перетворення («biometric-neural network transformations»).

В нечітких екстракторах та нечітких сховищах застосовуються коди з виявленням і виправленням помилок, що виникають в двійковому векторі через неточне відтворення біометричних даних. [9-10] Якість перетворень біометричних даних в ключ багато в чому визначається якістю використовуваних кодів, що виявляють та виправляють помилки. При обрахуванні криптографічного ключа обидва методи перетворювання потребують зберігання додаткових відкритих даних (відкритий хелпер від англ. public helper).

Шляхом створення нечітких екстракторів йдуть переважно закордонні дослідники. Нейромережеві перетворювачі біометрія-код запропоновані вітчизняними дослідниками. [11].В нейромережевих перетворювачах біометрія-код використовується штучна нейронна мережа, яка автоматично навчається на етапі реєстрації користувачів створювати його особистий ключ при пред’явленні біометричного образа «Свій». При цьому для випадкових біометричних даних, прикладів образів «Чужий» штучна нейронна мережа створює випадкову кодову комбінацію.

Таким чином, на даний момент існує два основні підходи до перетворення неоднозначного біометричного образа в криптографічний ключ користувача, в подальшому використовуваний в класичних механізмах аутентифікації: нечіткий екстрактор і нейромережевих перетворювач біометрія-код. Нечіткий екстрактор, являє собою аналог кодів, що виявляють й виправляють помилки у векторі двійкових біометричних параметрів. Нейромережевий перетворювач біометрія-код використовує навчену штучну нейронну мережу для формування точного значення ключа. Розглянемо данні підходи більш детально.

1.3 Безпека використання біометричних даних для аутентифікації

1.3.1 Зчитування відбитку пальця за допомогою оптичного та теплового сканеру

З початку двохтисячних років хакери відточили методологію виготовлення штучних силіконових копій по наявному рисунку. Якщо наклеїти тонку плівку на власний палець, то можна обійти практично будь-яку систему, навіть з додатковими мірами безпеки, такими як сенсорами, що вимірюють температуру прикладеного об’єкта і засвідчується, що до сканера прикладено палець живо людини, а не просто надрукований малюнок відбитку. Класичним керівництвом з виготовлення штучних відбитків вважається керівництво Цутому Мацумото від 2002 року. [12] Там детально пояснюється, як обробити відбиток пальця жертви за допомогою графітового порошку або парів цианоакрилата (суперклею), як потім обробити фотографію перед виготовленням форми і, нарешті, виготовити опуклу маску за допомогою желатину, латексного молочка або клею для дерева (рисунок 1.3).

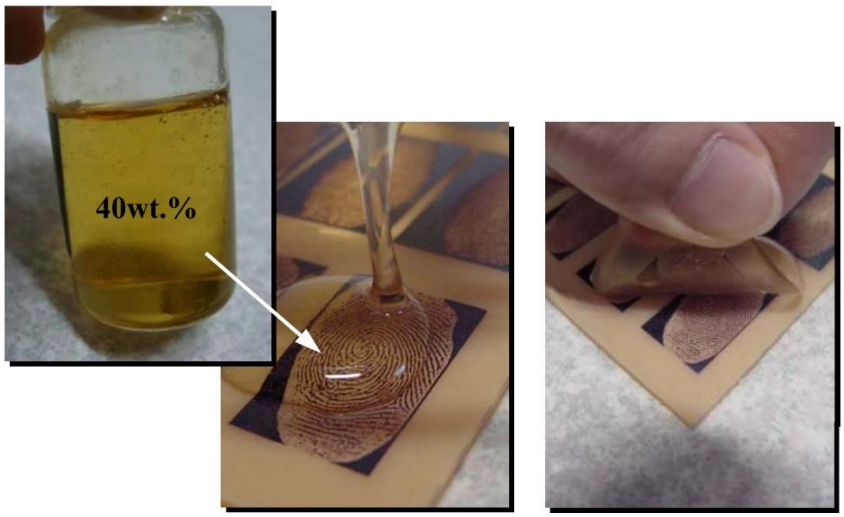

Рисунок 1.3 – Створення штучного відбитку за допомогою клею для дерева

Найбільша складність у цій процедурі - скопіювати справжній відбиток пальця. Кажуть, найякісніші відбитки залишаються на скляних поверхнях і дверних ручках. Але в наш час з'явився ще один спосіб: роздільна здатність деяких фотографій дозволяє відновити малюнок прямо з фотографії.

У 2017 році повідомлялося про проект дослідників з Національного інституту інформатики Японії. [13] Вони довели можливість відтворення малюнка відбитка пальця з фотографій, зроблених цифровим фотоапаратом з відстані в три метри. Ще в 2014 році на хакерській конференції Chaos Communication Congress продемонстрували відбитки пальців міністра оборони Німеччини, [14] відтворені за офіційними фотографіями з високою роздільною здатністю з відкритих джерел.

1.3.2 Створення збірного відбитку для будь-якого датчику

Відбитки пальців прийнято вважати одним з найнадійніших ідентифікаторів особистості, але вчені з Нью-Йоркського університету та університету штату Мічіган провели дослідження [15], яке ставить під сумнів дане твердження. В даній роботі було створено набір штучних відбитків – MasterPrints. Даний набір був створений штучним інтелектом на основі аналізу загальних елементів з відбитків великої кількості людей. За допомогою даного набору відбитків у 65% випадків вдавалось розблокувати мобільний пристрій волонтерів.

Перш за все це пов’язано з тим, що сканери на смартфонах доволі компактні і можуть зчитувати лише частину відбитку. Крім того, щоб відбиток було простіше зчитувати смартфони зазвичай запитують вісім-десять відбитків для порівняння, причому користувач нерідко залишає відбитки кількох пальців, що прямопропорційно збільшує ймовірність хибного підтвердження особистості. Таким чином телефон буде розблоковано якщо хоча б один із часткових відбитків користувача співпаде.

«Це схоже на ситуацію, коли у вас є 30 паролів, і шахраю потрібно підібрати лише один з них», - зазначив один з авторів дослідження професор інформатики Політехнічного інституту Нью-Йоркського університету Насир Мемон.

«Якщо мені потрібно взяти ваш телефон і скористуватися Apple Pay для покупок, і якщо я зможу влізти в один з десяти телефонів, це дуже непогана ймовірність», - вважає професор з системної інженерії Карлтонського університету Енді Адлер, що спеціалізується на вивченні біометричних систем безпеки в пристроях.

1.3.3 Сканер, що аналізує вени на пальці людини

Ніколас Дрейден – чоловік, що намагається продати цю технологію світові запевнює, що розташування вен на наших пальцях майже унікальне і шанс того, що це розташування співпаде з кимось іншим становлять 3,4 мільярди до одного, тож маючи картину цих вен ми можемо бути впевнені, що власних є тим за кого себе видає. Наразі ця технологія тестується в барі «Proud» у Лондоні.

На відміну від звичайного відбитку пальця вени не видно на фотографії, це є більш конфіденційною інформацією, яку важче здобути без відома власника відбитку. [16] Пристрій зображено на рисунку 1.4.

.

Рисунок 1.4– Пристрій для аналізу розташування вен на пальці людини

1.3.4 Використання знімку обличчя для ідентифікації

Перш за все даний спосіб ідентифікації не рекомендується для людей в яких є близнюк тієї ж статі, оскільки вірогідність помилкової ідентифікації дуже велика, але під цю категорію потрапляє лише 0,003% населення.

Тому для відображення ненадійності ідентифікації по знімку обличчя буде розглянуто цільову атаку на користувача шляхом виготовлення маски обличчя.

Для експерименту було обрано п’ять особистих пристроїв (смартфонів):: LG G7 ThinQ, Samsung S9, Samsung Note 8, OnePlus 6 та iPhone X.

Також на основі фотографій було створено 3Д-модель обличчя з гіпсу компанією Backface у Бермінгемі, Англія (рисунок 1,5), для експерименту фотографії були зроблені при гарному освітлені та з різних ракурсів, але зараз не є проблемою зробити чітке зображення обличчя людини з середньої або великої дистанції при наявності сучасної фототехніки та відповідних об’єктивів. Альтернативним шляхом отримання фотографій є соціальні мережі, оскільки з кожним роком якість камер у смартфонах підвищується і роздільна здатність зростає.

В ході даного експерименту виявилось, що гіпсового зліпку достатньо для чотирьох з п’яти протестованих пристроїв. Виключенням став iPhone X, оскільки він додатково робить не тільки об’ємну 3Д-модель обличчя, а й знімок обличчя в інфрачервоному діапазоні, тому неможливо обійти систему захисту у такий спосіб (рисунок 1,6).

Рисунок 1,5 – Гіпсова модель обличчя

Рисунок 1,6 – Апаратна складова FaceID

Економічна складова для даного експерименту становила чотириста доларів США. Також слід уточнити, що обладнання досить дороге і його вартість складає понад декілька тисяч доларів США. [17]

Так що такий спосіб захисту є доцільним для захисту інформації, що коштує менше, але недопустимим у випадку якщо в телефоні є корпоративні секрети, або об’єктом атаки є знаменитість, чиї особисті дані можуть коштувати більше.

Власне компанії, над апаратами яких проводилось дослідження зауважили, що технологія розпізнавання обличчя є зручною, але для вищого рівня безпеки все ж таки слід використовувати відбиток пальця або багатофакторну авторизацію.

1.3.5 Використання альтернативних біометричних даних для ідентифікації особистості

Окрім сканування відбитків пальців і розпізнавання обличчя в сучасних смартфонах поки що масово не використовуються інші методи біометричної ідентифікації. Хоча теоретична можливість є. Деякі з них проходять експериментальну перевірку, деякі впроваджені в комерційну експлуатацію в різних сферах, в тому числі сканування сітківки ока, верифікація по голосу і по малюнку вен на долоні, за зразком ДНК.

Відсутність даних технологій у повсякденному житті обумовлюється дорожнечею технологій, обладнання (як у випадку із скануванням сітківки), недостатньо швидкодією, великими габаритами сканеру (малюнок вен на лодоні).

1.3.6 Доцільність використання біометричної ідентифікації

Як і в економіці, в інформаційній безпеці теж є поняття економічної доцільності. Нехай стовідсоткового захисту не існує. Але захисні заходи співвідносяться з цінністю самої інформації. Загалом, принцип приблизно такий, що вартість зусиль по злом для хакера повинна перевищувати цінність для нього тієї інформації, яку він бажає отримати. Чим більше співвідношення - тим надійніший захист.

Однак у всіх методів біометричної захисту є одна фундаментальна уразливість: на відміну від пароля, свої біометричні характеристики практично неможливо замінити. Якщо ваші відбитки пальців злили у відкритий доступ - ви їх вже не зміните. Виходить, що це довічна вразливість.

Другою проблемою біометричних даних є обмеженість набору параметрів: в людини лише десять відбитків пальців рук, два ока та одне обличчя. На відміну від паролів, які можна зробити унікальними для кожного сервісу, із біометричними даними різні сервіси будуть оперувати одними й тими самими даними для ідентифікації і ненадійність хоча б одного з них ставить під загрозу всі інші.

В-третіх очевидно, що тенденція у виробників на зручність використання, а не на надійність, саме через це більшість методів є нестійкими до зламу, іншими словами в пріоритеті мінімізація хибних позитивних відхилень доступу.

Тож перш за все треба пам’ятати, що вартість зусиль на злом повинна перевищувати цінність здобутої інформації, тому доцільно використовувати багатофакторні системі аутентифікації, та зберігати свої біометричні дані і не довіряти їх сумнівним сервісам.

1.4 Принцип роботи нейромереж та основні шляхи їх навчання

Нейронна мережа - це мережа або ланцюг нейронів, або в сучасному розумінні штучна нейронна мережа, що складається із штучних нейронів або вузлів. [18] Таким чином, нейронна мережа - це або біологічна нейронна мережа, що складається з реальних біологічних нейронів, або штучна нейронна мережа для вирішення проблем штучного інтелекту.

Зв'язки біологічного нейрона моделюються як ваги. Позитивна вага відображає збудливий зв’язок, тоді як негативні значення означають гальмівні зв’язки. Усі входи модифікуються вагою та підсумовуються. Ця діяльність називається лінійною комбінацією. Нарешті, функція активації контролює

амплітуду вихідного сигналу. Прийнятий діапазон вихідних даних зазвичай становить від 0 до 1, але може бути наприклад і від -1 до 1.

Правило навчання або алгоритм навчання - це метод або математична модель, яка підвищує продуктивність штучної нейронної мережі і, як правило, це правило застосовується багаторазово по цілій мережі. [19] Це робиться шляхом поновлення ваг і рівнів упередженості мережі. Правило навчання може прийняти існуючі умови (ваги і зміщення) мережі і порівнює очікуваний результат і фактичний результат мережі, щоб дати нові та вдосконалені значення для ваг і зсуву. В залежності від складності конкретної моделі, яка моделюється, правило навчання мережі може бути настільки ж просто, як XOR входів або середньоквадратичної помилки, або це може бути результатом кількох диференціальних рівнянь. [20] Правило навчання є одним з факторів, який визначає, як швидко або наскільки точні можуть бути розроблені штучні мережі.

На сьогодні відомо три парадигми навчання нейронних мереж, в основу яких покладено особливості машинного навчання:

- навчання з вчителем (supervised learning);
- навчання без вчителя (unsupervised);
- навчання з підкріпленням (reinforcement leaning).

Навчання з вчителем (supervised learning) — передбачає, що для кожного вхідного вектору існує вектор вихідних значень. Разом ці два вектори називають навчальною парою, а множину навчальних пар — навчальною вибіркою. Процес навчання зводиться до почергового подавання на вхід нейронної мережі навчальних пар, вираховування похибки між дійсним і бажаним значенням нейронної мережі та корегування параметрів мережі в бік зменшення цієї похибки.

Навчання без вчителя (unsupervised)— один зі способів машинного навчання, при вирішенні яких випробовувана система спонтанно навчається виконувати поставлене завдання, без втручання з боку експериментатора. З точки зору кібернетики, є одним з видів кібернетичного експерименту. Як правило, це підходить тільки для задач, в яких відомий опис множини об'єктів

(навчальна вибірка), і необхідно виявити внутрішні взаємозв'язки, залежності, закономірності, що існують між об'єктами. Цей підхід часто протиставляється навчанню з учителем, коли для кожного об'єкта, що навчається, примусово задається «правильна відповідь», і потрібно знайти залежність між стимулами та реакціями системи.

Навчання з підкріпленням є проміжним варіантом двох попередніх парадигм . Замість «вчителя» в схему навчання вводиться блок «критика» , який відслідковує реакцію середовища на вхідний сигнал і опираючись на неї визначає евристичну похибку, яку покладено в процес навчання мережі. Вказані парадигми базуються на відповідних правилах навчання, які визначають основні особливості їх застосування.

1.5 Перетворення біометричних даних

1.5.1 Нечіткі екстрактори

Біометрична система аутентифікації на основі нечітких екстракторів також, як і всі розглянуті системи, потребує виконання етапу реєстрації користувача. На етапі біометричної реєстрації користувача створюється ключ key, який потім перетворюється в двійковий вектор KEY з використанням перешкодостійких кодів таких, як Боуза-Чоудхурі-Хоквінгема (БЧХ) та Ріда-Соломона:

$$KEY = Encode(key), \tag{1.1}$$

Більшість нечітких екстракторів базується на гамуванні. [21] Двійковий вектор KEY шифрується гамуванням, а в якості гами використовується двійковий вектор біометричних параметрів B, що вираховується відповідно до отриманої вибірки на етапі реєстрації:

$$T = B \oplus KEY, \tag{1.2}$$

де:

- знак $\oplus$ означає операцію побітового виключення АБО;
- T – двійковий вектор, результат виконання операції гамування.

У відкритому хелпері зберігаються T та Hash(key) – хеш-сума ключа key. Сам відкритий хелпер розміщується в базу даних відкритих хелперів зареєстрованих користувачів.

В процесі біометричної аутентифікації при пред'явленні користувачем біометричних даних знову обраховується двійковий вектор біометричних параметрів B'. Відновлення ключа key відбувається з використанням відкритого хелпера:

$$KEY' = T \oplus B' = KEY \oplus err, \tag{1.3}$$

де:

- err – похибка, що з'являється через різницю між двома двійковими векторами біометричних параметрів B та B', що зумовлюється неточністю відтворюваних біометричних даних.

Таким чином, ключ key' необхідно декодувати:

$$key' = Decode(KEY'), \tag{1.4}$$

Процедура аутентифікації вважається пройденою при виконанні рівності:

$$Hash(key') = Hash(key), \tag{1.5}$$

У випадку нерівності хеш-суми користувач не проходить перевірку його дійсності. На рисунках 1.7 - 1.8 зображено принцип роботи нечіткого екстрактора, що використовує коди БЧХ для виправлення похибки err.

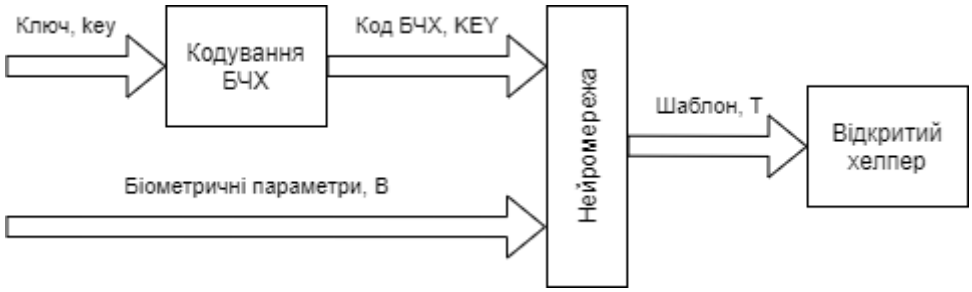

Рисунок 1.7 – Формування відкритого хелпера (реєстрація біометрії)

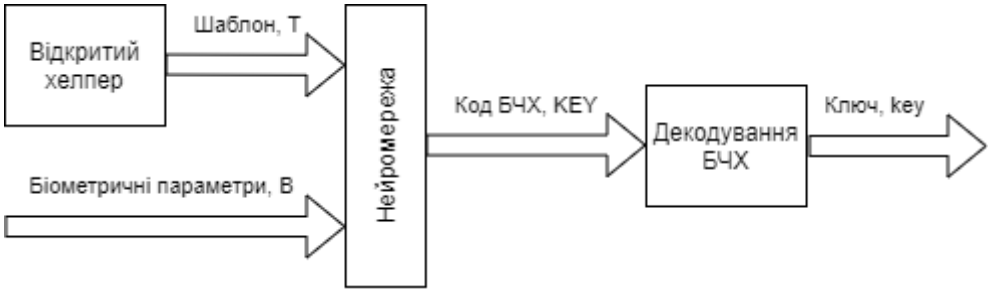

Рисунок 1.8 – Формування ключа нечітким екстрактором

Для того, щоб вважати перетворення з використанням нечіткого екстрактора надійними повинно виконуватися декілька умов. По-перше вхідний вектор двійкових біометричних параметрів B повинен мати рівноймовірний розподіл біт. По-друге вхідний вектор двійкових біометричних параметрів повинен містити незалежні (некореляційні) дані. По-третє дані T, які розміщуються у відкритому хелпері не повинні призводити до витікання криптографічного ключа. В-четверте при пред'явленні образів «Чужий» перетворювач повинен формувати випадкову кодову комбінацію.

Ймовірність похибок першого роду P_I, тобто здатність системи формувати ключ користувача при пред'явленні біометричного образа «Свій» суттєво залежить від похибки err. В одних з останніх робіт вітчизняних дослідників показано, що відновлення 25% помилок є недостатнім для повного відновлення ключа з вектору двійкових біометричних параметрів. Для забезпечення рівня P_I, менше ніж 0.1 дослідниками вводиться надмірність коригуючого коду, отже його позиції стають більш залежними. Завищення оцінки ймовірності похибки другого роду P_{II} розробниками систем біометричної аутентифікації зумовлюється тим, що не приймають до уваги кореляційні зв'язки біометричних параметрів.

Спроби використання кодів з виявленням й виправленням помилок призводять до багатократної надмірності двійкового вектору біометричних параметрів B. Доводиться застосовувати коди з 90% надмірністю, втрачаючи

велику частину низькоякісних параметрів (використовується не більш, ніж 10% вхідних параметрів, що мають достатню якість).

Можна відмітити декілька робіт, в яких доволі детально описується принцип побудови нечітких екстракторі, з'єднуючих біометрію та криптографію, тобто перетворюючих біометрію в особистий ключ користувача.

Таким чино, структура нечіткого екстрактора являє собою коригуючий код (БЧХ, Ріда-Соломона), який прагне виправити помилки, що виникають в біометричних параметрах, с ціллю дешифрування точного значення ключа з використанням відкритого хелпера. Основна складність, що виникає при побудові нечіткого екстрактора, полягає в створенні алгоритмів обрахування більш стійких до шумів біометричних параметрів рисунка відбитка пальця. Насьогодні загальноприйнятного підходу до створення такого алгоритму не існує.

1.5.2 Нейромережеві перетворювачі біометрія-код

Як вже було вище, нейромережеві перетворювачі біометрія-код використовують навчену штучну нейронну мережу для формування точного значення ключа. На етапі біометричної реєстрації користувача створюється ключ key і на його основі штучна нейрона мережа автоматично навчається без участі людини (вчителя).

Взагалі, стандарт передбачає одношарові та двошарові нейромережі. На рисунку 1.9 зображено загальна схема навчання.

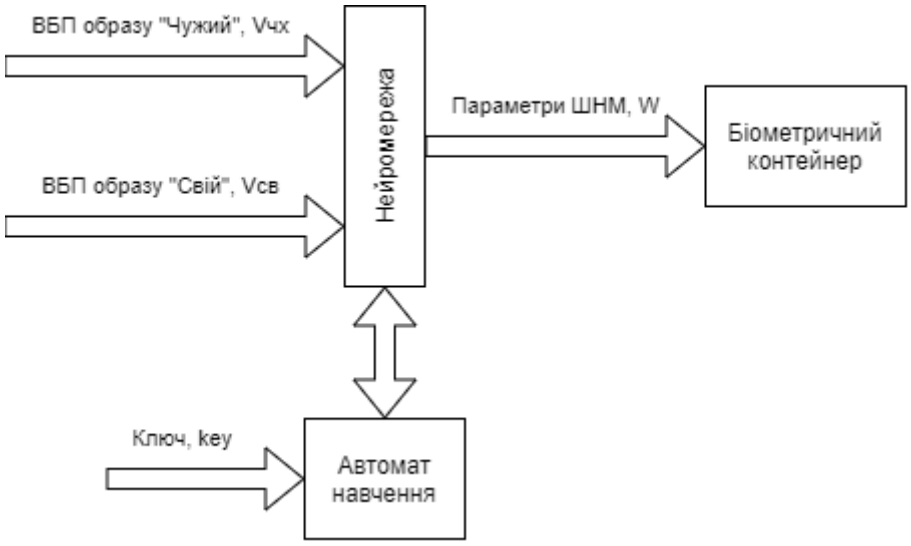

Рисунок 1.9 – Загальна схема методу автоматичного навчання нейромережевого перетворювача біометрія-код

Крім ключа для навчання нейромережі потребують вибірки (приклади) двох класів образів – «Свій» та «Чужий», представлені у вигляді векторів біометричних параметрів $V_{Св}$ та $V_{Чуж}$ відповідно. Передбачається навчання на безперервних та дискретних біометричних параметрах.

Формально штучна нейронна мережа описується матрицею вагових коефіцієнтів W або таблицями зв’язків нейронів та таблицями вагових коефіцієнтів, які вкладаються в біометричний контейнер для подальшого відновлення ключа. Процес навчання нейромережі необхідний для обрахування W такого, щоб при взаємодії на нього (подачі на вхід) прикладу образа «Свій» на виході нейромережі формувалось точне значення ключа key. Загальна схема нейромережевого перетворення біометрія-код зображена на рисунку 1.10. При впливі на нейромережу прикладами образа «Чужий» на виходах повинні формуватися випадкові вихідні коди.

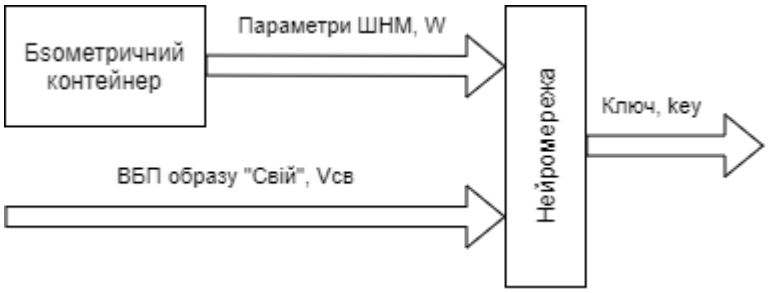

Рисунок 1.10 – Загальна схема нейромережевого перетворення біометрія-код

Перевагою нейромережевих перетворювачів біометрія-код вважаються хешуючі властивості нейромережі, які можуть бути значно посилені механізмом розмноження помилок проміжних кодів біометричних перетворень, заснованому на доданні по модулю два даних ще не використаних нейронів з фрагментами вже отриманого попередніми нейронами вихідного коду доступу або криптографічного ключа.

1.6 Висновки по розділу 1

Проведено огляд основних видів біометричних систем, який показав, що їх можна розділити на системи в яких обробляються біометричні шаблони, і системи, в яких обробляються біометричні контейнери.

Проведено аналіз закордонних і вітчизняних технологій перетворення біометричних даних в код аутентифікації (біометрія-код), який показав, що існує два головних напрямки рішення задачі. Перший напрямок характеризує використання нечітких екстракторів на основі кодів з виявленням і виправленням помилок. Другий напрямок характеризується застосуванням великих штучних нейронних мереж. Аналіз літератури показав, що дослідження нечітких екстракторів в основному проходять закордоном. В той час, як в Україні створений й активно розвивається напрямок нейромережевих перетворювачів біометричних даних людини в код доступу.

Розроблені вітчизняні стандарти високонадійної нейромережевої біометричної аутентифікації. Тоді як нечіткі екстрактори не є стандартизованою технологією. Стандартів, присвячених проектуванню й розробленню біометричних систем на основі нечітких екстракторів введено так і не було.

На основі нейромережевого перетворення й нечіткого екстрактора можуть бути розроблені засоби знеособлення біометричних даних, які дозволять зробити неможливим визначення належності біометричного контейнера конкретному суб’єкту. Тому актуальними задачами є розроблення методів знеособлення біометричних даних та їх моделювання.

2 ПЛАНУВАННЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ МЕТОДІВ ПЕРЕТВОРЕННЯ БІОМЕТРИЧНИХ ДАНИХ ОСОБИ ДЛЯ АУТЕНТИФІКАЦІЇ

2.1 Методи перетворення персональних даних

2.1.1 Метод введення ідентифікатора

До методів перетворення, відносять:

− метод введення ідентифікаторів, в якому частина персональних даних замінюється ідентифікаторами і створюється таблиця, в якій ідентифікатори відповідають вихідним значенням;

− метод зміни складу або семантики, в якому з використанням заміни результатів статистичної обробки, перетворення, узагальнення або видалення частини персональних даних, змінюється склад або семантика;

− метод декомпозиції, в якому безлічі персональних даних поділяються на кілька підмножин, які окремо зберігаються;

− метод перемішування, в якому групи значень або окремі значення атрибутів даних в масиві персональних даних переставляються.

Із застосуванням одного з методів знеособлення отримують знеособлені дані, які мають різні властивості, що здійснює всі види обробки персональних даних. Тому існують умови, при описі методів, які забезпечують здійснення певних властивостей і вимог.

Метод введення ідентифікаторів здійснюється шляхом заміни частини персональних даних, які дозволяють ідентифікувати суб'єкт і створювати таблиці відповідності.

Із застосуванням методу з'являється можливість отримувати знеособлені дані що мають такі властивості, як:

− повнота, в якому інформація, що дозволяє ідентифікувати суб'єкти персональних даних, переноситься в таблицю відповідності, а не видаляється;

− структурованість, при якій, після процедури знеособлення кожному ідентифікатору відповідає свій набір даних;

− семантична цілісність, при якій представлені дані переносяться в таблицю відповідності.

Можливість забезпечення анонімності відбувається при певних умовах вибору ідентифікаторів і персональних даних, які вони замінюють. метод стабільних до атак, які пов'язані з довідниками ідентифікаторів при непрямому персоніфіковані і до атак, пов'язаними з персоніфікацією, яке використовує інформацію з довідників ідентифікаторів. Крім того, при збільшенні обсягу даних, що знеособлюються, стійкість методу не підвищується.

Отримані в результаті застосування даного методу знеособлені дані, не володіють властивістю релевантності, так як в запиті і у відповіді на запит змінюється вид уявлення персональних даних, які були замінені ідентифікаторами.

Із застосуванням цього методу з'являється можливість зберігати в записах зв'язку між атрибутами знеособлених даних, які відповідають зв'язкам між атрибутами персональних даних

Застосування методу потрібно здійснювати при малій кількості атрибутів персональних даних і малому обсязі масиву персональних даних, так як обсяг довідників залежить від цих параметрів. Якщо часто вносити зміни до складу даних і значення атрибутів, то обчислювальна ефективність методу буде знижена.

2.1.2 Методи зміни складу чи семантики

Метод зміни складу або семантики здійснюється за допомогою додавання, зміни значень атрибутів персональних даних або видалення інформації, яка дозволяє розпізнати суб'єкт.

Із застосуванням методу з'являється можливість отримувати знеособлені дані, що володіють такими властивостями, як:

− структурованість, при якій не порушується зв'язок між окремими значеннями атрибутів персональних даних суб'єкта;

− анонімність, при якій до неоднозначності, при ідентифікації із застосуванням знеособлених даних, призводить видалення або додавання деяких даних.

Властивість повноти мають отримані знеособлені дані при проведенні змін в складі персональних даних, які зберігають дані. В отриманих знеособлених даних зменшується властивість повноти, якщо видалити частини інформації.

Якщо провести зміни в складі персональних даних, які зберігають семантику даних, то забезпечується семантична цілісність отриманих даних.

У тому числі забезпечуються такі властивості знеособлених даних, як:

− часткова релевантність, так як іноді отримують семантичну відповідність пошукового запиту і отриманої відповіді на запит;

− придатність, так як оператор може виконати обробку, яка не вимагає персоніфікації всього обсягу даних про суб'єктів.

Потрібно враховувати ймовірність проведення знеособлення із застосуванням даних атрибутів, при виділенні атрибутів персональних даних. Знеособлення може не виконатися, в разі простої зміни значень окремих атрибутів, так як виконається тільки зміна складу персональних даних.

Із застосуванням методу з'являється можливість частково зберігати в записах зв'язки між атрибутами знеособлених даних, які мають відповідності зв'язків між атрибутами персональних даних.

Даний метод застосовується в разі, коли можлива зміна складу і семантики, так, що завдання обробки персональних даних не вимагають персоніфікації, так як метод не має властивість оборотності при будь-яких змінах складу і семантики даних. інакше необхідно використовувати додаткову інформацію для проведення персоніфікації.

Також метод застосовується, коли оператор використовує автономно знеособлені дані, і коли несумісні з даними інших операторів.

2.1.3 Метод декомпозиції

Метод декомпозиції здійснюється за допомогою поділу безлічі атрибутів персональних даних на кілька підмножин і створення таблиць, які встановлюють зв'язок між підмножинами, з подальшим роздільним зберіганням записів, які відповідають підмножини цих атрибутів.

Із застосуванням методу з'являється можливість отримувати знеособлені дані, що володіють такими властивостями, як:

− повнота, при якій інформація про суб'єкти персональних даних переноситься в інше сховище, а не видаляється;

− структурованість, при якій зберігається зв'язок між записами в різних сховищах, яке дозволяє однозначно співставити їх;

− семантична цілісність, при якій не змінюється семантика і вид представлення даних про суб'єкт.

При складних зв'язках між сховищами і захисту сховищ від несанкціонованого доступу виконується анонімність, так як метод не стійкий до атак, які спрямовані на персоніфікацію шляхом аналізу даних з різних сховищ і непрямій персоніфікації.

Також забезпечуються такі властивості знеособлених даних, як:

− релевантність, так як можливо отримати семантичну відповідність пошукового запиту і отриманої відповіді на запит;

− придатність, так як оператор виконує обробку даних, які розташовуються в одному сховищі, як незалежно від іншого, так і при спільному їх використанні, без персоніфікації всього обсягу знеособлених даних.

Із застосуванням методу з'являється можливість зберігати в записах кожного сховища зв’язок між атрибутами знеособлених даних, які відповідають зв'язкам між атрибутами персональних даних.

Даний метод застосовується при значній кількості атрибутів персональних даних, але прі не частому внесення змін до даних і значення атрибутів.

2.1.4 Метод перемішування

Метод здійснюється за допомогою перемішування між собою окремих значень або груп значень атрибутів персональних даних.

Після застосування методу з'являється можливість отримувати знеособлені дані, що володіють такими властивостями, як:

- повнота, при якій зберігається вся інформація про суб'єкти персональних даних;
- структурованість, в якій, при персоніфікації, зв'язки між даними повністю відновлюються;
- семантична цілісність, при якій не змінюється семантика і вид представлення даних про особу;
- анонімність, при якій перемішуються дані кожного окремого атрибуту запису про суб'єкта, що не дозволяє без доступу до додаткової інформації визначити приналежність тих чи інших даних конкретного суб'єкта.

Також забезпечуються такі властивості знеособлених даних, як:

- релевантність, так як є можливість отримувати семантичну відповідність пошукового запиту і отриманої відповіді на запит;
- придатність, так як оператор може виконувати обробку як окремих записів про суб'єктів, так і всіх даних, без персоніфікації всього обсягу знеособлених даних, при наявності доступу до додаткової інформації.

Після застосування методу не зберігаються зв'язки в записах між атрибутами знеособлених даних, які відповідають зв'язкам між атрибутами персональних даних.

Даний метод застосовується при значній кількості атрибутів персональних даних і великому обсязі масиву персональних даних, так як стійкість методу до атак, спрямованих на персоніфікацію збільшується зі збільшенням зазначених

параметрів, а кількість додаткової інформації слабко залежить від обсягу масиву персональних даних.

2.2 Обчислення біометричних параметрів відбитка пальця

2.2.1 Двійкові біометричні параметри

Найбільш інформативними особливостями відбитка пальця є контрольні точки: закінчення гребнів, розгалуження гребнів. Взаємне розташування таких точок представляє унікальну інформацію, за якою можна однозначно ідентифікувати особу.

а) розгалуження б) закінчення

Рисунок 2.1 - Два типи особливих точок на папілярному малюнку відбитка пальця

Контрольні точки можна описати їх набором з 4 параметрів:

1) координата x;
2) координата y;
3) напрямок папілярної лінії в точці O;
4) тип точки t.

Сам відбиток пальця можна описати як безліч контрольних точок:

$$T = \{m_1, m_2, \dots, m_k\}, m_i = \{x_i, y_i, \theta_i, t_i\}, i = 1 \dots k, \quad (2.1)$$

де:

- m_i – контрольна точка, виявлена на зображенні відбитка пальця;

- k - число контрольних точок

Безліч контрольних точок T є біометричним шаблоном, який може бути використаний для порівняння з іншими зразками. Процес порівняння ведеться шляхом зіставлення різних пар контрольних точок, що містяться в шаблоні, і точок зразка. коефіцієнт відповідності двох відбитків пальців визначається за формулою:

$$K = \frac{D^2}{T \cdot S} \cdot 100\%, \quad (2.2)$$

де:

- D – кількість контрольних точок, що збіглися;
- T – кількість контрольних точок в шаблоні;
- S - кількість контрольних точок в зразку або ідентифікованому відбитку пальця.

Відбитки вважаються ідентичними при K більше 65%. Зберігання біометричного шаблону в базах даних несе в собі потенційну загрозу безпеці біометричних даних. втрата конфіденційності біометричного шаблону буде означати, що конфіденційність даних втрачена назавжди. Біометричний шаблон містить унікальні параметри людини в чистому вигляді і, опинившись в руках у зловмисника, дає йому можливість для здійснення злому біометричних систем і отримання цінних даних про людину.

Розташування контрольних точок відносно один одного може бути описано дискретним біометричним параметром у вигляді матриці позицій.

Малюнок відбитка пальця з виділенням фрагментів, де знаходяться особливі точки і відповідна їх розташуванню матриця представлений на рисунку 2.2.

Рисунок 2.2 - Малюнок відбитка пальця з виділенням фрагментів, де знаходяться особливі точки і відповідна їх розташуванню матриця

Можливе використання різних алгоритмів порівнянь пред'явленого малюнка відбитка пальця і еталонного малюнка. Найпростішим є порівняння матриць особливостей, представлене на правій частини малюнка 2.2. При необхідності дискретні біометричні параметри можуть бути переведені в безперервні.

2.2.2 Безперервні біометричні параметри

При розпізнаванні відбитків пальців за допомогою штучної нейронної мережі, унікальні особливості - сукупність контрольних точок і напрямки папілярних ліній можуть бути представлені просторово-частотними компонентами в горизонтальному і вертикальному вимірах. Дане рішення дозволяє отримати параметри, які описуються безперервною множиною можливих значень. Просторово-частотні компоненти широко застосовуються в методах стиснення зображень і обчислюються за допомогою двовимірних дискретних перетворень. Розглянемо двовимірні дискретні перетворення, які можуть бути застосовані для представлення відбитка пальця у вигляді набору параметрів і навчанні штучної нейронної мережі.

Відомим сімейством двовимірних дискретних перетворень є дискретне косинус-перетворення, перетворення Кархунена-Лоева, перетворення Уолша-Адамара, перетворення Хаара.

Дискретне косинус-перетворення (DCT, discrete cosine transform) є ортогональним перетворенням, тісно пов'язане з дискретним перетворенням Фур'є і достатньо вивчене. досить ефективно застосовується в таких методах стиснення зображень, як JPEG і MPEG.

Перетворення Кархунена-Лоева (KLT, Karhunen-Loeve transform) вважається теоретично найкращим з точки зору концентрації енергії. Однак його коефіцієнти залежать від вихідних даних, і їх обчислення проходить повільно. З цих причин метод KLT вельми рідко застосовується на практиці.

Перетворення Уолша-Адамара (WHT, Walsh-Hadamard transform) швидко обчислюється. Перетворення досить просто реалізується апаратно, так як використовується тільки додавання і віднімання значень даних. Однак за своїми властивостями виявляється трохи гірше за DCT.

Перетворення Хаара є вейвлетним перетворенням, яке використовується в методах стиснення зображень, в основному кольорових і чорно-білих з плавними переходами.

В якості функціоналів використовуються методи DCT і WHT, оскільки перевага, перш за все, віддається використанню відомих і простих в реалізації перетворенням.

Двовимірне DCT зображення позицій контрольних точок обчислюється за формулами:

$$D_{ij} = \frac{1}{\sqrt{2n}} C_i C_j \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} B_{uv} \cos(\frac{(2y+1)j\pi}{2n}) \cos(\frac{(2x+1)i\pi}{2n}) \quad (2.3)$$

$$C_f = \begin{cases} \frac{1}{\sqrt{2}}, & f = 0 \\ 1, & f > 0, \end{cases} \quad (2.4)$$

де:

- 0 менше або дорівнює i, j менше або дорівнює n-1;
- b_{uv} - зображення позицій контрольних точок розмірності n*n.

Метод DCT можна інтерпретувати за допомогою базису в n-мірному векторному просторі. Так для n дорівнює 16 виходить 256 базисів.

Двовимірне WHT обчислюється за формулою:

$$G_{ij} = \frac{1}{N} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} B_{xy} (-1)^{\sum_{i=0}^{m} [b_i(x)p_i(u) + b_i(y)p_i(v)]}, \quad (2.5)$$

де:

- n дорівнює 2^m- ступінь двійки;
- величина $b_i(u)$ дорівнює біту *i* в двійковому поданні цілого числа u, а $p_i(u)$ визначається за допомогою $b_i(u)$ p наступних рекурентних співвідношень:

$$\begin{aligned} p_0(u) &= b_{n-1}(u), \\ p_1(u) &= b_{n-1}(u) + b_{n-2}(u), \\ p_2(u) &= b_{n-2}(u) + b_{n-3}(u), \\ &\dots \\ p_{n-1}(u) &= b_1(u) + b_0(u). \end{aligned} \quad (2.6)$$

При n рівному 16 виходить 256 базисів WHT.

Перетворення матриці з 256 елементів дає стільки ж просторово-частотних коефіцієнтів, розмірність вектору безперервних біометричних параметрів повинна бути підібрана виходячи з узагальнюючої здатності штучної нейронної мережі.

Таким чином, відбиток пальця можна уявити просторово-частотними компонентами за допомогою двовимірного дискретного перетворення, що враховує зміну координат контрольних точок і їх взаємне розташування за двома вимірами (x, y). базисні зображення дають зрозуміти, що отримувані значення коефіцієнтів залежні від зміщення і повороту контрольних точок на зображенні, які виникають при скануванні відбитка пальця. Високе значення дисперсії біометричних параметрів може бути причиною низької узагальнюючої здатності штучної нейронної мережі, навченої на вибірці способу «Свій». Дисперсія біометричних параметрів обумовлена зміщенням контрольних точок на зображенні, одержуваному при скануванні відбитка пальця.

Для дослідження характеристик безперервних біометричних параметрів малюнка відбитка пальця була сформована база з 100 біометричних образів «Свій» і база «всі Чужі», що містить 100 відбитків пальців різних людей

Як правило, біометричні дані неточно відтворювані, і це призводить до варіації обчислюваних параметрів . Наприклад , деформація шкіри призводить до того , що частина особливостей може бути не виявлена на одному зображенні малюнка відбитка пальця і виявлена на іншому .Відомо , що похибка визначення математичних очікувань і стандартних відхилень біометричних параметрів впливає на помилку навчання .На малюнках 2.3 і 2.4 представлені гістограми розподілу різних параметрів (20-й, 30-й, 60-й і 81-й параметр) образів «Свій» і «всі Чужі» відбитка пальця, показаний графік нормального розподілу, побудований для обчислених значень математичного очікування і стандартного відхилення.

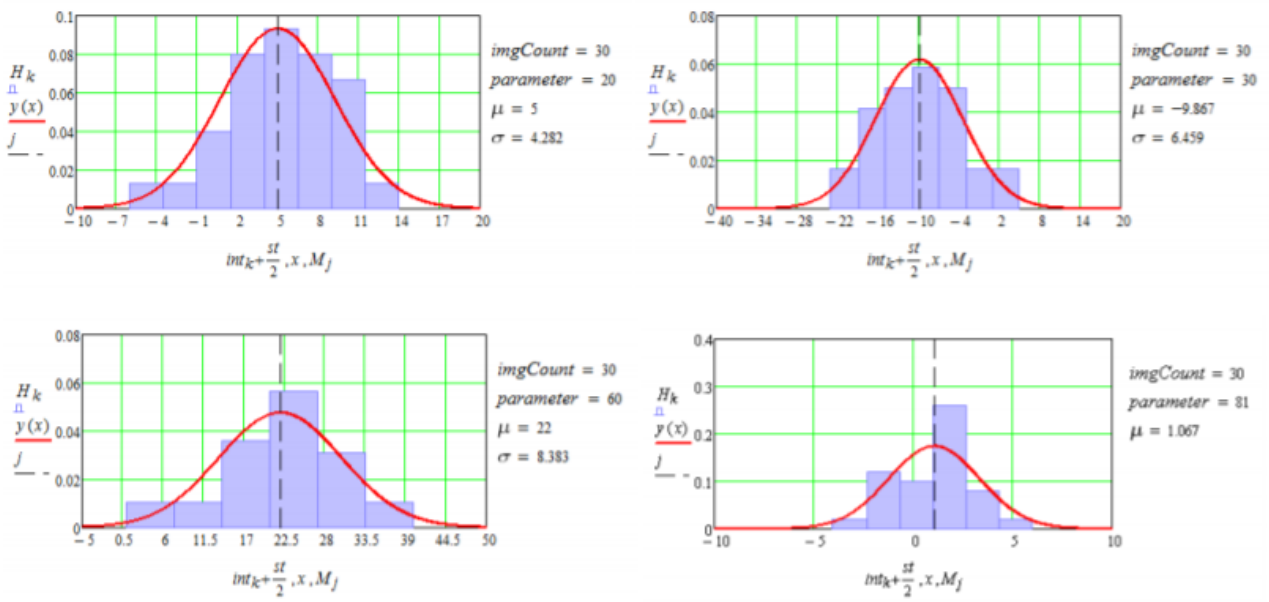

Рисунок 2.3 – Гістограма розподілу біометричних параметрів образу «Свій»

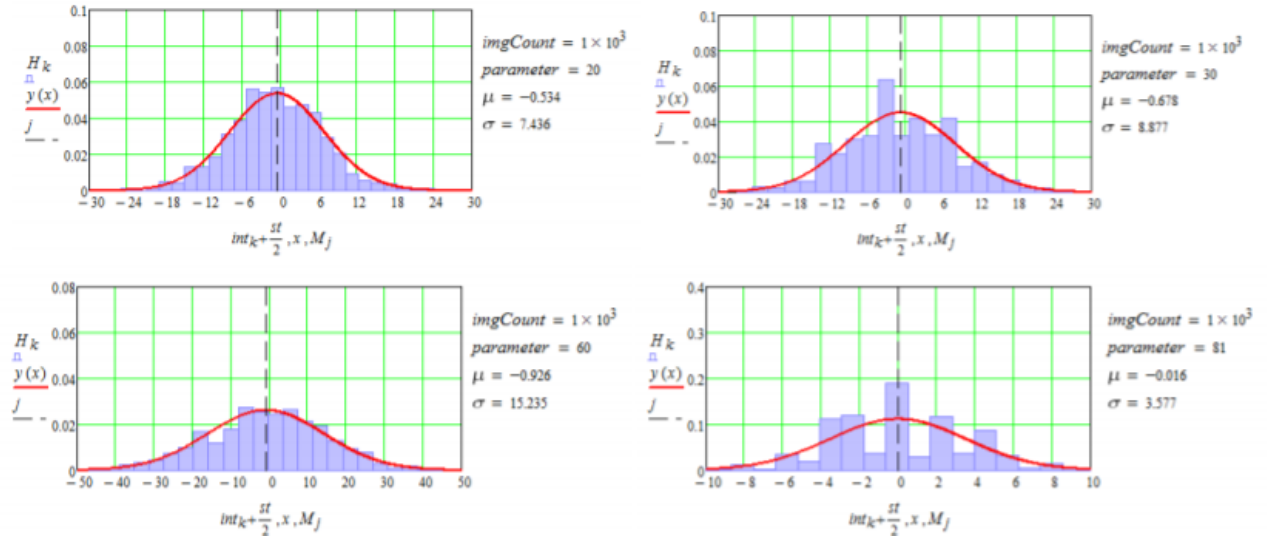

Рисунок 2.4 – Гістограма розподілу біометричних параметрів образу «всі Чужі»

На рисунку 2.5 представлений графік щільності ймовірності розподілу одного біометричного параметра для користувачів «Свій», «Чужий 1», «Чужий 2», «Чужий 3», «Чужий 4» і «все Чужі».

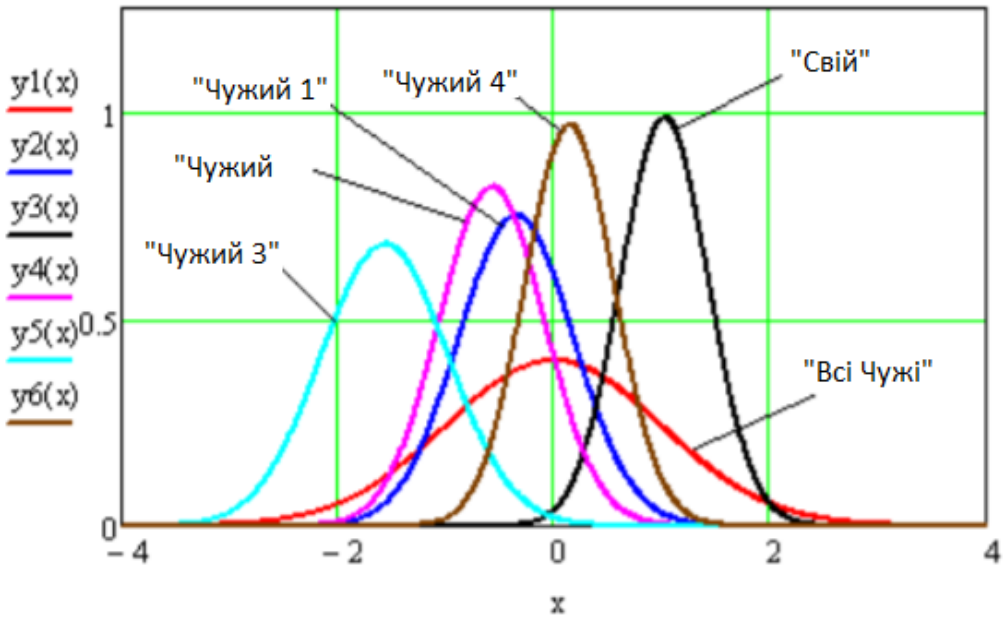

Рисунок 2.5 – Графік розподілу біометричного параметра різних образів «Свій», «Чужий 1», «Чужий 2», «Чужий 3», «Чужий 4» і «все Чужі».

З рисунка видно, що розподіл біометричних параметрів, що описується нормальним законом розподілу, істотно розрізняються . Математичні очікування параметрів різних образів виявляється рознесеними в області значень «Всі Чужі».

Кореляційний аналіз безперервних біометричних параметрів образів «Всі Чужі» малюнка відбитка пальця показав, що математичне очікування модулів коефіцієнтів парної кореляції $E\left(\left|r_{v_i,v_j}\right|\right)$ приблизно дорівнює 0,1. Оцінка кореляційних зв'язків біометричних параметрів образу «Свій» показала, що коефіцієнт кореляції між парами вхідних біометричних параметрів образу «Свій» є випадковою величиною, добре описуваною нормальним законом розподілу значень. І математичне очікування модулів парної кореляції $E\left(\left|r_{v_i,v_j}\right|\right)$становить значну величину від 0,2 до 0,6 (див. рисунок 2.6).

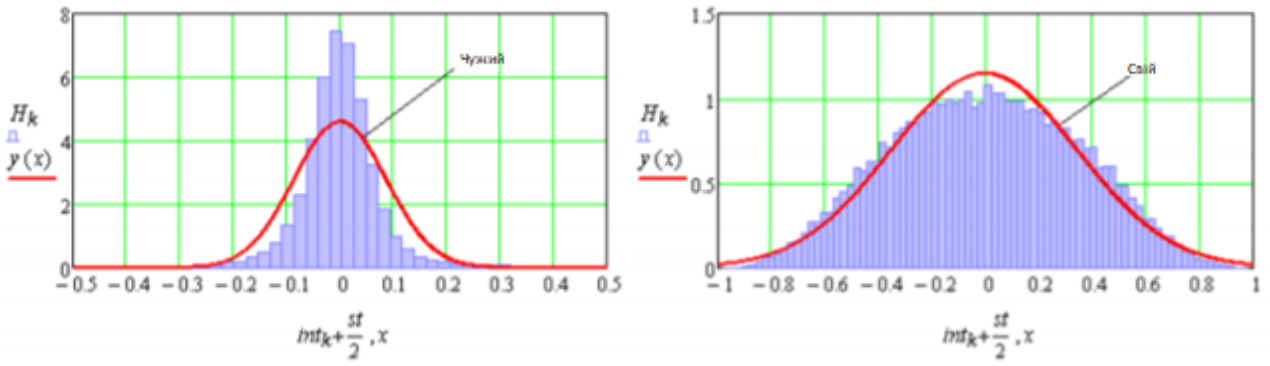

Рисунок 2.6 - Гістограми розподілів коефіцієнтів парної кореляції

2.3 Метод знеособлення персональних даних з використанням нейромережевого перетворювача біометрія – код

Аналіз методів знеособлення і їх алгоритмів показав, що процес можна описати як поділ безлічі атрибутів персональних даних суб'єкта P на дві підмножини - ідентифікуючих даних P_1 і знеособлених даних P_2:

$$P_1 = \{d_{11}, d_{12}, \dots, d_{1m}\}; \quad (2.7)$$

$$P_2 = \{d_{21}, d_{22}, \dots, d_{2n}\}; \quad (2.8)$$

$$P = P_1 \cup P_2. \quad (2.9)$$

Підмножина ідентифікуючих даних P_1 містить дані, за допомогою яких суб'єкт може бути однозначно ідентифікований. підмножина знеособлених даних P_2 не дозволяє однозначно ідентифікувати суб'єкт, тим самим забезпечує анонімність оброблюваних даних. персоніфікація полягає в об'єднанні двох підмножин, результат якого дозволяє визначити приналежність знеособлених даних конкретному суб'єкту.

Для збереження анонімності знеособлених даних пропонується замість персоніфікації використовувати біометричну аутентифікацію, для чого в підмножина знеособлених даних слід включити дані захищеного біометричного контейнера B. Схема методу знеособлення персональних даних з використанням нейромережевого перетворювача біометрія - код представлений на рисунку 2.7.

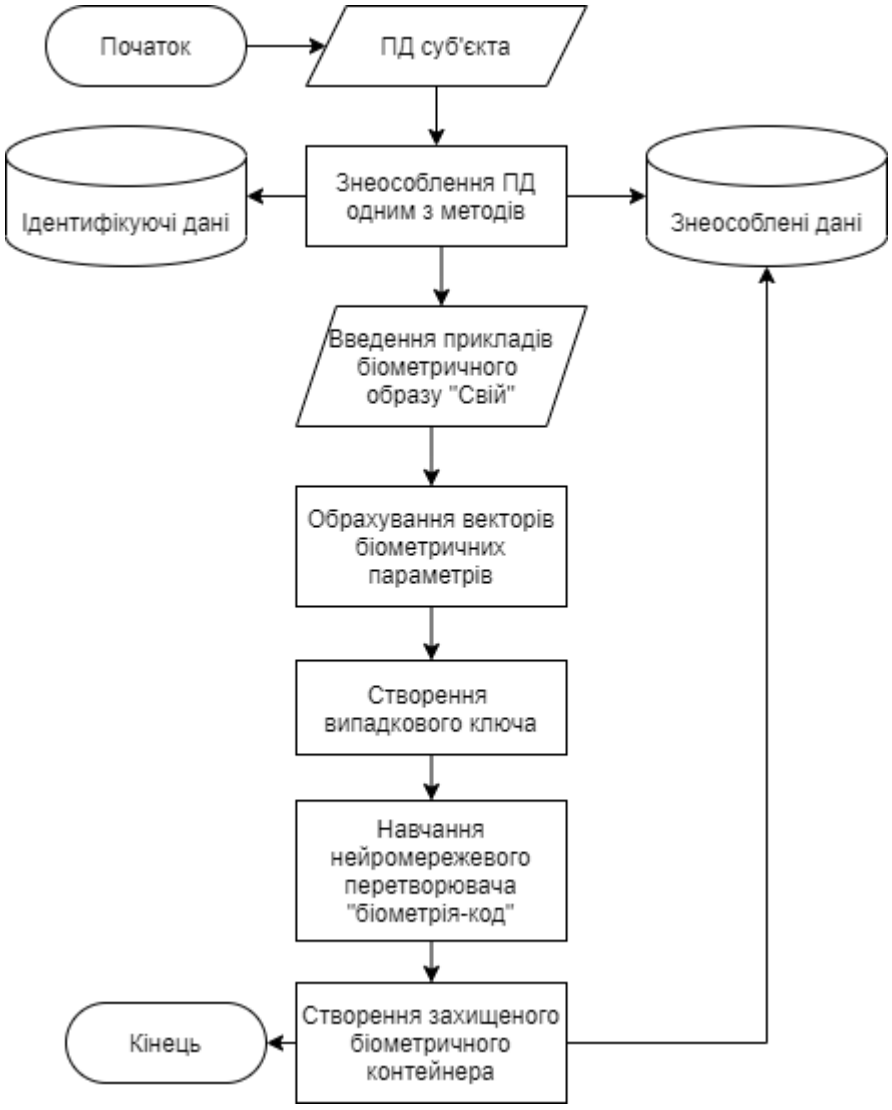

Рисунок 2.7 – Схема методу знеособлення персональних даних з використанням нейромережевого перетворювача

Отримані персональні дані суб'єкта знеособлюються за допомогою одного або декількох методів: методу введення ідентифікаторів, методу зміни складу або семантики, методу декомпозиції, методу перемішування. В результаті утворюються ідентифікатори і знеособлені дані. Біометричні дані знеособлюються за одним із методів. Для чого суб'єкт пред'являє кілька прикладів біометричного образу «Свій», тим самим формуючи навчальну вибірку. Потім обчислюються вектор біометричних параметрів і автоматично навчається штучна нейронна мережа на випадково згенерованому двійковому

ключі. Після закінчення навчання створюється захищений біометричний контейнер B, що містить ідентифікатор (псевдонім) суб'єкта Id, матрицю вагових коефіцієнтів W і Hash - значення хеш - функції від особистого ключа (коду доступу) суб'єкта. Вміст біометричного контейнера включається в знеособлені дані, що обробляються оператором інформаційною системою персональних даних, тобто B є підмножиною P_2.

Знеособлені дані є анонімними, що дозволяє запобігти несанкціонованому використанню персональних даних і знизити збиток від можливого розголошення відомостей про громадян, які обробляються оператором.

Для визначення приналежності знеособлених даних конкретному суб'єкту без доступу до ідентифікуючих даних і втрати анонімності відповідно до алгоритму, знеособлені дані знаходяться за ідентифікатором Id . Потім суб'єкт пред'являє зображення відбитка шляхом притискання до сканера пальця. Далі обчислюється вектор біометричних параметрів, а параметр W використовується для перетворення даного вектору в особистий ключ (код доступу) суб'єкта за допомогою штучної нейронної мережі. Алгоритм персоніфікації персональних даних з використанням нейромережевого перетворювача представлений на малюнку 2.8.

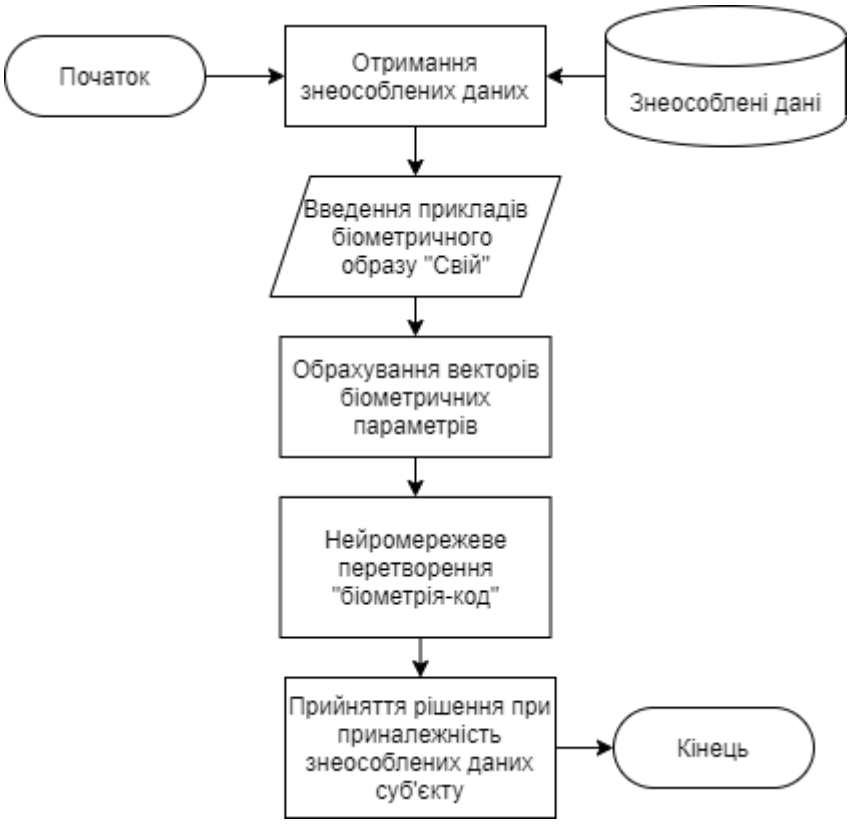

Рисунок 2.8 – Алгоритм персоніфікації персональних даних з використанням нейромережевого перетворювача

Рішення про приналежність даних визначається за результатом порівняння значення хеш - функції від сформованого ключа і значення Hash, що міститься в знеособлених даних. Якщо значення збігаються, то робиться висновок про те, що знеособлені дані належать суб'єкту. якщо значення не збігаються, то робиться висновок про неналежність знеособлених даних суб'єкту.

Таким чином, при B є підмножиною P_2 приналежність знеособлених даних суб'єкту може бути визначена без доступу до ідентифікуючих даних і процедури персоніфікації. За рахунок введення біометричної аутентифікації анонімність знеособлених даних зберігається. Доступ до ідентифікуючих даних може бути надано тільки обмеженому числу співробітників оператора при відсутності можливості здійснення біометричної аутентифікації суб'єкта.

2.4 Метод знеособлення персональних даних з використанням нечіткого екстрактора

Схема методу знеособлення ПДН з використанням нечіткого екстрактора представлено на рисунку 2.9.

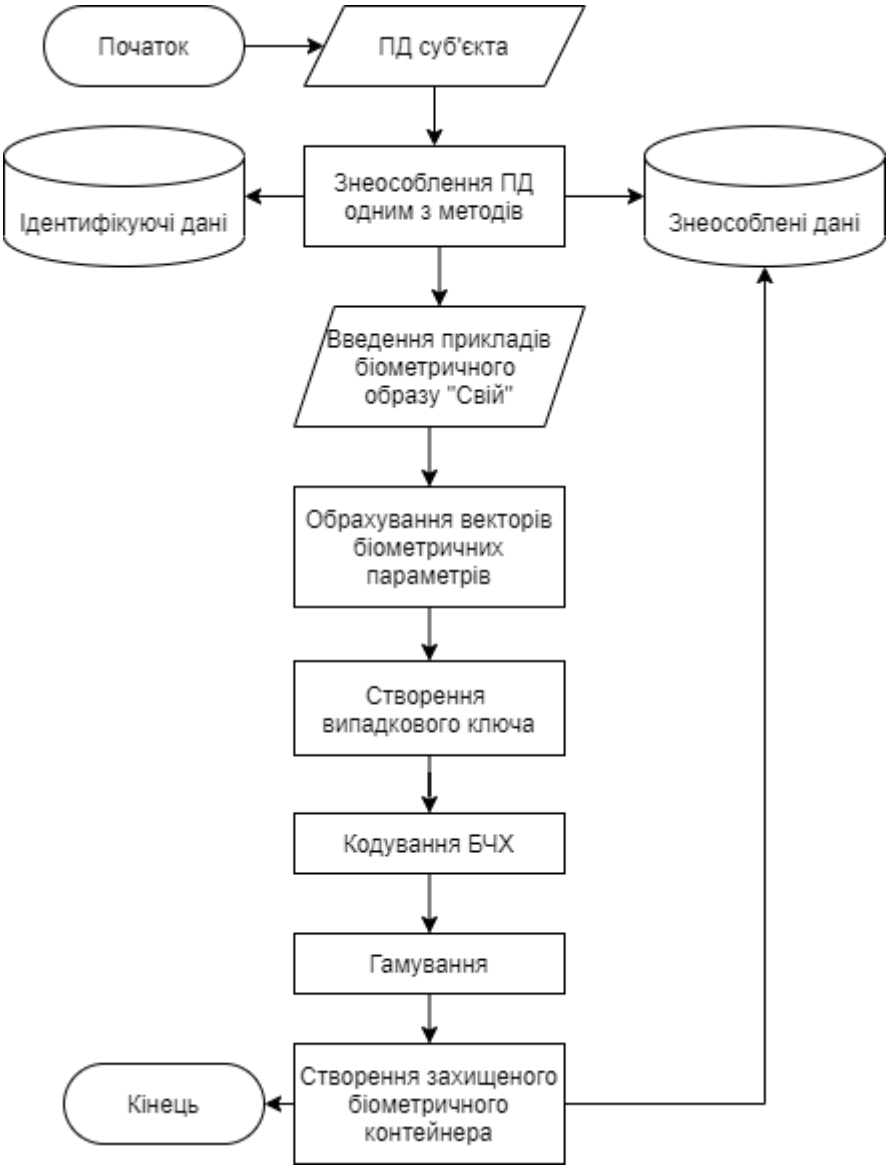

Рисунок 2.9 – Схема методу знеособлення персональних даних з використанням нечіткого екстрактора

Отримані ПДН суб'єкта знеособлюються за допомогою одного або декількох методів: методу введення ідентифікаторів, методу зміни складу або семантики, методу декомпозиції, методу перемішування. В результаті утворюються ідентифікують і знеособлені дані .Біометричні дані знеособлюються із використанням нечіткого екстрактора . Для чого суб'єкт пред'являє кілька прикладів біометричного способу «Свій». Потім обчислюються вектори двійкових біометричних параметрів . Далі створюється випадковий ключ, який потім перетвориться в двійковий вектор з використанням завадостійкого коду Боуза-Чоудхурі-Хоквінгема. Далі двійковий вектор шифрується гамуванням, а в якості гами використовується двійковий вектор, який вираховується з біометричних параметрів. В результаті створюється захищені біометричний контейнер B, що містить ідентифікатор (псевдонім) суб'єкта Id, результат гамування T і Hash - значення хеш - функції від особистого ключа (коду доступу) суб'єкта. Вміст біометричного контейнера включається до знеособлених даних , що обробляються оператором інформаційної системи персональних даних, тобто B є підмножиною P_2.

Для визначення приналежності знеособлених даних конкретному суб'єкту без доступу до ідентифікуючих даних і втрати анонімності відповідно до алгоритму персоніфікації персональних даних з використанням нечіткого екстрактора, зображено на рисунку 2.10. Знеособлені дані знаходяться за ідентифікатором Id, а потім суб’єкт пред’являє зображення відбитка пальця шляхом притискання його до сканера відбитків пальця. Обраховується двійковий вектор біометричних параметрів, потім виконується гумування з використанням гами T. Далі здійснюється декодування результату гамування, оскільки він містить похибку, внесену біометричними даними. Потім формується хеш-функція від ключа, що отримано внаслідок декодування.

Рішення про приналежність даних визначається за результатом порівняння значення хеш - функції від сформованого ключа і значення Hash,

що міститься в знеособлених даних. Якщо значення збігаються, то робиться висновок про те, що знеособлені дані належать суб'єкту. якщо значення не збігаються, то робиться висновок про неналежність знеособлених даних суб'єкту.

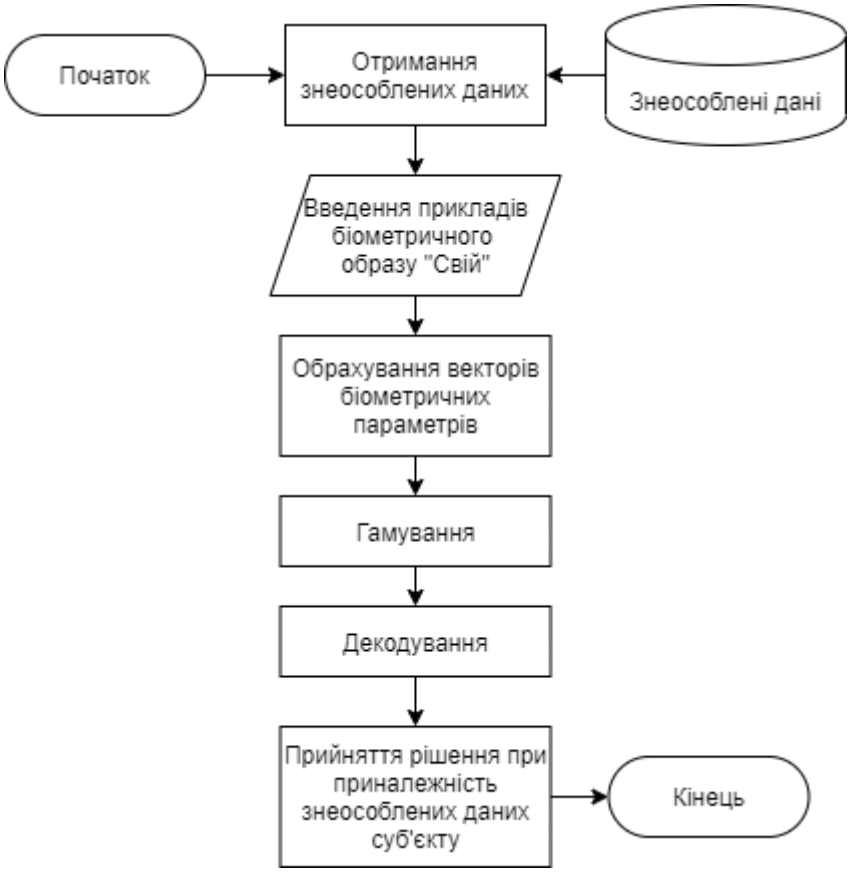

Рисунок 2.10 – Алгоритм персоніфікації персональних даних з використанням нечіткого екстрактора

2.5 Висновки по розділу 2

Існуючі методи перетворення персональних даних мають недолік, що полягає в необхідності використання при обробці операторами відомостей, що ідентифікують людину і що засвідчують його особистість .

Розроблено методи знеособлення персональних даних в біометричних системах. У першому методі використовується нейромережевий перетворювач біометрія-код. Другий метод відрізняється використанням нечіткого екстрактора.

Представлені методи знеособлення персональних даних, в яких отримані персональні дані суб'єкта знеособлюються за допомогою одного або декількох традиційних методів, а біометричні дані знеособлюються за допомогою нейромережевого перетворювача або нечіткого екстрактора.

Включення біометричного контейнера в знеособлені дані дозволяє обмежити доступ операторам до інформації, що ідентифікує людину, а приналежність знеособлених даних конкретного суб'єкта визначати за допомогою біометричної аутентифікації.

3 ПРОВЕДЕННЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ МЕТОДІВ ПЕРЕТВОРЕННЯ БІОМЕТРИЧНИХ ДАНИХ ОСОБИ ДЛЯ АУТЕНТИФІКАЦІЇ

3.1 Програмне середовище для збору біометричних даних

Програмне забезпечення призначене для організації збору біометричних даних (а саме відбитків пальців) людини. Забезпечує виконання наступних функцій:

- контроль коректного вводу зразків;
- завантаження бази даних, що була створена раніше.

Для повноцінної роботи ПЗ до комп’ютера було під’єднано сканер відбитків пальців Futronic FS80, та встановлено відповідні для нього драйвери.

3.1.1 Опис режимів роботи програми

Загальний вигляд програмного забезпечення формування баз відбитків зображено на рисунку 3.1.

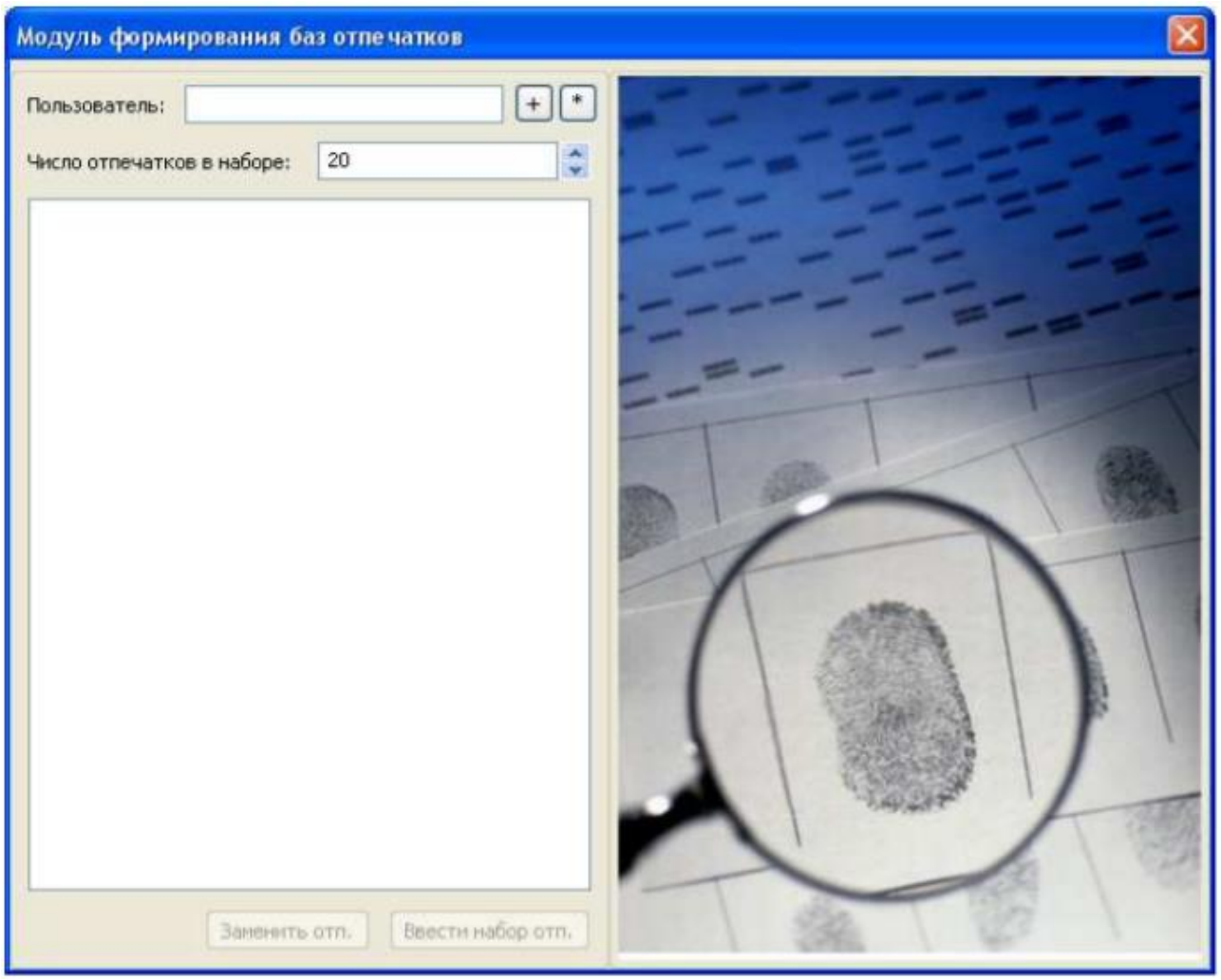

Рисунок 3.1 – Загальний вигляд програми

У лівій частині головного вікна програми розташовується область в яку вводить інформація о користувачі та його образах. В правій частині міститься вікно для огляду збережених образів баз відбитків.

Програмне забезпечення дозволяє створити нову базу біометричних даних або редагувати вже існуючу. Формування бази здійснюється шляхом додання облікового запису нового користувача і захвату зображення його відбитку пальця зі сканеру. Файли бази зберігаються в робочу папку програми.

3.1.2 Створення облікового запису нового користувача

Для створення нового запису о користувачеві необхідно в полі «користувач» ввести ПІБ користувача.

За замовчуванням поле «кількість образів» приймає значення 20. Максимальне значення образів, що вводяться для одного пальця дорівнює 8, а максимальне – 40.

Після натискання «+» запис о новому користувачі було додано в дерево каталогів. Дерево каталогів слугує для підвищення зручності навигації користувачів в середовищі й містить списки образів пальців користувачів, які необхідно ввести, або замінити.

Створення нового облікового запису о користувачеві зображено на рисунку 3.2.

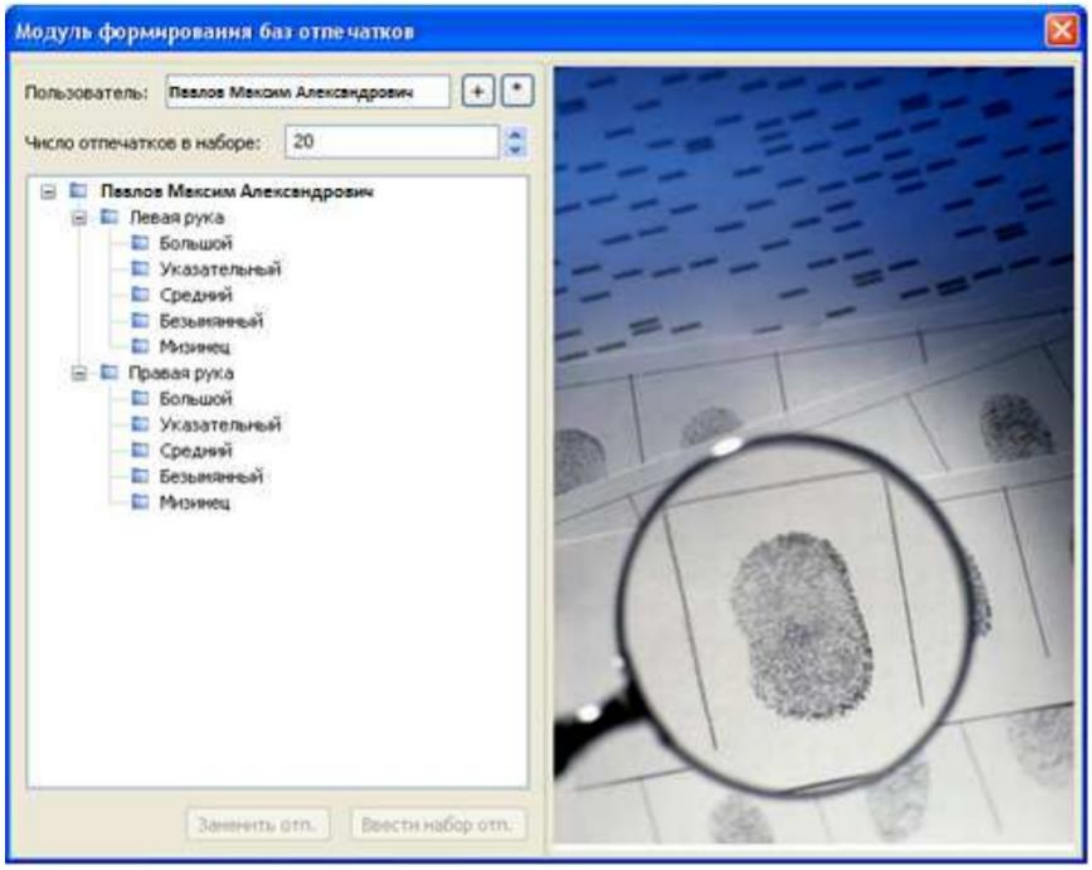

Рисунок 3.2 – Створення нового облікового запису користувача

3.1.3 Сканування біометричних даних

Після створення облікового запису о новому користувачеві у дерево каталогів баз відбитків було додано каталог «Ліва рука». Указано підкаталог, що відповідає назві пальця обраної руки. При натисканні «Ввести набір відбитків» запускається процедура зчитування набору образів обраного пальця, дивитись рисунок 3.3.

Аналогічні дії виконані при виборі підкаталогу «Права рука».

Рисунок 3.3 – Процедура захвату набору образів відбитку обраного пальця

В лівій частині вікна програми розташовується область, в якій відображається інформація о користувачеві та стан вводу його образів. Індикатор прогресу відображає хід виконання процесу формування бази. В правій частині вікна знаходиться область виводу зробленого зображення зі сканера зчитування відбитків, а також поле для виводу графічних повідомлень. Повідомлення інформують користувача о необхідності коректного розташування пальця на датчику.

Після появи повідомлення «Прикладіть палець до датчика» палець був розташований на сканері та притиснутий з невеликим зусиллям. Зроблене зображення відбитку повинно розташовуватись всередині рамки й переважно по центру. Якщо знімок відбитку не відбувається, то необхідно:

− притиснути палець сильніше при появі повідомлення «Притисніть палець сильніше», як зображено на рисунку 3.4;

− змістити палець при появі повідомлення «змістіть палець вліво» або «Змістіть палець вправо» в напрямі вказівної стрілки, що зображено на рисунку 3.5;

− повернути палець до вертикального положення при появі повідомлення «Поверніть палець вертикально», так щоб краї відбитку були умовно паралельні до рамки вікна й симетричні відносно один одного, зображено на рисунку 3.6.

Рисунок 3.4– Недостатньо притиснутий палець

Рисунок 3.5 – Зміщення розташованого пальця

Рисунок 3.6 – Невертикальне положення пальця

Для меншої деформації відбитку пальця при переміщенні рекомендується його послідовно відривати й притискати до сканера . Зсування й перекочування сильно деформує рисунок папілярних ліній, що може вплинути на якість бази.

Місцезнаходження пальця на сканері задовольняє вимогам місцерозташування, образ було створено. При завершені сканування заданого числа образів відбитку пальця з’явилось інформаційне повідомлення «Відбитки збережені», дивитись рисунок 3.7.

Рисунок 3.7 – Завершення сканування

Список зроблених образів відбитків пальця було сформовано у подкаталозі відповідного пальця. Для створення повної бази біометричних даних було проведено сканування для кожного пальця користувача.

3.1.4 Завантаження існуючої бази біометричних даних

При зміні (оновленні чи заміні образів) існуючих баз їх слід завантажити, натиснув на «*». Після натиску появиться попередження, зображене на рисунку 3.8.

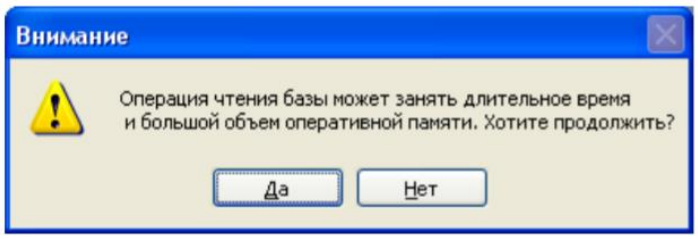

Рисунок 3.8 – Попереджуюче повідомлення

Зчитування великої бази потребую значних затрат часу та ресурсів комп'ютера, тому операцію можна відмінити натиснув «Ні». При натисканні на «Так» почнеться процедура завантаження бази.

Після завершення завантаження дерево каталогів заповниться раніше створеними базами користувачів й записами (образами), що містяться в них. Для перегляду образа відбитку пальця слідує обрати каталог користувача і відповідний підкаталог пальця руки, як зображено на рисунку 3.9.

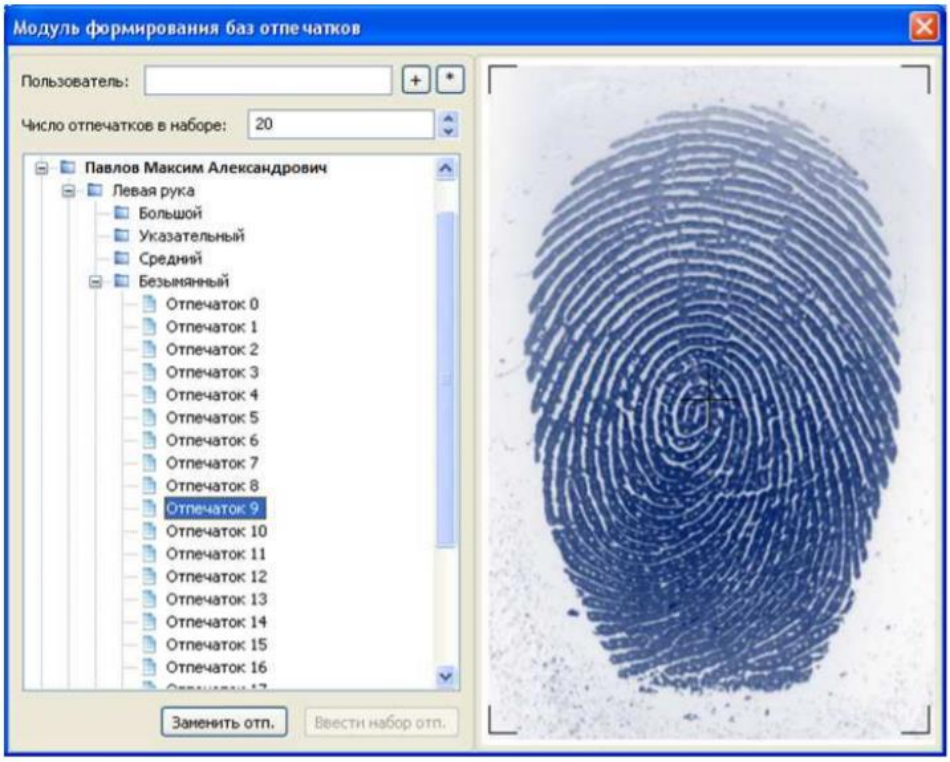

Рисунок 3.9 – Перегляд образа відбитку пальця

При натисканні «Ввести набір відбитків» запуститься процедура захвату образів обраного пальця.

Для заміни образа відбитку пальця слід обрати в дереві каталогів необхідний палець й натиснути «Замінити його». Після чого запуститься процедура захвату відповідного образу.

3.2 Побудова баз для експериментальних досліджень

Для моделювання нейромережевих перетворювачів та нечітких екстракторів, перетворюючих біометричні дані в код доступу необхідна експериментальна база біометричних образів. В рамках даного дипломного проекту розроблюються засоби знеособлення персональних даних з використанням нечітких екстракторів та нейромережевих перетворювачів, котрі використовують біометричні дані відбитків пальців людей для перетвореня в код доступу с ціллю наступної аутентифікації. Отже, необхідно сформувати базу відбитків пальців реальних людей. Розроблені моделі нечіткого екстрактора й нейромережевого перетворювача припускають, що для розрахунку надлишку коригуючого коду необхідно знати розподіл біометричних параметрів двох множин: образа «свій» та «Чужий». Образ «Свій» повинен реально відображати розподіл біометричних параметрів легітимного користувача, котрий пред’являє системі аутентифікації відбиток пальця. Образ «Чужий» повинен реально відображати розподіл біометричних параметрів несанкціонованого користувача, який намагається підібрати відбиток пальця легітимного користувача.

Для формування бази образів було використано програмне середовище збору відбитків пальців та сканер відбитків пальців. Базу відбитків пальців «Свій» було вирішено створити із 100 образів відбитків пальців по 20 прикладів кожного з них. В програмному середовищі збору відбитків пальців за допомогою сканера відбитків пальців було здійснено введення 20 образів для кожного із 100 відбитків пальців. Наявність великої кількості прикладів у вибірці дозволить розбити її на навчальну та тестову, що буде необхідно при порівнянні якості

прийнятих рішень нечітким екстрактором та нейромережевим перетворювачем біометрія-код.

У програмному середовищі збору відбитків пальців також було сформовано базу «Чужий» із 100 різноманітних відбитків пальців.

3.3 Проведення експериментальних досліджень

Моделювання перетворювача біометрія-код та нечіткого екстрактора проводилось за допомогою програми автоматизації статистичних розрахунків. В якості бази відбитків пальців образів «Свій» використовувалось 5 пальців руки. В програмному середовищі збору відбитків пальців за допомогою сканера відбитків пальців було здійснено введення 100 образів по 20 прикладів. При моделюванні база прикладів була розподілена на навчальну вибірку svTable та тестову вибірку testTable. Навчальна вибірка використовувалась для визначення маски і еталонного біометричного коду для біометричного образа «Свій». А тестова – для обчислення надлишку коригуючого коду Боуза-Чоудхури-Хоквингема. Після дискредитації неперервних біометричних параметрів отримані три таблиці (бази) svDTable, testDTable, chDTable.

При моделюванні нечіткого екстрактора параметр threshold (ймовірність похибки першого роду) функцій funcFindMask (svDTable, threshold) та funcEtalonKey (svDtable, threshold) приймав значення від 0.1 до 0.01 з кроком 0.01.

Для кожної вибірки svDTable, testDTable, chDTable отримано розподіл відстаней Хеммінга svHemming, testHemming, chHemming відповідно.

При моделюванні нечіткого екстрактора відсоток похибки в еталоному біометричному коді образу «Свій» розраховувався як відношення max(testHemming) і rows(etalonKey), результат якого використовується в якості вхідного параметра функції funcExcess(errPer).

При моделюванні нечіткого екстрактора довжина коду доступу, що формується нечітким екстрактором, обчислюється як відношення rows(etalonKey) і funcExcess(errPer).

Для кожного еталонного біометричного коду образа «Свій» розраховується ймовірність похибки другого роду, та існує ймовірність пропуску системою біометричної аутентифікації на основі нечіткого екстрактора «Чужого». Похибка розраховувалась як інтеграл від стандартної функції Mathcad dnorm (x, mu, sigma), яка повертає щільність ймовірності нормального розподілу з середнім mu та середньоквадратичним відхиленням sigma.

$$P_{II} = \int_{-\infty}^{\max(testHemming)} dnorm\big(x, mean(chHemming), stdev(chHemming)\big)dx, \qquad (3.1)$$

де:

− mean(chhemming) – стандартна функція Mathcad для обрахування математичного очікування;

− stdev(chHemming) – стандартна функція Mathcad для обрахування середньоквадратичного відхилення;

− x – безперервна змінна.

Текст програми для статистичної оцінки нейромережевих перетворень та нечітких екстракторів представлений в додатку Б.

На рисунку 3.10 зображено розподіл значень відстаней Хеммінга між біометричними кодами довжиною 66 біт.

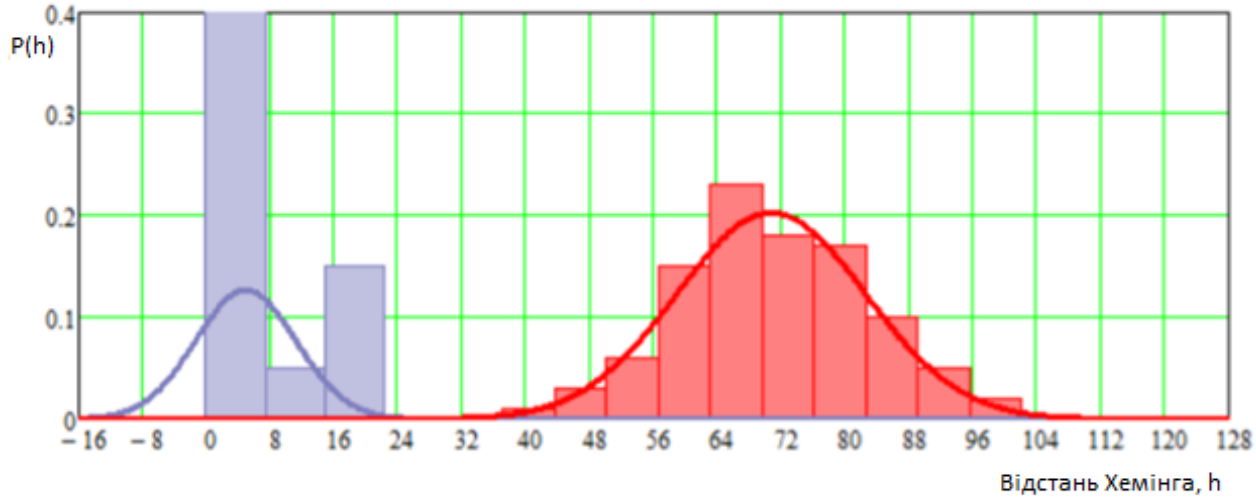

Рисунок 3.10 – Розподіл значень відстаней Хеммінга між біометричними кодами нечіткого екстрактора

Максимальне число похибок, виникаючих в еталонному біометричному коді дорівнює 6, що складає 9%. Це дозволяє використовувати в системі аутентифікації код доступу довжиною всього 13 біт.

За формулою 3.1 величина:

$$P_{II} \approx 3 \times 10^{-3} \tag{3.2}$$

Стійкість перетворювача біометрія-код на базі нечіткого екстрактора до атак підбору визначається як зворотна величина від P_{II}:

$$D = \frac{1}{P_{II}} = \frac{1}{3 \times 10^{-3}} = 333{,}3. \tag{3.3}$$

Моделювання нейромережевого перетворювача біометрія-код проводилось з використанням алгоритму автоматичного навчання і для цього були сформовані вибірки двох класів образів «Свій» та «Чужий», що представлені у вигляді векторів безперервних біометричних параметрів.

Перетворювач біометрія-код був реалізований як одношарова штучна нейронна мережа з m входами та n виходами, вона описується формулою 3.4:

$$key_i = \varphi(\sum_{j=1}^{m} w_{i,j} \cdot v_j + b_i) \quad (3.4)$$

де:

- key_i – розряд вихідного коду;
- v_i – неперервний біометричний параметр;
- $w_{i,j}$ – вагові коефіцієнти і-го нейрону;
- b_i – зміщення;
- φ – нелінейна (порогова передаточна) функція;
- m – число параметрів вектора біометричних параметрів (входів штучної нейронної мережі);
- n – число розрядів вихідного коду (число нейронів, виходів штучної нейронної мережі).

Нелілейна функція φ від суми x (лілейної комбінації вхідних біометричних параметрів) виглядає так:

$$\varphi(s) = \begin{cases} 1 & s \geq 0 \\ 0 & s < 0 \end{cases} \quad (3.5)$$

Навчання штучної нейронної мережі проходило на двійкових розрядах коду довжиною 256 біт, який генерувався випадково. Розрахунок вагових коефіцієнтів $w_{i,j}$ нейронів здійснювався пропорційно якості біометричних параметрів.

При навчанні кожного нейрону на біометричних параметрах образа «Свій» необхідно відштовхуватися від умови рівної ймовірності станів «0» та «1» для біометричних параметрів образа «Чужий», як показано на рисунку 3.14.

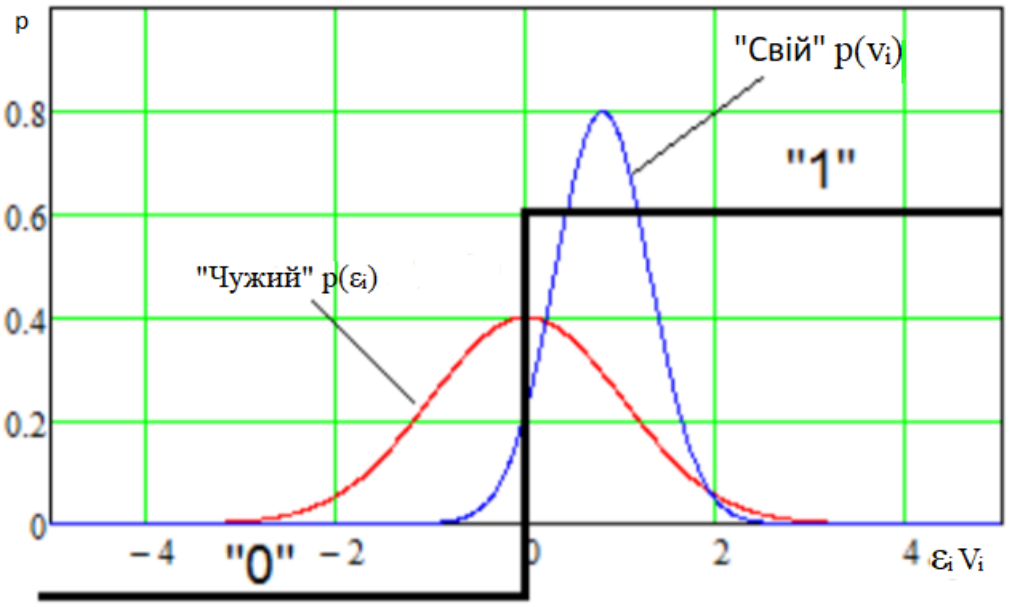

Рисунок 3.11 – Рівноімовірні значення «0» та «1»

Як видно з рисунку 3.11, розподіл значень і-го біометричного параметра $p(\varepsilon_i)$ образу «Чужий» має математичне очікування рівне нулю, тому якщо на виході нейрона необхідний стан «0», то знак вагового коефіцієнта при навчанні необхідно замінити на протилежний.

Якість роботи нейромережевого перетворювача біометрія-код оцінювалась як узагальнююча здатність навченої штучної нейронної мережі. Тобто можливість отримання чіткого значення коду key на виході штучної нейронної мережі при подачі на її вхід тестової вибірки, що отримує на вхід різні приклади.

Для обрахування ймовірності похибки другого роду P_{II} на вхід штучній нейронній мережі послідовно задавалось 100 різних відбитків пальців. На виході отримані кодові відклики штучної нейронної мережі один за одним порівнювались з кодом «Свій» в метриці Хеммінга. Порівняння кодів дає розділення значень відстаней Хеммінга, яке зображено на рисунку 3.15.

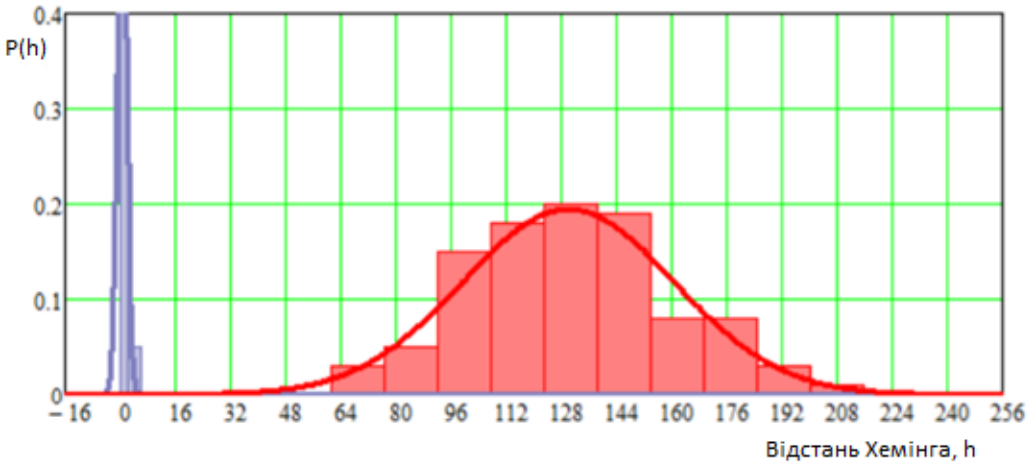

Рисунок 3.12 – Розподіл значень відстаней Хеммінга між біометричними кодами нейромережевих перетворень

Як видно з рисунку 3.12, мінімальною відстанню виявилось h_{min} рівне 64. Для чисельної оцінки P_{II} нейромережевого перетворення біометрія-код з розрядністю вихідного коду m>32 рекомендовано використовувати функцію закону нормального розподілу:

$$P_{II} \approx \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \cdot \int_{0}^{E(h)/\sigma(h)} \exp\left(-t^2/2\right) dt, \tag{3.6}$$

де:

- E(h) – математичне очікування розподілу значень відстаней Хеммінга;
- Σ(h) – середньоквадратичне відхилення розподілу значень відстаней Хеммінга;
- T – неперервна змінна.

Величина P_{II} приблизно дорівнює 0.0025. Стійкість нейромережевого перетворення «біометрія-код» до атак підбору визначається як зворотня величина від P_{II} :

$$D = \frac{1}{P_{II}} = \frac{1}{0{,}0025} \approx 400. \tag{3.7}$$

3.4 Аналіз експериментальних досліджень

Формально штучна нейронна мережа описується матрицею вагових коефіцієнтів, тому для формування коду доступу необхідно зберігати її в пам’яті. Нейромережевий біометричного контейнер - блок даних для збереження параметрів навченої штучної нейронної мережі.

При відкритому зберіганні матриці вагових коефіцієнтів в нейромережевому контейнері можливо проведення процедури відтворення невідомого біометричного образу «Свій». Процедура відтворення полягає в тому, що метрика Хеммінга дозволяє виділити найбільш близькі біометричні образи «Чужий» к біометричному образу «Свій» і використовувати їх для створення штучних прикладів шляхом схрещування. Результатом схрещування є друге покоління біометричних образів «Чужий». Таким чином, можливо отримати третє, четверте, п’яте покоління. З кожним новим поколінням відстань Хеммінга зменшується, як показано на рисунку 3.13. А матриця вагових коефіцієнтів обертається. Таким чином вдається отримати до 90% біометричних параметрів невідомого біометричного зразка. Недолік, що полягає у відновленні біометричних параметрів мають не тільки нейромережеві перетврювачі з відкритим зберіганням матриці вагових коефіцієнтів в біометричному контейнері, а й нечіткі екстрактори.

Як показано в роботах біометричний контейнер може бути захищений за допомогою механізму розмноження помилок. Цей механізм може бути побудований на основі оборотного шифрування матриці вагових коефіцієнтів. Для цього необхідно частину вихідного ключа, що формується за допомогою штучної нейронної мережі, використовувати для шифрування параметрів нейроів, значення на виході котрих ще не отримані.

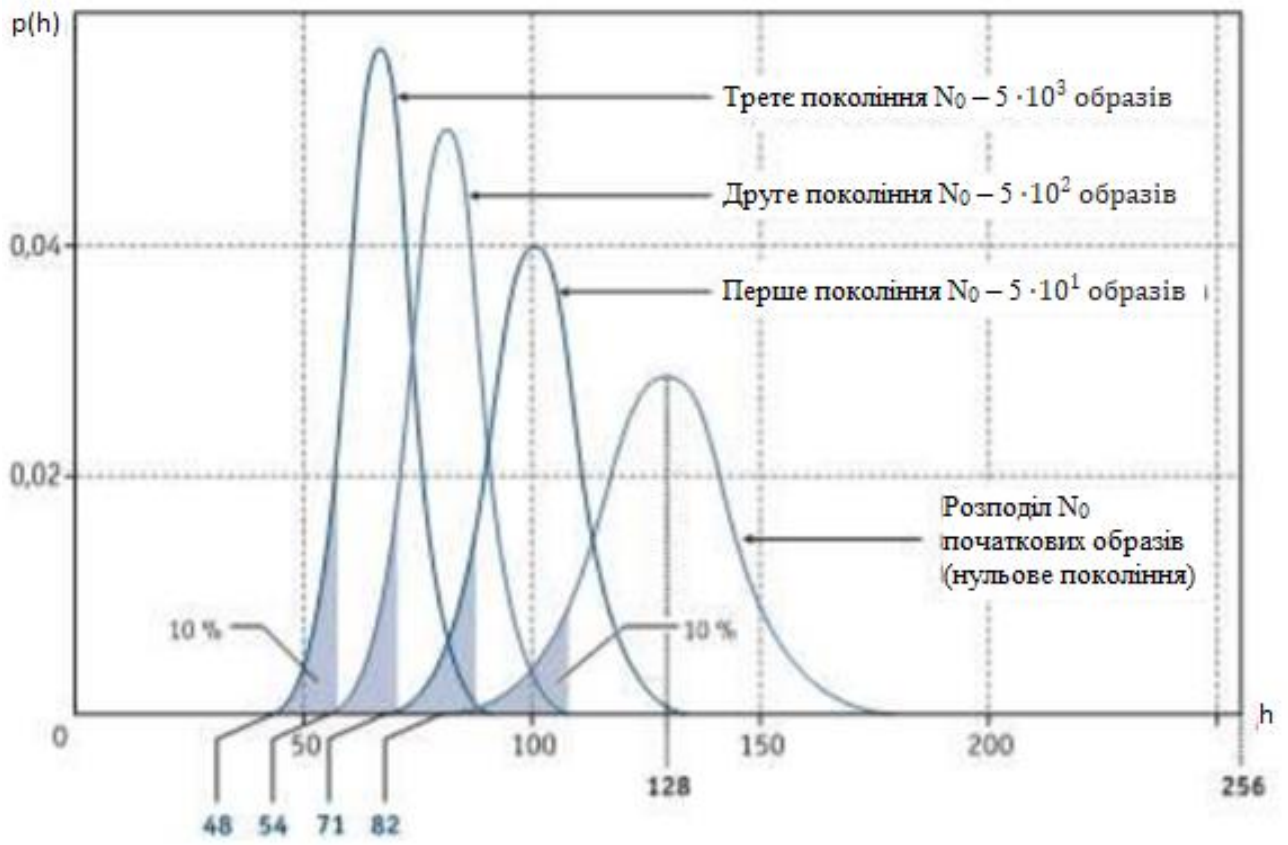

Рисунок 3.13 – Зменшення значень відстаней Хеммінга
Між кодами «Свій» та «Чужий» при селекції та схрещуванні

Таким чином, для біометричного образа «Свій» параметри нейронів будуть послідовно розшифровуватись, і на виході штучної нейронної мережі сформується точне значення ключа.

Інакше вийде при подачі на вході штучної нейронної мережі значень параметрів біометричного образу «Чужий». В цьому випадку після першого невірно сформованого біта на виході нейрона послідуючі розшифровування параметрів нейронів відбувається невірно, що призводить до хешуванню параметрів навченої штучної нейронної мережі (ефект розмноження помилок).

Ефект розмноження помилок при оцінці стійкості нейромережевого перетворення біометрії в код доступу з використанням захищеного біометричного контейнера дає інакший розподіл значень відстаней Хемінга.

Розподіл значень відстаней Хемінга між кодовими відкликами штучної нейронної мережі і кодом «Свій» представлено на рисунку 3.14.

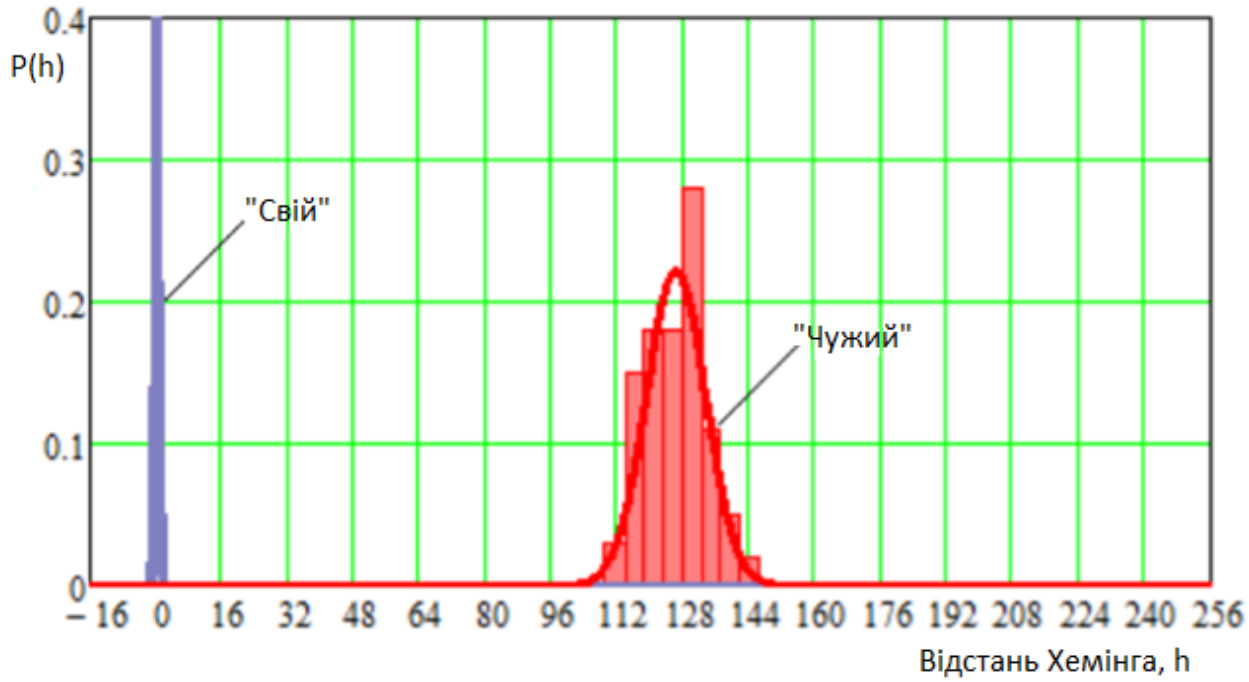

Рисунок 3.14 – Розподіл значень відстаней Хемінга між кодовими відкликами мережі і кодом «Свій»

Механізм розмноження помилок не дозволяє спостерігати реальний розподіл значень відстаней Хеммінга. Отже, отримання біометричних параметрів невідомого біометричного образу стає важно обчислювальною задачею.

Експериментальні дослідження засвідчили, що перетворювачі біометрія-код на основі нечітких екстракторів мають малу довжину коду доступу (16 біт та менше). Через неоднозначне відтворення даних людини необхідно виправляти близько 10% помилок в біометричному коді. Це призводить до високої надлишковості коригуючого коду, який перевищує 400%. Тому отримати код довільної довжини виявляється неможливим за допомогою нечіткого

екстрактора. А мала довжина коду створює загрозу його підбору звичайним перебором.

Нечіткі екстрактори не можуть забезпечити таємницю біометричних образів, оскільки просте гамування біометричного коду не захищає біометричні параметри.

Перевагою нейромережевих перетворювачів біометрія-код є можливість формування штучної нейронної мережі коду доступу довільної довжини. За рахунок вибору коду доступу довжиною 256 біт його підбор простим перебором стає важно обчислювальною операцією.

Стійкість до атак підбору біометричних даних для нейромережевого перетворювача і нечіткого екстрактора виявляється порівняльною. Проте за рахунок використання механізму розмноження помилок надійність нейромережевих перетворювачів підвищується.

Таким чином, нейромережеві перетворювачі біометрія-код доцільніше використовувати в системах біометричної аутентифікації, а також при розробленні безпечних протоколів біометрико-криптографічної аутентифікації особистості.

3.5 Висновки по розділу 3

Сформована база «Свій» шляхом зчитування 100 образів по 20 відбитків пальців у програмному середовищі збору відбитків пальців.

Сформована база «Чужий», що налічує 100 прикладів різних відбитків пальців.

Розроблена програма моделювання й статистичної оцінки нейромережевих перетворювачів та нечітких екстракторів.

Проведено моделювання та виконано порівняння вірогідносних характеристик нейромережевих перетворювачів та нечітких екстракторів із застосуванням баз «Свій» і «Чужий», отриманих в програмному середовищі збору відбитків пальців.

В ході моделювання встановлено, що при наявності більше 10% помилок у векторі двійкових біометричних параметрів використовувати нечіткий екстрактор для знеособлення біометричних даних неможливо, оскільки неможливо тримати коригуючий код. Нейромережеві перетворювачі позбавлені подібного недоліка, оскільки дозволяють формувати код довільної довжини, проте їх застосування також складне для відбитків пальців з низкою якістю папілярного рисунку.

ВИСНОВКИ

В ході виконання даної дипломної роботи було розглянуто сферу біометричної ідентифікації та аутентифікації користувачів.

Проведено аналіз закордонних і вітчизняних технологій перетворення біометричних даних в код аутентифікації (біометрія-код), який показав, що існує два головних напрямки рішення задачі. Перший напрямок характеризує використання нечітких екстракторів на основі кодів з виявленням і виправленням помилок. Другий напрямок характеризується застосуванням великих штучних нейронних мереж. Проаналізовано основні біометричні параметри, частоту їх використання та розповсюдженість. Проаналізовано методи перетворення біометричних даних користувача для аутентифікації особистості.

В роботі, шляхом знеособлення персональних даних користувача в біометричних системах підвищено ефективність системи аутентифікації користувача за рахунок перетворення біометричних даних в код доступу за допомогою нейромережі. Нейромережа працює з біометричним контейнером і дозволяє підвищити надійність і обмежити доступ операторів до інформації, яка може ідентифікувати людину. Ідентифікаційні дані конкретного суб'єкту визначаються за допомогою біометричної аутентифікації.

Проведено експериментальні дослідження методів перетворення біометричних даних особи для аутентифікації, в ході яких було встановлено, що використовувати нейромережеві перетворення значно простіше ніж нечіткі екстрактори, оскільки вони позбавлені недоліку - неможливості отримати коригуючий код через велику похибку у двійковому біометричному векторі біометричних параметрів. Також нейромережеві перетворення дозволяють формувати код доступу довільної довжини, а їх застосування може бути ускладнено лише при роботі з відбитками пальців, що мають низьку якість папілярного малюнку.

ПЕРЕЛІК ПОСИЛАНЬ

1 Barni M. Privacy-Preserving Fingercode Authentication [Електронний ресурс] / Mauro Barni – Режим доступу до ресурсу: http://piurilabs.di.unimi.it/Papers/awmm_2010.pdf

2 Spaun N. 2nd ieee international conference on biometrics theory applications and systems [Електронний ресурс] / Spaun – Режим доступу до ресурсу: https://cse.nd.edu/BTAS_08/

3 Fairhurst M. Seventh International Conference of Document Analysis and Recognition [Електронний ресурс] / Fairhurst – Режим доступу до ресурсу: https://ieeexplore.ieee.org/document/1227617/keywords#keywords

4 Schoutien B. Biometrics and their use in e-passports [Електронний ресурс] / Schoutien – Режим доступу до ресурсу: https://research.tue.nl/en/publications/biometrics-and-their-use-in-e-passports.

5 ISO/IEC 19785-1 [Електронний ресурс] – Режим доступу до ресурсу: https://www.iso.org/standard/66179.html.

6 ISO/IEC 19794-1 [Електронний ресурс] – Режим доступу до ресурсу: https://www.iso.org/ru/standard/50862.html.

7 Оглад біометричних методів ідентифікації особистості [Електронний ресурс] – Режим доступу до ресурсу: http://masters.donntu.org/2013/fknt/fomenko/library/article3.htm.

8 Biometric cryptosystems issues and challenges [Електронний ресурс] – Режим доступу до ресурсу: https://ieeexplore.ieee.org/document/1299169.

9 Fuzzy Extractors for Minutiae-Based Fingerprint Authentication [Електронний ресурс] – Режим доступу до ресурсу: https://link.springer.com/chapter/10.1007/978-3-540-74549-5_80.

10 Binary feature vector fingerprint representation from minutiae vicinities [Електронний ресурс] – Режим доступу до ресурсу: https://ieeexplore.ieee.org/document/5634488

11 Нейромережевий захист персональних біометричних даних [Електронний ресурс] – Режим доступу до ресурсу: https://azon.market/knigi/kompyyuternaya-literatura/neyrosetevaya-zaschita-personalnyih-biometricheskih-dannyih---volchihin-vi-mysh2567676?limit=100

12 Tcутома M. A Case Study for User Identificati [Електронний ресурс] / Матсумото Tcутома – Режим доступу до ресурсу: http://web.mit.edu/6.857/OldStuff/Fall03/ref/gummy-slides.pdf.

13 Reuters / [Електронний ресурс] – Режим доступу до ресурсу: https://www.reuters.com/video/watch/peace-signs-risk-fingerprint-theft-says-id370920514

14 Chaos Communication Congress [Електронний ресурс] – Режим доступу до ресурсу: https://www.theguardian.com/technology/2014/dec/30/hacker-fakes-german-ministers-fingerprints-using-photos-of-her-hands.

15 Roy A. MasterPrint: Exploring the Vulnerability of Partial Fingerprint-Based Authentication Systems [Електронний ресурс] / A. Roy, A. Ross, N. Memon – Режим доступу до ресурсу: https://ieeexplore.ieee.org/document/7893784/authors#authors.

16 Parsons A. Paying for a pint with your finger: The tech that could kill off cards [Електронний ресурс] / Adam Parsons – Режим доступу до ресурсу: https://news.sky.com/story/paying-for-a-pint-with-your-finger-the-tech-that-could-kill-off-cards-10780629.

17 Brewster T. We Broke Into A Bunch Of Android Phones With A 3D-Printed Head [Електронний ресурс] / Thomas Brewster – Режим доступу до ресурсу: https://www.forbes.com/sites/thomasbrewster/2018/12/13/we-broke-into-a-bunch-of-android-phones-with-a-3d-printed-head/?sh=2a17646c1330.

18 Hopfield J. Neural networks and physical systems with emergent collective computational abilities [Електронний ресурс] / Hopfield – Режим доступу до ресурсу: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC346238.

19 Wiley J. The Organization of Behavior [Електронний ресурс] / John Wiley – Режим доступу до ресурсу: https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.900930310

20 Ясницький Л. Ввеення в штучний інтелект [Електронний ресурс] / Л. Ясницький – Режим доступу до ресурсу: https://www.academia-moscow.ru/ftp_share/_books/fragments/fragment_17447.pdf

21 Dodis Y. Fuzzy extractors how to generate strong keys from biometrics and other noisy data [Електронний ресурс] / Yevgeniy Dodis – Режим доступу до ресурсу: https://link.springer.com/chapter/10.1007/978-3-540-24676-3_31