

**Сабодашко Д.В.**

Національний університет «Львівська політехніка»

**Хома Ю.В.**

Національний університет «Львівська політехніка»

**Хома В.В.**

Національний університет «Львівська політехніка»

## ДОСЛІДЖЕННЯ ЧАСОВОЇ СТІЙКОСТІ СИГНАЛУ ЕКГ ЯК БІОМЕТРИЧНОГО МАРКЕРА В СИСТЕМІ АВТЕНТИФІКАЦІЇ

*Унікальність електрокардіограми для кожного індивіда доведена багатьма дослідженнями. На цей час вирішено проблеми, пов'язані із особливістю роботи біометричних систем – простий відбір електричних потенціалів, ефективні методи кондиціонування ЕКГ-сигналу, надійні алгоритми інтелектуального аналізу прихованих індивідуальних характеристик електрокардіограми. Проте залишилося без відповіді питання довгочасової стабільності інформативних параметрів сигналів ЕКГ, що важливо для роботи біометричних систем автентифікації у реальних застосуваннях. Тому метою цієї роботи було – дослідити стійкість електрокардіограми як унікального біометричного маркера для автентифікації суб'єктів упродовж довгих інтервалів часу тривалістю місяці-роки. У статті представлено власну реалізацію системи біометричної автентифікації користувачів за їх електрокардіограмами, яку було використано для реєстрації ЕКГ-записів, їх комп'ютерного опрацювання, тренування і тестування моделей нейромережових класифікаторів. Верифікація роботи системи біометричної автентифікації здійснювалася на відкритій для доступу базі даних LBDS (Lviv Biometric Data Set), у якій зібрано 1 809 унікальних записів електрокардіограм тривалістю від 10 до 20 секунд. Записи здійснювалися впродовж двох років і належать 115 різним особам. Описано методику проведення досліджень, яка передбачала проведення двох експериментів. Перший експеримент націлений на верифікацію роботи системи біометричної автентифікації після одного-двох місяців, що минули з часу навчання нейромережового класифікатора, а другий – після одного-двох років. Обидва експерименти показали досить високі результати: точність автентифікації користувачів за записами їх ЕКГ-сигналів, зробленими через місяць та через два роки, утримується практично на одному рівні – відповідно 95,61% і 94,97%. Це дає підстави для висновку, що електрокардіограма є стабільним у часі біометричним маркером, який можна застосовувати у реальних системах біометричної автентифікації.*

**Ключові слова:** біометрична система, ЕКГ-записи, нейромережовий класифікатор, часова стабільність біометричних ознак ЕКГ-сигналу, точність автентифікації.

**Постановка проблеми.** Біоінформатика – це галузь науки, що займається дослідженням, розробкою і застосуванням обчислювальних методів і засобів для аналізу та інтерпретації біологічних, фізіологічних, медичних і поведінкових даних людини та інших біологічних об'єктів. Традиційно розрізняють такі напрями у біоінформатиці, як опрацювання біосигналів, розпізнавання медичних зображень, комп'ютерна геноміка і аналіз електронних медичних записів [1; 2]. Кожний напрям має свою специфіку, але спільним є видобування інформації із набору відібраних із об'єкта даних. Наприклад, у випадку біосигналів як змінного у часі процесу дані – це часові ряди, а у геноміці – це просторова зміна параметрів нуклеотидної послідовності, що визначає генетичний код.

У другій половині ХХ сторіччя спочатку електроніка, а згодом і комп'ютери стали застосовуватися для відбору і аналізу біомедичних сигналів. Комп'ютеризовані діагностичні системи вже давно стали надбанням не лише великих медичних закладів, але і звичайних поліклінік, де використовуються для цілей медичної діагностики пацієнтів.

Розвиток інформаційних технологій, удосконалення і здешевлення обчислювальної техніки, а, з іншого боку, нагромадження великого обсягу оцифрованих біомедичних даних спричинили суттєве розширення сфер застосування біоінформатики. На цей час шляхом опрацювання біосигналів можна розв'язувати різні задачі у «немедичних» сферах, наприклад, виконувати

біометричну автентифікацію користувачів інформаційних систем на основі електрокардіограм. З погляду практичного застосування сигналу ЕКГ як біометричного маркера важливо дослідити стійкість у часі тих інформативних ознак, за якими система приймає рішення про надання чи відмову у доступі.

**Аналіз останніх досліджень і публікацій.** Сигнали ЕКГ належать до найбільш відомих біосигналів і є носіями не лише цінної діагностичної інформації, але також і біометричної. В останні роки було підтверджено, що електрокардіограма (ЕКГ) є надійним біометричним маркером [3–5]. Важливо зазначити, що анатомічні відмінності серця між індивідуумами позначаються на формі сигналу ЕЕГ, роблячи його своєрідним індивідуальним почерком для людей, навіть серед близнюків і двійнят [6].

Проте принципи побудови і опрацювання сигналів ЕКГ, які застосовуються у сучасних цифрових електрокардіографах, не можна просто перенести на біометричні системи. Передовсім істотно відрізняються умови використання цих систем. Формування електрокардіографічних записів у медичній діагностиці (переважно у вигляді 12 відведень) відбувається в комфортних умовах, коли пацієнт розслаблено лежить на кушетці, а тривалість записів зазвичай не обмежена. У системах біометричної автентифікації застосовується простий відбір біопотенціалів, найчастіше із пальців рук [7; 8].

Такі зручні з погляду практичного застосування способи відбору біопотенціалів супроводжуються значно інтенсивнішими спотвореннями ЕКГ-записів, які мають різне походження, структуру і частотний діапазон. Зростає деструктивний вплив дрейфу ізоляції, наведення мережі живлення, електропотенціалів від роботи м'язів (інтерференція електроміографічного сигналу). Для їх придушення найчастіше застосовують методи цифрової фільтрації [9], хвилькову декомпозицію із відбором релевантних компонент [10]. Проте навіть у знешумленому сигналі можуть спостерігатися викиди (англ. – outliers), тобто аномальні відхилення вибірок у записі ЕКГ, наявність яких знижує достовірність результатів біометричної автентифікації.

Поширеним способом боротьби з викидами, що застосовувався багатьма дослідниками, було їх виявлення за певним правилом і подальше відкидання цілого сегменту, який містить аномальні відхилення окремих чи серії вибірок [11; 12].

Проте у застосуваннях біометрики тривалість записів є обмеженою і складає зазвичай до десяти серцевих циклів (сегментів). Тому відкидання сегментів з аномальними вибірками зменшує обсяг даних, які надходять на класифікатор системи, що теж негативно позначається на точності автентифікації. Для вирішення цієї суперечливості в роботах [14; 15] запропоновано не відкидати, а коригувати пошкоджені сегменти. Реалізація такого підходу дає змогу зберегти необхідний для надійної класифікації об'єктів обсяг вхідних даних, без потреби збільшення тривалості відбору ЕКГ.

Ще однією проблемою, важливою для біометричних застосувань, є забезпечення інваріантності до змін серцевого ритму (частоти серцевих скорочень) [16–18]. Переважне число досліджень спирається на бази даних ЕКГ сигналів, записаних упродовж одного сеансу, тому серцевий ритм для кожного суб'єкта має певне значення, яке класифікатор системи на етапі навчання зараховує як одну із інформативних ознак. Проте у роботі реальної системи неминуче виникатимуть ситуації, коли законний користувач одержить відмову у доступі через розбіжність значень частоти серцевих скорочень у моменти реєстрації (навчання моделі класифікатора) і запиту на автентифікацію (робочий режим класифікатора). Як показано у роботі [19], застосування часової нормалізації ЕКГ-сегментів в усіх записах є ефективним інструментом забезпечення інваріантності системи автентифікації до зміни серцевого ритму.

На шляху запровадження у практику систем біометричної автентифікації на основі ЕКГ важливо дослідити стабільність основних дискримінаційних ознак на довгих проміжках часу (місяці-роки), за якими класифікатор приймає рішення про надання чи відмову у доступі. Аналіз літературних джерел показав на відсутність публікацій із результатами подібних досліджень.

**Постановка завдання.** Метою роботи є дослідження часової стабільності ЕКГ-сигналів на довготривалих проміжках часу, а також оцінювання ступеня впливу варіативності інформаційних ознак електрокардіограми на точність автентифікації.

**Виклад основного матеріалу дослідження. Опис імплементації системи біометричної автентифікації.** Для виконання досліджень автори розробили власний варіант трирівневої системи біометричної автентифікації за ЕКГ (рис. 1).

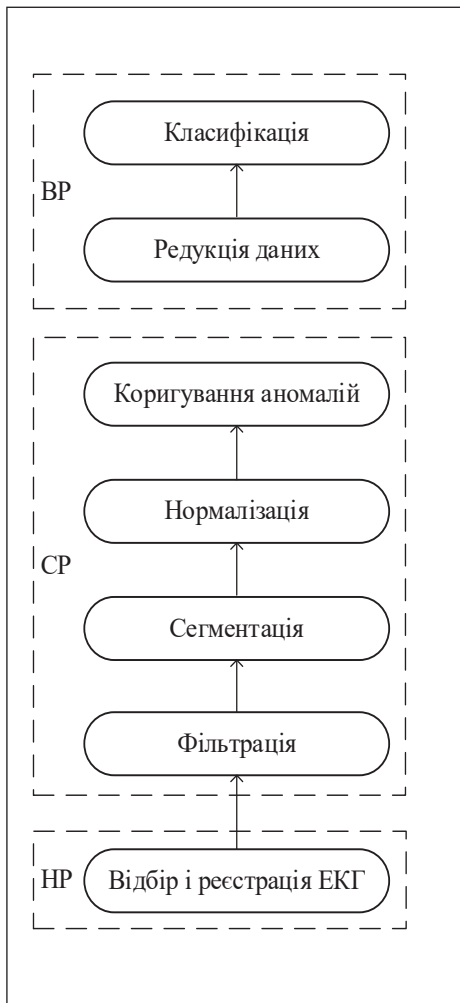


Рис. 1. Структурна схема біометричної системи автентифікації на основі ЕКГ-сигналу

На нижньому системному рівні (НР) для відбору і реєстрації електрокардіограм застосовано просту і доступну вимірювально-обчислювальну систему на базі таких апаратних засобів, як e-Health Sensor Platform V2.0 (рис. 2,а) та Arduino Uno (рис. 2,б), із відповідним програмним забезпеченням на мові Python. Для формування сигналу ЕКГ здійснювався відбір електричних потенціалів із пальців правої та лівої руки трьома електродами (рис. 2, в).

Цей спосіб реєстрації електрокардіограми за термінологією медичної діагностики відповідає першому стандартному відведенню (I-lead).

До основних компонент апаратної обчислювальної платформи Arduino належать плата мікроконтролера з елементами вводу / виводу та середовище розробки Arduino IDE. Платформа Arduino може використовуватися як для створення автономних інтерактивних об'єктів, так і підключатися до програмного забезпечення, яке виконується на комп'ютері [20]. Апаратна платформа e-Health Sensor Shield V2.0 дає змогу використовувати Arduino для біометричних і медичних цілей. За допомогою десяти сенсорів можна вимірювати різні біомедичні параметри і сигнали [21].

З нижнього рівня оцифровані за допомогою 8-бітного АЦП із частотою дискретизації 277 Гц ЕКГ-сигнали передаються через COM-порт на ПК з використанням бібліотеки PySerial. Кожне вимірювання тривало близько 10 с., відповідно, кожний запис зазвичай містить не менше 10-ти серцевих циклів.

На ПК реалізуються функції середнього системного рівня (СР), націленого на кондиціонування ЕКГ-сигналу методами цифрового опрацювання сигналів, а також верхнього (ВР), де виконується інтелектуальний аналіз даних засобами машинного навчання. Корисний ЕКГ-сигнал має низький рівень, а тому особливо вразливий на спотворення такого характеру:

- низькочастотний складник, спричинений диханням, рухами тіла, потовиділенням, що впливає на імпеданс контакту «тіло-електрод»;
- зосереджена гармонічна завада 50 Гц (наведення від мережі живлення);
- широкосмугові завади, зумовлені інтерференцією електричних потенціалів м'язів, внутрішніми шумами електронних пристроїв вимірювального тракту, побічним електромагнітним випромінюванням.

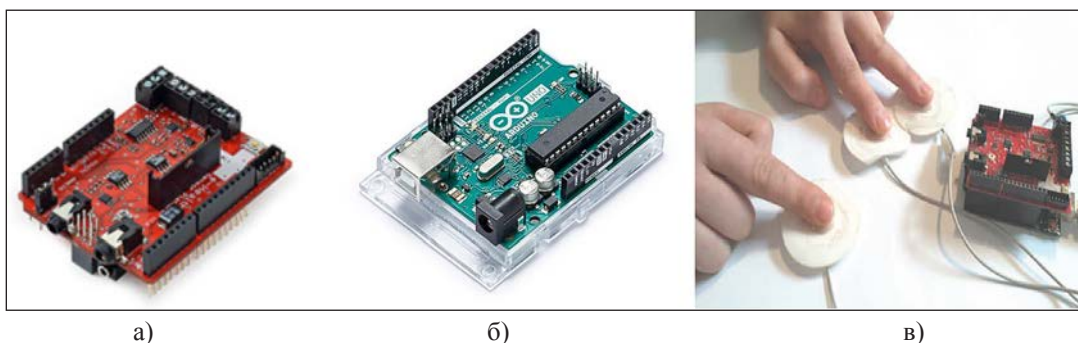


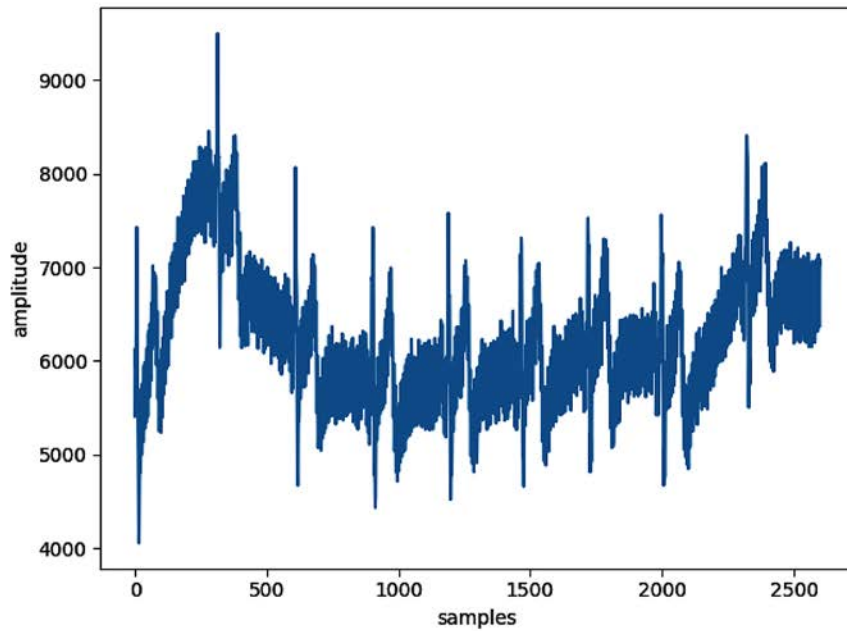
Рис. 2. Мікросхеми вимірювального тракту біометричної системи автентифікації

Вигляд зареєстрованого сирого ЕКГ-сигналу наведено на рис. 3, а. Для придушення зазначених вище шумів застосовано смуговий рекурсивний фільтр Баттерворта, а результат для візуальної оцінки показано на рис. 3, б.

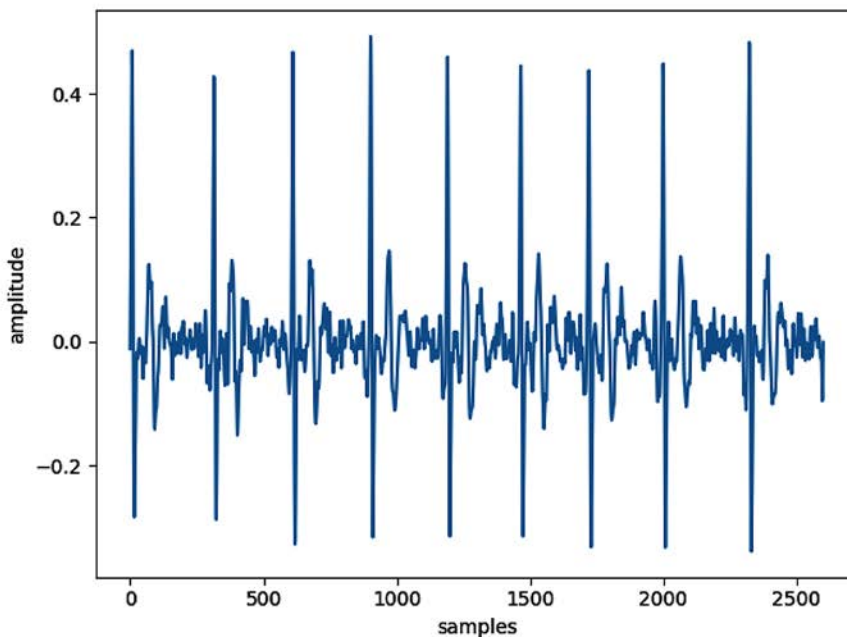
Далі відфільтрований сигнал піддається сегментації та нормалізації за амплітудою. Сегментацію проведено на основі відомого алгоритму Гамільтона шляхом детекції R-піків у записі (рис. 4). Це уможливило виділення сегментів, еквівалентних окремим циклам серцебиття. Сегментований

ЕКГ-сигнал далі унормовується за амплітудою (до масштабу  $\pm 1$ ).

Для використання електрокардіограми як біометричного маркера у реальних системах автентифікації важливо забезпечити інваріантність класифікатора до змін серцевого ритму. Для цього на наступному етапі застосовано темпоральну нормалізацію за алгоритмом, детально представленим у роботі [19]. Ідея алгоритму полягає у приведенні всіх сегментів із записів кожного суб'єкта до однакової фіксованої тривалості. З



а)



б)

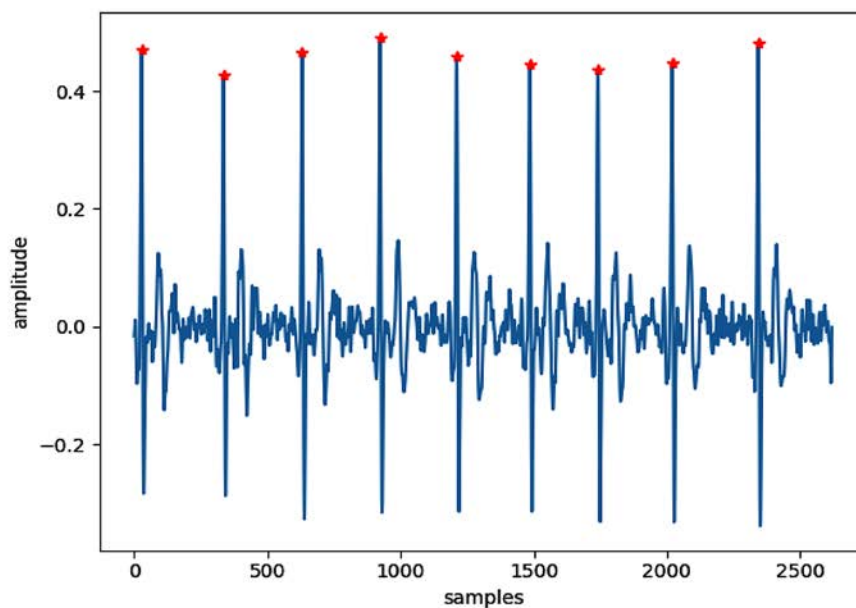
Рис. 3. Вимірний ЕКГ-сигнал із шумами до (а) та після (б) фільтрації



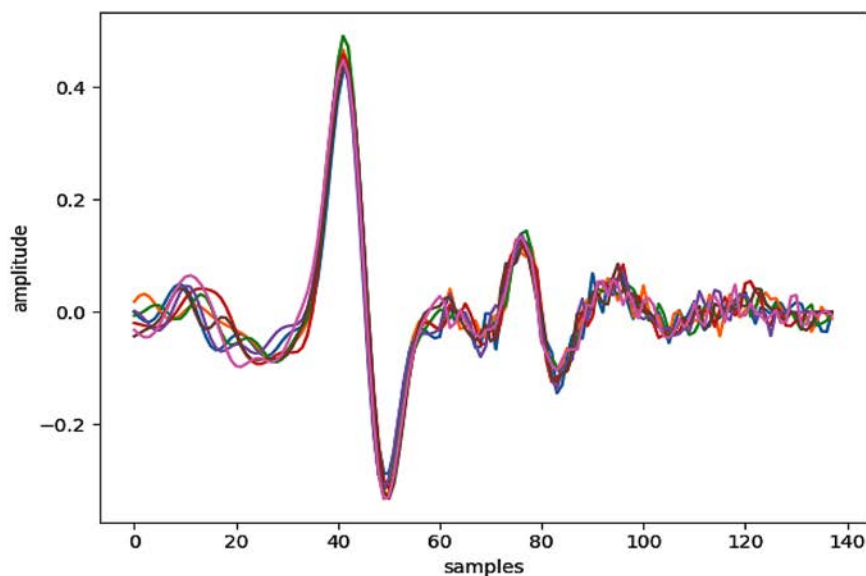
цією метою кожен сегмент розбивається на дві частини: перший – від початку Р зубця до кінця QRS-комплексу, а другий – від початку ST-сегмента до кінця U зубця. Перша частина сегмента мало залежить від зміни серцевого ритму, тому залишається незмінною, а друга піддається стисненню чи розтягненню так, щоб тривалість цілого нормалізованого сегмента точно становила  $2/3$  секунди (90 уд./хв.).

Навіть після знешумлення у ЕКГ-записах, а отже, і в унормованих за амплітудою і часом

сегментах можуть залишаються спотворення у вигляді аномальних відхилень окремих чи серії вибірок. Щоб точність автентифікації не знижувалася, застосовано статистичний метод виявлення та коригування аномалій, детально описаний у [22]. Даний метод заснований на ковзаючому вікні, яке обчислює евклідову відстань між вибірками всередині вікна та відповідними усередненими на всіх сегментах значеннями. Якщо відстань перевищує встановлений поріг, всі вибірки у поточному вікні розглядаються як



a)



б)

Рис. 4. ЕКГ-запис з відміченими R-піками (а) і унормовані за рівнем сегменти ЕКГ-сигналу (б)

аномальні, а тому підлягають заміні на усереднені значення.

У машинному навчанні для збільшення продуктивності роботи класифікатора часто зменшують розмірність даних. Для цього використано метод головних компонент (Principle Component Analysis – PCA) як один із найпростіших, але ефективних алгоритмів редукції розмірності даних [23].

Заключним етапом опрацювання електрокардіограм у системі біометричної автентифікації є їх класифікація. Класифікатор побудовано на основі багатошарової штучної нейронної мережі (перцептрона). Нейрони кожного шару з'єднуються з нейронами попереднього і наступного шарів за принципом «кожний з кожним». Вхідний шар розробленого нейромережевого класифікатора складається із 30 нейронів, що відповідає числу дискрипторів (інформаційних ознак), відібраних за алгоритмом PCA із вибірок кожного сегменту ЕКГ-сигналу. Розмір двох прихованих шарів становить 256 і 128 нейронів, а розмір вихідного шару визначається числом зареєстрованих у біометричній системі користувачів. У нейромережевому класифікаторі застосовано такі гіперпараметри: функція активації (гіперболічний тангенс), алгоритм оптимізації (Adam), коефіцієнт L2 регуляризації (0,0001), точність оптимізації ( $10^{-5}$ ), чисельна стабільність ( $10^{-8}$ ).

**Методика виконання експериментів.** У дослідженнях використовувалася база даних електрокардіограм Lviv Biometric Data Set (LBDS), одержаних на описаному у попередньому пункті обладнанні, яка на момент написання статті мстила 1809 ЕКГ-записів від 115-ти осіб [24]. Важливо зазначити, що у базі зібрано записи сигналів ЕКГ із кількох сеансів вимірювань, проведених упродовж двох років. На цей час база електрокардіограм LBDS є однією із найбільших, як за числом осіб, так і за обсягом записів. Набір даних із цієї бази зорієнтований на біометричні застосування та неодноразово використовувався для різних досліджень [5; 14; 25].

Для оцінювання ефективності роботи розробленої системи біометричної автентифікації використано такі загальноприйняті у машинному навчанні метрики [14; 26]:

– помилка першого виду FRR (False Reject Rate) – ймовірність того, що система автентифікації не зможе автентифікувати справжнього користувача (або часто говорять, що система приймає «свого» за «чужого»);

– помилка другого роду FAR (False Accept Rate) – ймовірність того, що система автентифікації автентифікує незареєстрованого користувача (тобто прийме «чужого» за «свого»);

– точність (Accuracy) – відношення числа вірно розпізнаних об'єктів до загального їх числа.

Для детальнішої інтерпретації результатів досліджень використано матрицю помилок (confusion matrix), яка має розмірність  $N \times N$  ( $N$  – число класів). Стовпці відповідають істинним значенням, а рядки – результатам роботи класифікатора.

У статті наведено результати досліджень, отриманих авторами в процесі двох експериментів. Мета першого експерименту – дослідження середньострокової стійкості електрокардіограми як біометричного маркера. Для цього з бази даних електрокардіограм було виокремлено підмножину ЕКГ-записів, інтервал між якими становив один-два місяці. Отриманий набір даних містить записи ЕКГ 20-ти осіб. Його було розділено на навчальний та тестовий набори так, щоб проміжок між останнім вимірюванням навчального набору та першим вимірюванням тестового становив як мінімум один місяць. Для кожної особи у табл. 1 наведено дані про число записів у навчальному і тестовому наборах.

Результати першого експерименту наведено у табл. 2. Для демонстрації ефекту темпоральної нормалізації експерименти проведено в наступній конфігурації:

– без застосування алгоритму темпоральної нормалізації. У навчальному та тестовому наборі збережено оригінальне значення серцевого ритму;

– з застосування алгоритму темпоральної нормалізації. Значення серцевого ритму усіх ЕКГ-записів навчального та тестового наборів приведено до значення, еквівалентного 90 ударам серця за хвилину.

У багатьох дослідженнях здійснюють випадковий поділ на навчальний і тренувальний набори даних. Це зумовлено тим, що зазвичай набори даних, які використовуються у цих роботах, містять невеликий обсяг вимірювань, виконаних впродовж короткого проміжку часу. Набір даних, який використовується у цій роботі, дає можливість розділити тренувальний і тестовий набори за часом вимірювання (це повинно репрезентувати використання біометричної системи в реальних умовах повсякденного використання). Тому в цій роботі також представлено порівняння цих двох варіантів поділу на навчальний і тестовий набори даних: випадковий та за часом вимірювання.

Вища точність автентифікації за випадкового розподілу пояснюється тим, що ЕКГ-записи як у навчальному, так і тестовому наборі було виміряно на короткому проміжку часу. Результати у варіанті часового поділу наборів ЕКГ-записів, тобто коли тестовий набір виміряно на один-два місяці пізніше, ніж навчальний, також є достатньо високими за умови проведення темпоральної нормалізації серцевого ритму (див. останню колонку табл. 2). Для згаданої конфігурації системи автентифікації матрицю помилок наведено на рис. 5, де її головна діагональ матриці репрезентує випадки вірно автентифікованих користувачів за їх ЕКГ-сигналами. Як видно за нормалізації серцевого ритму, мало місце 9 помилок на 205 ЕКГ-записах (проти 24 без такої нормалізації), а також відпо-

відно є кращими результати за метриками точності, FAR і FRR.

Варто також зазначити, що система здатна розпізнати кожного користувача (відсутні нульові елементи на головній діагоналі матриці помилок). Це свідчить про те, що імплементована система сконфігурована належним чином та може достовірно автентифікувати її користувачів.

Метою другого експерименту було – дослідити стійкість електрокардіограми як біометричного маркера на тривалішому часовому горизонті. Для цього з бази електрокардіограм було виокремлено підмножину записів, інтервал між якими становив приблизно два роки. Отриманий набір даних містить ЕКГ-записи 20-ти осіб (як і у першому експерименті). Його було розділено на навчальний

Таблиця 1

Опис набору даних першого експерименту

Номер суб'єкта	Навчальний набір		Тестовий набір		Інтервал, дні
	Кількість вимірювань	Дата останнього вимірювання	Кількість вимірювань	Дата першого вимірювання	
user4	11	2017-12-01	11	2018-02-12	73
user9	11	2017-10-25	11	2017-11-22	28
user13	10	2017-12-01	10	2018-02-12	73
user14	11	2017-12-01	11	2018-02-12	73
user15	11	2017-12-01	2	2018-02-12	73
user54	17	2017-12-01	11	2018-02-12	73
user55	11	2017-12-19	11	2018-02-01	44
user73	9	2017-10-25	10	2017-11-22	28
user74	8	2017-10-25	17	2017-11-22	28
user75	9	2017-10-25	11	2017-11-22	28
user76	11	2017-10-25	11	2017-11-22	28
user78	11	2017-10-25	11	2017-11-22	28
user80	21	2017-10-25	22	2017-11-22	28
user81	11	2017-10-25	11	2017-11-22	28
user82	11	2017-10-25	10	2017-11-22	28
user84	11	2017-10-25	5	2018-01-24	91
user85	11	2017-11-22	7	2018-01-24	63
user86	11	2017-11-22	6	2018-01-24	63
user87	17	2017-11-22	8	2018-01-24	63
user88	11	2017-11-22	9	2018-01-24	63

Таблиця 2

Результати стійкості ЕКГ як біометричної характеристики упродовж 1-2 місяців

Поділ навчального і тестового наборів	Без темпоральної нормалізації		Із темпоральною нормалізацією	
	Випадковий	Відповідно до табл. 1	Випадковий	Відповідно до табл. 1
Обсяг тестового набору	205			
Число помилок	3	24	2	9
Точність	0.9854	0.8829	0.9902	<b>0.9561</b>
FAR	0.0009	0.0106	0.0007	<b>0.0031</b>
FRR	0.011	0.0399	0.0092	0.0376

і тестовий набори так, щоб проміжок між останнім вимірюванням навчального набору та першим вимірюванням тестового становив два роки. У табл. 3 наведено відомості про число ЕКГ-записів у навчальному і тестовому наборах даних для кожної особи. Зазначимо, що користувачі у першому і другому експериментах не збігаються на 100% через відсутність у базі даних їх ЕКГ-записів за обидвома періодами.

Результати другого експерименту наведено у табл. 4. Помітно, що при випадковому розподілі ЕКГ-записів на навчальний та тестовий набори система також показала високу точність автентифікації. Але і результати автентифікації за тестовим набором ЕКГ-сигналів, що виміряні два роки пізніше за навчальний набір, також є достатньо

Рис. 5. Матриця помилок експерименту 1 для конфігурації наборів ЕКГ із табл. 1

Таблиця 3

Опис набору даних другого експерименту

Номер суб'єкта	Навчальний набір		Тестовий набір		Інтервал, дні
	Кількість вимірювань	Дата останнього вимірювання	Кількість вимірювань	Дата першого вимірювання	
user4	32	2018-02-12	15	2019-12-27	683
user9	35	2017-11-22	15	2019-12-20	758
user13	34	2018-02-12	15	2019-11-01	627
user14	33	2018-02-12	11	2019-12-27	683
user15	17	2018-02-12	6	2019-11-08	634
user17	21	2018-02-12	10	2019-12-09	665
user73	25	2018-01-24	9	2019-12-20	695
user74	31	2018-01-24	7	2019-12-20	695
user75	20	2017-11-22	10	2019-12-20	758
user78	26	2018-01-24	12	2019-12-20	695
user80	50	2018-01-24	6	2019-12-20	695
user81	30	2018-01-24	9	2019-12-20	695
user84	16	2018-01-24	7	2019-12-20	695
user85	18	2018-01-24	10	2019-12-20	695
user86	17	2018-01-24	10	2019-12-20	695
user87	25	2018-01-24	10	2019-12-20	695
user88	20	2018-01-24	6	2019-12-20	695
user96	7	2018-01-24	8	2019-12-20	695
user97	10	2018-02-12	9	2019-12-27	683
user99	11	2018-02-12	14	2019-12-27	683

Таблиця 4

Результати стійкості ЕКГ як біометричної характеристики упродовж 1-2 років

Поділ навчального і тестового наборів	Без темпоральної нормалізації		Із темпоральною нормалізацією	
	Випадковий	Відповідно до табл. 3	Випадковий	Відповідно до табл. 3
Обсяг тестового набору	199			
Число помилок	2	45	0	<b>10</b>
Точність	0.9899	0.7739	1.0	<b>0.9497</b>
FAR	0.0006	0.0177	0.0	<b>0.0025</b>
FRR	0.0096	0.0839	0.0	0.0433



user_4	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
user_9	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
user_13	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
user_14	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
user_15	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
user_17	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0
user_73	0	0	0	0	0	0	8	0	1	0	0	0	0	0	0	0	0	0	0
user_74	0	0	0	0	0	0	0	5	0	1	0	0	0	0	0	0	0	0	0
user_75	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0
user_78	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0
user_80	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0
user_81	0	0	0	0	0	0	1	0	0	0	0	8	0	0	0	0	0	0	0
user_84	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0
user_85	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
user_86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	1	0	0	0
user_87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0
user_88	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	5	0	0
user_96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0
user_97	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	7
user_99	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	13

Рис. 6. Матриця помилок експерименту 2 для конфігурації наборів ЕКГ із табл. 3

високими. Як видно із рис. 6, мало місце 10 помилок на 199 ЕКГ-записах (проти 45 без нормалізації серцевого ритму).

Таким чином, на основі результатів двох експериментів можна зробити висновок, що електрокардіограма є стабільним у часі біометричним маркером – точність автентифікації за тестовими записами ЕКГ-сигналів, зробленими через місяць та через два роки є на одному рівні – відповідно 95,61% і 94,97%. Близькими також є значення помилок: FAR - 0.003/10.0025 і FRR - 0.0376/0.0433.

**Висновки.** Електрокардіограма несе в собі інформацію не лише про стан і особливості роботи серця людини, але також є визнаним і надійним біометричним маркером для кожного індивіда. Отже, ЕКГ-сигнали можуть застосовуватися у системах біометричної автентифікації та контролю доступу, хоча ще не набули комерційного поширення, подібно до таких біомаркерів як відбитки пальців чи голос. Для застосування у реальних системах важливо дослідити стабільність інформативних параметрів ЕКГ-сигналу, за

якими класифікатор приймає рішення, а саме – належить чи ні поточна електрокардіограма заявленому суб'єкту. Як показав аналіз публікацій, подібних досліджень ще не проводилося.

Автори розробили прототип системи біометричної автентифікації за ЕКГ-сигналами, а також детально описали застосовані програмно-апаратні засоби відбору і реєстрації електрокардіограм, а також алгоритмів їх кондиціонування сигналу та інтелектуального аналізу. Також у деталях розписано мету і методику виконання двох експериментів.

Дослідження мали на меті валідувати дві гіпотези:

- підтвердити і порівняти відтворюваність біометричного маркера на основі ЕКГ на різних часових горизонтах, а саме впродовж одного місяця і двох років;
- обґрунтувати доцільність нормалізації серцевого ритму (пульсу) в системах біометричної автентифікації.

З метою реалізації поставлених цілей авторами було використано власну базу ЕКГ-записів Lviv Biometric Data Set (LBDS), яка в момент написання статті містила 1809 ЕКГ-записів 115-ти осіб, зібраних впродовж двох років. На цей час це один з найбільших наборів даних для біометричних застосувань, як за числом унікальних осіб і окремих записів, так і за часовим інтервалом, впродовж якого збиралися ці дані.

Проведені дослідження засвідчили, що ЕКГ є стабільним маркером для побудови систем автентифікації, на основі якого система здатна адекватно розпізнавати користувачів упродовж тривалого часу без необхідності проміжних перекалібрувань системи.

Крім того, слід відзначити, що застосування алгоритму темпоральної нормалізації ЕКГ сигналу засвідчило свою ефективність. Його застосування дало змогу підвищити точність ідентифікації на 7% (з 88,29% до 95,61%) – на горизонті в один місяць та аж на 17% (з 77,39% до 94,97%) – на горизонті в два роки.

#### Список літератури:

1. Seonwoo M., Byunghan L., Sungroh Y. Deep learning in bioinformatics. *Briefings in Bioinformatics*. Vol. 18(5), 2017, p. 851–869.
2. Miotto R., Wang F., Wang S., Jiang X., Dudley J.T. Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*. Vol. 19(6), 2018, p. 1236–1246.
3. Pinto J.R., Cardoso J.S., Lourenço A. Evolution, Current Challenges, and Future Possibilities in ECG Biometrics. *IEEE Access*. Volume 6, 2018, p. 34746–34776.
4. Fratini, A., Sansone, M., Bifulco, P., Isarelli, M. Individual identification via electrocardiogram analysis. *BioMedical Engineering OnLine* (2015), 23 p.
5. Pelc M., Khoma Y., Khoma V. ECG Signal as Robust and Reliable Biometric Marker: Datasets and Algorithms Comparison. *Sensors*. 2019, 19(10), 2350, p. 1–8.
6. Belgacem, N.; Fournier, R.; Nait-Ali, A.; Bereksi-Reguig, F. A novel biometric authentication approach using ECG and EMG signals. *J. Med. Eng. Technol.* 2015, 39, 226–238.

7. Matos, A.C., Lourenco, A., Nascimento, J. Embedded system for individual recognition based on ECG Biometrics. Conf. on Electronics, Telecom. and Computers. Procedia Technology 17, 2014, p. 265–272.
8. Albulbul, A. Evaluating Major Electrode Types for Idle Biological Signal Measurements for Modern Medical Technology. Bioengineering. 2016, 20(3), p. 1–10.
9. Wieclaw, L., Khoma, Y., Falat, P., Sabodashko, D., & Herasymenko, V. (2017). “Biometric Identification From Raw ECG Signal Using Deep Learning Techniques”. 9th IEEE Intern. Conf. on Intelligent Data Acquisition & Advanced Computing Syst: 21-23 September 2017, Bucharest, Romania, p. 129–133.
10. Lugovaya, T.S. “Biometric human identification based on electrocardiogram”. Saint-Petersburg, Russian Federation, Publ. “LETI”, 2005.
11. Hodge, V., & Austin, J. (2004). A Survey of Outlier Detection Methodologies. Artificial Intelligence Review, 22(2), p. 85–126.
12. Chan, A.D.C., Hamdy, M.M., Badre, A., et al. (2008). Wavelet distance measure for person identification using electrocardiograms. *IEEE Trans. Instrum. Meas.* Vol. 57, No. 2, 248–253.
13. Lourenco, A., Plácido da Silva, H., Carreiras, C. (2013). Outlier detection in non-intrusive ECG biometric system. In International Conference Image Analysis and Recognition. Berlin, Heidelberg, 43–52.
14. Хома, В.В., Хома, Ю.В., Сабодашко, Д.В., Хома П.П. Автоенкодера для опрацювання промахів сигналів ЕКГ у системі біометричної автентифікації. *Штучний інтелект.* № 1–2, 2019, с. 101–110.
15. Karpinski, M., Khoma, V., Dudykevych, V., Khoma, Y., & Sabodashko, D. Autoencoder Neural Networks for Outlier Correction in ECG-Based Biometric Identification. IEEE 4th Int. Symp. on Wireless Syst. (IDAACS-SWS), 2018 p. 210–215.
16. Tawfik M.M., Selim H., Kamal T. Human identification using time normalized QT signal and the QRS complex of the ECG. In Proc. 7th Int. Symp. Commun. Syst. Netw. Digit. Signal Process. (CSNDSP), Jul. 2010, pp. 755–759.
17. Lourenço A., Silva H., Fred A. Unveiling the biometric potential of finger-based ECG signals. *Comput. Intell. Neurosci.* vol. 2011, p. 5, Jan. 2011.
18. Kim K.S., Yoon T.H., Lee J.W., Kim D.J., Koo H.S. A robust human identification by normalized time-domain features of electrocardiogram. In: 27th annual international conference of the engineering in medicine and biology society, 2005 IEEE-EMBS 2005. 2005. p. 1114–7.
19. Дудикевич В.Б., Хома В.В., Чекурін В.Ф., Хома Ю.В., Сабодашко Д.В. Нормалізація сигналів ЕКГ для застосування в системах біометричної ідентифікації. Вчені записки Таврійського національного університету ім. В.І. Вернадського. Серія: Технічні науки, 2019, Том 30 (69), ч. 1 № 4, с. 49–56.
20. Arduino UNO & Genuino UNO Pi. URL: <https://www.arduino.cc/en/Main/arduinoBoardUno>.
21. e-Health Sensor Platform V2.0 for Arduino and Raspberry Pi. URL: <https://www.cooking-hacks.com/documentation/tutorials/ehealth-biometric-sensor-platform-arduino-raspberry-pi-medical>.
22. Khoma, V., Pelc, M., Khoma, Y., & Sabodashko, D. “Outlier Correction in ECG-Based Human Identification”. “Biomedical Engineering and Neuroscience”. *BCI Adv. in Intel. Syst. & Comp.* 2018, Vol. 720. p. 11–22. Springer.
23. Метод головних компонент. URL.: [https://uk.wikipedia.org/wiki/Метод\\_головних\\_компонент](https://uk.wikipedia.org/wiki/Метод_головних_компонент).
24. Lviv Biometric Data Set. Available online. URL: <https://github.com/YuriyKhoma/Lviv-Biometric-Data-Set>.
25. Khoma V., Pelc M., Khoma Y. Artificial Neural Network Capability for Human Being Identification based on ECG. *Proc. 23rd Intern. Conf. on Methods and Models in Automation and Robotics*, Miedzyzdroje, 27-30 August 2018, p. 479 – 482.
26. Хома Ю.В. «Класифікація віброартрографічних сигналів з використанням хвилькового перетворення і технологій машинного навчання» *Вісник НУ «Львівська політехніка» – «Інформаційні системи та мережі»*. 2019. Вип. 5. с. 40–52.

**Sabodashko D.V., Khoma Yu.V., Khoma V.V. STUDY OF TIME STABILITY OF ECG SIGNAL AS A BIOMETRIC MARKER IN AUTHENTICATION SYSTEM**

Many studies has already proven that electrocardiogram is unique for individual person. Currently, the most issues related to biometric systems design are solved – simple analog-front end for measuring of electrical potentials, effective methods of the ECG signal conditioning, reliable algorithms for the intellectual analysis for individual hidden patterns disclosure from the electrocardiogram. Long-term stability of the ECG signal informative parameters is important for the operation of biometric authentication systems in real-world applications, however this question still remains opened.

Therefore, the purpose of this work was to investigate the stability of an electrocardiogram as a unique biometric marker for human authentication over long time intervals of months/years.

The article presents its own design and implementation of the ECG-based system for human biometric authentication, which was used for the registration of ECG records, their computer processing, training and

*testing of models of neural network classifiers. The biometric authentication system was verified on an open Lviv Biometric Data Set (LBDS) dataset, which collected over 1800 unique electrocardiogram records lasting from 10 to 20 seconds. The recordings were made over a two-year period and are belong to 115 different individuals. Research methodology for two sets of experiments is described in the paper. The first experiment is aimed to verify systems reliability after one or two months, science neural network classifier has been trained, while the second experiments is aimed to verify reliability for a much longer timeline of two years. Both experiments have shown rather good results: the accuracy of authentication with one month and two years interval is kept practically at the same level – 95,61% and 94,97% respectively. This leads to the conclusion that the electrocardiogram is a time-stable biometric marker that can be used in real-world biometric applications.*

**Key words:** *biometric system, ECG records, neural network classifier, time stability of biometric markers, biometric authentication, FAR, FRR.*