

## АНОТАЦІЯ

**Сабодашко Д.В.** Вдосконалення методів і засобів біометричної автентифікації на основі електрокардіограми. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 125 «Кібербезпека» (12 – Інформаційні технології). – Національний університет «Львівська політехніка», Львів, 2021.

У дисертаційній роботі розв'язано актуальну науково-прикладну задачу у галузі кібербезпеки - покращення характеристик системи біометричної автентифікації за сигналом електрокардіограми на основі раціонального поєднання технологій цифрового оброблення сигналів і машинного навчання, що підвищує рівень захищеності ресурсів на об'єктах інформаційної діяльності.

Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку літературних джерел та додатків.

В *першому розділі* розглянуто основні режими роботи і подано порівняльну характеристику сучасних систем біометричної автентифікації. Проведено порівняння найпоширеніших біометричних маркерів за допомогою формалізованих критеріїв (універсальність, унікальність, постійність, вимірюваність, продуктивність, прийнятність, стійкість до обману, ціна, тощо). Представлено детальний опис електрокардіограми (ЕКГ) як біометричного маркера в системах розпізнавання, показано його переваги і проблеми на шляху практичного застосування в системах автентифікації. Проаналізовано відомі підходи опрацювання ЕКГ-сигналу на основі виділення характерних точок (fiducial points) та без такого виділення (non-fiducial point), тобто на основі інтелектуального аналізу повного набору вибірок ЕКГ-сигналу. Сформульовано завдання дисертаційного дослідження.

У *другому розділі* розглянуто особливості процесу автентифікації за ЕКГ-сигналом. Формалізовано структуру біометричної системи розпізнавання.

Наведено детальний опис і функції кожного із структурних елементів. Розглянуто перспективні підходи до покращення технічних і експлуатаційних характеристик біометричної системи ЕКГ-автентифікації. Передовсім, обґрунтовано доцільність введення в ланцюг опрацювання електрокардіограми двох додаткових компонент:

- компонента темпоральної нормалізації ЕКГ-сигналу, що покликана забезпечити інваріантність результатів автентифікації до зміни частоти серцевого ритму, тим самим підвищивши достовірність роботи біометричної системи;
- компонента виявлення та коригування артефактів у ЕКГ-сигналі, яка за допомогою інструментарію статистики або машинного навчання підвищує точність і швидкодію системи біометричної автентифікації.

Подано методики оцінювання ефективності методів і засобів біометричної автентифікації на основі ЕКГ-сигналу. Представлено сформовану автором базу записів електрокардіограм (Lviv Biometric Dataset), яка на момент написання дисертації містила 1809 записів виміряних у 115 суб'єктів на часовому горизонті понад два роки. Базу Lviv Biometric Dataset викладено у відкритий доступ, поряд із іншими базами ЕКГ-записів.

*Третій розділ* спрямовано на розроблення моделей та методів для покращення характеристик біометричних систем автентифікації на основі ЕКГ.

Спрощений і зручний для систем автентифікації відбір ЕКГ-потенціалів із пальців лівої і правої рук призводить до зниження якості запису. Частотні смуги корисного сигналу і завад перекриваються, тому після цифрової фільтрації у ЕКГ-записах спостерігаються залишкові артефакти. Описані у літературних джерелах підходи спираються на виявлення і відкидання фрагментів ЕКГ з аномальними відхиленнями.

У роботі вперше запропоновано не відкидати, а виправляти фрагменти з аномальними відхиленнями, що важливо для збереження необхідного обсягу даних для класифікатора і скорочення часу відбору ЕКГ. Розроблено підхід до

виправлення залишкових артефактів у ЕКГ-сигналах, який складається із трьох етапів:

1. формування референційного образу біометричного маркера;
2. виявлення фрагментів ЕКГ-сигналу із промахами;
3. заміна цих фрагментів на відповідні значення із референційного образу.

Застосування такого підходу дає змогу суттєво підвищити достовірність результатів автентифікації.

Запропоновано і досліджено два методи формування референційного образу ЕКГ-маркера для виправлення залишкових артефактів у ЕКГ-сигналах у системі біометричної автентифікації:

- на основі невимогливі до обчислювальних ресурсів формальної статистичної моделі;

- на базі нечіткої нейромережевої моделі, що дає змогу зменшити похибки автентифікації першого і другого роду відповідно у 4 та 3 рази.

Для етапу виявлення фрагментів із промахами виконано дослідження впливу гіперпараметрів (тривалість ковзного вікна і поріг допустимого відхилення вибірки) на точність автентифікації, що дало змогу знайти оптимальні їх значення за критерієм мінімальної похибки.

Розроблено та апробовано інструментарій для темпоральної нормалізації ЕКГ-сигналу. Серед сучасних методів класифікації (метод опорних векторів, лінійний дискримінантний аналіз, k-найближчих сусідів, дерева рішень, нейронні мережі, тощо) здійснено вибір оптимального для побудови системи автентифікації. Досліджено придатність біометричної системи автентифікації до масштабування, а саме визначено вплив збільшення числа користувачів на точність автентифікації.

У *четвертому розділі* імплементовано біометричну систему автентифікації з покращеними характеристиками на основі використання розроблених автором моделей і методів. Зокрема, щоб показати можливість імплементування біометричної системи ЕКГ-автентифікації на пристроях з обмеженими

обчислювальними ресурсами імплементовано біометричну систему на базі мікрокомп'ютера Raspberry Pi 3B. Проведено дослідження швидкодії імплементованої біометричної системи, за результатами якого сформовано рекомендації для імплементції біометричних систем автентифікації як на основі персональної робочої станції, так і на базі мікрокомп'ютера Raspberry Pi.

Досліджено часову стабільність ЕКГ-сигналів на довготривалих проміжках часу (роки, місяці), а також визначено ступінь впливу варіативності інформативних ознак електрокардіограми на точність автентифікації. Результати досліджень засвідчили, що ЕКГ є стабільним маркером і може застосовуватися у реальних системах автентифікації, причому система здатна адекватно розпізнавати користувачів упродовж тривалого часу без необхідності проміжного калібрування. Таким чином, доведено високий потенціал і перспективність електрокардіограми, як біометричного маркера, для побудови надійних систем автентифікації.

Наведено перелік можливих сфер застосування та опис прикладних застосувань для біометричних систем автентифікації за ЕКГ-сигналом.

**Ключові слова:** біометрична автентифікація, електрокардіограма, цифрове оброблення сигналу, машинне навчання, нейронні мережі, автоенкодера, корекція аномалій, темпоральна нормалізація, Raspberry Pi.

## ABSTRACT

***Sabodashko D.V. Improvement of methods and means of biometric authentication based on electrocardiogram.*** – Manuscript.

Thesis paper for achievement of the scientific degree Doctor of Philosophy in the specialty 125 “Cybersecurity” (12 – Information technologies). – Lviv Polytechnic National University, Lviv, 2021.

The thesis paper solved the relevant scientific and applied task in the sphere of cybersecurity – improvement of characteristics in biometric authentication systems based on electrocardiogram signal via the rational combination of digital signal processing technologies and machine learning, which increases the resources protection level on the information activity objects.

The thesis paper consists of introduction, four chapters, conclusions, list of references and appendixes.

*The first chapter* examines the main working modes and provides the comparative characteristics of modern biometric authentication systems. Comparison of the most widely spread biometric markers using the formalized criteria (scalability, uniqueness, steadiness, productivity, acceptability, stability, forge resistance, price, etc.) was performed. Detailed description of electrocardiogram (ECG) as a biometric marker in the recognition systems was presented, its advantages and problems on the way of practical application in the authentication systems were revealed. Known approaches of ECG-signal processing based on fiducial points extraction and without such an extraction (non-fiducial point), thus on the basis of the intellectual analysis of the full set of ECG-signal data selection were analyzed. The task of the thesis research was stated.

In *the second chapter* the peculiarities of authentication process based on ECG-signal are studied. The structure of biometric recognition system is formalized. The detailed description and functions of each of the structure elements are provided. Perspective approaches to the improvement of the technical and operating

characteristics for ECG-authentication biometric system are studied. First of all, the expediency of introduction into the electrocardiogram processing pipeline of two additional components is reasoned:

- component of ECG-signal temporal normalization, which is supposed to provide invariability of the authentication results as to the change of the heart rate, thus increasing the authenticity of the biometric system work;
- component detecting and correcting anomalies in the ECG-signal, which using the statistical or machine learning instruments increase the accuracy and the speed of the biometric authentication.

Techniques evaluating efficiency of methods and means of biometric authentication based on ECG-signal are provided. Electrocardiogram records dataset (Lviv Biometric Dataset) formed by the author is presented, which at the moment the thesis was written contained 1809 records measured from 115 subjects on the time horizon of over two years. Lviv Biometric Dataset is open sourced together with other ECG-records datasets

*The third chapter* is directed onto development of models and methods for improvement of ECG-based biometric authentication systems.

The simplified and friendly to authentication systems selection of ECG-signals from the fingers of the left and right hands lead to decrease in record quality. Frequency bands of the useful signal and obstacles are overlaid, thus after the digital filtering the residual artifacts remain the in the ECG-signal. The approaches described in the references are drawn upon the detection and removal of ECG fragments with anomaly deviations.

It was first suggested in the work not to remove, but to correct the fragments with the anomaly deviations, which is important for preservation of the data amount for the classifier and shortening of time for ECG selection. Approach for correction of the residual artifacts in the ECG-signals was developed, it consists of three stages:

- 1) formation of reference pattern for the biometric marker;
- 2) detection in the ECG-signal of fragments with anomalies;

3) substitution of these fragments with the corresponding values from the reference pattern.

Application of such an approach allows to substantially increase the validity of the authentication results.

Two methods for reference pattern formation of the ECG-marker were suggested for correction of the residual artefacts in the ECG-signals for systems of biometric authentication:

- based on the undemanding as to the calculation resources formal statistical model;
- based on the fuzzy neural network model allowing to decrease the type I and type II authentication errors in 4 and 3 times correspondingly.

For the purposes of stage detecting fragments with anomalies the research on influence of hyperparameters was performed (duration of the sliding window and the threshold of the acceptable deviation of the dataset) for accuracy of the authentication, which allowed to find their optimal values according to the criterion of the minimal error.

Instruments for temporal normalization of ECG signal were designed and tested. The selection was performed among the modern classification methods (support vector machines, linear discriminant analysis, k-nearest neighbors algorithm, decision trees, neural networks, etc.) for optimal design of the authentication system. Scalability of the biometric authentication system was researched, namely the influence of users increase on the authentication accuracy was studied.

Biometric authentication system with improved characteristics based on usage of the models and methods designed by the author was implemented in the *fourth chapter*. Among others, to show the implementation possibility of the biometric system with ECG-authentication on the devices with limited calculation resources the biometric system based on Raspberry Pi 3B was implemented. The speed of the implemented biometric authentication system was verified, and based on the obtained results the recommendations for implementation of the biometric authentication systems on the stand-alone PC as well as on the Raspberry Pi microcomputer were formed.

Long term (years, months) temporal stability of the ECG-signals was studied, as well as the influence of the variability of the informative features from electrocardiogram on the authentication accuracy was defined. Research results certified that the ECG was a stable marker and might be applied in real authentication systems, at this the system was able to adequately recognize the users during the long time without the necessity of the intermediate calibration. Therefore, the high potential and prospects of the electrocardiogram as a biometric marker for the design of the reliable authentication systems are proved.

The list of possible application spheres and the description of the applied usage for biometric authentication systems based on the ECG-signal are provided.

**Keywords:** biometric authentication, electrocardiogram, digital signal processing, machine learning, neural networks, autoencoders, anomaly correction, temporal normalization, Raspberry Pi.