63rd FITCE International Congress
26 - 28 September 2024, Kraków, Poland

Privacy-preserving Framework for Automated Detection of Arrhythmias from ECG Data

Kacper Gil[1]; Andres Vejar[2]

Institute of Telecommunications, AGH University of Krakow, Kraków, Poland

[1]kagil@student.agh.edu.pl, [2]avejar@agh.edu.pl,

Abstract

The integration of machine learning in biomedical engineering applications is crucial to ensure user data security and privacy. This work explores anonymization and differential privacy frameworks to reduce the risk of biometric identification. The differential privacy method is explored to filter biosignal data without compromising diagnostic trustworthiness. The proposed approach for privacy-preserving arrhythmia detection uses a machine learning diagnostic system that reduces discrepancies between prepossessed and raw data, maintaining high diagnostic precision while improving privacy. The application is evaluated using a control model to analyze the accuracy difference when using privacy-preserving input data.

Introduction

Automated diagnostic systems allow reducing the load on health facilities and contribute to improving the quality of care at home. Such systems require tracking of several biosignals to monitor the health status of patients. It is important to consider privacy-enhanced methods in the diagnostic system, given that these signals, like electroencephalogram (EEG) and or electrocardiogram (ECG) can reveal the identities of patients using biometric identification methods. An ideal feature of an automated diagnostic system is the ability to ensure privacy by design[1, 2], where privacy should be built into technology. Important elements to consider are the minimization of the user data, the controllability of personal data, the transparency about the system operation, the control on which authorized entities can have data access and also to secure the segregation of the data. In practice, many privacy-enhancing techniques (PET) are available in the literature. Jordan et al. [3], specified three general categories: (1) Algorithmic PETs like homomorphic encryption, differential privacy, and zero-knowledge proofs. (2) Architectural PETs, such as federated learning and multi-party computation. (3) Augmentation PETs for example in synthetic data and digital twinning. In this work, we focus on Algorithmic PET via differential privacy methods.

Biometric Identification

Biometric Identification consists of two phases, an enrollment phase and a verification phase. The enrollment is the process of registering a source of biometric data jointly with its associated identification index, with the possibility of including other diverse biometric data, e.g. fingerprints and face image. The data stored are generally processed to obtain a set of features that are characteristic to one person, that is, the biometric template data. The verification phase consists of matching the template data into new data. This phase can be challenging, because biometric data can vary from measurement to measurement. Biometric identification using ECG [4, 5] can be achieved directly or in conjunction with other sources of biometric data.

Arrhytmia Detection

Arrhythmia is a medical condition characterized by an irregular heartbeat, also classified as tachycardia or bradycardia if the heart beats too fast or too slow, respectively. Alternatively, the irregularity can display no

pattern; in such cases it is called fibrillation. Factors of increased risk of arrhythmia include cardiovascular disease, heart surgery, and cardiomyopathy that implies changes in heart structure. Other causes not related to the heart are electrolyte imbalances, medications, and certain stimulants. Personal lifestyle also plays a role in the incidence rate of heart irregularities. High levels of stress, smoking, and physical exertion are the most common. Generally, arrhythmias manifest only as palpitations, light dizziness, and shortness of breath. However, in more severe cases it can lead to fainting and can even be life-threatening. The diagnosis procedure involves the use of an ECG, usually taken over a period of 24-48, with the help of a Holter monitor. Several arrhytmia detection methods [6, 7, 8], can be found in the literature, where the physionet ECG dataset is an important benchmark for machine learning methods [9]. More advanced data sets are also available, for example, the 12 leads ECG data set [10].

System Description

In Figure 1 the diagram of the proposed approach to preserving the privacy of the arrhythmia detection system is presented. This research examines a machine learning diagnostic system in which raw ECG biosignals (x) undergo client-side pre-processing to become a filtered signal (u). Subsequently, this signal is utilized by the diagnostic system (g) at the diagnostic center. The goal of this system is to reduce the discrepancy between the results of the preprocessed (g) and a raw data classifier (f), $(f(x) \approx g(u))$, thus maintaining high diagnostic precision while improving privacy. The application is tested with the control model f that is not privacy preserving, to compare the accuracy level of the arrythmia detection.

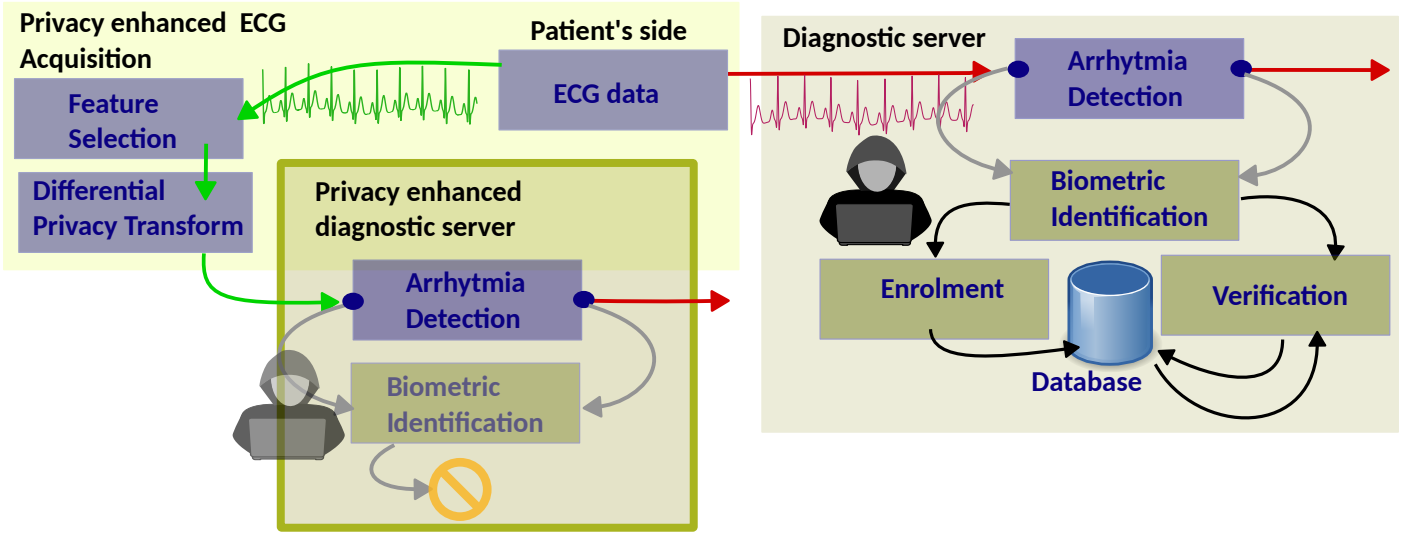

Figure 1: System diagram considering raw and privacy-enhanced arrhytmia detection. In the left side is presented the proposed privacy enhanced arrhytmia diagnostic system. A compromised, raw ECG arrhytmia diagnostic system is depicted in the right side of the diagram.

In this application we consider two stages of privacy enhanced ECG acquisition on the patient's side. The aim of these stages is to gather the portion of user data that is relevant for the machine learning classifier. The first stage *Feature Selection* consist on the selection of the most useful temporal characteristics, that are encoded into a signal u_c. The second stage *Differential Privacy Transform* considers a controlled transformation $u = T(u_c) + \eta$, where T is the transformation function and $\eta \sim N(0, \sigma^2)$ is the noise incorporated to achieve the differential privacy goals of the machine learning model.

Conclusions and Further Work

In this work, a privacy-preserving framework for the detection of arrhytmia is presented. The framework considers a *Privacy enhanced ECG Acquisition* on the patient's side, useful for remote diagnostic in homecare, and a *Privacy enhanced diagnostic server* that provides the automated diagnostic service. The ongoing work considers a validation of the results with standard ECG biosignal databases. One of the objectives of this work is to promote the application of privacy-enhancing technologies in the early stages of automated diagnostic systems. Further work will consider the design of automated diagnostic systems with the joint goal of security

and privacy. In addition, the use of a large set of sensors (e.g. temperature, pulse oxymetry, EEG, and EMG) and different target pathologies for detection can be explored as an extension of the proposed approach.

Acknowledgments

This research was supported by the National Research Institute, grant number POIR.04.02.00-00-D008/20-01, on "National Laboratory for Advanced 5G Research" (acronym PL-5G) as part of the Measure 4.2 Development of modern research infrastructure of the science sector 2014–2020 financed by the European Regional Development Fund.

References

[1] Peter Schaar. "Privacy by Design". In: *Identity in the Information Society* 3.2 (Apr. 2010), pp. 267–274. ISSN: 1876-0678. DOI: 10.1007/s12394-010-0055-x. URL: http://dx.doi.org/10.1007/s12394-010-0055-x.

[2] Anders Nordgren. "Privacy by Design in Personal Health Monitoring". In: *Health Care Analysis* 23.2 (Aug. 2013), pp. 148–164. ISSN: 1573-3394. DOI: 10.1007/s10728-013-0262-3. URL: http://dx.doi.org/10.1007/s10728-013-0262-3.

[3] Sara Jordan, Clara Fontaine, and Rachele Hendricks-Sturrup. "Selecting Privacy-Enhancing Technologies for Managing Health Data Use". In: *Frontiers in Public Health* 10 (2022). ISSN: 2296-2565. DOI: 10.3389/fpubh.2022.814163. URL: https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.814163.

[4] Adrian D. C. Chan, Mohyledin M. Hamdy, Armin Badre, and Vesal Badee. "Person Identification using Electrocardiograms". In: *2006 Canadian Conference on Electrical and Computer Engineering*. 2006, pp. 1–4. DOI: 10.1109/CCECE.2006.277291.

[5] Jianbo Xu, Tianhui Li, Ying Chen, and Wenxi Chen. "Personal Identification by Convolutional Neural Network with ECG Signal". In: *2018 International Conference on Information and Communication Technology Convergence (ICTC)*. 2018, pp. 559–563. DOI: 10.1109/ICTC.2018.8539632.

[6] Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network". In: *Nature Medicine* 25.1 (2019), p. 65.

[7] Runchuan Li, Xingjin Zhang, Honghua Dai, Bing Zhou, and Zongmin Wang. "Interpretability Analysis of Heartbeat Classification Based on Heartbeat Activity's Global Sequence Features and BiLSTM-Attention Neural Network". In: *IEEE Access* 7 (2019), pp. 109870–109883. DOI: 10.1109/ACCESS.2019.2933473.

[8] Sajad Mousavi and Fatemeh Afghah. "Inter-and intra-patient ECG heartbeat classification for arrhythmia detection: a sequence to sequence deep learning approach". In: *arXiv preprint arXiv:1812.07421* (2018).

[9] Gari Clifford, Chengyu Liu, Benjamin Moody, Li-wei Lehman, Ikaro Silva, Qiao Li, Alistair Johnson, and Roger Mark. "AF Classification from a Short Single Lead ECG Recording: the Physionet Computing in Cardiology Challenge 2017". In: *2017 Computing in Cardiology Conference (CinC)*. CinC2017. Computing in Cardiology, Sept. 2017. DOI: 10.22489/cinc.2017.065-469. URL: http://dx.doi.org/10.22489/CinC.2017.065-469.

[10] Erick A Perez Alday et al. "Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020". In: *Physiological Measurement* 41.12 (Dec. 2020), p. 124003. DOI: 10.1088/1361-6579/abc960. URL: https://dx.doi.org/10.1088/1361-6579/abc960.