Risk Assessment of Atrial Fibrillation: a Failure Prediction Approach

Jelena Milosevic[1], Andreas Dittrich[1], Alberto Ferrante[1], Miroslaw Malek[1],
Camilo Rojas Quiros[2], Rubén Braojos[2], Giovanni Ansaloni[2], David Atienza[2]

[1]ALaRI, Faculty of Informatics, Università della Svizzera italiana, Lugano, Switzerland
[2]Embedded Systems Laboratory, École Polytechnique Fédérale de Lausanne, Switzerland

Abstract

We present a methodology for identifying patients who have experienced Paroxysmal Atrial Fibrillation (PAF) among a given subject population. Our work is intended as an initial step towards the design of an unobtrusive portable system for concurrent detection and monitoring of chronic cardiac conditions.

The methodology comprises two stages: off-line training and on-line analysis. During training the most significant features are selected using machine learning methods, without relying on a manual selection based on previous knowledge. Analysis is done in two phases: feature extraction and detection of PAF patients. Light-weight algorithms are employed in the feature extraction phase, allowing the on-line implementation of this step on wearable sensor nodes. The detection phase employs techniques borrowed from the field of failure prediction. While these algorithms have found extensive application in diverse scenarios, their application to automated cardiac analysis has not been sufficiently investigated to date.

The proposed methodology is able to correctly classify 68% of the test records in the PAF Prediction Challenge database, performing comparably to state of the art offline algorithms. Nonetheless, the proposed method employs embedded signal processing for the critical feature extraction step, which is executed on resource-constrained body sensor nodes. This allows for a real-time and energy-efficient implementation.

1. Introduction

Wireless Body Sensor Nodes (WBSNs) are low-power devices able to capture bio-signals, such as electrocardiograms (ECGs), for an extended period of time. Data recorded by these devices is of paramount clinical importance in the assessment of numerous heart-related conditions. Among them, the prediction of Paroxysmal Atrial Fibrillation (PAF) episodes and the risk stratification of PAF-prone patients is particularly challenging, as it requires an analysis *before* the episodes to allow timely medical assistance. Real-world solutions must also take into account the constrained resources (e.g. battery, computational power, memory) of wearable devices, striving for optimized light-weight algorithms.

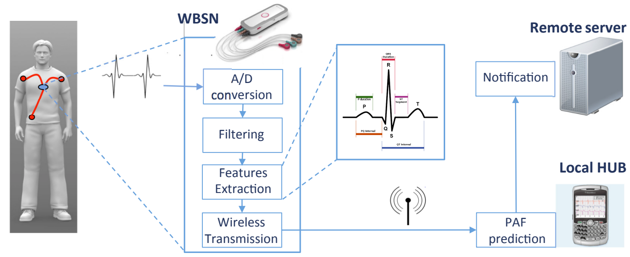

Figure 1. Application scenario: a WBSN performs ECG feature extraction, while a local hub predicts/detects dangerous cardiac episodes.

In this paper, we describe a methodology for identification of PAF patients and discuss experimental results. The considered scenario is shown in Figure 1. The wearable device monitors the ECG of a patient, filters the acquisitions and extracts relevant features. These features are further elaborated on by a local hub (e.g. a smartphone) which performs a predictive assessment of the PAF risk of the subject. Results are then transmitted to a remote server and/or displayed locally for medical evaluation. The illustrated distributed system minimizes the transmission bandwidth from the WBSN to the local hub because only relevant features are transmitted over the power-hungry transmission link, as opposed to the full set of acquired data [1]. The illustrated distributed environment requires carefully tailored algorithms to perform the different application phases. While effective methodologies for PAF prediction are proposed in the literature, they either employ computationally complex algorithms to derive classification features (bispectrum and non-linear features in [2]) or a variety of sensors to acquire multiple bio-signals in addition to ECG (e.g., blood pressure and pulse oximetry in [3]), thus negatively impacting the wearability of

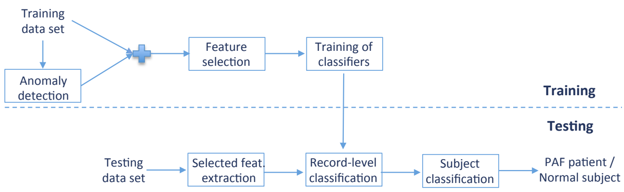

Figure 2. Methodology for designing the classification.

the resulting systems. Our framework instead only considers the time-domain characteristics of each heartbeat, extracted from ECG acquisitions. These features are collected using the methodology proposed in [4], which can be executed on low-power platforms with reduced memory (tens of KBytes) and clock speed (few MHz).

The remaining part of this paper is organized as follows: Section 2 explains the proposed methodology. Section 3 discusses obtained results while Section 4 concludes the paper.

2. Detection Methodology

As shown in Figure 2, the design of classification algorithms for PAF prediction is divided in a training and a testing phase. Both phases are performed off-line on different sets of recorded data. The same classification algorithms used in testing are also used in the final on-line application, which execute within the computational constraints imposed by the application scenario described in Figure 1.

Four main steps are performed during the training phase: feature extraction, system observation, feature selection and model estimation [5]. All time-domain parameters derived from the feature extraction step (e.g., the peak and boundary of the characteristic waves composing each heartbeat) are evaluated in the system observation phase to discover correlations among them. A subset of features is then selected based on their significance as predictors of PAF. Finally, during model estimation different classification algorithms and selected features are employed to discriminate ECG records of PAF and non-PAF subjects. The resulting trained predictor is then employed during the test phase.

In the remaining part of this section, we explain in detail how the four steps introduced above are performed.

2.1. Feature Extraction

Data used in our experiments are taken from the *PAF Prediction Challenge* database [6]. Half of the records present in the database are acquired from patients prone to PAF (either immediately before fibrillation or at a time distant from a fibrillation), the other half from healthy subjects. No records of ongoing PAFs are considered, as we aim at discerning PAF-prone patients before the occurrence of a fibrillation. The database provides a training set, consisting of 50 records of PAF patients and 50 records of healthy subjects and a test set of 50 records. Excerpts are 30 minutes long, concurrently acquired from two leads. The challenge considered in this paper is to distinguish between subjects that experienced PAF in the past (*group A*) and others that did not (*group N*).

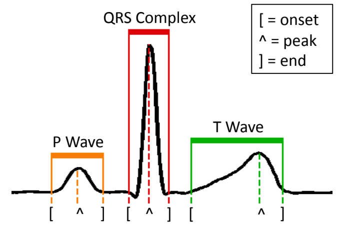

Figure 3. Fiducial points of a ECG heartbeat.

Before feature extraction, the database is pre-processed using the morphological filtering technique described in [4]. Morphological filtering allows to retain useful information from acquired signals while effectively eliminating noise originating from multiple sources such as low-frequency baseline wandering (due to respiration and perspiration) and higher-frequency components (due to muscular activity). The implementation is optimized for the execution on resource-constrained embedded devices, only requiring 4KB of memory and 15% of the duty cycle of a state of the art WBSN running at 6MHz [4]. A further step to enhance the quality of the acquired data is to combine signals from different inputs (leads) before the feature extraction step. In this study, we employ a Root Mean Square (RMS) combination of the two signals provided in the PAF-prediction database.

An initial set of features is then derived from the recordings focusing on time-domain characteristics, which are calculated on-node with little computational effort. The employed algorithm, based on the Digital Wavelet Transform (DWT) [4], retrieves the interval between two successive heartbeats (R-R interval) and the fiducial points of each heartbeat: the start, peak and end of each characteristic wave (Figure 3). Similarly to the filtering step, DWT delineation is also executed directly on WBSNs, requiring 18KB of memory and 5% of the duty cycle of the target platform [4]. Delineation is performed both on the two individual signals provided by the PAF prediction database and on their RMS combination. Only delineated points are transmitted to the local hub, greatly reducing the transmission bandwidth over the energy-hungry wireless link.

The presence and position of fiducial points drives the construction of a comprehensive feature set. For each heartbeat, we consider the presence or absence of the P and T waves. We also include in the feature set the duration of each wave and the inter-wave intervals, as well as the distance between two consecutive heartbeats. Furthermore, we consider the amplitude of each wave, in terms of area of the waves and height of their peaks. The feature set is further enriched by aggregating the retrieved data at different resolutions. For each of the parameters described above, we calculate their mean and standard deviation, both from the start of the record and considering sliding windows of 5, 50, 150, 200 and 250 heart-beats. Moreover, we annotate the pNN20 and pNN50 metrics, which are the proportions of successive heart-beat intervals differing more than 20 and 50 ms with respect to all detected heartbeats.

2.2. System Observation, Feature Selection and Model Estimation

After feature extraction, the features described in Section 2.1 are observed and analyzed, with the goal of discovering the most significant ones for discerning PAF and non-PAF subjects. Limiting the number of features allows not only to lower the computational complexity of the detection/prediction algorithms, but in many cases also increases their accuracy by filtering out misleading parameters. A classifier on single records is used, as shown in Figure 2. In the following *Subject Classification*, subjects are classified as PAF patients (*group A*) or not (*group N*) depending on the percentage of anomalous records.

Classifiers require a training phase. In our case, as a first step of this process we identify records in the training set that have peculiar characteristics, which can be connected to PAF patients. To mark these records for training, we use an algorithm for anomaly detection named Local Outlier Factor (LOF) [7]. In contrast to other anomaly detection techniques that classify records just as outliers or not, LOF gives an "outlier score" for each record, measuring its local deviation with respect to its neighbors. LOF scores are used to mark characteristic anomalous records of PAF patients. All the other records are marked as regular. The classifiers are then trained to discern between annotated records.

Different methods are considered to reduce the set of significant features used in the classifiers: Correlation Attribute Evaluation, Gain Ration Attribute Evaluator and Info Gain Attribute Evaluator [8]. The selected features extracted from the training set are then employed to train different classification algorithms (J48, Multilayer Perceptron, JRip, Logistic Regression, Naive Bayes, OneR, PART, Random Forest, REPTree and SVM) [8] for a comparative evaluation. Once training is completed, these classifiers are used to classify anomalous records in a different data set. By counting the percentage of anomalous records, subjects can be classified as belonging to group A or N. The considered threshold is decided empirically during the training phase.

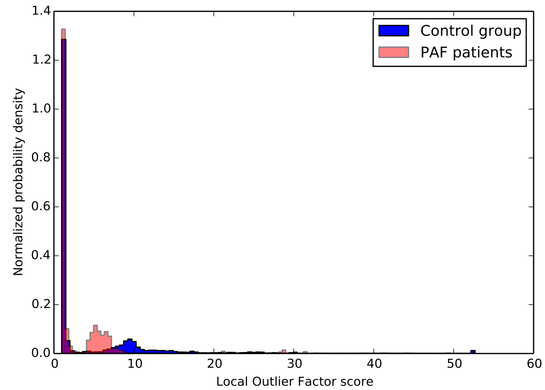

Figure 4. Histogram of anomalies detected in group A (PAF-prone) and group N (healthy) subjects.

Table 1. Most significant features ranked by relevance

Relevance	Feature name
0.568169	Presence/absence of P and T waves
0.243048	Duration of the QRS complex (avg. on 50 heartbeats)
0.241538	Duration of the QRS complex (avg. on 100 heartbeats)
0.240357	Start of QRS complex
0.237312	Duration of P wave (avg. on 200 heartbeats)
0.236797	Duration of P wave (avg. on 150 heartbeats)
0.233854	Duration of QRS complex (avg. on 150 heartbeats)
0.232693	Duration of P wave (avg. on 250 heartbeats)
0.229858	Duration of P wave (avg. on 100 heartbeats)
0.223869	Duration of QRS complex (avg. on 150 heartbeats)

3. Experimental Results

The annotation of training data as well as the training of classifiers, the feature selection and finally, the classification itself is performed using the Weka toolset [9]. Weka supports all previously mentioned algorithms, either directly or by means of external plugins. The LOF algorithm follows the implementation described in [7].

The distribution of obtained LOF scores for healthy subjects and PAF patients is shown in the histogram of Figure 4. It shows the normalized number of records assigned to a specific LOF score. The distribution of scores is different for the two cases. We are interested in the records that are outliers for PAF-prone (group A) subjects. Therefore, based on the observed clusters in Figure 4, we consider the records with an LOF score between 4.1 and 7.3 as abnormal.

Among the feature selection methods mentioned in Section 2.2, the best results are obtained by Correlation Attribute Evaluation. This algorithm measures the correlation between each attribute and the considered class. By means of this method, the features have been ranked by relevance and the ten most relevant ones have been selected

Table 2. Accuracy of the J48 and Multilayer Perceptron Classifiers

	J48	Multilayer Perceptron
Correctly classified subjects	34/50 (68%)	34/50 (68%)
Correctly classified group A subjects	20/28 (71%)	19/28 (68%)
Precision on group N subjects	0.64	0.63
Recall on group N subjects	0.64	0.68
Precision on group A subjects	0.71	0.73
Recall on group A subjects	0.71	0.68

out of the 95 available ones (see Table 1). By changing the number of selected features, the trade-off between accuracy and computational resources required by the classification algorithm can be tuned. The threshold on the number of anomalous records that is required to identify a subject as belonging to group *A* has been set to 2% empirically by testing different thresholds between 1 and 10%: the ones above 2% provide decreased precision, while the ones below 2% provide decreased recall.

The best performing classifiers with respect to the number of correctly classified data are J48 and MultilayerPerceptron. The classification results for these classifiers are shown in Table 2. Our approach has a classification performance of 68% in the two cases considered, which is in line with other state of the art methods [10]. The main advantage of the developed framework lies in the on-node implementation of the feature extraction step, which greatly decreases the energy requirements of the sensor nodes by decreasing its transmission bandwidth, leading to a power-efficient implementation. The proposed system is therefore suitable for unobtrusive long-term monitoring, paving the way for novel and more efficient solutions in e-health monitoring systems and personalized healthcare.

4. Conclusions

We described how a machine learning approach derived from failure prediction methods can be used in the scenario of PAF risk assessment. Our approach has the following benefits: First, it is able to assess the predictive value of multiple complex parameters as PAF markers. Second, the correlation between features can be detected, effectively pruning the feature set by retaining only the most significant ones. Third, the proposed framework is realized on systems based on wireless body sensor nodes, by employing an embedded DSP for morphological ECG filtering and DWT delineation. Experimental evidence shows that the accuracy of our on-line approach is in line with off-line state of the art methods.

References

[1] Chen F, Lim F, Abari O, Chandrakasan A, Stojanovic V. Energy-aware design of compressed sensing systems for wireless sensors under performance and reliability constraints. Circuits and Systems I Regular Papers IEEE Transactions on March 2013;60(3):650–661.

[2] Mohebbi M, Ghassemian H. Prediction of paroxysmal atrial fibrillation based on non-linear analysis and spectrum and bispectrum features of the heart rate variability signal. Computer Methods and Programs in Biomedicine 2012;105(1):40 – 49.

[3] Kwon K, Hwang H, Kang H, Woo KG, Shim K. A remote cardiac monitoring system for preventive care. In Consumer Electronics (ICCE), 2013 IEEE International Conference on. Jan 2013; 197–200.

[4] Braojos Lopez R, Ansaloni G, Atienza Alonso D, Vallejos R, Javier F. Embedded Real-Time ECG Delineation Methods: a Comparative Evaluation. In IEEE 12th International Conference on BioInformatics and BioEngineering (BIBE 2012), volume 1. 2012; 99–104.

[5] Salfner F, Lenk M, Malek M. A survey of online failure prediction methods. ACM Computing Surveys March 2010; 42(3):10:1–10:42.

[6] PhysioBank. Webpage, January 2014. URL `http://www.physionet.org/physiobank`.

[7] Breunig MM, Kriegel HP, Ng RT, Sander J. Lof: Identifying density-based local outliers. ACM SIGMOD Record 2000;29(2):93–104.

[8] Witten I, Frank E, Hall M. Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques. The Morgan Kaufmann Series in Data Management Systems. Elsevier Science, 2011.

[9] Weka 3: Data mining software in java. Webpage, January 2014. URL `http://www.cs.waikato.ac.nz/ml/weka`.

[10] Computers in Cardiology Challenge 2001 Top Scores. Webpage, 2001. URL `http://www.physionet.org/challenge/2001/top-scores.shtml`.

Acknowledgments

This work was partially supported and funded by the Hasler Foundation under the Project HFP (no. 13058), EC FP7 FET Phidias project (no. 318013) and ObeSense (no. 20NA21 143081) RTD project evaluated by the Swiss NSF and funded by Nano-Tera.ch with Swiss Confederation financing. The paper reflects only the view of the authors.

Address for correspondence:

Alberto Ferrante
ALaRI, Faculty of Informatics, Università della Svizzera italiana
Via G. Buffi, 13 – 6904 Lugano – Switzerland
Tel.: +41-058-666-4558
Email: alberto.ferrante@usi.ch