A Novel Atrial Fibrillation Prediction Algorithm Applicable to Recordings from Portable Devices

Zhi Li[1], Harm Derksen[2], Jonathan Gryak[1],Hamid Ghanbari[4], Pujitha Gunaratne[6], and Kayvan Najarian[1,3,5]

Abstract— **Atrial Fibrillation (AFib) is by itself a strong risk factor for many life-threatening heart diseases. An estimated 2.7 to 6.1 million people in the United States have AFib. With the aging of the U.S. population, this number is expected to increase. In this preliminary study, a heart rate-duration criteria region is proposed to automatically label symptomatic AFib events using recordings from portable ECG monitors. A Markov Chain algorithm is implemented to classify prediction intervals that are 2 minutes before the symptomatic AFib events. The method yields an overall accuracy value of 82% with 0.91 AUC.**

Keywords - Atrial Fibrillation, Machine Learning, Markov Chain, Portable ECG Device

I. INTRODUCTION

Atrial Fibrillation (AFib) is the most common sustained cardiac arrhythmia. As of 2014, between 2.7 million and 6.1 million American adults are affected by AFib, a figure that is projected to double over the next 25 years [1]. The prevalence of AFib increases with age, from approximately 1% in patients younger than 60 years old to more than one third of the patients aged 80 years or older. In particular, the lifetime risk of developing AFib is approximately 25% in those who have reached the age of 40 [1], [2].

Symptoms of AFib range from nonexistent to severe. Patients with AFib are at an increased risk of frequent hospitalizations, hemodynamic abnormalities, and thromboembolic events, which result in significant morbidity and mortality. AFib is associated with a 5-fold increased risk of stroke and is one of the components in the Framingham stroke risk prediction profile [3]. AFib is also associated with a 3-fold increased risk of heart failure and a 2-fold increased risk of both dementia and mortality [2]. It can thus be suggested that prediction of abnormal heart rates or AFib events may help prevent life-threatening conditions like stroke or heart failure. Until now, most of the existing monitoring systems take place in hospital and require long-term hospitalization. However, in order to monitor heart rates in daily life, more portable devices should be used to streamline the data collection procedure. AFib detection and prediction algorithms based on portable devices would greatly advance clinical decision making and provide better patient care.

This study presents a new approach to annotate ECG signals from portable devices and predict AFib events using signals that are several minutes before the events.

II. DATA

Physionet/CinC 2017 database is used for the AFib detection part of the study [4]. There are 738 annotated AFib episodes and 5360 normal episodes in the Physionet/CinC database. The database used in the AFib prediction part of this study includes 5 participants from a pilot study with persistent AFib. Each patient enrolled in the study has about a month of recorded ECG signal and there are a total of 108 days of continuous ECG signals. The ECG signals were collected using portable ECG monitors (BodyGuardian® Heart, Preventice Solutions). The recordings were digitized at 256 samples per second.

III. METHODS

The method section is divided into four major parts. The first part describes pre-processing and noise detection. The second part describes annotation of the signals. The third part describes signal extraction for the analysis. In the final part the algorithm for signal encoding and the Markov Chain model for classification are outlined.

A. Data Pre-processing

During the signal pre-processing step, a fourth-order Butterworth bandpass filter with cutoff frequencies of 0.5 and 40 Hz is first applied to the raw ECG signal to remove noise, after which a double median filter with orders equal to 0.2 and 0.6 times the sampling frequency is applied to remove baseline wandering. To calculate heart rate and beat duration, the Pan–Tompkins algorithm for QRS detection is used [5] [6].

After R peak detection, a stepwise noise detection method is performed. There are 3 criteria for noise detection. First, it finds the percentage of missing signal in a defined time window (300 seconds), and checks if the missing signal percentage is above a certain threshold (15%). If the percentage of the missing signal is above the threshold, the signal is classified as being noisy by the first criteria. The majority of the signals used in this study will pass the first filter. Secondly, if the signal passes the first criteria then it will check for missing spikes by a non-linear filter analysis. Thirdly, if the signal passes both the first and the second

[1] Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA
[2] Department of Mathematics, University of Michigan, Ann Arbor, MI, USA
[3] Department of Emergency Medicine, University of Michigan, Ann Arbor, MI, USA
[4] Department of Internal Medicine - Cardiology, University of Michigan, Ann Arbor, MI, USA
[5] Michigan Center for Integrative Research in Critical Care University of Michigan, Ann Arbor, MI, USA
[6] Toyota Motor North America, Ann Arbor, MI, USA

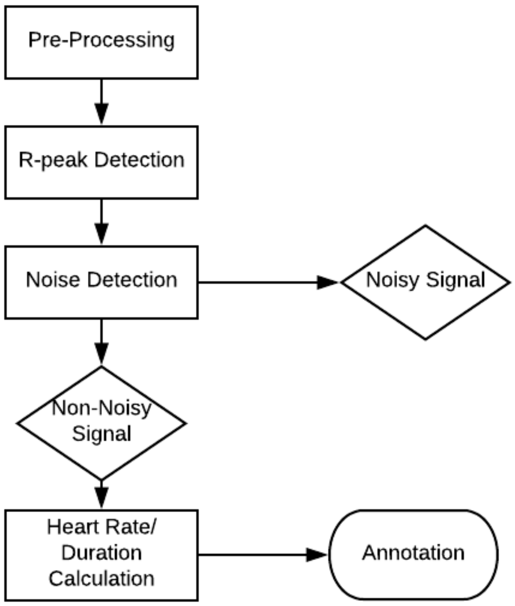

Fig. 1: Pre-Processing and Annotation Steps

criteria, the third criteria will check if the percentage of missing R peaks is above a certain threshold (15%) in the current window. The window of signal will be classified as noisy if it fails any of the three criteria. Noisy signals will not be used for subsequent annotation. Signals that are too close to the noisy signal will also be excluded from the analysis. Further details about signal extraction will be given in Section C.

Heart rate is calculated based on the R peaks detected in the previous step in a time interval. Duration is calculated by counting the number of consecutive intervals that a particular heart rate spans. For now, the time interval is set to 30 seconds and counted up to 6 intervals, which produces a duration ranging from 30 to 180 seconds.

B. Annotation

The increasing importance of a deeper understanding of AFib has lead to intense investigations in recent years. Despite such efforts, there is still no consensus within the various contemporary guidelines on treatment decisions, e.g., for events such as stroke [7]. Such difficulty could be partly traced back to the definition of AFib itself. In [2], AFib is defined to be a cardiac arrhythmia with three characteristics: ‘absolutely’ irregular RR intervals, no distinct P waves, and variable atrial cycles of length <200 ms (>300 bpm). However, the qualitative nature of these characteristics implies that AFib detection and diagnosis relies heavily on human input, which impedes the development of automated treatment decision making algorithms based on AFib data. In order to improve the situation, a heart rate-duration criteria line is proposed to annotate symptomatic AFib events.

Portions of signals with extremely high heart rate y_2 that last for a short time period x_1, and a relatively lower heart rate y_1 that lasts for a longer time period x_2, are both considered to be indicative of symptomatic AFib. Using a straight line passing through the points (x_1,y_2) and (x_2,y_1), the lower bound for heart rates in the criteria region between durations x_1 and x_2 are determined. If any of the heart rates on the duration grid is above the heart rate-duration criteria line, then that point of the signal lies in the criteria region and will be classified as symptomatic. The lower duration limit of 30s is chosen in accordance with the definition of an AFib event provided by the 2014 AHA guideline [1]. The lower heart rate limit is set to 110 bpm, as this rate has proven to be an effective treatment target for AFib [8]. The higher heart rate limit is set at 160 bpm since a rapid heart rate is more likely to cause symptoms. Rapid heart rate in AFib may also have an untoward effect on cardiac function, resulting in tachycardia-induced cardiomyopathy [9]. Using this annotation method it becomes possible to capture the portions of the signal that correspond to symptomatic AFib with sufficiently high severity, either in the form of extremely high heart rate over a span of 30s, or moderately high heart rate stretching over 180s. For this preliminary study, a heart rate-duration criteria region that lies above the line passing through the points $(30, 160)$ and $(180, 110)$ was used. A total number of 279 events are labeled by this criteria.

C. Pre-event Signal Extraction

The aim of this study to predict the onset of a symptomatic AFib event using pre-event signals. In order to address the predicative power of our method, the events that happen within 8 minutes after a previous event will be excluded, because the prediction interval should not be too close to the end of the previous signal. The events that happen within 8 minutes of a noisy signal will also be excluded, with the aim to ensure prediction interval is out of the noisy signal range. This leaves 87 AFib events for analysis.

After excluding the AFib events and noisy signals (defined in section III.A), a period of pre-AFib signal that is 2 minutes long and happens at least 2 minutes before each AFib event is extracted, resulting in a total number of 87 2-minute long pre-AFib prediction intervals that happen at least 2 minutes before the annotated symptomatic events. Figure 2 shows an example of the classified ECG signal. After annotation, there are 3 AFib events: A, B and C, and a noisy signal D. Event A is excluded because it is too close to noise D, event B is excluded because it happens right after event A. Event C is what remains. Then the portion E of the signal, which is 2 minutes long with a 2-minute gap before event C, is extracted.

D. Signal Encoding and Markov Chain

A Markov Chain algorithm for classifying AFib prediction intervals has been developed. The input of the Markov Chain algorithm is a sequence of probability vectors

$$\begin{bmatrix} p_1 \\ q_1 \end{bmatrix}, \begin{bmatrix} p_2 \\ q_2 \end{bmatrix}, \begin{bmatrix} p_3 \\ q_3 \end{bmatrix}, \begin{bmatrix} p_4 \\ q_4 \end{bmatrix}, \ \ldots$$

for peak (p_i) vs. non-peak (q_i). To obtain the probability vectors, an additional round of signal processing techniques is applied. The various techniques involve:

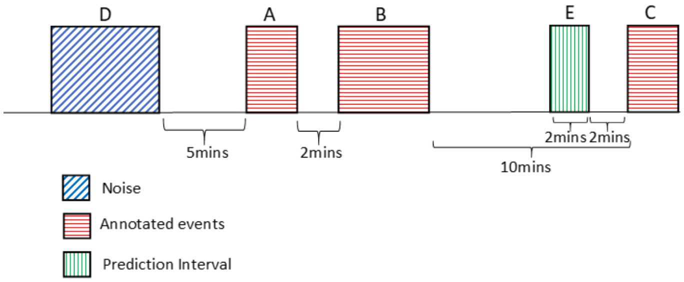

Fig. 2: Prediction Interval Extraction

1) *Baseline removal* The baseline removal step will remove baseline drift from the ECG signal and enhance the R peaks in comparison to the T and P waves.
2) *Normalization* The ECG signal is normalized to have value between 0 and 1.
3) *Non-linear filtering* Apply a non-linear filter to enhance R-peaks and suppress other peaks.
4) *Re-sampling of data* The signal is discretized and then re-sampled due to dimensionality considerations.
5) *Thresholding* A thresholding procedure is applied to systematically to convert the signal into probability vectors.

Previous studies have shown that Markov Chain models are able to detect AFib events [10]. Here we propose a Markov Chain algorithm not only for AFib detection but also for the prediction of symptomatic AFib events. The ability of the algorithm for detecting AFib events has been tested using the annotated PhysioNet/CinC 2017 data.

The next step is to test the algorithms' ability on the prediction of symptomatic events using real-life data rather than detection. The Markov Chain algorithm is performed on the training data to get the transition matrices which are then applied on the testing data for prediction.

IV. Result

The result section begins by showing the algorithms' ability on detection of annotated AFib episodes using Physionet/CinC 2017 data. In the second part of the result section the results for AFib prediction using the real life ECG signals recorded using portable devices are outlined.

A. AFib Detection

There are 738 annotated AFib episodes and 5360 normal episodes in the Physionet/CinC 2017 database. Each signal is a 30-60s long ECG. The data is split into training and testing sets. The training set has 643 AFib episodes and 4455 normal episodes, the testing set has 95 AFib episodes and 905 normal episodes. The algorithm has correctly classified 80 out of 95 AFib episodes with sensitivity of 84.2% and 844 out of 905 normal episodes with specificity of 93.3% in the testing set. The F1 score is 0.7447 and the area under curve (AUC) is 0.9542.

B. AFib Prediction

After examining the ability of the algorithms for detecting AFib episodes, their performance on predicting AFib events

TABLE I: Confusion Matrix for AFib Detection

		Annotated Label		
		AFib	Normal	Total
Prediction	AFib	80	61	141
	Normal	15	844	859
	Total	95	905	1000

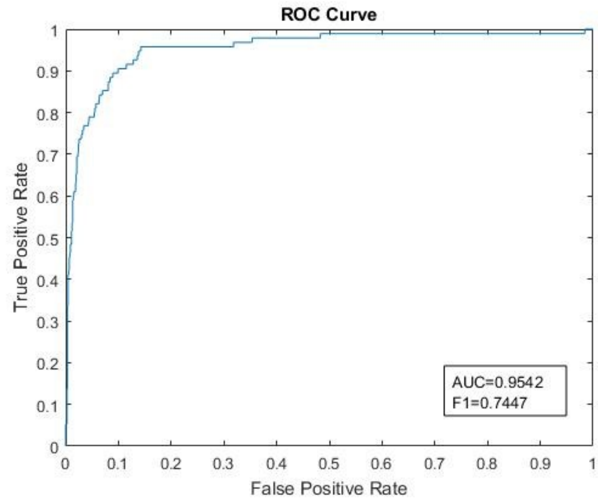

Fig. 3: ROC Detection Physionet/CinC 2017

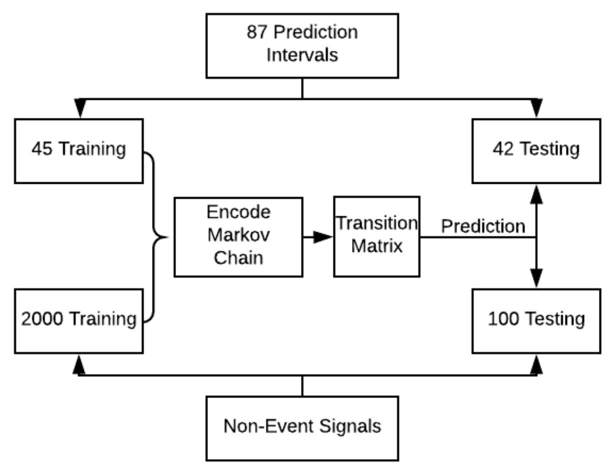

Fig. 4: Training and Testing Datasets

on real-life ECG data recorded by portable devices is also tested. By the methods described in the method section, 87 2-minute pre-event intervals for symptomatic event prediction have been extracted. In addition to this, other 52968 2-minute no-event intervals are also extracted, which are outside the AFib and noisy signal range. For the training part, 45 2-minute pre-event intervals and 2000 2-minute non-AFib intervals are randomly selected. The encoding and Markov Chain algorithms are applied on the training data to obtain the transition matrices. The transition matrices are then applied to the testing set to obtain the predictions. The process is summarized in Figure 4. Using the training data, the algorithm generates 782 transition states for pre-AFib

TABLE II: Confusion Matrix for AFib Prediction

		Annotation		
		Pre-AFib	Non-AFib	Total
Prediction	Pre-AFib	36	20	56
	Non-AFib	6	80	86
	Total	42	100	142

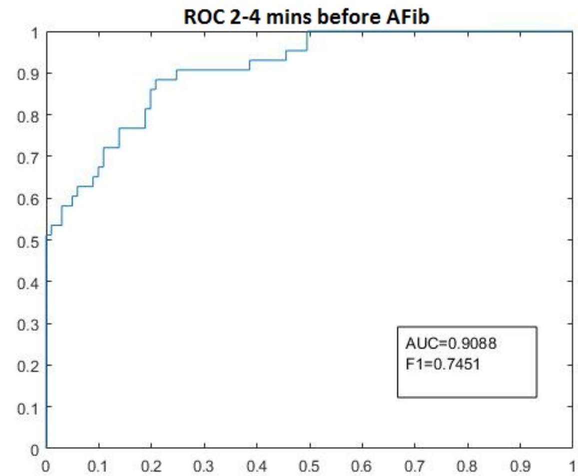

Fig. 5: ROC Prediction

intervals and 63273 transition states for non-AFib events. Applying these transition states to the testing data, 36 out of 42 pre-AFib intervals have been correctly identified with a sensitivity of 86% and 80 out of 100 non-AFib intervals, a specificity of 80%. The overall accuracy is 82%. The F1 score is 0.7451 with an AUC of 0.9088.

V. Discussion

In this study, a heart rate-duration criteria region is proposed to classify symptomatic AFib. AFib is irregular and often characterized with high heart rate. The duration of an AFib episode can be short or long and there is to date no consensus on the definition of an AFib based on heart rate and duration. This annotation method is based on a straight line passing through two points, one point with relative low heart rate that lasts for a long period of time and one point with high heart rate that lasts for a short period of time. By using this heart-rate duration criteria region above this straight line, it is possible to capture AFib events that behave in various forms instead of a single heart rate or duration cut-off. However, this preliminary study has not demonstrated the correlation between our criteria region and symptomatic data. The criteria region used in this preliminary study may not be the optimal one for classifying AFib symptoms. For the next step of the project, more symptomatic data will be gathered and the criteria region will be examined based on the classification performance of symptomatic events. Using the correlation between different criteria regions and symptomatic events, it should be possible to identify an optimal criteria region that identifies AFib symptoms.

A total of 108 days of continuous ECG signal from 5 participants are used in this study. Although the sample size of the total ECG signals is adequate, the signals are obtained from a small number of participants. The variations in the signals might be limited. For future study, more participants will be enrolled, and thus will enable the ability of algorithms in prediction of symptomatic AFib events to be examined with more variations.

VI. Conclusion

AFib is a common type of arrhythmia that by itself is a strong independent risk factor for many life-threatening diseases. This preliminary study proposes a heart rate-duration criteria region to annotate the symptomatic AFib events and use pre-processing, encoding techniques and a Markov Chain algorithm to predict ECG signals which occur 2 minutes prior to the symptomatic events. This preliminary result has an overall accuracy of 82% for prediction with 0.91 AUC. The method using real-life ECG signals from portable devices has proven to be robust in prediction and can advance clinical decisions.

Acknowledgement

The research work in this paper was funded by Toyota Motor North America.

References

[1] Craig T January, L Samuel Wann, Joseph S Alpert, Hugh Calkins, Joaquin E Cigarroa, Jamie B Conti, Patrick T Ellinor, Michael D Ezekowitz, Michael E Field, Katherine T Murray, et al. 2014 aha/acc/hrs guideline for the management of patients with atrial fibrillation: a report of the american college of cardiology/american heart association task force on practice guidelines and the heart rhythm society. *Journal of the American College of Cardiology*, 64(21):e1–e76, 2014.

[2] Developed with the special contribution of the European Heart Rhythm Association (EHRA), Endorsed by the European Association for Cardio-Thoracic Surgery (EACTS), Authors/Task Force Members, A John Camm, Paulus Kirchhof, Gregory YH Lip, Ulrich Schotten, Irene Savelieva, Sabine Ernst, Isabelle C Van Gelder, et al. Guidelines for the management of atrial fibrillation: the task force for the management of atrial fibrillation of the european society of cardiology (esc). *European heart journal*, 31(19):2369–2429, 2010.

[3] Philip A Wolf, Ralph B D'agostino, Albert J Belanger, and William B Kannel. Probability of stroke: a risk profile from the framingham study. *Stroke*, 22(3):312–318, 1991.

[4] Gari Clifford, Chengyu Liu, Benjamin Moody, L Lehman, Ikaro Silva, Qiao Li, A Johnson, and R Mark. Af classification from a short single lead ecg recording: The physionet computing in cardiology challenge 2017. *Computing in Cardiology*, 44, 2017.

[5] Jiapu Pan and Willis J Tompkins. A real-time qrs detection algorithm. *IEEE transactions on biomedical engineering*, (3):230–236, 1985.

[6] Hooman Sedghamiz. Matlab implementation of pan tompkins ecg qrs detector. 03 2014.

[7] Peter Brønnum Nielsen, Torben Bjerregaard Larsen, Flemming Skjøth, Thure Filskov Overvad, and Gregory YH Lip. Stroke and thromboembolic event rates in atrial fibrillation according to different guideline treatment thresholds: a nationwide cohort study. *Scientific reports*, 6:27410, 2016.

[8] Isabelle C Van Gelder, Hessel F Groenveld, Harry JGM Crijns, Ype S Tuininga, Jan GP Tijssen, A Marco Alings, Hans L Hillege, Johanna A Bergsma-Kadijk, Jan H Cornel, Otto Kamp, et al. Lenient versus strict rate control in patients with atrial fibrillation. *New England Journal of Medicine*, 362(15):1363–1373, 2010.

[9] Valentin Fuster, Lars E Rydén, David S Cannom, Harry J Crijns, Anne B Curtis, Kenneth A Ellenbogen, Jonathan L Halperin, Jean-Yves Le Heuzey, G Neal Kay, James E Lowe, et al. Acc/aha/esc 2006 guidelines for the management of patients with atrial fibrillation. *Circulation*, 114(7):e257–e354, 2006.

[10] Rodrigo Varejão Andreão, Bernadette Dorizzi, and Jérôme Boudy. Ecg signal analysis through hidden markov models. *IEEE Transactions on Biomedical engineering*, 53(8):1541–1549, 2006.