Predicting Episodes of Atrial Fibrillation using RR-Intervals and Ectopic Beats

Dilranjan S. Wickramasuriya, Calvin A. Perumalla, and Richard D. Gitlin, *Life Fellow, IEEE*

***Abstract*— Atrial fibrillation (AF) is the most common cardiac arrhythmia and is associated with an increased risk of stroke and mortality. Research into the prediction of AF onset has been motivated by the necessity to develop better pacing therapy to reduce the incidence of AF and maintain the heart's normal sinus rhythm. In this paper, we address a similar problem to that posed in the Computers in Cardiology Challenge 2001, but develop patient specific models to distinguish between epochs of ECG located far away from AF rhythms and those located just prior to the onset of those episodes. Our approach is validated using a publicly available dataset.**

I. Introduction

Atrial fibrillation (AF) is a condition where the heart is unable to contract effectively and is the most common type of cardiac arrhythmia. It is associated with an increased risk of embolic stroke and affects more than 2.2 million people in the United States [1][2]. AF may be classified as paroxysmal, persistent or permanent based on the duration of the fibrillatory rhythms. In the paroxysmal case, intermittent AF episodes that terminate on their own occur and put patients at risk of being undiagnosed with the arrhythmia.

The Computers in Cardiology Challenge 2001 was created to encourage research into predicting the onset of AF [3]. One of the underlying clinical objectives was to help develop technology that could possibly stabilize a patient's normal sinus rhythm and prevent the onset of an AF episode with different pacing mechanisms. The competition comprised of two parts. The first part of the challenge was to diagnose the incidence of AF by developing classification techniques to distinguish between 30-minute ECGs of patients who had the arrhythmia and those who did not. The second part was to predict the onset of paroxysmal AF. Given pairs of 30-minute ECGs recorded from the same patient, the problem was to identify which of the two occurred just prior to the commencement AF and which one was from further away. The challenge has since been revisited by others in order to improve the accuracy of some of the original entries submitted to the contest.

Our research group is developing a novel vectorcardiogram monitoring device [4] for long-term signal acquisition. The device has the ability to compensate for errors induced when placed incorrectly on a patient's chest and will feature wireless connectivity to a user's smartphone. Accumulating continuous heart signal recordings will better equip healthcare providers to monitor patients remotely, help them avoid unnecessary hospital visits and also opens up the possibility of developing algorithms tailored to each patient (i.e., personalized) to predict impending cardiac events. For instance, Lindsberg *et al.* [5] point out the necessity of long-term monitoring for diagnosing infrequent paroxysmal AF, and the need for automated event detection especially if episodes are asymptomatic. In this paper, we present a patient specific approach to predict the occurrence of impending episodes of AF, similar to the problem posed in the second part of the Computers in Cardiology Challenge 2001, and validate it using a publicly available dataset.

The authors are with the Department of Electrical Engineering, University of South Florida, Tampa FL 33620, USA. (e-mail: {dilranjanw, calvin4}@mail.usf.edu, richgitlin@usf.edu)

II. Literature Review

In this section we review some of the literature describing entries to the Computers in Cardiology Challenge 2001 to predict the onset of AF. There were 28 pairs of ECGs in the test set and scores were calculated as a percentage of the number of correctly classified pairs. Lynn and Chiang [6] used state-space domain features of heart rate variability (HRV). They classified return and difference maps using k-nearest neighbors and scored 64% in predicting AF onset. Zong *et al.* [7] used the number of isolated premature atrial complexes (PACs) appearing in ECGs modified with a weight to favor complexes that had most recently occurred to predict AF episodes. Their rule-based classifier scored 79%. Langley *et al.* [8] extracted features based on the frequency of ectopic beat occurrence in RR-intervals. A moving average of the RR-intervals was maintained and an ectopic beat was flagged if an interval under consideration was less than 20% of the average. They predicted the onset of AF by counting the number of atrial and ventricular ectopic beats and obtained a score of 61%. Schreier *et al.* [9] applied beat classification to the ECG signals and extracted regular and premature heartbeats. They calculated the correlation between a representative template of the p-wave in either group with those extracted from test signals. Diagnosis was performed using a statistical test and they scored 71% in the prediction problem where a weighted measure of the correlation coefficients was used. De Chazal and Heneghan [10] developed a number of features and used Linear Discriminant Analysis for classification. The types of features they used included time and frequency domain features of RR-intervals, features extracted from p-wave amplitudes and the frequency content of p-waves. Their proposed method received 68% on the prediction challenge. Maier *et al.* [11] used features extracted from RR-intervals and the number of ectopic beats and evaluated different classification approaches. Their entries to the contest received a score of 68%.

Thong *et al.* [12] revisited the original challenge and proposed a predictor based on the number of PACs not

978-1-5090-4179-4/17/$31.00 ©2017 IEEE

followed by a regular RR-interval, the occurrence of atrial bigeminy and trigeminy rhythms and the appearance of short durations of paroxysmal atrial tachycardia. Their rule-based system scored 89% outperforming [7] that had scored highest in the 2001 contest. It must be noted that these authors and subsequent researchers had access to additional data that the original competitors did not. More recently, Hilavin and Kuntalp [13] used complex correlation measures of HRV and obtained a score of 72% on the same dataset. Chesnokov *et al.* [14] trained a neural network to distinguish between distant and pre-AF records using the power spectral density of HRV signals and obtained a sensitivity of 72.7% on test data. Arotaritei and Rotariu [15] considered the appearance of successive PACs, the Teager entropy and root mean square of successive differences in RR-intervals and achieved a sensitivity of 52%.

In a study of Holter recordings seeking to determine if paroxysmal AF had a common mode of initiation in all patients, Hnatkova *et al.* [16] concluded that there was no consistent mode of AF onset across a population having the arrhythmia and even within an individual. They claim that "there is marked variation in the rate and rhythm before onset not only between patients, but also between episodes of AF in an individual patient, and no clearly identifiable pattern was seen." Their conclusions caused us to seek to develop patient specific classification models instead of global classifiers that could predict AF onset for everyone.

III. Methodology

A. Data

The Long-Term Atrial Fibrillation Database (LTAFDB) from PhysioBank [17][18] contains 84 ECG recordings from patients with sustained or paroxysmal AF. Each recording is approximately 24 hours in duration and contains two ECG signals recorded at a sampling frequency of 256 Hz. For the purpose of this research we needed to distinguish ECGs that occurred shortly before the commencement of episodes of AF from ECG signals that were sufficiently far away from such abnormal rhythms. We randomly selected 100 ECG epochs, each being 2 minutes in duration, from every record, located at least 10 minutes away from an AF rhythm and labeled this collection of ECGs as the normal or distant set. Thereafter, we selected ECG epochs of identical duration from the same records but which terminate 0, 1, 2 and 3 minutes respectively, prior to the start of each episode of AF and labeled it as the pre-atrial fibrillation (pre-AF) set. If an epoch contains another AF episode it is discarded. We only selected recordings containing at least 20 episodes of AF in order to have enough data for both classes to train patient specific models.

B. Feature Extraction

Several of the original entries in the 2001 contest utilized features based on HRV and we used beat annotations provided in the LTAFDB to extract RR-interval time series for the ECGs. A cubic spline interpolated trend line was subtracted from these series to center them about zero. We manually inspected the RR-interval data to identify good features that were correlated with the target class. We chose eight statistics to characterize variations in heart rate during each 2-minute interval. After normalizing the data by dividing by the maximum absolute value, outliers were defined as the points greater than three standard deviations away from zero. The number of outliers along with their maximum, minimum, mean and median were the first five statistics. The median and root mean square value of the series without the outliers were also extracted. The final statistic was the number of data points in the series that is equivalent to the number of heartbeats during the epoch.

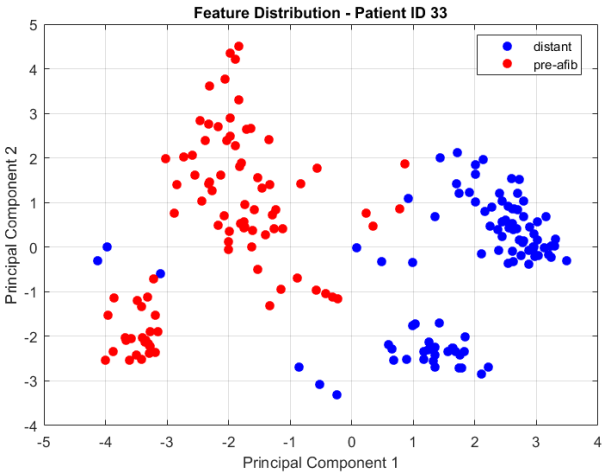

Figure 1. Feature Distribution of distant and pre-AF data points for patient/record 33 for 0-minute prediction horizon

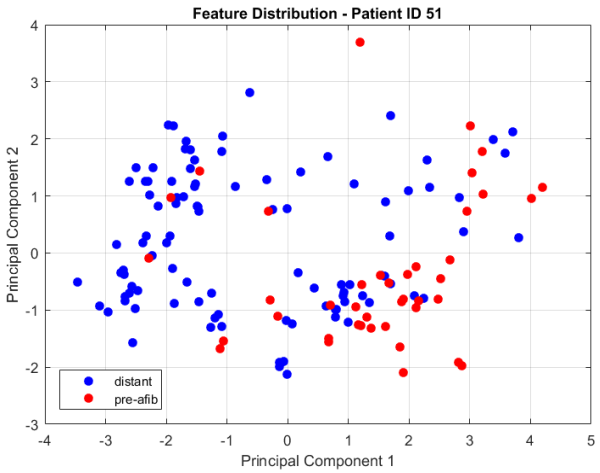

Figure 2. Feature distribution of distant and pre-AF data points for patient/record 51 for 1-minute prediction horizon

Autoregressive (AR) modelling has been used previously in ECG signal classification during AF [19] and we also extracted four AR coefficients to capture the variation within the RR-interval time series.

Annotations accompanying the original ECG signals indicate locations where abnormal beats and abnormal rhythm changes occurred. The most frequently appearing abnormal beat types include PACs and premature ventricular contractions while sinus bradycardia, ventricular tachycardia, atrial bigeminy, supraventricular tachycardia and ventricular bigeminy are among some of the abnormally occurring

TABLE 1. SPECIFICITY (SPEC.), SENSITIVITY (SENS.) AND ACCURACY (ACC.) OF CLASSIFICATION BETWEEN DISTANT AND PRE-ATRIAL FIBRILLATION ECGS FOR EACH PATIENT/RECORD FOR DIFFERENT PREDICTION HORIZONS

Patient/ Record ID	Total AF Episodes	Time Between Termination of ECG Epoch and Next Atrial Fibrillation Episode (Prediction Horizon)											
		0 minutes			*1 minute*			*2 minutes*			*3 minutes*		
		Spec.	*Sens.*	*Acc.*	*Spec.*	*Sens.*	*Acc.*	*Spec.*	*Sens.*	*Acc.*	*Spec.*	*Sens.*	*Acc.*
00	21	0.99	1.00	0.992	0.98	0.647	0.932	0.98	0.643	0.939	0.99	0.75	0.964
01	35	0.96	0.971	0.963	0.92	0.889	0.913	0.94	0.739	0.902	0.93	0.667	0.89
10	31	0.98	1.00	0.985	0.97	0.65	0.917	0.94	0.588	0.889	0.97	0.385	0.903
15	91	0.94	0.989	0.963	0.90	0.727	0.839	0.93	0.714	0.866	1.00	0.00	0.763
23	37	0.94	0.973	0.949	0.90	0.481	0.811	0.94	0.25	0.806	1.00	0.00	0.877
25	69	0.88	0.957	0.911	1.00	0.00	0.645	1.00	0.00	0.694	1.00	0.00	0.719
26	89	0.80	0.966	0.878	0.75	0.739	0.746	0.83	0.627	0.762	0.92	0.429	0.775
32	32	1.00	1.00	1.00	1.00	0.647	0.949	1.00	0.40	0.945	1.00	0.00	0.917
33	90	0.97	1.00	0.984	0.84	0.26	0.588	0.79	0.352	0.608	0.74	0.279	0.544
39	78	0.94	0.987	0.961	0.85	0.403	0.679	0.91	0.212	0.671	0.99	0.091	0.715
42	24	1.00	0.958	0.992	0.94	0.143	0.842	0.92	0.154	0.832	0.98	0.00	0.899
45	32	0.93	0.875	0.917	0.92	0.696	0.878	0.96	0.75	0.938	0.96	0.455	0.91
51	45	0.98	0.956	0.972	0.90	0.561	0.801	0.91	0.432	0.781	0.97	0.441	0.836
53	47	0.87	0.702	0.816	0.93	0.313	0.78	0.99	0.083	0.815	1.00	0.00	0.82
62	38	0.98	0.974	0.978	0.90	0.00	0.652	0.91	0.105	0.688	0.94	0.294	0.776
74	96	0.86	0.927	0.893	0.82	0.50	0.708	0.97	0.025	0.70	1.00	0.00	0.763
100	89	0.98	1.00	0.989	0.90	0.439	0.732	0.99	0.22	0.766	0.99	0.121	0.774
101	90	0.98	1.00	0.989	0.88	0.671	0.79	0.93	0.607	0.814	0.95	0.537	0.83
102	23	1.00	0.957	0.992	0.99	0.00	0.839	-	-	-	-	-	-
105	20	0.98	0.80	0.95	1.00	0.00	0.885	0.97	0.20	0.90	0.99	0.20	0.918
112	146	0.844	0.979	0.938	0.766	0.96	0.884	0.797	0.935	0.872	0.781	0.937	0.858
115	44	1.00	1.00	1.00	0.98	0.25	0.82	0.97	0.05	0.817	0.96	0.00	0.814
119	26	0.98	0.885	0.96	0.95	0.632	0.899	0.95	0.632	0.899	0.95	0.143	0.851
120	21	0.99	0.905	0.975	1.00	0.00	0.833	1.00	0.00	0.847	1.00	0.133	0.887
121	70	0.84	1.00	0.906	0.87	0.298	0.662	1.00	0.00	0.69	1.00	0.00	0.735
204	102	0.81	1.00	0.906	0.89	0.754	0.836	0.86	0.63	0.788	0.90	0.366	0.745
Mean	57.15	0.939	0.952	0.952	0.913	0.448	0.802	0.935	0.374	0.809	0.956	0.249	0.823

rhythms. The number of these individual types of beats and rhythms occurring in each epoch were also incorporated into the features. When concatenated, all of the above attributes form a feature vector with 27 values corresponding to each 2-minute ECG.

B. Feature Selection

A Mann-Whitney ranksum test [20] was used to select the best features that distinguish between the distant class and the subset of ECGs belonging to the pre-AF class which terminate right at the commencement of an episode (i.e., 0 minutes before the onset of fibrillatory rhythm) for each patient. The test indicates whether the particular feature being considered is distributed with different medians for each class. Features for which the p-value was less than 0.05 were selected to perform classification. Then ranksum tests were repeated for the remaining subsets of ECGs terminating 1, 2 and 3 minutes prior to the AF episodes to select the best set of features that distinguish each of them from features belonging to the distant set. In total, four separate classifiers were trained for each patient/record. We label the time between the end of an ECG epoch and the commencement of a fibrillatory rhythm, being either 0, 1, 2 or 3 minutes, as the prediction horizon. Principal Component Analysis (PCA) was performed using the eigenvalues of the feature covariance matrix to project the data onto a lower dimensional subspace for visualization. The Principal Components (PCs) are linear combinations of the features and have no physical interpretation by themselves. Feature distributions for two different patients for different prediction horizons projected onto the first two PCs have been shown in Fig. 1 and Fig. 2.

C. Classification

We evaluated Support Vector Machines (SVMs), Decision Trees and Logistic Regression for classification. We selected SVMs due to their superior accuracy and used the LibSVM package in Weka [21] with 10-fold cross-validation. We calculated sensitivity, specificity and overall accuracy for each of the four classifiers trained per patient. Sensitivity measures the ratio between the number of correctly predicted data points belonging to the pre-AF class and the actual number of data points belonging to that class. Specificity is a similar measure but is calculated for data points belonging to the distant class. The overall accuracy is the percentage of the total number of correct predictions. These values have been shown for each patient/record in Table I.

IV. RESULTS AND DISCUSSION

Sensitivity is above 85% for almost all patients when the SVM performs classification between ECGs that terminate right at an AF episode and ECGs belonging to the distant class. As shown in Fig. 3, sensitivity tends to drop on average as the prediction horizon increases and ECGs further away from the episode make up the pre-AF class. The occasional small increases in sensitivity, after an initial drop, at subsequent stages (corresponding to longer prediction horizons) could indicate a slight overfit on the part of the SVMs. It is likely that as we move further away from

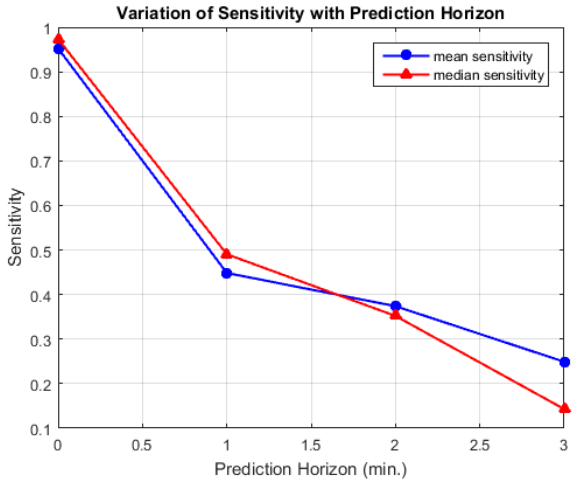

Figure 3. Variation of mean and median classification sensitivity with prediction horizon for all patients/records

fibrillatory rhythms, RR-intervals and ectopic beats do not contain much information regarding an impending abnormal rhythm. Moreover, each patient experiences a different number of AF events during Holter monitoring and consequently the data points belonging to the pre-AF class is very low for certain records. As the number of distant data points in a record will typically be much larger, a class imbalance occurs and a statistical classifier will tend to favor the majority class to improve accuracy, often resulting in low sensitivities. For patient 102, none of the features show a separation in medians for certain prediction horizons and hence the ranksum tests yield no good features to feed into a classifier. Furthermore, features belonging to both classes projected onto the first two principal components gradually begin to overlap as the prediction horizon increases.

V. Conclusions

Atrial fibrillation is associated with an increased risk of ischemic stroke and can potentially go undiagnosed until a more serious ailment occurs. The Computers in Cardiology Challenge 2001 aimed at sparking interest in predicting the occurrence of these sporadic rhythms. In this paper, we considered a similar version of the original problem, with shorter signals, where we attempted to distinguish between 2-minute epochs of ECG occurring prior to fibrillatory rhythms and those that occur sufficiently further away. Additionally, our method is more general as it seeks to classify between all the data points belonging to a particular patient at once instead of merely deciding which ECG occurs just prior to the AF episode when presented with two ECGs at a time.

Classification accuracy is high for most patients when classifying between the distant signals and those that terminate just prior to fibrillatory rhythms. However, as the distance or prediction horizon increases to a few minutes, sensitivity drops significantly. Future work would include incorporating p-wave morphologies into the features to improve sensitivity and extending the prediction horizon. A final product would incorporate this prediction algorithm running on an embedded platform such as proposed in [4].

References

[1] P. A. Wolf, J. B. Mitchell, C. S. Baker, W. B. Kannel, and R. B. D'Agostino, "Impact of atrial fibrillation on mortality, stroke, and medical costs," *Arch. Intern. Med.*, vol. 158, pp. 229–234, Feb. 1998.

[2] W. M. Feinberg, J. L. Blackshear, A. Laupacis, R. Kronmal, and R. G. Hart, "Prevalence, age distribution, and gender of patients with atrial fibrillation," *Arch. Intern. Med.*, vol. 155, pp. 469–473, Mar. 1995.

[3] G. Moody, A. Goldberger, S. McClennen, and S. Swiryn, "Predicting the onset of paroxysmal atrial fibrillation: the Computers in Cardiology Challenge 2001," *Comput. Cardiol. 2001*, vol. 28, pp. 113–116, 2001.

[4] G. E. Arrobo, C. A. Perumalla, Y. Liu, T. P. Ketterl, R. D. Gitlin, and P. J. Fabri, "A novel vectorcardiogram system," in *Proc. IEEE Int. Conf. e-Health, Net. Appl. Serv.*, Oct. 2014, pp. 188–192.

[5] P. J. Lindsberg, L. Toivonen, and H. Diener, "The atrial fibrillation epidemic is approaching the physician's door: will mobile technology improve detection?," *BMC Med.*, vol. 12, no. 180, Sep. 2014.

[6] K. S. Lynn and H. D. Chiang, "A two-stage solution algorithm for paroxysmal atrial fibrillation prediction," *Comput. Cardiol. 2001*, vol. 28, pp. 405–407, 2001.

[7] W. Zong, R. Mukkamala, and R. G. Mark, "A methodology for predicting paroxysmal atrial fibrillation based on ECG arrhythmia feature analysis," *Comput. Cardiol. 2001*, vol. 28, pp. 125–128, 2001.

[8] P. Langley, D. di Bernado, J. Allen, E. Bowers, F. E. Smith, S. Vecchietti, and A. Murray, "Can paroxysmal atrial fibrillation be predicted?," *Comput. Cardiol. 2001*, vol. 28, pp. 121–124, 2001.

[9] G. Schreier, P. Kastner, and W. Marko, "An automatic ECG processing algorithm to identify patients prone to paroxysmal atrial fibrillation," *Comput. Cardiol. 2001*, vol. 28, pp. 133–135, 2001.

[10] P. de Chazal and C. Heneghan, "Automated assessment of atrial fibrillation," *Comput. Cardiol. 2001*, vol. 28, pp. 117–120, 2001.

[11] C. Maier, M. Bauch, and H. Dickhaus, "Screening and prediction of paroxysmal atrial fibrillation by analysis of heart rate variability parameters," *Comput. Cardiol. 2001*, vol. 28, pp. 129–132, 2001.

[12] T. Thong, J. McNames, M. Aboy, and B. Goldstein, "Prediction of atrial fibrillation by analysis of atrial premature complexes," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 561–569, April 2004.

[13] I. Hilavin and M. Kuntalp, "Prediction of paroxysmal atrial fibrillation onset by using ECG," in *Proc. Int. Symp. Inno. Int. Sys. Appl.*, July 2012, pp. 1–3.

[14] Y. V. Chesnokov, A. V. Holden, and H. Zhang, "Distant prediction of paroxysmal atrial fibrillation using HRV data analysis," *Comput. Cardiol. 2007*, vol. 34, pp. 455–458, 2007.

[15] D. Arotaritei and C. Rotariu, "Automatic prediction of paroxysmal atrial fibrillation in patients with heart arrhythmia," in *Proc. Int. Conf. Exp. Elec. Pow. Eng.*, Oct. 2014, pp. 549–552.

[16] K. Hnatkova, J. E. P. Waktare, Francis D. Murgatroyd, X. Guo, X. Baiyan, A. J. Camm, and M. Malik, "Analysis of the cardiac rhythm preceding episodes of paroxysmal atrial fibrillation," *Am. Heart J.*, vol. 131, no. 6, pp. 1010–1019, June 1998.

[17] S. Petrutiu, A. V. Sahakian, and S. Swiryn, "Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans," *Eurospace*, vol. 9, pp. 466–470, May 2007.

[18] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, pp. e215–e220, June 2000.

[19] K. Padmavathi and K. Sri Ramakrishna, "Classification of ECG signal during atrial fibrillation using autoregressive modeling," *Procedia Comput. Sci.*, vol. 46, pp. 53–59, 2015.

[20] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Ann. Math. Stat.*, vol. 18, no. 1, pp. 50–60, 1947.

[21] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *SIGKDD Explor. Newsl.*, vol. 11, pp. 10–18, Nov. 2009.