

A Risk and Incidence Based Atrial Fibrillation Detection Scheme for Wearable Healthcare Computing Devices

Redjem Bouhenguel, Imad Mahgoub
Department of Computer, Electrical Engineering & Computer Science
Florida Atlantic University
Boca Raton, Florida, USA
rbouheng@fau.edu, mahgoubi@fau.edu

Abstract-Today small, battery-operated electrocardiograph devices, known as Ambulatory Event Monitors, are used to monitor the heart's rhythm and activity. These on-body healthcare devices typically require a long battery life and moreover efficient detection algorithms. They need the ability to automatically assess atrial fibrillation (A-Fib) risk, and detect the onset of A-Fib from EKG recordings for further clinical diagnosis and treatment. The focus of this paper is the design of a real-time early detection algorithm cascaded with an A-Fib risk assessment algorithm. We compare accuracy of machine learning schemes such as J48, Naïve Bayes, and Logistic Regression and choose the best algorithm to classify A-Fib from EKG medical data. Though all three algorithms have similar accuracy, the Logistic Regression model is selected for its easy portability to mobile devices. A-Fib risk factor is used to determine a monitoring schedule where the detection algorithm is triggered by the age dependent A-Fib incidence rate inside a circadian prevalence window. The design may provide a great public health benefit by predicting A-Fib risk and detecting A-Fib in order to prevent strokes and heart attacks. It also shows promising results in helping meet the needs for energy efficient real-time A-Fib monitoring, detecting and reporting.

Keywords-Algorithms, classification, arrhythmia, atrial fibrillation, wearable computing, real-time monitoring, logistic regression model of atrial fibrillation.

I. INTRODUCTION

Atrial fibrillation (A-Fib) is the most common cardiac arrhythmia [1] [2] [3]. The American College of Cardiology and the American Heart Association define A-Fib as a supraventricular tachyarrhythmia characterized by uncoordinated atrial activation accompanied by the deterioration of atrial mechanical function. A-Fib is responsible for approximately 15 percent of the strokes occurring in people with A-Fib. The cost to treat A-Fib in the United States exceeds \$6.4 billion per year [4]. Electrocardiograph portable devices are used to record the heart's rhythms and monitor arrhythmia however they are plagued by technological challenges such as energy constraints, process optimization problems, data security risks and interference [5]. Future wearable computing devices [6] require the ability to not only continuously monitor but also efficiently detect, analyze and report cardiac arrhythmia.

This paper presents the design of a Risk and Incidence Based Atrial Fibrillation Detection Scheme to be used in a

wearable computing application. Section I briefly illustrates the need for future wearable computing devices to possess the ability to not only continuously monitor but also efficiently detect, analyze and report cardiac arrhythmia. Section II highlights the related work. Section III describes arrhythmia monitoring and detection devices issues. Section IV introduces the data mining of arrhythmia and the first episode of A-Fib. Section V describes telemetry used in current devices, describes the incidence rate of A-Fib and the accuracy of A-Fib clinical diagnosis. Subsection VI establishes the A-Fib risk and detection models. Section VII is the conclusion.

II. RELATED WORK

The focus of this paper is the implementation of a risk assessment algorithm and the design of an incidence based A-Fib detection scheme for wearable healthcare computing devices. Related work in biomedicine and information technology introduced various algorithms for diagnosing and detecting different types of arrhythmia, and developed cardiovascular disease prediction algorithms. The Framingham heart study [7] developed a risk score to calculate individual's risk of developing atrial fibrillation and a development framework for researcher. The work by [8] developed a prediction model to detect tachycardia and send alerts to a designated care center for appropriate medical action. The research funded by the Health Technology Assessment Program addresses the accuracy of electrocardiogram (EKG) for the diagnosis of A-Fib and the potential risk of A-Fib misinterpretation errors [9]. A mobile medical device, dubbed HeartSaver [10] was developed to monitor the onset of atrial fibrillation and other cardiac pathologies. Other related work deals with the classification of arrhythmia and the performance of machine learning algorithms such as OneR, J48 and Naïve Bayes [11] but does not address logistic regression covered in this paper. The feasibility of EKG data collection by wireless sensors networks is derived in [12]. The duration and incidence rate of A-Fib are estimated in [13] with A-Fib predictors derived in [14].

III. CURRENT ARRHYTHMIA MONITORING AND DETECTION DEVICES ISSUES

Today, cardiac arrhythmia is diagnosed under the supervision of a physician, through the use of various diagnostic methods and tools. Patients visit health centers to

receive and get hooked up to devices, which are then carried by the patient and activated when arrhythmia symptoms are present. Current healthcare monitoring solutions are designed to work over a scheduled or pre-programmed period of time. Monitoring becomes ineffective for patients who experience infrequent symptoms outside the scheduled period and/or of very short duration. Additionally, the procedure becomes impractical when the patient is incapacitated during symptomatic periods. Certified technicians and doctors in remote medical centers review and analyze the data before a full report is generated and communicated to the patient. Furthermore ambulatory monitors such as Holter monitors, event monitors, and telemetry are not energy efficient and require long battery life.

IV. DATA MINING ARRHYTHMIA

Data mining or Knowledge Discovery in Databases is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data [15]. It uses machine learning, statistical and visualization techniques to design and develop algorithms that are capable of inducing knowledge from the data. Data mining is not an exact science. Human interaction is sometimes required to decipher ambiguities during the four phases of data mining process: data collection, data pre-processing, data mining and information evaluation and interpretation. Few machine learning algorithms and statistical approaches have been applied to cardiac arrhythmia classification [16] [17] [18].

A. Data Cleaning and Data Preprocessing

Biomedical data is highly distributed and often uncontrollably generated. Data may contain information that simply does not make sense and requires cleaning. Data cleaning is defined as a preprocessing step, and is essential in data mining to ensure accuracy, completeness, and consistency of data [21]. Before proceeding with data cleaning, understanding the data and how it was gathered helps eliminate outliers and data corruption. The dataset is partitioned using cross-validation. The training set is used to train the learning algorithm, and the induced decision rules are tested on the test set. The model is to be first built and evaluated using 10-fold cross validation on the fit data set, and then validated using the test data set. In 10-fold cross-validation, the dataset is divided into 10 subsets of (approximately) equal size. The dataset is split 10 times, each time leaving out one of the subsets to use for testing. The basic idea is to use 90% of the dataset to build a model and 10% to test the performance of the model [22].

B. Classification and Analysis Environment

The Waikato Environment for Knowledge Analysis software environment for Machine Learning (a.k.a. WEKA) [23] is used to analyze the dataset and classify the presence and absence of A-Fib. WEKA contains tools for data pre-processing, classification, regression, clustering, association

rules, and visualization. It is also well-suited for developing new machine learning schemes.

C. Incidence Rate of A-Fib

Among all arrhythmia, A-Fib is the most frequently diagnosed and affects 2.5 million people in the United States or close to 1% of the total population [4]. The Manitoba study [24] concluded that the incidence of A-Fib is 0.13 to 0.36 for people between 25 and 60 years old, 5.7 per 1,000 person-years after age 60, and 9.7 per 1,000 person-years after age 70. The Framingham Heart study [25] and other studies draw attention to the significance of the higher frequency of A-Fib with advancing age [26]. Patients with A-Fib have a 1.5-2 fold increase in mortality rate when compared with the general population as suggested by Framingham Heart study data [27]. Early recognition of A-Fib is difficult because most people are not aware of this silent rhythm disturbance [28]. Today, frequent monitoring and screening of patients allow for early detection of arrhythmia.

D. Clinical Diagnosis Accuracy of A-Fib

At least one-third of the A-Fib episodes go undetected [9] because either people do not get screened often or A-Fib diagnosis is missed by a general practitioner or practice nurse. Few studies have addressed the misdiagnosis of A-Fib from an electrocardiogram (EKG) and the potential risk of A-Fib misinterpretation errors. Knight et al. [29] concluded that A-Fib is more often misdiagnosed by internists than cardiology fellows and cardiologists. Mant et al. [30] discovered that general practitioners correctly detected A-Fib 80% (true positive) of the time when interpreting 12-lead EKG data and misinterpreted 8% (false positive) of sinus rhythm cases as A-Fib. One of the major misdiagnosis confuses A-Fib with atrial flutter [29] [31].

E. Predictors of A-Fib

A-Fib is the most prevalent arrhythmia in the United States and accounts for more than 750,000 strokes per year [32]. According to classification guidelines used by cardiologists and electro-physiologists, for the management of patients with A-Fib [14], after the first A-Fib is detected, there are mainly four types of A-Fib: Paroxysmal, persistent, longstanding persistent, and permanent. A-Fib is termed progressive. Once a patient is diagnosed with a paroxysmal A-Fib he or she will eventually migrate to persistent A-Fib. Similarly, a patient diagnosed with persistent A-Fib will drift to longstanding persistent A-Fib and in time to permanent A-Fib [33]. The EKG waves and intervals explained below are used to describe the heart electrical:

The QRS interval (see Figure 1) is the duration of the ventricular muscle depolarization. The P wave is a record of the electrical activity or the sequential activation (depolarization) through the right and left atria. The PR interval is the time interval measured from the beginning of

the P wave (atrial depolarization) to the onset of the QRS complex (ventricular depolarization). The RR interval is the

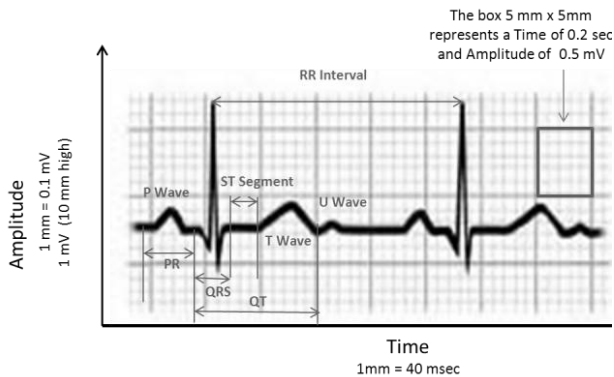


Figure 1: Typical EKG wave and intervals.

duration of the ventricular cardiac cycle; it is an indicator of the ventricular rate. The PP interval is the duration of the atrial cycle; it is an indicator of the atrial rate.

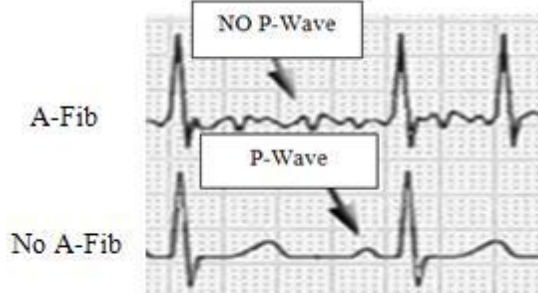


Figure 2: A-Fib depicted by the absence of the P-Wave.

We acknowledge detecting A-Fib is difficult and requires a more intense research, however one of the strong indicator of A-Fib presence is the absence of P waves on the EKG plot (see Figure 2) and an erratic noise like activity in their place combined with irregular R-R intervals [34][33][32]. Sometimes when the heart rate is too fast, irregular RR intervals may be difficult to determine [24]. Wide QRS complexes may be present with rapid ventricular response.

V. TELEMETRY DATA ANALYSIS

Telemetry is widely accepted in healthcare for remotely collecting and sending vital data to a monitoring station for analysis and interpretation of all types of arrhythmia in outpatients. Today, when prescribed by a physician, telemetry may be applied continuously for few days in the hope of capturing episodes of A-Fib. Telemetry may also be user-triggered by the patient as soon as he or she feels symptoms of A-Fib (such as heart palpitations). Using triggered events to start an A-Fib telemetry monitoring device runs the risk of missing the first 30 seconds of A-Fib.

Moreover, triggering events might not be possible if the user is incapacitated.

The telemetry model continuously senses EKG signals, transmits EKG data, receives EKG records, and reports EKG information to a healthcare center for further diagnostics and analysis by a doctor or a healthcare specialist. The telemetry report includes all positive and negative results. We assume that telemetry EKG interpretations are conducted by a cardiologist or a cardiophysio-physiologist who are trained experts at EKG readings; thus all judgments of what constitutes A-Fib are going to be assumed to be as accurate as possible. Unfortunately not every physician is a cardiologist, so general practitioners are often the first to interpret EKG readings during a general screening evaluation. General practitioners introduce human errors when interpreting EKG readings [30].

VI. A-FIB RISK AND DETECTION MODELS

Several clinical methods have been applied to treat arrhythmia in people, but these medical interventions and clinical treatments come after the fact and are expensive. Moreover, they do not come without risks to the patients [34]. There would be a greater positive public health impact from predicting arrhythmia risk and detecting arrhythmia to prevent strokes and heart attacks. Few machine learning algorithms and statistical approaches have been applied in medical applications; for example, classification of EKG arrhythmias using neural networks [35], EKG arrhythmia classification based on logistic model tree [16], and analysis of EKG signals using self-organizing maps (SOM). In this paper we concentrate on the design of a real-time early detection algorithm cascaded with an A-Fib risk assessment algorithm. We compare accuracy of machine learning schemes such as J48, Naïve Bayes, and Logistic Regression and choose the best algorithm to classify A-Fib from EKG medical data.

A. Developing A-Fib Risk Factor

The risk of developing A-Fib may depend on several factors—some associated with lifestyle and some from heredity. Many of these factors behave nonlinearly, complicating accurate A-Fib risk assessment in people. Standardizing the prediction of A-Fib from mere clinical diagnoses is difficult [36]. Few studies have addressed the misdiagnosis of A-Fib from an electrocardiogram (EKG) [9] [29] and the potential risk of A-Fib misinterpretation errors. Data mining techniques and statistical methods such as the Cox proportional hazards model [37] and the logistic regression model are used in many epidemiological studies.

The Cox Proportional Hazards Model is a multivariate statistical method used to compare survival in two different groups and determines the contribution of different variables on survival. The Framingham Heart study in the United States and the Prospective Cardiovascular Münster (PROCAM) study in Europe used the Cox model to develop

standardized risk factor assessments that may complement clinical practice. The Cox proportional-hazards regression [37] is used to analyze the effect of risk factors on survival. The probability of the onset of A-Fib is called the hazard. The following covariates and their corresponding coefficients responsible for predicting A-Fib risk in people aged between 45 and 95 years old are extracted from the Framingham Heart Study [27]: Age, Age², Gender, Body Mass Index (BMI), Systolic Blood Pressure (SBP), Treatment for Hypertension (TH), Significant Heart Murmur (SHM), Prevalent Heart Failure (PHF), Gender*age², and Age*PHF, PR Interval (PR_{interval}). We can express the hazard or risk of getting A-Fib at time t as:

$$H(t) = H_0(t) * e^{\sum_{i=1}^k \beta_i X_i}$$

We can linearize this model by dividing both sides of the equation by $H_0(t)$ and then taking the natural logarithm of both sides:

$$\ln\left(\frac{H(t)}{H_0(t)}\right) = \sum_{i=1}^k \beta_i X_i$$

$$\sum_{i=1}^k \beta_i X_i = \beta_1 Gender + \beta_2 Age + \beta_3 BMI + \beta_4 SBP + \beta_5 TH + \beta_6 SHM + \beta_7 PHF + \beta_8 Age^2 + \beta_9 Gender * Age^2 + \beta_{10} Age * SHM + \beta_{11} Age * PHF + \beta_{12} PR_{interval}$$

The quantity $H_0(t)$ is the baseline or underlying hazard function. It is practically the probability of getting A-Fib when all the other covariates are set equal to zero. The baseline hazard function is analogous to the intercept in linear regression. The regression coefficients β_1 to β_{12} provide the model with the proportional change or contribution from each covariate. The derived Cox proportional hazards equation is described below:

$$\begin{aligned} \ln(H(t)/H_0(t)) = & 1.994060 Gender + 0.150520 Age \\ & + 0.019300 BMI + 0.006150 SBP \\ & + 0.424100 TH + 3.795860 SHM \\ & + 9.428330 PHF - 0.000380 Age^2 \\ & - 0.000280 Gender * Age^2 \\ & - 0.042380 Age * SHM - 0.123070 Age \\ & * PHF + 0.070650 PR_{interval} \end{aligned}$$

Where $H_0(10) = 0.96337$ is the 10 year baseline survival or cumulative hazard function at time $t = 10$ years extracted from the Framingham Heart study [27]. The values of the means for each covariate are tabulated in Figure 3:

Covariate	Xbar	Covariate	Xbar
Gender	0.4464	SHM	0.0281
Age	60.9022	PHF	0.0087
BMI	26.2861	Age ²	3806.90
SBP	136.1674	Gender*Age ²	1654.66

TH	0.2413	Age*SHM	1.8961
PR _{Interval}	16.3901	Age*PHF	0.61

Figure 3: A-Fib risk covariates coefficients

For example, we calculate the risk factor of a male person who is 70 years old, weighing 70 kg, with a body mass index of 22.96, a systolic blood pressure of 130, with no hypertension, a PR interval measuring 16 ms, with no significant heart murmur, and no previous heart failure. Comparing to the mean values of the 10 year study from the Framingham Heart study we get:

$$A = \sum_{i=1}^k \beta_i X_i - \sum_{i=1}^k \beta_i X_{bar_i} = 11.669 - 10.786 = 0.883$$

$$D = e^{0.883} = 2.418$$

The predicted risk factor is:

$$k = 1 - H_0^D = 1 - 0.96337^{2.418} = 0.0863$$

The predicted Risk Factor is 0.0863 compared to a risk for a person of the same age and gender with BMI 20 to 24.9, Normal SBP (120 to 129), No Treatment for Hypertension, PR Interval 16, No significant murmur or prevalent heart failure.

B. A-Fib Detection Model

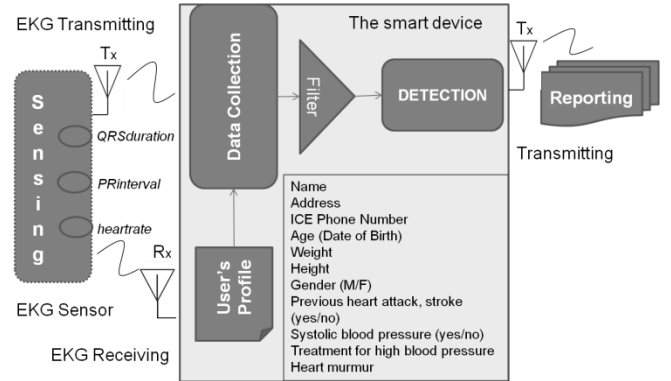


Figure 4: Overview of a wearable computing diagram.

Figure 4 shows an overview of a wearable computing diagram. Typically the general detection A-Fib model discovers the first episode of A-Fib by sensing EKG signals through a portable, low-power, wireless two-lead EKG system [12] [38], transmitting EKG data to a GSM/EDGE cell phone, receiving EKG records into a cell phone, detecting, and reporting when the detection algorithm detects the first 30 seconds of A-Fib. Data mining techniques and tools allow us the freedom to experiment

with various features to observe their effect on the A-Fib detection model. Figure 5 describes the features selected to predict A-Fib:

	Variable	Description	Value
1	age	Age in years , linear	real
2	Age ²	Age ² in years ²	real
3	Gender	Gender (0 = male; 1 = female) , nominal	{0, 1}
4	BMI	Kg/m ² , Linear	real
5	QRSduration	Average of QRS duration in msec., linear	real
6	PRinterval	Average duration between onset of P and Q waves in msec., linear	real
7	heartrate	Number of heart beats per min, linear	real
	class	{A-Fib present, A-Fib absent}	binary

Figure 5: A-Fib attributes.

The dataset used in our analysis was extracted from the Machine Learning Repository at University of California, Irvine [19], MIT-BIH Atrial Fibrillation database [20] and from data donated and corroborated by a cardiologist. The dataset describes the attributes for diagnosing cardiac A-Fib where each instance or patient is classified into two categories: presence of cardiac A-Fib and absence of cardiac A-Fib. The resulting dataset contains 304 records including 80 A-Fib cases, 224 non-A-Fib cases, 7 attributes and 2 classes (A-Fib Present, A-Fib Absent). The cardiologist's classification is used as a reference.

Three machine learning techniques, J48, Naïve Bayes algorithms, and regression analysis are explored to test for the detection of the presence or absence of A-Fib: a 7-attribute case and a 10-fold cross validation are used. The differences in accuracies from all three machine learning algorithms are not significant J48 at 96.71 %, Naïve Bayes at 96.38 %, and Logistic Regression at 97.37 %. In this paper, logistic regression is selected for its direct predictive simple computation and accuracy. Logistic regression [39] determines the relative effect of independent variables x_i on the dependent variable Y or class and their statistical significance. This effect is usually explained in terms of odds ratios where the odds of an event x occurring with probability p is defined as: odds (p) = p / (1-p) where p is the probability of the presence of the disease [40]. The logit transformation described in Figure 6 is defined as the natural log of odds,

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$$

$$p(Y = 1|x) = \frac{1}{1 + e^{-\text{logit}(p)}}$$

Where:

$$\text{logit}(p) = \beta_0 + \sum_{i=1}^k \beta_i x_i$$

$x_i = (x_1, x_2, \dots, x_k)$ is the covariate vector and β_i ($i = 1, 2, \dots, k$) denotes the coefficients of the k predictors. Fitting a logistic regression model to a given data implies deriving estimates of the coefficients β_i that maximize the likelihood of the model.

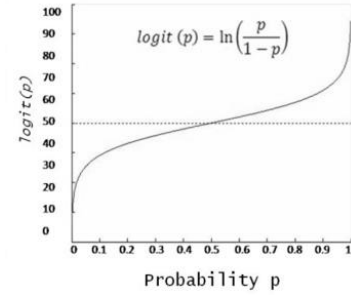


Figure 6: Logit function

The outcomes of the Logistic Regression include all True Positive and False Positive results. They may be triggered at A-Fib incidence rates reported in the Manitoba studies [24] where the incidence of A-Fib is 0.13 to 0.36 for people between 25 and 60 years old, 5.7 per 1,000 person-years after age 60, and 9.7 per 1,000 person-years after age 70 . A-Fib is predicted present if probability p (A-Fib is Present | age, age², gender, BMI, QRSduration, PRinterval, heartrate) > 0.5

Otherwise, A-Fib is absent.

Where:

$$\text{logit}(p) = -41.175 + 0.820 \text{ age} - 0.006 \text{ age}^2 + 4.737 \text{ Gender} - 0.047 \text{ BMI} + 0.098 \text{ QRSduration} - 0.178 \text{ PRinterval} + 0.066 \text{ Heartrate}$$

$$\text{and } p = 1 / (1 + e^{-\text{logit}(p)})$$

1) Evaluating Classifier Performance

Given an EKG record, a binary classification has four possible outcomes or rates: True negative (TN), False Positive (FP), True Positive (TP), and False Negative (FN). Detection rates are measured in terms of sensitivity and specificity [40]. When considering the results of a particular test in two populations, one population with an A-Fib, the other population without A-Fib, the distribution of the test results will overlap, as shown in Figure 7.

For every possible cut-off point in the test there are cases with A-Fib that are correctly classified as positive (TP = True Positive fraction); cases with A-Fib that are incorrectly classified as negative (FN = False Negative

fraction), cases without A-Fib that are correctly classified as negative (TN = True Negative fraction), and cases without A-Fib that are incorrectly classified as positive (FP = False Positive fraction).

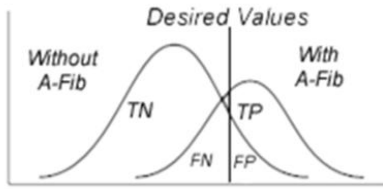


Figure 7: Test results

Both the overall classification accuracy and the overall classification error defined below may be used to evaluate the performance of the classifier:

$$\text{Overall Error rate} = \frac{FP+FN}{TP + TN+ FP+FN} = 2.63\%$$

$$\text{Overall Accuracy} = \frac{TP+TN}{TP + TN+ FP+FN} = 97.37\%$$

but when the costs of misclassifications of the different classes are uneven, this measure may be unacceptable. In order to take into account the unevenness of misclassification costs when evaluating a classifier, area under the Receiver Operating Characteristic (or ROC) curve is explored.

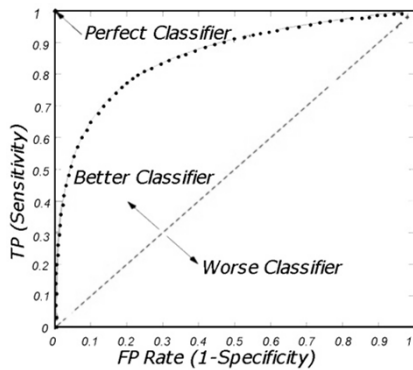


Figure 8: Specificity versus Sensitivity Curve

ROC curves have been used in biomedical informatics [41] to express the sensitivity versus specificity of classifiers. The ROC curve plot displays the False Positive rate on the X-axis (1- Specificity) and the True Positive rate (Sensitivity) on the Y-axis as shown in Figure 8. Each point on the ROC curve represents a sensitivity/(1-specificity) pair corresponding to a particular decision threshold. The area under the ROC curve measures how well a particular parameter can distinguish between two diagnostic groups (such as presence of a disease/ absence of A-Fib). The bigger the area is and the closest to 1, the better the

classifier performance. The area under the ROC curve for the derived logistic regression model is 0.986.

Figure 9 shows the interpretation of the confusion matrix with the A-Fib predicted class represented by the columns of the matrix, and the actual class represented by the rows of the matrix.

Confusion Matrix		Actual A-Fib	
		A-Fib Positive (Present)	A-Fib Negative (Absent)
Predicted A-Fib	A-Fib Present. (Positive)	TP=79 Sensitivity = $\frac{TP}{TP+FN}$ 98.8%	FP=7 (Type I Error) FP = $\frac{FP}{TN + FP}$ 3.1%
	A-Fib Absent (Negative)	FN=1 (Type II Error) FN $\frac{FN}{TP + FN}$ 1.3 %	TN=217 Specificity TN = $\frac{TN}{TN + FP}$ 96.9%

Figure 9: Confusion matrix of A-Fib Logistic Regression

The A-Fib detection algorithm is triggered by the onset of A-Fib. The incidence rate of A-Fib is higher in older people [24]. Suggested studies [31] reveal that clinical measurement of sensitivity (True Positive rate) of 80% and specificity (True Negative rate) of 92% when A-Fib is diagnosed by internists and general practitioners instead of cardiologists. Our logistic regression classification of A-Fib has a measurement of sensitivity of 98.8% and specificity of 96.9 %. The false positive results, usually interpreted as false alarms, contribute to wasted or needless energy spent in transmitting inaccurate information. In this analysis the logistic regression algorithm has a False Positive rate of 3.1% (see Fig. 9, Confusion matrix of A-Fib Logistic Regression).

C. Applying a Risk and Incidence Based A-Fib Detection Model

A-Fib monitoring devices may become impractical when they run out of battery energy, an undesirable condition when the patient is incapacitated during symptomatic periods. Typical monitoring and detection healthcare wearable body network devices have limited energy and therefore limited monitoring duration [42]. The implementation of a Risk and Incidence Based A-Fib Detection Scheme in such devices alleviates the aforementioned challenges. For instance, A-Fib risk factors may be classified in three categories made up of risk ranges such as $k < 0.05$, $0.05 < k < 0.15$, $k > 0.15$. Knowing the A-Fib risk factor of a patient allows one to prescribe an A-Fib monitoring and detection schedule inside an appropriate circadian prevalence timing window [42] (see Figures 10

and 11). A high A-Fib risk factor may suggest more frequent monitoring and wider circadian window compared to a low A-Fib risk factor.

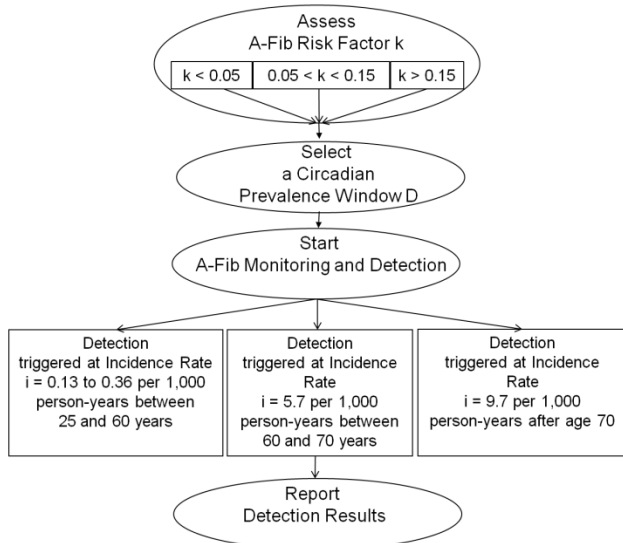


Figure 10: Overview of an efficient wearable computing device

Because A-Fib is not a common occurrence [30], a result is reported only when there is an actual occurrence of A-Fib.

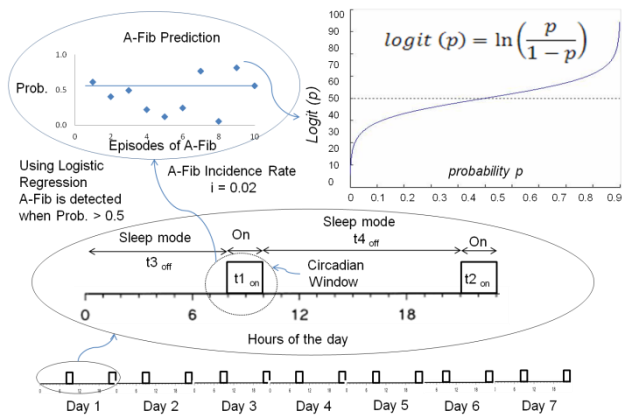


Figure 11: A-Fib Episodes inside a circadian prevalence window

In Figure 11, the A-Fib logistic regression model detects the first episode of A-Fib by continually monitoring EKG signals, detecting, and reporting when the first 30 seconds of continuous A-Fib occurs. The width of the circadian monitoring window depends on the A-Fib risk value and varies within a 24 hour period. Monitoring may continue beyond the 24 hour period. After the first 30 seconds of A-Fib is detected, monitoring may proceed to detect paroxysmal, persistent, long standing persistent and permanent A-Fib, which may require monitoring for days or weeks. The realization of a Risk and Incidence Based A-Fib

Detection Model may be a good fit to an energy constraint monitoring and detection model.

Using the two-lead EKG Alive Technologies Heart Monitoring Device A102D7 [38], we compare the energy consumed by a Risk and Incidence Based A-Fib Detection Scheme to the energy consumed by a telemetry model. We assume the telemetry model continuously monitors and transmits EKG signals during a 24-hour period. The device monitors and transmits EKG signals via Bluetooth to a MacBook. We realize a preliminary energy savings of 89.7% when we use an A-Fib incidence rate of 0.02 and continuously monitoring the onset of A-Fib during a cumulative 4-hour circadian prevalence window [42]. Similarly, we realize a preliminary energy savings of 38.2% when using an A-Fib incidence rate of 0.02 and continuously monitoring the onset of A-Fib during a 24-hour window (see Figure 12).

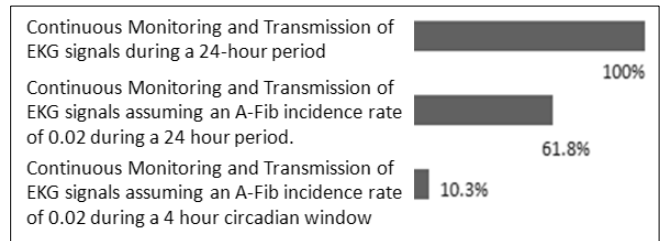


Figure 12: Comparing the energy consumed in different scenarios

The authors plan to implement a risk and incidence based atrial fibrillation detection scheme in a wearable device and further validate the results in a clinical setting.

VII. CONCLUSION

In this paper, we design a risk and incidence based atrial fibrillation detection scheme to alleviate the abovementioned problems in energy constrained wearable computing devices. We recommend an A-Fib Risk factor assessment to determine a risk category and implement a monitoring and a detection schedule by using a circadian prevalence window. The detection is triggered based on age dependent incidence rates. Studies [24] [25] suggest if the detection algorithm is as accurate as the cardiologist's accuracy of interpreting EKG readings then the design shows promising results in meeting the energy needs of monitoring, detecting and reporting A-Fib required in wearable computing healthcare applications.

REFERENCES

- [1] <http://www.OmniMedicalSearch>.
- [2] Domanski MJ. The epidemiology of atrial fibrillation. *Coronary Artery Disease* 1995;6:95–100.
- [3] "Atrial Fibrillation (for Professionals)". American Heart Association, Inc. 2008-12-04.

- [4] Coyne KS, Paramore C, Grandy S, Mercader M, Reynolds M, Zimetbaum P. Assessing the Direct Costs of Treating Nonvalvular Atrial Fibrillation in the United States. *Value Health*. Sep-Oct 2006.
- [5] R. Bouhenguel, I. Mahgoub, and M. Ilyas, "Bluetooth Security in Wearable Computing Applications," Proceedings of the IEEE 5th International Symposium on High Capacity Optical Networks & Enabling Technologies (HONET 2008), Penang, Malaysia, Page(s) 182 – 186, November 2008.
- [6] L.Schwiebert, S. Gupta & J. Weinmann, "Research challenges in Wireless networks of Biomedical Sensors". ACM SIGMOBILE 7/01 Rome Italy; 2001.
- [7] Renate B Schnabel MD et al., Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study, *The Lancet*, Volume 373, Issue 9665, Pages 739 - 745, 28 February 2009.
- [8] Khalil,I.;Sufi,F., Mobile device assisted remote heart monitoring and Tachycardia prediction, *Information Technology and Applications in Biomedicine*, 2008. ITAB 2008.
- [9] Bogun F, Anh D, Kalahasty G, et al. Misdiagnosis of Atrial Fibrillation and its Clinical Consequences. *Am J Med*. 2004
- [10] Ziad Sankari, Hojjat Adeli, HeartSaver: A mobile cardiac monitoring system for auto-detection of atrial fibrillation, myocardial infarction, and atrio-ventricular block, *Computers in Biology and Medicine* Volume 41, Issue 4 , Pages 211-220, April 2011.
- [11] Soman, T. and Bobbie, P.O., "Classification of Arrhythmia Using Machine Learning Techniques," *WSEAS Transactions on Computers*, Vol.4, issue 6, June 2005, pp. 548-552.
- [12] Thaddeus R. F. Fulford-Jones, Gu-Yeon Wei, Matt Welsh, A Portable, Low-Power, Wireless Two-Lead EKG System, Div. of Engineering and Applied Sciences, Harvard University, Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco, CA, USA Sept 1-5, 2004.
- [13] Ana Ruigómez, Saga Johansson, Mari-Ann Wallander, Luis Alberto, García Rodríguez, Incidence of Chronic Atrial Fibrillation in General Practice and Its Treatment Pattern, *Journal of Clinical Epidemiology* 55; 2002.
- [14] Peter Reinelt, Georg Karth, Alexander Geppert and Gottfried Heinz, Incidence and Type of Cardiac Arrhythmias in Critically Ill Patients: a Single Center Experience in a Medical-Cardiological ICU, *Intensive Care Medicine*, Volume 27, Number 9, 1466-1473.
- [15] G. Piatetsky-Shapiro and W. J. Frawley, *Knowledge Discovery in Databases*, AAAI Press, Menlo Park, CA, 1991.
- [16] Mahesh, V et al., EKG arrhythmia classification based on logistic model tree *Journal of Biomedical Science and Engineering (JBISE)* Vol: 2 Issue: 6 ISSN: 1937-6871 Date: 10/2009 Start Page: 405
- [17] H. Altay Guvenir, Burak Acar, Gulsen Demiroz, Ayhan Cekin "A Supervised Machine Learning Algorithm for Arrhythmia Analysis "Proceedings of the Computers in Cardiology Conference,Lund, Sweden, 1997.
- [18] Weiss, S. I., and Kulikowski, C. *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Networks, Machine Learning, and Expert Systems*. San Francisco, Calif.:Morgan Kaufmann, 1991.
- [19] UCI Machine Learning Repository, University of California, Irvine, California, USA.
- [20] PhysioBank, ANSI/AAMI EC13 Test Waveforms, <http://physionet.org/physiobank/database/aami-ec13/>
- [21] P. Bajcsy, J. Han, L. Liu, J. Yang, Survey of Biodata Analysis from a Data Mining Perspective, *Advanced Information and Knowledge Processing*, 2005.
- [22] Goutte, C., "Note on free lunches and cross-validation," *Neural Computation*, 1997.
- [23] WEKA <http://www.cs.waikato.ac.nz/~ml/weka/index.html>
- [24] Andrew D. Krahn, MD, Jut-e Manfreda, MD, Robert B. Tate, MSc, Francis A.L. Mathewson, MD, T. Edward Cuddy, *The Natural History of Atrial Fibrillation: Incidence, Risk Factors, and Prognosis in the Manitoba Follow-Up Study*, MD, Winnipeg, Canada, *The American Journal of Medicine*, May 2005.
- [25] Wolf PA, Kannel WB, McGee DL, et al. Duration of Atrial Fibrillation and Imminence of Stroke: The Framingham Study. *Stroke*. 1983;14(5):664-667.
- [26] Furberg CD, Psaty BM, Manolio TA, et al., Prevalence of Atrial Fibrillation in Elderly Subjects (The Cardiovascular Health Study), *Am J Cardiol*, 1994.
- [27] Lloyd-Jones DM, Wang TJ, Leip EP, et al. Lifetime Risk for Development of Atrial Fibrillation: The Framingham Heart Study. *Circulation* 2004.
- [28] Abdel Latif A, Messinger-Rapport BJ. Should Nursing Home Residents with Atrial Fibrillation be Anticoagulated? *Cleve Clin J Med*. Jan 2004
- [29] Knight BP, Michaud GF, Strickberger SA, et al. Electrocardiographic Differentiation of Atrial Flutter From A-Fib by Physicians. *J Electrocardiol* 1999.
- [30] Jonathan Mant, David A Fitzmaurice, F D Richard Hobbs, Sue Jowett, Ellen T Murray, Roger Holder, Michael Davies, Gregory Y H Lip, Accuracy of Diagnosing Atrial Fibrillation on Electrocardiogram by Primary Care Practitioners and Interpretative Diagnostic Software: Analysis of Data from Screening for Atrial Fibrillation in the Elderly (SAFE) Trial, *BMJ* 2007.
- [31] Shiyovich A, Wolak A, Yacobovich L, Grosbard A, Katz A., Accuracy of Diagnosing Atrial Flutter and Atrial Fibrillation from a Surface Electrocardiogram by Hospital Physicians: Analysis of Data from Internal Medicine Departments, Cardiology Department, Soroka University Medical Center, Ashkelon, Israel, *Am J Med Sci*. 2010.
- [32] Perez, Marco V; Dewey, Frederick E; Marcus, Rachel; Ashley, Euan A; Al-Ahmad, Amin A; Wang, Paul J, Electrocardiographic predictors of atrial fibrillation, *American Heart Journal*, ISSN 0002-8703, 2009, Volume 158, Issue 4, pp. 622 - 628.
- [33] HRS/EHRA/ECAS Expert Consensus Statement on Catheter and Surgical Ablation of Atrial Fibrillation: Recommendations for Personnel, Policy, Procedures and Follow-Up http://www.hrsonline.org/News/Media/press_releases/CSAblation.cfm
- [34] Fuster V, Rydén LE, Cannom DS, et al. (2006). "ACC/AHA/ESC 2006 Guidelines 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 European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001.
- [35] Bani-Hasan et al., Electrocardiogram signals identification for cardiac arrhythmias using prony's method and neural network, *IEEE Engineering in Medicine and Biology Society. Conference* Vol: 1 ISSN: 1557-170X Date: 2009 Start Page: 1893.
- [36] Tu JV, Naylor CD. Clinical prediction rules. *J Clin Epidemiol*. 1997;50:743-744.
- [37] D.R. Cox, Regression Models and Life-Tables, journal articles in statistics and medicine published in 1972.
- [38] Alive Heart and Activity Monitor, Alive Technologies Pty. Ltd, <http://www.alivetec.com/products.htm>
- [39] David W. Hosmer, Stanley Lemeshow, *Applied logistic regression*, John Wiley and sons, second edition, 2000.
- [40] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, San Francisco, 2001.
- [41] Lasko T-A, Bhagwat J-G, Zou K-H, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. *J Biomed Inform* 2005;38:404-15.
- [42] R. Bouhenguel, I. Mahgoub, M Ilyas, Window-Based Energy-Aware Model for Real-Time Detection and Reporting of Progressive Development of Cardiac Atrial Fibrillation in Wearable Computing, Proceedings of the IEEE 5th International Symposium on High Capacity Optical Networks & Enabling Technologies (HONET 2011), Riyadh, Saudi Arabia, 2011.