Predicting the near-future impact of daily activities on heart rate for at-risk populations

Gordana Velikic
ECE, CFH
University of Rochester
Rochester, NY, USA
gvel@ece.rochester.edu

Joseph Modayil
Dep. of Computing Science,
University of Alberta
Alberta, Canada
jmodayil@cs.ualberta.ca

Mike Thomsen
SUNY Geneseo
Geneseo, NY USA
roberto7327@gmail.com

Mark Bocko, Alice Pentland
Center for Future Health
University of Rochester
Rochester, NY USA
bocko@ece.rochester.edu

Abstract **--In this paper we demonstrate the ability to predict changes to heart rate due to changes in levels of activity, up to an hour into the future. Activity levels are calculated from data collected by a worn accelerometer for a person performing daily activities outside a laboratory enviroment. People with congestive heart failure must take care not to excessively stress their heart. This can be a challenge due to the difficulty of predicting how much stress an activity is exerting on the heart. We propose to model the relationship between motion and heart rate and thus to enable the prediction of heart rate changes prior to performing an activity. We explored three methods to predict current and future heart rate from activity level: a continuous state Kalman Filter, two simple linear models, and a nonlinear model given in the literature [5]. The results from healthy subjects and subjects with congestive heart failure show that using the proposed models, the heart rate can be predicted an hour into the future using accelerometer data.**

Keywords: Heart Rate, activity level, nonlinear models, Kalman, Filter, prediction

I. INTRODUCTION

The simplest, most informative ECG feature for heart health is the heart rate (HR), measured as the number of beats per minute (bpm). When a person is engaged in activity, an increase in HR reflects the additional work performed by the heart to meet increased demands on the body. A person's HR varies with age, gender, physical fitness, experienced emotions, physical strain, and disease. Also, HR variations correlate with autonomic nervous system fluctuations, and studies have shown a connection between HR variations and patient mortality/morbidity [1]. Particular effort has gone in the area of modeling heart rate response during exercise and recovery after exercise, as it is believed that such knowledge would be beneficial to predicting heart disease mortality [2]. The models predominantly consist of feedforward and feedback components [3], or use a static nonlinearity cascaded at the input of a linear system and performed well for short duration exercises [4]. An exception is a nonlinear model by Cheng at al [5], in which they associated a transient response with the intensity of the exercise, measured as the speed of a treadmill, thus making the model applicable to long duration exercises. Critchley and al. showed the human brain activity that responds to emotional facial expressions can be used to predict differential heart rate response [6]. In the same vein, measured response of beta-adrenergic receptors to brain stimuli may be used to predict heart rate [7]. However, all these variables are not easily accessible in a non-laboratory environment. A major focus of cardiac health research is the detection of possible catastrophic events [8], thus centering research goals on the detection of anomalies that already exist [9]. Cardiac health may begin to deteriorate long before a person experiences heart related problems. The moment of change is not detectable from the ECG signal taken during a patient's ambulatory exam. Continuous data collection opens possibility to a personalized model that enables comparison between the actual and predicted state of cardiac health. A personalized model could potentially be used to monitor and detect minor but substantial changes in a person's continually monitored cardiac health that would be otherwise imperceptible during brief clinical observations. This model could help cardiac patients to get input to the efficacy of their treatment sooner, or assist in early diagnostics of otherwise healthy people. Advances in technology are enabling continuous monitoring, particularly in patients at risk. However, continuous monitoring of the ECG signal may involve additional problems such as allergies to contact gel, the number of electrodes needed, and wires. The problem can be resolved by mapping the ECG features to other more easily observed variables from compact wireless sensors. Although subtle ECG differences might not be directly visible in short observations from simpler sensors like accelerometers, the ability to acquire data over a long period may reveal substantial information about a person's heart health. Potentially, a personalized model can add to quality of life of chronic patients with an additional insight to consequences of an activity before it is performed.

In this preliminary study, we examine whether accelerometers can be used to extract and predict a person's heart rate during structured activities, and also how far into the future we can reliably predict heart rate from unstructured activity.

II. DATA COLLECTION AND PROCESSING

We used two sets of devices to collect data. In one we employed an Alive Heart monitor positioned on a hip of a subject, a device which records both ECG with a sampling rate of 300Hz and acceleration with a sampling rate of 75Hz. The ECG data is taken across the heart with one electrode placed on the top of the bone directly between the collarbones and below the throat. A second electrode is placed on the lower left rib area, so that an imaginary line that connects two electrodes passes directly through the heart. The second set consisted of a three lead commercial holter device, with a sampling rate 255Hz, and the ADXL330 accelerometer (sampling rate 20Hz) mounted on a TMoteSky in custom made casing, which is worn on the hip. Although the holter device has three channels, for our analyses we used only the channel that corresponded to the electrode placement described above.

We collected 20 hours of continuous data from ten subjects with congestive heart failure and ten healthy controls, performing their everyday activities. We also collect data from healthy subjects performing structured activities during 20 minutes: slow walk, brisk walk, climbing stairs, jog, and run, averaging three minutes per activity with a minute rest in a standing position. The signals were additionally processed to accommodate for different sampling rates. For each subject we had multiple runs, not less than three, collected at least two weeks apart.

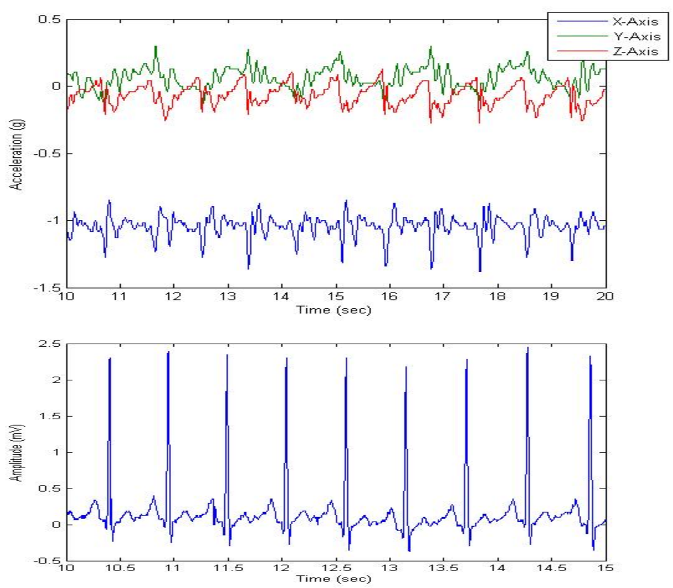

Fig.1: Upper plot: *A 5 second sample of ECG data taken from the Alive Heart Monitor. The sharp peaks are known as the R peaks and correspond to one heart beat. Lower plot: A 20 second sample of 3-axis acceleration data recorded by the Alive Heart Monitor during a walk. The accelerometer x-axis is oriented vertically and captures the force of gravity.*

Acceleration data is recorded for 3 axes: x, y, and z. A sample of acceleration data is shown in Fig. 1, upper plot. Analysis of each of three axes of acceleration is unreliable in cases where an accelerometer is loosely secured to a subject, and also unnecessary for this application. Therefore the magnitude of acceleration *a,* is calculated from the three acceleration components.

$$a = \sqrt{x^2 + y^2 + z^2} \qquad (1)$$

The line integral of the acceleration vector is then computed to measure the total changes in acceleration over a time window of one second.

$$ActivityLevel = \sum_{i=t-1}^{t} \sqrt{\Delta a^2 + \Delta t^2} \qquad (2)$$

where Δa is the change in acceleration over a small time change Δt, equal to one sample. The activity level provides an estimate of the average energy involved in movement occurring over a short period of time. It shows when the subject is stationary and when the subject is moving, along with the intensity of this movement. Figure 2 shows the activity level associated with a variety of different activities.

We applied an algorithm given in [10], to detect R-peaks from ECG signal (Fig. 1, lower plot) and to calculate the HR as:

$$HR = \frac{60}{t_{R-R}}[bpm] \qquad (3)$$

Although there may be simpler algorithms to detect the most prominent peak of the ECG signal, we decided to use this algorithm because it works well with noisy data collected from non-stationary subjects.

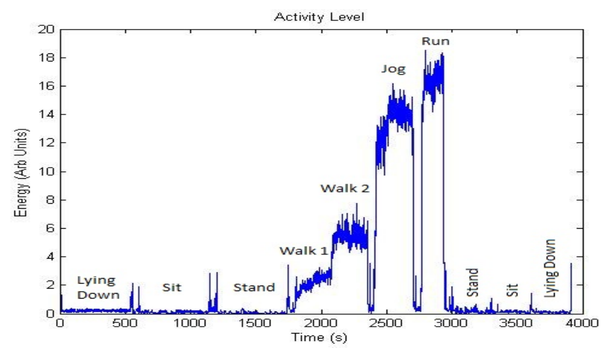

Fig. 2: *Activity level for Multiple Activities. The three stationary activities at the beginning and the end show close to zero activity level as expected. The moving activities are clearly distinct from stationary activities and also distinct from each other. The two walks performed at different speeds have different activity levels associated with them as do the jog and run. This activity level method provides a clear picture of the relative difference between different activities.*

III. MODELS

We employed machine learning techniques to capture the underlying relationship (*f*) between heart rate (*x*) and activities (*u*) from examples $\{(x_1,u_1), (x_2,u_2)\ldots\}$. The methodology is standard: split the data into training and testing sets, use training data to find a function f, and then use testing data to evaluate *f*. All models are personalized, i.e. trained and tested per each subject, except for the nonlinear model given by Cheng at al. [5].

The Kalman filter is described by following general equations:

$$x_{t+1} = Ax_t + Bu_t + n_t \qquad (4)$$

$$y_t = Cx_t + v_t \qquad (5)$$

Here *x* is the hidden state, *y* is the observed heart rate, A and B are general form matrices, *n* and *v* are normally distributed Gaussian noise parameters, *u* is activity level, and *t* is the current time step. The state *x* contains the unobserved true heart rate and C is a fixed matrix (not learned) that selects this component. We considered four different models.

Model 1: The hidden state consists of the person's true (unobserved) heart rate.

$$x_t = h_t \qquad (6)$$

$$h_{t+1} = a(h_t - \overline{r}) + \overline{r} + bu_t + n_t \qquad (7)$$

where h_t is heart rate, and $\overline{r}$ is an absolute resting heart rate

Model 2: The hidden state contains the true (unobserved) heart rate, and the person's resting heart rate.

$$x_t = \binom{h_t}{r_t} \qquad (8)$$

$$h_{t+1} = a_1(h_t - r_t) + r_t + b_1 u_t + n_{1t} \qquad (9)$$

$$r_{t+1} = a_2(r_t - \overline{r}) + \overline{r} + b_2 u_t + n_{2t} \qquad (10)$$

where r_t is a slowly varying resting heart rate

Model 3: The heart rate is a constant.

$$y_t = \overline{r} + n_3 \qquad (11)$$

Model 4: Heart rate is an affine function of activity only.

$$y_t = a_3 u + b_3 + n_4 \qquad (12)$$

Models 1 and 2 are simple models of a heart, but seem to work well. Models 3 and 4 are used as additional baselines. The parameters are determined using linear regression from the training data.

Nonlinear Model: Cheng *et al.* [5] present a nonlinear model that is defined similarly to the models above, but is designed to work across multiple subjects for a single activity - walking on a treadmill over an extended period of time. The primary benefit of their model is the ability to model a nonlinear HR response to different walking speeds, *i.e.* walking at 5 km/h and 6 km/h changes the HR by a constant offset while walking at 7 km/h yields a linear change in HR over time. Using v to denote the treadmill speed, $z_t = h_t - \overline{r}$ for the elevation of heart rate over baseline, and q_t for the long term perturbation of the HR, their system model is as follows:

$$dz_t = -a_1 z_t + a_2 q_t + a_2 v_2 \qquad (13)$$

$$dq_t = -a_3 q_t + a_4 \sigma(z_t) z_t \qquad (14)$$

where $\sigma(z_t) = (1 + e^{-z_t - a_5})^{-1}$ is a nonlinear sigmoid function. The parameters for this patient-independent model were found by the Levenberg–Marquardt method from six healthy patients across three trials. As such, this model is neither personalized nor directly applicable to a person's life off a treadmill. For more details and parameter values see [5].

IV. RESULTS

We tested the models on both the 20 minute runs with structured activities and 20 hour runs for a person's daily life, see Fig. 3 and 4. The Kalman filter model parameters were estimated from the training data and evaluated on the separate test data. The nonlinear model was only applicable to the 20 minute run, for which velocity information could be estimated from the accompanying GPS data. Furthermore for the nonlinear model as subjects had widely varying resting heart rates, we adjusted the resting heart rate for each individual from the training data.

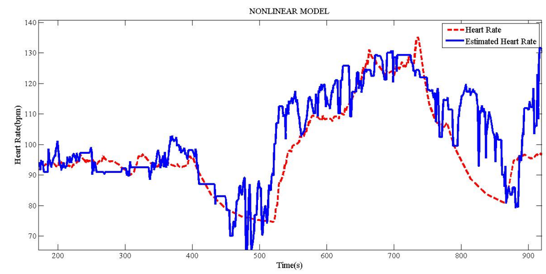

Fig. 3. *Nonlinear model applied to a 20 minute data, zoomed in during walk and brisk walk of a healthy female subject. It is easy to observe that predicted and real heart rate have high correlation.*

To evaluate the predictive performance of the models, we assume the model state variables are updated at each time step with the observed heart rate and activity level, and then the models are used to predict the patient's heart rate at some point in the future given only the activity level. The tables show the average RMSE from computing the difference at each point in time.

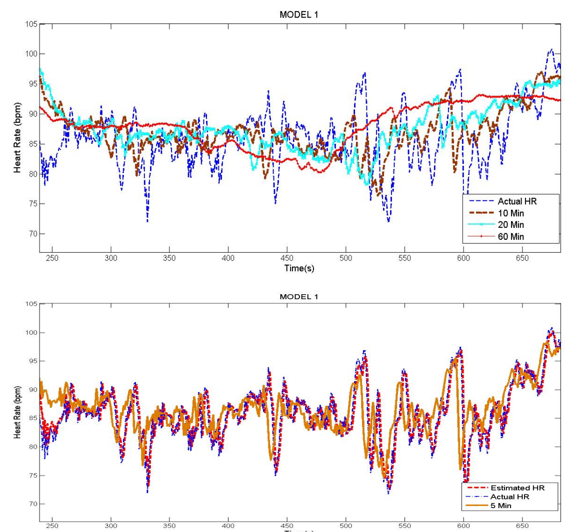

Fig. 4. *Zoomed detail of the cardiac patient using 20 hour data and model 1. It is easy to observe that predictions are highly correlated with HR. The horizontal axis is time in minutes, and the vertical is HR [bpm]. For clarity the same segment is zoomed with lines showing prediction 10, 20 and 60 minutes in the future (upper plot), and zero and 5 minutes (lower plot).*

The results from the 20 hour run are shown in Tables 1 and 2. Model one and two, substantially outperform baselines three and four. Predictions for the cardiac patients are comparable to those for the healthy subjects, showing that this approach is appropriate for the target patient population.

TABLE I. RMSE of 20 Hour Kalman models for Healthy Subjects.

	0 Min	5 Min	10 Min	20 Min	60 Min
Model 1	2.86	4.76	5.92	7.78	13.00
Model 2	2.93	5.34	7.36	10.65	19.31
Model 3	18.41	18.41	18.49	18.54	18.76
Model 4	17.46	17.54	17.59	17.68	17.99

In Table 3. are given results for the 20 minute trials. We see that Models 1 and 2 outperform the nonlinear model. Figure 3, shows that the nonlinear model can do a good job of walks and brisk walks, but it does not generalize to the broader set of activities tested here.

TABLE 2: RMSE of 20 Hour Kalman Models for Cardiac Patients.

	0 Min	5 Min	10 Min	20 Min	60 Min
Model 1	3.03	3.78	4.89	6.57	11.15
Model 2	3.16	4.49	6.19	8.54	13.08
Model 3	15.78	15.84	15.87	15.97	16.28
Model 4	14.33	14.41	14.45	14.55	15.03

The results in Table 3 for nonlinear model are restricted to the non-strenuous activities, otherwise the RMSE jumps from 15.03 to over 40bpm.

TABLE 3: RMSE of 20 minute trials with structured activities.

	0 Min	5 Min	10 Min
Model 1	2.52	7.54	10.1
Model 2	2.52	7.56	10.76
Model 3	10.35	11.63	12.73
Model 4	10.9	12.26	13.42
Nonlinear model	15.03	15.03	15.03

V. DISCUSSION

This study has shown different methods to predict a subject's heart rate based on activity level. With the exception of the nonlinear model, the parameters of the models are personalized, i.e. calculated per subject, instead of being averaged per demographic group. The nonlinear model was optimized for average healthy subjects with limited activity levels, and as expected this model yielded questionable results for non-healthy subjects and activities that are out of scope of brisk walk. However, it can be seen from the RMSE that the performance of the model does not vary with the depth of future prediction. We should keep in mind that we applied parameters given in [5], and the only deviation from the given model is adaption of resting heart rate per subject, which was set to 74 bpm in [5]. This leaves room for improvements, such as broadening the spectrum of activity levels and personalization of the parameters instead of averaging. We expect that our future work will prove this model to be very powerful for our research goals. The Kalman filter in Models 1 and 2 is effective for predicting heart rate from activity level, for both healthy population and cardiac patients. The real goal is to predict the consequences of an activity before it is performed. The use of a holter is for verification, and even a low-cost pulse watch could be deployed in practice. Planning and models are limited, so distant forecasts are bound to have error. Still, the ability to perform any prediction has potential benefits. This is a promising direction for future research. The Kalman models presented above are linear, and the heart is known to be non-linear. A hidden Markov model can account for some non-linear aspects, but our preliminary results showed that continuous nature of the Kalman models is more suitable for our application needs.

There are many studies involved with activity classification methods. These techniques can be applied to the acceleration data, and would certainly prove useful in a long-term study with a subject performing undocumented activities. However, the activity level method provides enough information to pinpoint changes between certain activities, and clearly shows difference between stationary and moving activities.

VI. CONCLUSION

New technologies are opening the door for many uses of predictive models. An at-risk patient can wear sensors from which data can be processed in real-time via a cell-phone. From a retrospective perspective it aids patient-caregiver interaction, from an introspective perspective it enables status monitoring for patients, and from prospective perspective it can answer a patient's question: what are the consequences of walking for an hour?

Further, continuous monitoring enables personalized models, and detection of changes in model parameters per subject, which may aid to either to detect or foresee future heart problems.

VII. REFERENCES

[1] De Godoy MF, Takakura IT, Corrêa PR, Machado MN, Miranda RC, Brandi AC, "Preoperative nonlinear behavior in heart rate variability predicts morbidity and mortality after coronary artery bypass graft surgery", Med Sci Monit. 2009; 45 (3), pp117-122

[2] C. R. Cole, E. H. Blackstone, F. J. Pashkow, C. E. Snader, M. S.Lauer, "Heart rate recovery immediately after exercise as a predictor of mortality," The New England Journal of Medicine, vol. 341, no. 18, pp.1351–1357, 1999.

[3]R. L. Hughson, "Regulation of blood flow at the onset of exercise by feed forward and feedback mechanisms," Canadian Journal of AppliedPhysiology, vol. 28, no. 5, pp. 774–787, 2003.

[4]S. W. Su, L. Wang, B. G. Celler, A. V. Savkin, Y. Guo, "Identification and control for heart rate regulation during treadmill exercise," IEEE Trans. Biomed. Eng., vol. BE-54, pp. 1238–1246, Jul. 2007.

[5]T Cheng, A. Savkin, B. Celler, S. Su, and L. Wang, "Nonlinear Modeling and Control of Human Heart Rate Response During Exercise With Various Load Intensities", IEEE Transactions on Biomedical Engineering, Vol.55, No.11, Nov. 2008

[6]H. D. Critchley, P. Rotshtein,Y. Nagai, J. O'Doherty, C. J. Mathias, R. J. Dolan, "Activity in the human brain predicting differential heart rate responses to emotional facial expressions", pp. 751-762, Neuro Image 24, 2005.

[7]P. J. Mills, J. E. Dimsdale, M. G. Ziegler, C. C. Berry, R.D. Bain, "Beta-Adrenergic Receptors Predict Heart Rate Reactivity to a Psychosocial Stressor", Psychosomatic Medicine 52, 1990. pp. 621-623

[8]Zhanpeng Jin; Yuwen Sun; Cheng, A.C., "Predicting cardiovascular disease from real-time electrocardiographic monitoring: An adaptive machine learning approach on a cell phone", Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, Minneapolis, M, 3-6 September, 2009.

[9]Karaham R., "Prediction of the Behavior of the Human Heart Using Methods of Non-Linear Dynamics", thesis, Texas Tech University, May 1996.

[10]J. Pan and W. Tompkins, "A Real-Time QRS Detection Algorithm", IEEE Transactions on Biomedical Engineering, Vol. BME-32, No. 3, March 1985.