2012 32nd International Conference on Distributed Computing Systems Workshops

# How Hard Am I Training? Using Smart Phones to Estimate Sport Activity Intensity

Igor Pernek
Faculty of Health Sciences
University of Maribor
Email: igor.pernek@uni-mb.si

Gregor Stiglic
Faculty of Health Sciences
University of Maribor
Email: gregor.stiglic@uni-mb.si

Peter Kokol
Faculty of Health Sciences
University of Maribor
Email: kokol@uni-mb.si

***Abstract*—Smart phones are increasingly being used to track and recognize different types of activity. However, the task of using smart phones to infer the intensity of sport activities has not received a lot of attention yet. Therefore, we study how off-the-shelf smart phones with built-in accelerometers can be used to estimate the intensity of recreational sport activities. We focus on finding the most appropriate model along with a set of high level acceleration features that could be used to predict heart rate during a sport activity on a resource constrained smart phone device. We collect more than 300 minutes of acceleration and heart rate data from five subjects playing badminton and evaluate four different numeric prediction models using different combinations of acceleration features in terms of correlation between the actual and predicted heart rate and the heart rate estimation error. The evaluations show that linear regression provides good intensity inference accuracy (correlation coefficient: 0.86; mean absolute error: 15.52 beats per minute) and is, considering its low computational demands, the most feasible to be implemented on a smart phone device.**

## I. Introduction

Regular physical activity is important for maintaining a healthy style of life. Scientific literature suggests that there is irrefutable evidence of the effectiveness of regular physical activity in the prevention of several chronic diseases such as cardiovascular diseases, diabetes, cancer, hypertension, obesity, and osteoporosis [1]. Previous studies have also shown that regular tracking of exercises can motivate people to be more physically active [2] and that smart phones are a viable platform for tracking activity and dietary data [3].

During the last years different solutions have been proposed utilizing off-the-shelf smart phones for physical activity tracking. Most of the previous research focused on using smart phones to recognize type and duration of different daily activities [4], [5]. However, considerably less attention has been devoted to utilizing smart phones to infer exercise intensity. It has been shown that sufficient exercise intensity is important for improving health and fitness [6]. Additionally, different exercise intensities have different health benefits. E.g. moderate intensity training develops basic endurance and aerobic capacity, hard intensity training improves the cardiovascular system, etc. [6], [7]. Therefore, providing information on the intensity of a training is important for optimal training benefits.

Currently, exercise intensity can be measured as a percentage of the Maximum number of Heart Beats per minute [8]. However, heart rate monitors are external wearable sensors that have to be connected to smart phones using a wireless communication protocol such as ANT+ or bluetooth. This makes them inconvenient to use, as they have to be bought separately from the phone. Additionally, not a lot of phones currently support the ANT+ protocol, which additionally limits the choice of heart rate monitors. On the other hand, accelerometers are built into most of the off-the-shelf smart phones and will become even more ubiquitous in the next smart phone generations [9].

Therefore, we investigated if and how smart phones with built-in accelerometers can be used to estimate the intensity of a training session. We did this by trying to model a person's heart rate as a function of different acceleration features during a sport activity. To evaluate how accurately the training intensity could be modelled, we collected heart rate and acceleration data from five individuals playing badminton. We built four different numeric prediction models and compared the results in terms of correlation between the actual and predicted heart rate and the mean absolute error in beats per minute (BPM). Although collected data is badminton specific, a similar approach could be used for other sports such as basketball, volleyball, tennis, squash, etc.

The contributions of the paper are twofold: (1) to evaluate if off-the-shelf smart phones are a feasible platform for estimating sport activity intensity without the use of any additional external sensors (such as heart rate monitors) and (2) to assess the accuracy of different sets of features and numeric prediction models for the task of heart rate prediction.

## II. Data collection and analysis

In this section we describe the data collection and analysis process. Prior to that we provide some information about hard- and software used. Data was collected using a Sony Ericsson Xperia Active Android phone and a Garmin ANT+ heart rate chest monitor. All the data processing and feature extraction was performed in R [10]. For the machine learning tasks University of Waikato's Weka machine learning toolbox was used [11].

### A. Data collection

To collect acceleration and heart rate data first a smart phone sensor capturing application had to be developed. Due to its openness and rapid prototyping capabilities Android was

1545-0678/12 $26.00 © 2012 IEEE
DOI 10.1109/ICDCSW.2012.34

Fig. 1. Smart phone placement position

chosen as our target platform. The application developed was able to connect to an ANT+ heart rate monitor and fetch the computed heart rate (in beats per minute) and the heart beat count with a frequency of 4 Hz. While the computed heart rate was used to provide information about the heart activity and training intensity, the beat count was only used to verify the consistency of the captured signal. Additionally, acceleration in direction of x-, y-, and z-axis ($A_x$, $A_y$, $A_z$) was captured along with the heart rate information. Acceleration values were sampled at the approximate frequency of 80 Hz. All the sensor readings were saved to a comma-separated file to be later transferred to a personal computer for analysis.

Five healthy subjects (four male, one female; 26 years old), all regularly playing badminton on recreational basis, were asked to wear a smart phone and a heart rate monitor during a session of badminton. During the game the phone was put into a Jabra Sport pouch and strapped bellow the knee of the non-dominant leg of the subject (Fig. 1). Such placement was selected as the subjects reported it was less disturbing than pelvis and arm placements. A total of 377.28 minutes of badminton data was collected. On average, we collected $79.49 \pm 5.01$ minutes of data for each subject. However, preliminary analysis of the collected data showed some parts of the heart rate data were inconsistent as some heart rate samples were missing. Presumably, this happened because the heart rate monitor was temporarily misplaced during the game. We therefore first cleaned the data, removing the corrupted heart rate sections. After the data cleaning process, we were left with a total of 304.19 minutes. The average duration of captured data for each subject was $59.89 \pm 10.78$. In the next section we explain the data processing and feature extraction steps.

### B. Feature extraction

Prior to the feature extraction step all acceleration and heart rate data was pre-processed and put into one second frames. We did this to neutralize the effect of irregular accelerometer sampling rate, a common limitation of most of the smart phone architectures. For each second of data arithmetic mean of heart rate and accelerometer data was calculated to get a regularly sampled signal. This also had an effect of low-pass filtering the signal and removing unwanted high frequency spikes. Additionally, we figured users were unlikely to wear the smart phone in exactly the same manner, most probably placing it in the pouch with different orientations. Therefore, total acceleration $A_t$ (Eq. 1) of the phone was obtained and used instead of values from individual axes.

$$A_t = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{1}$$

Features were then extracted from the re-sampled acceleration signal using a sliding window approach. Six different window sizes were used to extract features for different time frames. The window sizes selected were as follows: 60, 50, 40, 30, 20, and 10 frame window size. Due to our preprocessing steps, one frame represented one second of the data. Different window sizes allowed us to capture acceleration activity information for different intervals in the past. However, for all windows sizes one frame overlap was used. This simply meant that along with heart rate information for each second we calculated acceleration features for the last 60, 50, 40, 30, 20, and 10 seconds.

Nine different features were extracted for each window size, given a total of 54 features. We have selected only features, which have already been proven to be useful for accelerometer based activity recognition and can efficiently be calculated online on a smart phone device. In [12] an excellent overview of the possible features for context recognition from accelerometer data, along with their computational costs, is provided. Due to lower computational costs only time domain features were considered. The features extracted were: arithmetic mean of $A_t$, standard deviation of $A_t$, max of $A_t$, min of $A_t$, range of $A_t$, RMS of $A_t$, correlation between $A_x$ and $A_y$, correlation between $A_x$ and $A_z$, correlation between $A_y$ and $A_z$.

However, using all of the 54 features for activity intensity estimation on a smart phone would be too time-consuming and computationally intensive. Thus, we decided to use a feature selection method to identify a subset of features that delivers good intensity prediction rate while reducing the complexity of the overall process. In the section to follow we describe the feature selection method and the protocol used to obtain the optimal number of features.

### C. Feature selection

To select a subset of our original feature space we have used the RReliefF [13] feature selection method. RReliefF is an extension to the Relief method, which searches for several nearest neighbours, is robust to noise, and handles continuous data. The RReliefF feature selection algorithm was chosen as it does not assume the conditional independence of the attributes upon the target variable and is therefore appropriate for problems which involve much feature interaction. We expect our features to be heavily conditionally dependant as some of the window sizes used for feature extraction overlap in great extent. E.g. a major part of the 60 frame window are frames from the 50 frame window, a major part of the 50 frame window are frames from the 40 frame window, etc.

The following protocol was used to select the optimal number of features. We first ranked the features according to their ReliefF value in the descending order. Afterwards, we selected different subsets of features starting with the first two most important features and one by one adding the next feature by importance. Consequently the first subset of features contained the two most important features according to their ReliefF value, the second subset contained the three most important features, etc. Each subset of features was used to perform training of four different numeric prediction models (described in the next section) and to evaluate the prediction accuracy of the model.

### D. Intensity estimation models

Four different machine learning algorithms were used to infer sport activity intensity. We treated the intensity prediction task as a regression problem, predicting the subject's heart rate for each frame using different features. Therefore, the selected machine learning algorithms all had to be able to deal with numerical class data. The following models were selected for numerical prediction of the heart rate data: multilayer perceptron, linear regression, regression tree, and support vector regression. For the sake of completeness this section briefly describes the models used in our experiments.

Multilayer perceptron (MP) [14] is an artificial neural network based algorithm, which uses backpropagation to classify instances. It consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. The nodes are all sigmoid, except for when the class is numeric in which case the output nodes become unthresholded linear units. In our experiments we used a MP algorithm with two hidden layers, each containing three nodes.

Linear regression (LR) is one of the simplest machine learning models. It fits a straight line through a set of points in such way that it makes sum of squared residuals of the model as small as possible. Simple linear regression therefore picks only one parameter that results in the lowest squared error.

To evaluate the regression tree [15] classifier, Weka's REPTree algorithm was used. REPTree is a fast decision tree learner that builds a regression tree using information gain/variance and prunes it using reduced-error pruning. However, in our case, pruning was turned off as we were not as much interested in an interpretable model, as we were interested in a model with high prediction accuracy.

SMOreg [16] algorithm implemented in Weka, was chosen to evaluate the support vector regression approach. SMOreg implements sequential minimal optimization algorithm for training a support vector regression using polynomial or RBF kernels. In our case a Gaussian kernel with $\gamma = 0.01$ was chosen to support non-linear relations in the model.

Before introducing the results it has to be pointed out that no parameter tuning of the induction algorithms was carried out. Therefore, the results obtained are likely slightly underestimated and could most probably be improved by choosing the most appropriate set of parameters for each model.

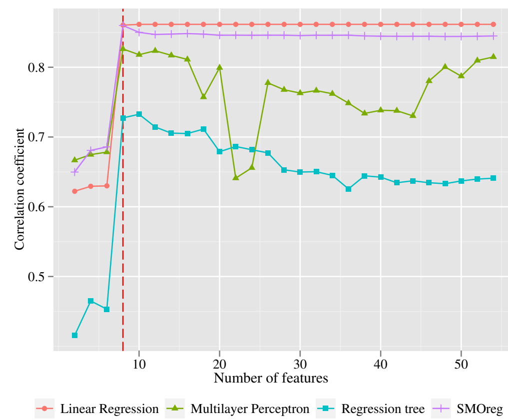


Fig. 2. Correlation coefficient for different models and number of features

## III. Results

To evaluate sport intensity models leave-one-subject-out protocol was used. We therefore created five pairs of different training and test datasets, each training dataset containing badminton data for all users except one, whose data was put into the test dataset. Using such protocol allowed us to always test the models on data different from the data they were trained on. This resulted in more generalized and robust models with less overfitting.

The performance of intensity estimation was evaluated using two different numeric prediction evaluation measures: *correlation coefficient (CC)* and *mean absolute error (MAE)*. CC measures the statistical correlation between actual and predicted values and ranges from 1 for perfectly correlated results, through 0 when there is no correlation, to -1 when the results are perfectly correlated in the negative direction. CC as close to 1 as possible is therefore desired. On the other side, MAE measures the average magnitude of individual errors without taking into account their signs and should therefore be lower for better performance.

Fig. 2 and 3 present the CC and MAE measures for different number of features and different numeric prediction models. They show that the MAE and CC measures for all models considerably improve when the eight feature (60 frame RMS) is added. Additionally, they show that adding more features does not provide significantly better results and in some cases even degrades them (e.g. negative peak in Fig 2 for 22 features using MP). Thus, we decided to interpret the results using only the first eight highest ranking features selected by the ReliefF method. However, as linear regression operates only with one parameter, due to its high predicting value only the 60 frame RMS feature was used in the linear regression model. This is consistent with our goal of finding a model that would be feasible to implement on a resource constrained smart phone device.


Fig. 3. Mean prediction error for different models and number of features

TABLE I
HEART RATE PREDICTION RESULTS FOR 8 FEATURES

| | Correlation coefficient | Mean absolute error |
|---|---|---|
| Linear Regression | $0.86 \pm 0.1$ | $15.52 \pm 3.76$ |
| Multilayer Perceptron | $0.83 \pm 0.1$ | $14.66 \pm 3.77$ |
| REPtree | $0.73 \pm 0.06$ | $19.51 \pm 1.76$ |
| SMOreg | $0.86 \pm 0.1$ | $15.6 \pm 3.73$ |

Table I shows that LR and SMOreg are doing equally good in terms of CC and that MP is not far behind. However, the performance of REPtree is much lower, both, in terms of CC and MAE. Additionally, the tree built by the REPtree algorithm contains a large number of nodes (approximately 1600) and is therefore computationally very expensive and not feasible to implement on a smart phone device. Considering the MAE, again, LR and SMOreg yield similar results, with MP performing slightly better this time.

The results clearly show that LR, MP, and SMOreg all provide good performance on heart rate and consequently training intensity estimation. The CC around 0.85 shows high correlation between actual heart rate and heart rate predicted from accelerometer data. Additionally, the MAE of around 15 beats per minute represents less than 8 percent of the maximal heart rate of the subjects performing the experiments, which is inside the margin of individual hear rate activity zones, and can therefore be used to estimate the training intensity zones during the training (e.g. the aerobic, anaerobic zone, etc.).

However, not all of three algorithms are equally computationally intensive. While SMOreg and MP are more expensive to compute, LR model contains only a simple formula, which could easily be implemented in a smart phone environment. We therefore conclude that LR would be the most feasible to implement on a smart phone device to provide real time training intensity estimation.

## IV. CONCLUSION

In this paper we studied how accelerometer equipped off-the-shelf smart phones could be employed to infer sport activity intensity information. We collected acceleration data along with heart rate information for five test subjects and evaluated different numeric prediction models in terms of correlation between actual and predicted heart rate and mean absolute error in beats per minute. We discovered that acceleration information provides useful information for inferring heart rate activity and consequently training intensity. Additionally, we concluded that a simple linear regression model along with the RMS of acceleration for the last 60 seconds is the most feasible model to be implemented in a constrained smart phone environment.

We are aware that the data collected is biased as all of our testing subjects were younger and well trained. Thus, we would like to extend this work in the future by including additional subjects of different ages and with different levels of physical fitness. We additionally plan to validate how our approach scales to the other types of sport activities.

## REFERENCES


[1] D. E. R. Warburton, C. W. Nicol, and S. S. D. Bredin, "Health benefits of physical activity: the evidence." *Canadian Medical Association Journal*, vol. 174, no. 6, pp. 801–9, Mar. 2006.

[2] E. Schlenk, J. Dunbar-Jacob, S. Sereika, T. Starz, A. Okifuji, and D. Turk, "Comparability of daily diaries and accelerometers in exercise adherence in fibromyalgia syndrome." *Measurement in Physical Education and Exercise Science*, vol. 4, no. 2, pp. 133–4, 2000.

[3] M. B. Conroy, K. Yang, O. U. Elci, K. P. Gabriel, M. A. Styn, J. Wang, A. M. Kriska, S. M. Sereika, and L. E. Burke, "Physical Activity Self-Monitoring and Weight Loss: 6-Month Results of the SMART Trial," *Medicine & Science in Sports & Exercise*, vol. 43, no. 8, pp. 1568–1574, 2011.

[4] I. Anderson, J. Maitland, S. Sherwood, L. Barkhuus, M. Chalmers, M. Hall, B. Brown, and H. Muller, "Shakra: Tracking and Sharing Daily Activity Levels with Unaugmented Mobile Phones," *Mobile Networks and Applications*, vol. 12, no. 2, pp. 185–199, Jun. 2007.

[5] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*. Springer, 2004, pp. 1–17.

[6] M. Pollock, G. Gaesser, and J. Butcher, "The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults," *Medicine & Science in Sports & Exercise*, vol. 30, no. 6, pp. 975–91, 1998.

[7] S. Edwards, *The Heart Rate Monitor Book*. Polar Electroinc, 1993.

[8] J. Achten and A. E. Jeukendrup, "Heart rate monitoring: applications and limitations," *Sports Medicine*, vol. 33, pp. 517–38, 2003.

[9] R. Cozza, C. Milanesi, and T. Nguyen, "Market Share Analysis: Mobile Devices, Worldwide," Gartner Group Research Report, Tech. Rep., 2011.

[10] R. D. C. Team, "R: A language and environment for statistical computing," Vienna, Austria, 2001. [Online]. Available: http://www.r-project.org/

[11] M. Hall, F. Eibe, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–8, 2009.

[12] D. Figo, P. C. Diniz, D. R. Ferreira, and J. a. M. P. Cardoso, "Preprocessing techniques for context recognition from accelerometer data," *Personal and Ubiquitous Computing*, vol. 14, no. 7, pp. 645–62, Mar. 2010.

[13] M. Robnik-Sikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Machine learning*, vol. 53, no. 1, pp. 23–69, 2003.

[14] S. K. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification." *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, vol. 3, no. 5, pp. 683–97, Jan. 1992.

[15] A. Karalič, "Employing linear regression in regression tree leaves," in *10th European conference on Artificial intelligence*, Aug. 1992, pp. 440–441.

[16] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, Aug. 2004.